



MPSI

Exercices d'algèbre et de probabilités

David Delaunay



RÉSUMÉS DE COURS



MÉTHODES



3 NIVEAUX D'EXERCICES :

- apprentissage
- entraînement
- approfondissement



CORRIGÉS DÉTAILLÉS

PAS À PAS

Exercices d'algèbre et de probabilités

Collection Prépas scientifiques

Dirigée par Olivier Rodot

C. ANTONINI, Algèbre MP/MP*

N. BASBOIS et P. ABBRUGIATI, Algèbre MPSI/PCSI, 2^e édition

G. COSTANTINI, Analyse MPSI/PCSI, 2^e édition

K. DAO DUC et D. DELAUNAY, Probabilités

D. DELAUNAY, Exercices d'analyse MP/MP*

D. DELAUNAY, Exercices d'analyse MPSI

D. DELAUNAY, Exercices d'algèbre et de probabilités MP/MP*

D. DELAUNAY, Exercices d'algèbre et de probabilités MPSI

O. RODOT, Analyse MP/MP*

Chez le même éditeur

T. RIBEYRE, Chimie PC/PC*

M.-A. SCHOTT, J. VALENTIN, G. MAGADUR, S. CLÈDE, A.-L. LEFEVRE,

A. ALTMAYER-HENZIEN, Chimie PCSI/MPSI



MPSI

Exercices d'algèbre et de probabilités

David Delaunay



RÉSUMÉS DE COURS

MÉTHODES

3 NIVEAUX D'EXERCICES :

- apprentissage
- entraînement
- approfondissement

CORRIGÉS DÉTAILLÉS
PAS À PAS

Pour toute information sur notre fonds et les nouveautés dans votre domaine de spécialisation, consultez notre site web : **www.deboecksuperieur.com**

© De Boeck Supérieur s.a., 2017
Rue du Bosquet, 7 B-1348 Louvain-la-Neuve

1^{ère} édition, 2017
1^{er} tirage, 2017

Tous droits réservés pour tous pays.

Il est interdit, sauf accord préalable et écrit de l'éditeur, de reproduire (notamment par photocopie) partiellement ou totalement le présent ouvrage, de le stocker dans une banque de données ou de le communiquer au public, sous quelque forme et de quelque manière que ce soit.

Imprimé aux Pays-Bas.

Dépôt légal :

Dépôt légal France : juin 2017

Dépôt légal Belgique : 2017/13647/091

ISBN : 978-2-8073-0613-4

La pratique d'exercices est essentielle à l'apprentissage du cours de mathématiques : il n'est pas de meilleure façon de mémoriser et de comprendre un théorème que d'en faire usage !

Cet ouvrage regroupe sur 13 chapitres 401 exercices portant sur le programme d'algèbre et de probabilités en classe de MPSI. Il respecte strictement le programme en cours et vient compléter l'ouvrage d'*analyse* que l'on retrouvera dans la même collection.

Chaque chapitre commence par un rappel des principales définitions et des résultats essentiels du cours. Il se poursuit avec des exercices aux corrigés détaillés regroupés sur trois niveaux :

- *Les exercices d'apprentissage* servent à l'acquisition des concepts fondamentaux du cours. Ce sont souvent des sujets faciles où j'ai choisi volontairement de ne faire figurer que peu de technicité.
- *Les exercices d'entraînement* permettent de poursuivre l'acquisition du cours, trois niveaux d'étoiles servent à anticiper leur difficulté. Ces sujets ont été choisis pour leur intérêt, leur classicisme ou ont été inspirés par des questions rencontrées aux écrits et aux oraux des différents concours.
- *Les exercices d'approfondissement* sont les plus ambitieux, ils nécessitent souvent de passer par une phase de recherche ou entrent en résonance avec d'autres chapitres du programme. Ces sujets sont inspirés de questions rencontrées aux concours les plus ambitieux.

Les corrections des exercices sont accompagnées de *méthodes*. Celles-ci servent à souligner les idées récurrentes ou bien à mettre en exergue la démarche qui va être suivie pour résoudre la question posée. Le lecteur pourra prendre appui sur celles-ci pour amorcer une résolution ou pour reprendre la main lors de sa lecture d'une correction. Afin d'aider le lecteur dans son étude, il est fait référence aux théorèmes utilisés lors de leurs premiers usages. Les notes de bas de pages complètent les résolutions en présentant des démarches alternatives ou font le lien avec d'autres sujets présents dans l'ouvrage.

Je remercie vivement Olivier RODOT d'avoir initié ce projet, François PANTIGNY pour son expertise TeXnique et Pierrick SOLEILLANT pour sa relecture attentive ainsi que les corrections apportées.

Je dédicace cet ouvrage à ma fille Libi.

David DELAUNAY

CHAPITRE 1

Ensembles et applications

1.1 Rudiments de logique

Définition

|| Une *assertion* est une phrase mathématique syntaxiquement correcte.

Dans un cadre axiomatique¹ donné, une assertion est susceptible d'être vraie ou fausse.

Dans ce qui suit, \mathcal{P} , \mathcal{Q} et \mathcal{R} désignent des assertions.

1.1.1 Négation

Définition

|| On appelle *négation* d'une assertion \mathcal{P} , l'assertion « $\text{non}(\mathcal{P})$ » définie comme étant vraie lorsque \mathcal{P} est fausse et inversement.

Les assertions \mathcal{P} et $\text{non}(\text{non}(\mathcal{P}))$ ont mêmes valeurs de vérité.

1.1.2 Conjonction et disjonction

Définition

|| On appelle *conjonction* de deux assertions \mathcal{P} et \mathcal{Q} , l'assertion « \mathcal{P} et \mathcal{Q} » définie comme étant vraie si \mathcal{P} et \mathcal{Q} le sont toutes les deux et fausse lorsqu'au moins l'une des deux assertions l'est.

1. Le cadre axiomatique usuel est celui de Zermelo et Fraenkel avec axiome du Choix (ZFC). Celui-ci est *présumé* non contradictoire ce qui signifie l'absence d'assertions à la fois vraies et fausses... .

Définition

On appelle *disjonction*¹ de deux assertions P et Q , l'assertion « P ou Q » définie comme étant vraie lorsqu'au moins l'une des deux assertions P ou Q est vraie et fausse lorsque les deux assertions le sont.

En notant $P \sim Q$ pour signifier que les deux assertions P et Q ont les mêmes valeurs de vérité, on vérifie les propriétés :

Idempotences :

$$(P \text{ et } P) \sim P \text{ ainsi que } (P \text{ ou } P) \sim P;$$

Commutativités :

$$(P \text{ et } Q) \sim (Q \text{ et } P) \text{ ainsi que } (P \text{ ou } Q) \sim (Q \text{ ou } P);$$

Associativités :

$$(P \text{ et } (Q \text{ et } R)) \sim ((P \text{ et } Q) \text{ et } R) \text{ et l'on note simplement } (P \text{ et } Q \text{ et } R),$$

$$(P \text{ ou } (Q \text{ ou } R)) \sim ((P \text{ ou } Q) \text{ ou } R) \text{ et l'on note simplement } (P \text{ ou } Q \text{ ou } R);$$

— *Distributivités*² :

$$(P \text{ et } (Q \text{ ou } R)) \sim ((P \text{ et } Q) \text{ ou } (P \text{ et } R)),$$

$$(P \text{ ou } (Q \text{ et } R)) \sim ((P \text{ ou } Q) \text{ et } (P \text{ ou } R));$$

— *Lois de Morgan* :

$$(\text{non}(P \text{ et } Q)) \sim (\text{non}(P) \text{ ou } \text{non}(Q)),$$

$$(\text{non}(P \text{ ou } Q)) \sim (\text{non}(P) \text{ et } \text{non}(Q)).$$

1.1.3 Implications et équivalences**Définition**

On définit l'assertion d'*implication* « $P \Rightarrow Q$ » comme étant vraie lorsque la véracité de P impose la véracité de Q . Plus précisément, l'assertion « $P \Rightarrow Q$ » est vraie lorsque P est vraie et Q vraie ou bien lorsque P est fausse.

L'assertion « $P \Rightarrow Q$ » n'est fausse que lorsque P est vraie et Q fausse. On vérifie que les assertions « $P \Rightarrow Q$ » et « $\text{non}(P) \text{ ou } Q$ » ont mêmes valeurs de vérité.

Théorème 1 (Contraposition)

$$(P \Rightarrow Q) \sim (\text{non}(Q) \Rightarrow \text{non}(P)).$$

Théorème 2 (Négation d'une implication)

$$(\text{non}(P \Rightarrow Q)) \sim (P \text{ et } \text{non}(Q)).$$

1. Le « ou » du langage commun est souvent *exclusif* comme dans l'expression « fromage ou dessert ». Le « ou » mathématique est *inclusif*.

2. On évitera d'écrire $(P \text{ et } Q \text{ ou } R)$ car l'interprétation n'en est pas claire : doit-on comprendre $((P \text{ et } Q) \text{ ou } R)$ ou $(P \text{ et } (Q \text{ ou } R))$?

Définition

On définit l'assertion d'*équivalence* « $\mathcal{P} \Leftrightarrow \mathcal{Q}$ » comme étant vraie lorsque les assertions \mathcal{P} et \mathcal{Q} ont mêmes valeurs de vérité.

Théorème 3 (Double implication)

$$(\mathcal{P} \Leftrightarrow \mathcal{Q}) \sim ((\mathcal{P} \Rightarrow \mathcal{Q}) \text{ et } (\mathcal{Q} \Rightarrow \mathcal{P})).$$

1.1.4 Quantificateurs

Lorsque la valeur de vérité d'une assertion \mathcal{P} dépend d'un paramètre x , cette assertion peut être notée $\mathcal{P}(x)$ afin de souligner cette dépendance.

Définition

On définit l'assertion de *quantification universelle* « $\forall x \in E, \mathcal{P}(x)$ » comme étant vraie lorsque l'assertion $\mathcal{P}(x)$ est vraie pour tout élément x de l'ensemble E et fausse sinon.

On définit l'assertion de *quantification existentielle* « $\exists x \in E, \mathcal{P}(x)$ » comme étant vraie lorsque l'assertion $\mathcal{P}(x)$ est vraie pour au moins un élément x de E et fausse sinon.

Les symboles \forall et \exists se lisent « quel que soit » et « il existe ». On définit aussi l'assertion « $\exists! x \in E, \mathcal{P}(x)$ » comme étant vraie lorsque l'assertion $\mathcal{P}(x)$ est vérifiée pour un élément x de E et un seul :

$$(\exists! x \in E, \mathcal{P}(x)) \sim \left(\exists x \in E, \mathcal{P}(x) \text{ et } (\forall x \in E, \forall x' \in E, (\mathcal{P}(x) \text{ et } \mathcal{P}(x')) \Rightarrow x = x') \right)$$

Dans les phrases quantifiées précédentes, la lettre x revêt un rôle muet.

Théorème 4 (Négation d'une phrase quantifiée)

$$\text{non}(\forall x \in E, \mathcal{P}(x)) \sim \exists x \in E, \text{non}(\mathcal{P}(x))$$

$$\text{non}(\exists x \in E, \mathcal{P}(x)) \sim \forall x \in E, \text{non}(\mathcal{P}(x)).$$

Toute assertion commençant par « $\exists x \in \emptyset$ » est assurément fausse, peu importe ce qui suit. Par négation, toute assertion débutant par « $\forall x \in \emptyset$ » est vraie.

1.1.5 Raisonnements

On peut vérifier une assertion \mathcal{P} :

Par *déduction*¹ :

On détermine une assertion vraie \mathcal{Q} telle que « $\mathcal{Q} \Rightarrow \mathcal{P}$ » soit vraie.

— Par *disjonction de cas* :

On détermine une assertion \mathcal{Q} telle que « $\mathcal{Q} \Rightarrow \mathcal{P}$ » et « $\text{non}(\mathcal{Q}) \Rightarrow \mathcal{P}$ » soient toutes les deux vraies.

1. Ou *modus ponens*.

— Par l'*absurde* :

On détermine une assertion Q fausse telle que « $\text{non}(\mathcal{P}) \Rightarrow Q$ » soit vraie.

On peut vérifier une implication « $\mathcal{P} \Rightarrow Q$ » :

— Par *enchaînement d'implications* :

On détermine une assertion R telle que « $\mathcal{P} \Rightarrow R$ » et « $R \Rightarrow Q$ » soient toutes les deux vraies (et l'on peut enchaîner les assertions intermédiaires).

— Par *contraposition* :

On établit « $\text{non}(Q) \Rightarrow \text{non}(\mathcal{P})$ ».

1.2 Ensembles

On appelle *ensemble* toute collection d'objets appelés *éléments* de cet ensemble. Pour signifier l'appartenance d'un élément x à un ensemble E , on écrit $x \in E$.

Deux ensembles sont dit *égaux* lorsqu'ils sont constitués des mêmes éléments.

L'ensemble ne contenant aucun élément est appelé *ensemble vide* et est noté¹ \emptyset .

1.2.1 Inclusion

Définition

On dit qu'un ensemble F est *inclus* dans un ensemble E si tous les éléments de F sont éléments de E . On note alors $F \subset E$.

Théorème 5 (Double inclusion)

Si E est inclus dans F et F inclus dans E , les ensembles E et F sont égaux.

Définition

Tout ensemble F inclus dans un ensemble E est appelé *sous-ensemble* de E . Plus communément, on dit que F est une *partie* de E .

Les parties de E constituent un ensemble noté $\wp(E)$ et appelé *ensemble des parties* de E . Si $\mathcal{P}(x)$ est une assertion dépendant d'un paramètre x , on note²

$$\{x \in E \mid \mathcal{P}(x)\}$$

le sous-ensemble constitué des éléments de E qui vérifient l'assertion $\mathcal{P}(x)$.

1.2.2 Opérations sur les parties d'un ensemble

Soit A , B et C trois parties d'un ensemble E .

1. La notation $\{\}$ doit être considérée comme caduque tandis que la notation $\{\emptyset\}$ ne décrit pas l'ensemble vide mais un ensemble à un élément qui est l'ensemble vide...

2. On dit que cette partie est définie en *compréhension* : ses éléments ne sont pas explicitement décrits mais sont déterminés par la vérification d'une propriété.

Définition

On appelle *complémentaire* de la partie A de E l'ensemble noté $\complement_E A$ formé des éléments de E qui ne sont pas dans A :

$$\complement_E A \stackrel{\text{def}}{=} \{x \in E \mid x \notin A\}.$$

S'il n'y a pas d'ambiguïté sur l'ensemble E contenant la partie A étudiée, on emploie aussi la notation \bar{A} pour désigner le complémentaire de A dans E .

On vérifie $\complement_E(\complement_E A) = A$ (ou encore $\bar{\bar{A}} = A$) ainsi que le renversement des inclusions par passage au complémentaire :

Théorème 6

$$A \subset B \implies \complement_E B \subset \complement_E A \quad \text{c'est-à-dire} \quad \bar{B} \subset \bar{A}.$$

Définition

On appelle *union* des parties A et B l'ensemble noté $A \cup B$ formé des éléments de E qui appartiennent à au moins l'une des deux parties :

$$A \cup B \stackrel{\text{def}}{=} \{x \in E \mid x \in A \text{ ou } x \in B\}.$$

Les parties A et B sont toutes deux incluses dans l'union $A \cup B$. Aussi, $A \cup B$ est inclus dans toute partie C qui contient à la fois les parties A et B .

Définition

On appelle *intersection* des parties A et B l'ensemble noté $A \cap B$ formé des éléments de E qui appartiennent aux deux parties :

$$A \cap B \stackrel{\text{def}}{=} \{x \in E \mid x \in A \text{ et } x \in B\}.$$

L'intersection $A \cap B$ est incluse dans chacune des deux parties A et B . Aussi, $A \cap B$ contient toute partie C à la fois incluse dans A et dans B .

Lorsque l'intersection des parties A et B est vide, celles-ci sont dites *disjointes*.

Par conjonctions, disjonctions et négations, on vérifie les propriétés suivantes :

Idempotences : $A \cup A = A$ et $A \cap A = A$;

- *Commutativités* : $A \cup B = B \cup A$ et $A \cap B = B \cap A$;

Neutralités : $A \cup \emptyset = A$ et $A \cap E = A$;

*Associativités*¹ : $A \cup (B \cup C) = (A \cup B) \cup C$ et $A \cap (B \cap C) = (A \cap B) \cap C$;

- *Distributivités* : $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ et $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

- *Lois de Morgan*² : $\complement_E(A \cup B) = \complement_E A \cap \complement_E B$ et $\complement_E(A \cap B) = \complement_E A \cup \complement_E B$.

1. Par ces propriétés d'associativité, il est légitime d'écrire $A \cup B \cup C$ et $A \cap B \cap C$ sans préciser de parenthèses organisant le calcul. En revanche, écrire $A \cap B \cup C$ est ambigu et doit être évité.

2. On peut aussi énoncer $\overline{A \cap B} = \overline{A} \cup \overline{B}$ et $\overline{A \cup B} = \overline{A} \cap \overline{B}$.

Définition

On définit l'ensemble *différence* $A \setminus B$ comme étant constitué des éléments de E qui sont dans A sans être dans B .

On observe $A \setminus B = A \cap \complement_E B = A \cap \overline{B}$.

1.2.3 Produit cartésien

Etant donné deux éléments a et b , on définit l'objet *couple* (a, b) de sorte que deux couples (a, b) et (a', b') sont égaux si, et seulement si, $a = a'$ et $b = b'$.

Définition

On appelle *produit cartésien* de deux ensembles E et F , l'ensemble $E \times F$ formé des couples (a, b) pour a parcourant E et b parcourant F .

Lorsque $E = F$, on note plus légèrement E^2 au lieu de $E \times E$.

Si E , F et G désignent trois ensembles, on note $E \times F \times G$ l'ensemble $(E \times F) \times G$. Ses éléments se nomment des *triplets* et se visualisent sous la forme (a, b, c) avec a dans E , b dans F et c dans G .

Plus généralement, on définit par récurrence le produit cartésien de $n + 1$ ensembles à partir du produit cartésien de n ensembles :

$$E_1 \times \cdots \times E_n \times E_{n+1} \stackrel{\text{def}}{=} (E_1 \times \cdots \times E_n) \times E_{n+1}.$$

Un élément de l'ensemble $E_1 \times \cdots \times E_n$ se nomme un *n-uplet*. On parle aussi de *multiplet*. Lorsque les ensembles E_1, \dots, E_n sont égaux, on note E^n au lieu de $E_1 \times \cdots \times E_n$.

1.3 Applications

Soit E , F et G des ensembles.

1.3.1 Définition**Définition**

On appelle *application* (ou *fonction*) f au départ de E et à valeurs dans F l'association à chaque élément x de E d'un unique élément y de F . L'élément y est appelé *valeur prise* par l'application f sur l'élément x , on le note $f(x)$.

On écrit $f: E \rightarrow F$ pour signifier que f est une application de E vers F . On note $\mathcal{F}(E, F)$ (ou F^E) l'ensemble constitué des applications de E vers F .

Deux applications f et g au départ de E et à valeurs dans F sont dites *égales* lorsqu'elles prennent les mêmes valeurs en tout point, c'est-à-dire

$$f(x) = g(x) \quad \text{pour tout } x \in E.$$

Il est commun de définir une application f en précisant une « démarche calculatoire » permettant de déterminer sans ambiguïté la valeur prise par f sur chaque élément de

l'ensemble de départ. Par exemple, on définit l'application *identité* de E par

$$\text{Id}_E : \begin{cases} E \rightarrow E \\ x \mapsto x. \end{cases}$$

Aussi, si A désigne une partie de E , on définit la *fonction indicatrice* de A comme étant l'application $\mathbf{1}_A : E \rightarrow \{0, 1\}$ déterminée par

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{sinon.} \end{cases}$$

Lorsqu'une application $f : E \rightarrow F$ est définie par une formule du type $x \mapsto f(x)$, étudier sa bonne définition consiste en la résolution des deux problèmes suivants :

- vérifier que pour chaque valeur de x dans E , le « calcul » de $f(x)$ est possible ;
- vérifier que la valeur obtenue par ce calcul est élément de F .

1.3.2 Prolongements et restrictions

Soit une application $f : E \rightarrow F$ et des ensembles \bar{E} et \bar{F} vérifiant $E \subset \bar{E}$ et $F \subset \bar{F}$.

Définition

On dit qu'une application $\bar{f} : \bar{E} \rightarrow \bar{F}$ prolonge f si elle vérifie :

$$\bar{f}(x) = f(x) \quad \text{pour tout } x \in E.$$

Soit une application $f : E \rightarrow F$ et des parties A et B incluses dans E et F respectivement vérifiant

$$f(x) \in B \quad \text{pour tout } x \in A. \tag{*}$$

Définition

On définit la *restriction* de f au départ de A et à valeurs dans B comme étant l'application $f|_A^B : A \rightarrow B$ déterminée par

$$f|_A^B(x) = f(x) \quad \text{pour tout } x \in A.$$

La condition (*) est nécessaire à la bonne définition de la restriction. Lorsque $B = F$, cette condition est évidemment respectée et la restriction simplement notée $f|_A$.

1.3.3 Composition

Soit deux applications $f : E \rightarrow F$ et $g : F \rightarrow G$.

Définition

On appelle *composée* de g par f l'application $g \circ f$ de E vers G déterminée par

$$(g \circ f)(x) =^{\text{f}} g(f(x)) \quad \text{pour tout } x \in E.$$

On vérifie aisément $\text{Id}_F \circ f = f$ et $f \circ \text{Id}_E = f$. Au surplus, si h désigne une application de G vers un ensemble H , on a la propriété d'associativité $(h \circ g) \circ f = h \circ (g \circ f)$. Par celle-ci, on peut écrire $h \circ g \circ f$ sans avoir la nécessité de préciser le parenthésage organisant le calcul.

1.3.4 Familles

Soit I un ensemble et une application $a: I \rightarrow E$.

Définition

En notant a_i la valeur $a(i)$ pour chaque indice i dans I , l'application a est appelée *famille d'éléments de E indexée par I* et est notée $(a_i)_{i \in I}$.

La notion de famille ne se distingue pas de la notion d'application. On parle de famille plutôt que d'application lorsque ce sont les valeurs prises a_i qui nous intéressent plus que l'association $i \mapsto a(i)$.

On peut ainsi introduire la famille $(2k)_{k \in \mathbb{N}}$ des entiers pairs, la famille $(p_n)_{n \in \mathbb{N}^*}$ des nombres premiers (en convenant de noter p_n le n -ième nombre premier), etc.

Si $(A_i)_{i \in I}$ désigne une famille de parties de E , on définit l'*union* et l'*intersection* des éléments de cette famille par

$$\bigcup_{i \in I} A_i \stackrel{\text{def}}{=} \{x \in E \mid \exists i \in I, x \in A_i\} \quad \text{et} \quad \bigcap_{i \in I} A_i \stackrel{\text{def}}{=} \{x \in E \mid \forall i \in I, x \in A_i\}.$$

1.3.5 Injection, surjection, bijection

Définition

On dit qu'une application $f: E \rightarrow F$ est *injective* lorsque celle-ci ne prend jamais deux fois la même valeur :

$$\forall (x, x') \in E^2, \quad x \neq x' \implies f(x) \neq f(x').$$

Par contraposition, f est injective lorsque

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

Par passage à la négation, f n'est pas injective si

$$\exists (x, x') \in E^2, \quad f(x) = f(x') \text{ et } x \neq x'.$$

Définition

On dit qu'une application $f: E \rightarrow F$ est *surjective*¹ lorsque celle-ci prend toutes les valeurs de son ensemble d'arrivée :

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

Par passage à la négation, f n'est pas surjective si

$$\exists y \in F, \forall x \in E, \quad f(x) \neq y.$$

1. La notion de surjectivité est intimement liée aux ensembles entre lesquels l'application opère : on ne peut qualifier une application $x \mapsto f(x)$ de surjective sans préciser quels sont les ensembles de départ et d'arrivée.

Définition

On dit qu'une application $f: E \rightarrow F$ est *bijective*¹ si f prend toutes les valeurs de son ensemble d'arrivée une fois et une seule :

$$\forall y \in F, \exists !x \in E, f(x) = y.$$

On peut alors introduire son *application réciproque* $f^{-1}: F \rightarrow E$ définie de sorte que

$$\forall (x, y) \in E \times F, y = f(x) \iff x = f^{-1}(y).$$

Théorème 7

Soit une application $f: E \rightarrow F$. On a équivalence entre :

- (i) f est bijective ;
- (ii) f est injective et surjective ;
- (iii) il existe une application $g: F \rightarrow E$ vérifiant $g \circ f = \text{Id}_E$ et $f \circ g = \text{Id}_F$.

De plus, si tel est le cas, g est l'application réciproque de f .

Enfin, on peut énoncer les résultats de composition :

Théorème 8

Soit deux applications $f: E \rightarrow F$ et $g: F \rightarrow G$.

Si f et g sont injectives, la composée $g \circ f$ l'est aussi.

Si f et g sont surjectives, la composée $g \circ f$ l'est aussi.

Si f et g sont bijectives, la composée $g \circ f$ l'est aussi et $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

1.3.6 Images directes et images réciproques

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

Définition

On appelle *image directe* par l'application f d'une partie A de E , le sous-ensemble de F constitué des éléments possédant un antécédent dans A :

$$f(A) = \{y \in F \mid \exists x \in A, f(x) = y\}.$$

Plus légèrement, l'ensemble $f(A)$ est constitué des valeurs prises par f sur les éléments de A . On écrit :

$$f(A) = \{f(x) \mid x \in A\}.$$

En particulier, $f(E)$ désigne l'ensemble des valeurs prises par f . Celui-ci s'appelle l'*image* de f et est noté $\text{Im}(f)$.

1. Ici aussi il est essentiel de spécifier entre quels ensembles l'application opère lorsque l'on affirme celle-ci bijective.

Définition

On appelle *image réciproque* par l'application f d'une partie B de F , l'ensemble constitué des éléments de E dont l'image est dans B :

$$f^{-1}(B) \stackrel{\text{def}}{=} \{x \in E \mid f(x) \in B\}.$$

L'emploi de la notation f^{-1} dans l'écriture $f^{-1}(B)$ ne présume pas que l'application soit bijective¹.

1.4 Relations binaires

Soit E un ensemble.

Définition

On appelle *relation binaire* \mathcal{R} sur l'ensemble E toute propriété vérifiée par certains couples (x, y) d'éléments de E et fausse pour les autres.

Lorsqu'un couple (x, y) vérifie la relation \mathcal{R} , on écrit $x \mathcal{R} y$. On écrit $x \not\mathcal{R} y$ sinon.

1.4.1 Propriétés remarquables

Les relations binaires sur un ensemble E peuvent être nombreuses et variées. Parmi toutes celles-ci, les plus intéressantes vérifient une ou plusieurs des propriétés qui suivent.

Définition

Soit \mathcal{R} une relation binaire sur l'ensemble E . On dit que

- \mathcal{R} est *réflexive* lorsque : $\forall x \in E, x \mathcal{R} x$;
- \mathcal{R} est *symétrique* lorsque : $\forall (x, y) \in E^2, x \mathcal{R} y \implies y \mathcal{R} x$;
- \mathcal{R} est *antisymétrique* lorsque : $\forall (x, y) \in E^2, (x \mathcal{R} y \text{ et } y \mathcal{R} x) \implies x = y$;
- \mathcal{R} est *transitive* lorsque : $\forall (x, y, z) \in E^3, (x \mathcal{R} y \text{ et } y \mathcal{R} z) \implies x \mathcal{R} z$.

Seule l'égalité vérifie ces quatre propriétés.

1.4.2 Relation d'ordre

Définition

On appelle *relation d'ordre* toute relation binaire à la fois réflexive, antisymétrique et transitive.

La relation \leq est une relation d'ordre² sur la droite réelle \mathbb{R} .

L'inclusion sur $\wp(E)$ ou la divisibilité dans \mathbb{N}^* sont aussi des relations d'ordre.

1. Si f est bijective, la notation $f^{-1}(B)$ devient ambiguë : s'agit-il d'une image réciproque par f ou d'une image directe par f^{-1} ? En fait, ces deux interprétations se confondent dans ce cas.

2. La relation \geq est aussi une relation d'ordre mais pas la relation $<$ car non réflexive : cette dernière s'appelle un *ordre strict*.

Définition

Une relation d'ordre \mathcal{R} sur un ensemble E est qualifiée de *totale* lorsque tous les éléments de E sont deux à deux *comparables* :

$$\forall (x, y) \in E^2, \quad x \mathcal{R} y \quad \text{ou} \quad y \mathcal{R} x.$$

Si non l'ordre est dit n'être que *partiel*.

La relation \leq sur la droite réelle est une relation d'ordre totale. La relation de divisibilité sur \mathbb{N}^* est une relation d'ordre partielle : les entiers 2 et 3 ne sont pas comparables au sens de la divisibilité puisque aucun ne divise l'autre.

1.4.3 Relation d'équivalence**Définition**

On appelle *relation d'équivalence* toute relation binaire à la fois réflexive, symétrique et transitive.

Une relation d'équivalence se comprend souvent comme une égalité *modulo* certains critères. En réunissant entre eux les éléments qui se correspondent pour une relation d'équivalence donnée, on définit le concept suivant :

Définition

Soit \mathcal{R} une relation d'équivalence sur l'ensemble E . On appelle *classe d'équivalence* d'un élément x de E pour la relation \mathcal{R} la partie $\text{Cl}(x)$ formée des éléments qui sont en relation avec x :

$$\text{Cl}(x) = \{y \in E \mid x \mathcal{R} y\}.$$

Les classes d'équivalence sont souvent notées \bar{x} , \hat{x} , \tilde{x} , etc.

Théorème 9

Si \mathcal{R} est une relation d'équivalence sur l'ensemble E alors

- a) $\forall x \in E, x \in \text{Cl}(x)$;
- b) $\forall (x, y) \in E^2, x \mathcal{R} y \implies \text{Cl}(x) = \text{Cl}(y)$;
- c) $\forall (x, y) \in E^2, x \not\mathcal{R} y \implies \text{Cl}(x) \cap \text{Cl}(y) = \emptyset$.

Ainsi, les classes d'équivalence ne sont jamais vides et deux classes d'équivalence distinctes sont disjointes. De plus, tout élément y d'une classe d'équivalence $\text{Cl}(x)$ détermine entièrement celle-ci puisque $\text{Cl}(y) = \text{Cl}(x)$: on dit que les éléments d'une classe d'équivalence sont des *représentants* de celle-ci.

1.5 Exercices d'apprentissage

E et F désignent des ensembles.

1.5.1 Phrases quantifiées

Exercice 1

Soit f une fonction de \mathbb{R} vers \mathbb{R} . Que signifient les phrases quantifiées suivantes ?

- (a) $\exists M \in \mathbb{R}, \forall x \in \mathbb{R}, f(x) \leq M$.
- (b) $\forall M \in \mathbb{R}, \exists x \in \mathbb{R}, f(x) > M$.
- (c) $\forall x \in \mathbb{R}, x > 0 \implies f(x) > 0$.
- (d) $\forall x \in \mathbb{R}, f(x) = 0 \implies x = 0$.
- (e) $\exists A \in \mathbb{R}, \exists C \in \mathbb{R}, \forall x \in \mathbb{R}, x \geq A \implies f(x) = C$.

Solution

méthode

On comprend une phrase quantifiée en lisant celle-ci de la droite vers la gauche et en interprétant successivement chaque quantificateur.

- (a) La portion « $\forall x \in \mathbb{R}, f(x) \leq M$ » signifie que la fonction f est majorée par M . La phrase complète se comprend :

« La fonction f est majorée¹ »

- (b) La portion « $\exists x \in \mathbb{R}, f(x) > M$ » signifie que la fonction f prend une valeur au moins égale à M . Ceci valant pour tout M , la phrase complète peut s'exprimer :

« La fonction f prend des valeurs arbitrairement grandes. »

- (c) L'implication « $x > 0 \implies f(x) \geq 0$ signifie que, si x est positif², la valeur $f(x)$ est aussi positive. La phrase complète s'exprime

« La fonction f prend des valeurs positives sur \mathbb{R}_+ . »

- (d) L'implication « $f(x) = 0 \implies x = 0$ » signifie que si f s'annule en x alors x est nul. Ceci valant pour n'importe quel x , la phrase complète signifie :

« La fonction f ne peut³ s'annuler qu'en 0. »

- (e) La portion « $\forall x \in \mathbb{R}, x \geq A \implies f(x) = C$ » signifie que f est constante égale à C au delà de A . La portion « $\exists C \in \mathbb{R}, \forall x \in \mathbb{R}, x \geq A \implies f(x) = C$ » affirme que f est constante au delà de A . La phrase complète se comprend :

« Pour des valeurs assez grandes de la variable⁴, la fonction f est constante. »

1. Les paramètres M et x exprimant cette phrase quantifiée sont muets et ne doivent donc pas apparaître dans l'expression de l'interprétation de celle-ci.

2. Sans plus de précision, *positif* se comprend au sens large : « positif ou nul ».

3. Ce qui ne signifie pas pour autant que la fonction s'annule en 0.

4. Plus légèrement, on dira que la fonction f est constante au voisinage de $+\infty$.

Exercice 2

Soit $f : E \rightarrow F$ une application. Dans chaque cas, donner la différence de sens entre les deux assertions proposées ?

(a) « $\forall x \in E, \exists y \in F, y = f(x)$ » et « $\exists y \in F, \forall x \in E, y = f(x)$ ».

(b) « $\forall y \in F, \exists x \in E, y = f(x)$ » et « $\exists x \in E, \forall y \in F, y = f(x)$ ».

Soit $\mathcal{P}(x, y)$ une assertion dépendant d'un couple (x, y) élément de $E \times F$.

(c) Laquelle des deux assertions suivantes entraîne l'autre ?

$$\text{« } \exists x \in E, \forall y \in F, \mathcal{P}(x, y) \text{ » et « } \forall y \in F, \exists x \in E, \mathcal{P}(x, y) \text{ ».}$$

Pourquoi ne sont-elles généralement pas équivalentes ?

Solution**méthode**

|| Intervertir \forall et \exists transforme significativement le sens d'une phrase quantifiée.

(a) La première phrase signifie que la fonction f prend une valeur en tout point : ceci est une propriété vraie pour n'importe quelle fonction définie de E vers F . La seconde phrase signifie que la fonction f prend la même valeur en tout point, autrement dit, elle est constante.

(b) La première phrase signifie que la fonction f atteint toute valeur de son ensemble d'arrivée, autrement dit, elle est surjective. La seconde phrase signifie qu'il existe une valeur de la variable qui prend par f toutes les valeurs de F : cette propriété est notamment fausse dès que F possède au moins deux éléments, une fonction ne prend qu'une seule valeur en tout point !

(c) S'il existe une valeur de x pour laquelle l'assertion $\mathcal{P}(x, y)$ est vérifiée pour tout y , il suffit, pour chaque y , de reprendre cette valeur de x pour vérifier $\mathcal{P}(x, y)$. Autrement dit

$$\left(\exists x \in E, \forall y \in F, \mathcal{P}(x, y) \right) \implies \left(\forall y \in F, \exists x \in E, \mathcal{P}(x, y) \right).$$

Cependant, la réciproque peut être fausse car dans la seconde phrase, la valeur de x est susceptible de dépendre de y . Ce n'est pas le cas dans la première phrase où celle-ci est *uniforme*, c'est-à-dire la même pour chaque valeur de y .

Exercice 3

Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une application. Écrire les négations des phrases quantifiées suivantes :

(a) $\exists M \in \mathbb{R}, (\forall x \in \mathbb{R}, f(x) > M)$ ou $(\forall x \in \mathbb{R}, f(x) \leq M)$.

(b) $\forall x \in \mathbb{R}, f(x) \geq 0 \implies x \geq 0$.

(c) $\forall (x, y) \in \mathbb{R}^2, x \leq y \implies f(x) \leq f(y)$.

(d) $\forall a \in \mathbb{R}, \forall \varepsilon > 0, \exists \alpha > 0, \forall x \in \mathbb{R}, |x - a| \leq \alpha \implies |f(x) - f(a)| \leq \varepsilon$.

Solution**(a) méthode**

- || La négation d'une assertion est mécanique :
- les « et » deviennent « ou » et inversement ;
 - les « \forall » deviennent « \exists » et inversement (Th. 4 p. 5).

La négation de la phrase proposée s'exprime

$$\forall M \in \mathbb{R}, (\exists x \in \mathbb{R}, f(x) < M) \text{ et } (\exists x \in \mathbb{R}, f(x) > M).$$

La phrase initiale signifie que la fonction f est minorée ou majorée, sa négation que la fonction n'est ni minorée, ni majorée.

(b) méthode

- || La négation de « $P \implies Q$ » est « P et non(Q) » (Th. 2 p. 4).

La négation de la phrase proposée s'écrit

$$\exists x \in \mathbb{R}, f(x) > 0 \text{ et } x < 0.$$

La phrase initiale signifie que la fonction f ne prend des valeurs positives que sur \mathbb{R}_+ . Sa négation correspond à l'existence d'une valeur strictement négative de la variable sur laquelle la fonction est positive.

(c) La négation cherchée est

$$\exists (x, y) \in \mathbb{R}^2, x \leq y \text{ et } f(x) > f(y).$$

La phrase initiale signifie que la fonction f est croissante. Sa négation ne signifie pas que la fonction est décroissante puisque le « \forall » est devenu « \exists » !

(d) méthode

- || Il n'est pas nécessaire de comprendre¹ le sens d'une phrase quantifiée pour en exprimer la négation !

La négation cherchée s'écrit

$$\exists a \in \mathbb{R}, \exists \varepsilon > 0, \forall \alpha > 0, \exists x \in \mathbb{R}, |x - a| \leq \alpha \text{ et } |f(x) - f(a)| > \varepsilon.$$

Exercice 4

Soit $f: \mathbb{R} \rightarrow \mathbb{R}$ une application. Signifier à l'aide de phrases quantifiées les affirmations suivantes :

- (a) La fonction f est la fonction nulle.
- (b) La fonction f s'annule.
- (c) La fonction f ne s'annule que sur \mathbb{R}_+ .
- (d) La fonction f s'annule au plus une fois.

1. Cependant, cette phrase quantifiée a bien un sens : elle exprime que la fonction f est continue.

Solution

(a) Il s'agit d'une affirmation *universelle* : « $\forall x \in \mathbb{R}, f(x) = 0$ ».

(b) Il s'agit d'une affirmation *existentielle* : « $\exists x \in \mathbb{R}, f(x) = 0$ ».

(c) **méthode**

|| L'affirmation est de nature hypothétique : « si la valeur $f(x)$ est nulle alors x est élément de \mathbb{R}_+ ». On l'exprime par une implication.

On écrit « $\forall x \in \mathbb{R}, f(x) = 0 \implies x \in \mathbb{R}_+$ » ou, la forme équivalente obtenue par contraposition (Th. 1 p. 4), « $\forall x \in \mathbb{R}, x < 0 \implies f(x) \neq 0$ ». On peut aussi écrire, mais c'est plus alambiqué, « $\forall x \in \mathbb{R}, f(x) \neq 0$ ou $x > 0$ ».

(d) Il s'agit à nouveau d'une affirmation de nature hypothétique. On peut l'exprimer par la phrase « $\forall (x, y) \in \mathbb{R}^2, (f(x) = 0 \text{ et } f(y) = 0) \implies x = y$ ».

1.5.2 Raisonnements

Exercice 5

Montrer que $\sqrt{2}$ est un nombre irrationnel.

Solution

On veut établir l'impossibilité d'écrire $\sqrt{2}$ comme un nombre rationnel.

méthode

|| On raisonne par l'absurde¹.

Par l'absurde, supposons $\sqrt{2}$ rationnel. Il est alors possible d'écrire $\sqrt{2}$ sous la forme d'une fraction irréductible p/q avec p et q entiers. En élevant au carré et en organisant les membres, on obtient $p^2 = 2q^2$. L'entier p^2 est donc pair et p l'est aussi. On peut alors écrire $p = 2k$ avec k entier et l'égalité $p^2 = 2q^2$ se simplifie en $2k^2 = q^2$. On en déduit que l'entier q^2 est pair et donc q l'est aussi. Ainsi, les entiers p et q sont tous deux pairs. C'est absurde puisque la fraction p/q est supposée irréductible !

Finalement, $\sqrt{2}$ est nombre irrationnel.

Exercice 6

(a) Vérifier que $x^2 + x + 1$ est strictement positif quelle que soit la valeur du réel x .

(b) Vérifier que lorsque n est un entier naturel, le nombre $\frac{n(n+1)}{2}$ l'est aussi.

(c) Vérifier que lorsque le produit de deux réels est nul, l'un des facteurs est nul.

1. Lors d'un raisonnement par l'absurde, il importe d'introduire celui-ci en écrivant « Par l'absurde » et de conclure ... à une absurdité !

Solution

(a) En terme quantifié, la propriété voulue s'exprime

$$\forall x \in \mathbb{R}, \quad x^2 + x + 1 > 0.$$

méthode

|| Pour démontrer une propriété universelle, on introduit une valeur arbitraire du paramètre en rédigeant « Soit x dans E ». On établit ensuite la propriété voulue pour cette valeur de x désormais fixée.

Soit x dans \mathbb{R} . On écrit le trinôme $x^2 + x + 1$ sous forme canonique et l'on conclut

$$x^2 + x + 1 = \underbrace{\left(x + \frac{1}{2}\right)^2}_{\geq 0} + \frac{3}{4} \geq \frac{3}{4} > 0.$$

(b) On veut établir la propriété

$$\forall n \in \mathbb{N}, \quad \frac{n(n^2 + 1)}{2} \in \mathbb{N}.$$

Soit n dans \mathbb{N} .

méthode

|| Par disjonction de cas, on discute selon la parité de n .

Cas : n pair. On peut écrire $n = 2k$ avec $k \in \mathbb{N}$ et alors

$$\frac{n(n^2 + 1)}{2} = k(n^2 + 1) \in \mathbb{N}.$$

Cas : n impair. On peut écrire $n = 2k + 1$ avec $k \in \mathbb{N}$ et alors

$$\frac{n(n^2 + 1)}{2} = n(2k^2 + 2k + 1) \in \mathbb{N}.$$

Dans les deux cas, la propriété voulue est vérifiée.

(c) Introduisons des notations permettant de préciser l'étude : considérons x et y deux réels tels que $xy = 0$, on veut établir

$$x = 0 \quad \text{ou} \quad y = 0.$$

méthode

|| Pour montrer une disjonction « \mathcal{P} ou \mathcal{Q} », il est usuel de supposer non(\mathcal{P}) et de vérifier que l'on a alors nécessairement \mathcal{Q} .

Supposons x non nul. On peut introduire son inverse $1/x$ et réaliser le calcul suivant :

$$y = \frac{1}{x} \times xy = \frac{1}{x} \times 0 = 0.$$

Ainsi, lorsque x n'est pas nul, y l'est nécessairement.

Exercice 7

- (a) Montrer que tout nombre rationnel peut s'écrire comme somme de deux nombres irrationnels.
- (b) Montrer que, pour tout entier naturel, il existe un nombre premier¹ qui lui est strictement supérieur.
- (c) À quelle condition un réel peut-il s'écrire à la fois comme la somme et le produit des deux mêmes réels ?

Solution

(a) En termes quantifiés, la propriété voulue s'exprime

$$\forall x \in \mathbb{Q}, \exists (a, b) \in (\mathbb{R} \setminus \mathbb{Q})^2, \quad x = a + b.$$

Soit x un nombre rationnel. Il s'agit d'établir l'existence d'un couple (a, b) tel que voulu.

méthode

On peut établir une existence en exhibant un élément convenable, ici, un couple solution.

Exploitons l'irrationalité de $\sqrt{2}$ vue précédemment et posons $a = x - \sqrt{2}$ et $b = \sqrt{2}$. Le réel a est irrationnel car, par l'absurde, s'il était rationnel $\sqrt{2} = x - a$ le serait aussi par différence de deux nombres rationnels. Le réel b est aussi irrationnel et l'on a évidemment $x = a + b$.

Finalement, l'existence du couple (a, b) est établie.

(b) En notant \mathcal{P} l'ensemble des nombres premiers, la propriété souhaitée s'exprime

$$\forall n \in \mathbb{N}, \exists p \in \mathcal{P}, \quad p > n.$$

Soit n un entier naturel. On ne sait pas exprimer de nombres premiers arbitrairement grands, il est donc difficile d'exhiber un nombre premier solution.

méthode

|| On peut montrer une existence en constatant que l'inexistence est impossible.

Par l'absurde, s'il n'existe pas de nombres premiers strictement supérieurs à n , c'est que tous les nombres premiers sont inférieurs à n . Il n'existe alors qu'un nombre fini de nombres premiers. Ceci est absurde car contredit le théorème d'Euclide sur l'infinité des nombres premiers.

(c) Soit $t \in \mathbb{R}$. On cherche à quelle condition sur t , il existe des réels a et b tels que l'on puisse écrire

$$t = a + b \quad \text{et} \quad t = ab.$$

1. Un nombre premier est un entier $p \geq 2$ dont les seuls diviseurs positifs sont 1 et lui-même : 2, 3, 5, 7, 11, 13, ... sont les premiers nombres premiers. Le théorème d'Euclide (Th. 16 p. 92) assure l'existence d'une infinité de nombres premiers. Cette notion sera approfondie dans le chapitre 3.

Il n'est pas évident d'exhiber des réels a et b donnant cette écriture et l'énoncé suggère que celle-ci n'est peut-être d'ailleurs pas toujours possible.

méthode

On raisonne par « analyse-synthèse ». Lors de la phase d'analyse, on suppose la propriété vraie et l'on étudie les implications de celle-ci. Lorsque cette étude paraît suffisante, on aborde la phase de synthèse où l'on vérifie la propriété dans le contexte fourni par l'analyse.

Analyse : Supposons $t = a + b$ et $t = ab$ pour un certain couple (a, b) de réels. Les réels a et b sont les deux racines de l'équation $(x - a)(x - b) = 0$ d'inconnue x réelle. En développant, cette équation s'écrit encore $x^2 - tx + t = 0$. Pour que cette équation possède deux racines réelles, il est nécessaire que son discriminant soit positif ce qui donne $t^2 - 4t \geq 0$. Cette condition semblant suffisante, on peut aborder la synthèse.

Synthèse : Supposons $t^2 - 4t \geq 0$. L'équation $x^2 - tx + t = 0$ possède deux racines réelles :

$$a = \frac{t - \sqrt{t^2 - 4t}}{2} \quad \text{et} \quad b = \frac{t + \sqrt{t^2 - 4t}}{2}.$$

Pour celles-ci, on vérifie par le calcul $t = a + b$ et $t = ab$.

En résumé, un réel t est somme et produit des deux mêmes réels si, et seulement si, $t^2 - 4t \geq 0$ soit encore $t \in]-\infty ; 0] \cup [4 ; +\infty[$.

Exercice 8

(a) Soit $a \in \mathbb{R}$. Etablir l'implication

$$(\forall \varepsilon \geq 0, |a| \leq \varepsilon) \implies a = 0.$$

(b) Soit $a \in \mathbb{R}$. Etablir l'implication

$$(\forall \varepsilon > 0, |a| \leq \varepsilon) \implies a = 0.$$

(c) Soit x et y deux réels. Établir l'équivalence

$$x^2 + y^2 = 0 \iff x = 0 \text{ et } y = 0.$$

Solution

(a) **méthode**

On peut montrer une implication « $\mathcal{P} \implies \mathcal{Q}$ » en supposant \mathcal{P} et en vérifiant alors \mathcal{Q} .

Supposons $|a| \leq \varepsilon$ pour tout $\varepsilon > 0$. Cette propriété vaut en particulier pour $\varepsilon = 0$ et donc $|a| \leq 0$. On en déduit immédiatement $a = 0$.

(b) **méthode**

On peut montrer une implication « $\mathcal{P} \implies \mathcal{Q}$ » en établissant sa contraposée « $\text{non}(\mathcal{Q}) \implies \text{non}(\mathcal{P})$ » (Th. 1 p. 4).

Par contraposée, montrons : « $a \neq 0 \implies \exists \varepsilon > 0, |a| > \varepsilon$ ».

Supposons $a \neq 0$. Pour $\varepsilon = |a|/2$, on vérifie $\varepsilon > 0$ et $|a| > \varepsilon$. Ceci détermine une valeur de ε convenable établissant l'existence voulue.

(c) **méthode**

On peut établir une équivalence par enchaînement d'équivalences mais aussi en raisonnant par double implication (Th. 3 p. 5). Lorsque l'une d'elles est facile, on peut alors se focaliser sur l'implication difficile.

Raisonnons par double implication.

(\implies) Si $x = y = 0$, il est immédiat que $x^2 + y^2 = 0$.

(\impliedby) Supposons $x^2 + y^2 = 0$. Puisque y^2 est positif, on peut écrire l'encadrement

$$0 < x^2 \leq x^2 + y^2 = 0$$

et affirmer $x^2 = 0$. On en déduit $x = 0$ puis $y = 0$.

1.5.3 Opérations dans $\wp(E)$

Exercice 9

Soit A , B et C trois parties d'un ensemble E .

(a) Montrer $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$.

(b) On suppose $A \cap B \subset A \cap C$ et $A \cup B \subset A \cup C$. Montrer $B \subset C$.

(c) On suppose $A \setminus B = C$. Montrer $A \cup B = B \cup C$.

(d) On suppose $A \cap B = B \cap C = C \cap A$ et $A \cup B = B \cup C = C \cup A$. Montrer que les trois ensembles A , B et C sont égaux.

Solution

(a) **méthode**

On peut montrer l'égalité de deux ensembles en vérifiant qu'il est équivalent d'appartenir à l'un ou à l'autre.

Soit $x \in E$. Par définition, on a

$$x \in A \setminus (B \cap C) \iff x \in A \text{ et } x \notin (B \cap C).$$

Par négation de « $x \in (B \cap C)$ »,

$$x \notin (B \cap C) \iff x \notin B \text{ ou } x \notin C.$$

On peut alors reprendre l'équivalence précédente et poursuivre par distributivité

$$\begin{aligned} x \in A \setminus (B \cap C) &\iff x \in A \text{ et } (x \notin B \text{ ou } x \notin C) \\ &\iff (x \in A \text{ et } x \notin B) \text{ ou } (x \in A \text{ et } x \notin C) \\ &\iff (x \in A \setminus B) \text{ ou } (x \in A \setminus C) \\ &\iff x \in (A \setminus B) \cup (A \setminus C). \end{aligned}$$

Les parties $A \setminus (B \cap C)$ et $(A \setminus B) \cup (A \setminus C)$ étant constituées des mêmes éléments, elles sont égales¹.

(b) méthode

|| On montre une inclusion en choisissant un élément arbitraire dans le premier ensemble et en établissant que celui-ci appartient au second.

Soit x un élément de B . Procédons par disjonction de cas.

Cas : $x \in A$. L'élément x appartient à A et B donc à $A \cap B = A \cap C$ et donc à C .

Cas : $x \notin A$. L'élément x appartient à B donc à $A \cup B = A \cup C$. Il n'est cependant pas élément de A , c'est donc qu'il est élément de C .

Dans les deux cas, x est élément de C et ainsi, tout élément de B est nécessairement élément de C . Finalement, la partie B est incluse dans C .

(c) méthode

|| On peut montrer l'égalité de deux ensembles par double inclusion (Th. 5 p. 6).

B est évidemment inclus dans $A \cup B$ et $C = A \setminus B$ est inclus dans A donc dans $A \cup B$. On en déduit une première inclusion $B \cup C \subset A \cup B$.

Inversement, soit x un élément de $A \cup B$. Procédons par disjonction de cas.

Cas : $x \in B$. L'élément x appartient évidemment² à $B \cup C$.

Cas : $x \notin B$. L'élément x appartient nécessairement à A mais il n'est pas élément de B , il est donc élément de $A \setminus B = C$ donc de $B \cup C$.

Dans les deux cas, on peut affirmer que x est élément de $B \cup C$ et conclure à la seconde inclusion $A \cup B \subset B \cup C$. Par double inclusion, on a l'égalité demandée.

(d) méthode

|| Les hypothèses portées par les parties A , B et C sont symétriques, il suffit d'établir une inclusion pour conclure.

Soit x un élément de A . Supposons par l'absurde qu'il n'est pas élément de B . Il n'est donc pas élément de $A \cap B$ et donc pas élément de $A \cap C$. Il est cependant élément de A , il n'est donc pas élément de C . L'élément x n'appartient ni à B , ni à C , il n'appartient donc pas à $B \cup C$. Cependant, il appartient à l'union $A \cup B$ qui est supposée égale à $B \cup C$. C'est absurde.

On vient ainsi d'établir $A \subset B$. Par symétrie, on peut aussi affirmer $B \subset C$ et $C \subset A$ puis conclure que ces trois parties sont égales.

1. On peut aussi obtenir cette égalité par opérations en écrivant $A \setminus (B \cap C) = A \cap \overline{B \cap C}$ puis en exploitant que le complémentaire d'une intersection est l'union des complémentaires.

2. N'affirmons pas de suite que $A \cup B$ est inclus dans $B \cup C$ dans le cas où x appartient à B : cela n'a pas de sens. On pourra affirmer l'inclusion seulement lorsque toutes les situations de la disjonction de cas seront résolues !

1.5.4 Applications

Exercice 10

On considère la fonction $f: \mathbb{R} \rightarrow \mathbb{R}$ définie par $f(x) = x^2$. Déterminer

- (a) $\text{Im}(f)$ (b) $f([-1; 2])$ (c) $f^{-1}([1; 2[)$.

Solution

(a) méthode

|| Ne pas confondre l'image et l'ensemble d'arrivée d'une application.

L'ensemble d'arrivée de l'application f est la droite réelle et son image correspond à l'ensemble des valeurs qu'elle prend, ici la demi-droite des réels positifs.

(b) méthode

|| $f(A)$ est l'ensemble des valeurs prises par f sur A .

Sur $[-1; 2]$, la fonction f est continue, décroît de 1 à 0 avant de croître de 0 à 4 : elle prend ses valeurs dans $[0; 4]$ et toute valeur de $[0; 4]$ est une valeur prise par f sur $[-1; 2]$. On peut affirmer

$$f([-1; 2]) = [0; 4].$$

(c) méthode

|| Il n'est pas besoin de savoir f bijective pour introduire $f^{-1}(B)$: cette partie est simplement l'ensemble des antécédents éventuels des éléments de B .

On détermine l'ensemble des antécédents des éléments de $[1; 2[$ en recherchant les réels x vérifiant $1 \leq x^2 < 2$. En discutant selon le signe de x , on obtient

$$f^{-1}([1; 2[) =]-\sqrt{2}; -1] \cup [1; \sqrt{2}[.$$

Exercice 11

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

- (a) Soit A et A' deux parties de E . Montrer

$$A \subset A' \implies f(A) \subset f(A').$$

- (b) Soit B et B' deux parties de F . Établir

$$B \subset B' \implies f^{-1}(B) \subset f^{-1}(B').$$

Solution

(a) On suppose $A \subset A'$ et l'on veut établir l'inclusion $f(A) \subset f(A')$: on introduit un élément dans le premier ensemble et l'on montre qu'il appartient nécessairement au second. Soit y un élément¹ de $f(A)$.

méthode

Pour exploiter ou caractériser l'appartenance à une image directe, on prend appui sur l'équivalence suivante :

$$y \in f(A) \iff \exists x \in A, f(x) = y.$$

Il existe x dans A tel que $y = f(x)$. Or x est aussi élément de A' et y peut se comprendre comme une valeur prise par f sur A' . Ainsi, tout élément de $f(A)$ est aussi élément de $f(A')$: on a l'inclusion $f(A) \subset f(A')$.

(b) On suppose $B \subset B'$ et l'on veut établir l'inclusion $f^{-1}(B) \subset f^{-1}(B')$. Soit x un élément de $f^{-1}(B)$.

méthode

Pour exploiter ou caractériser l'appartenance à une image réciproque, on s'appuie² sur l'équivalence suivante :

$$x \in f^{-1}(B) \iff f(x) \in B.$$

On a $f(x) \in B$ et donc $f(x) \in B'$ car on a supposé $B \subset B'$. On a ainsi immédiatement x élément de $f^{-1}(B')$ et l'on peut conclure $f^{-1}(B) \subset f^{-1}(B')$.

Exercice 12

Soit $f: \mathbb{N} \rightarrow \mathbb{N}$ et $g: \mathbb{N} \rightarrow \mathbb{N}$ les applications déterminées par :

$$f(k) = 2k \quad \text{et} \quad g(k) = \begin{cases} k/2 & \text{si } k \text{ est pair} \\ (k-1)/2 & \text{si } k \text{ est impair.} \end{cases}$$

- (a) Étudier la bonne définition des applications f et g .
- (b) Étudier les injectivités des applications f et g .
- (c) Étudier les surjectivités des applications f et g .
- (d) Préciser les applications $g \circ f$ et $f \circ g$. Sont-elles bijectives ?

1. Puisque cet élément est une valeur prise par f , il est « naturel » de le noter y plutôt que x .

2. Ignorant si f bijective, il est impossible d'affirmer que x s'écrit $f^{-1}(y)$ pour un certain y dans B .

Solution**(a) méthode**

On observe qu'une application $f: E \rightarrow F$ est bien définie en vérifiant que, pour chaque x dans E , la valeur $f(x)$ est parfaitement définie¹ et est élément de F .

La bonne définition de l'application f ne revêt aucune difficulté : pour chaque $k \in \mathbb{N}$, la valeur $2k$ est parfaitement définie et est élément de \mathbb{N} .

La bonne définition de l'application g est à peine plus délicate : pour chaque $k \in \mathbb{N}$, si k est pair, la valeur $k/2$ est bien définie et est un entier naturel et, si k est impair, il en est de même pour la valeur $(k-1)/2$.

On peut visualiser les premières valeurs prises par ces deux fonctions dans les tableaux suivants

k	0	1	2	3	4	...
$f(k)$	0	2	4	6	8	...

k	0	1	2	3	4	...
$g(k)$	0	0	1	1	2	...

(b) méthode

Pour montrer l'injectivité d'une application f , on introduit x et x' dans son ensemble de départ et l'on suppose $f(x) = f(x')$ afin d'établir $x = x'$.

Soit k et k' dans \mathbb{N} . On suppose $f(k) = f(k')$, c'est-à-dire $2k = 2k'$. En simplifiant par 2, on obtient directement $k = k'$. On peut conclure que f est injective.

méthode

Pour montrer qu'une application n'est pas injective, il suffit d'exhiber deux valeurs x et x' distinctes ayant même image.

On a $g(0) = 0$ et $g(1) = 0$: l'application g n'est pas injective.

(c) méthode

Pour montrer qu'une application n'est pas surjective, il suffit d'exhiber une valeur de l'ensemble d'arrivée qui n'est pas une valeur prise.

Les valeurs prises par f sont toutes des entiers pairs, la valeur 1 appartient à l'ensemble d'arrivée mais n'est pas une valeur prise par f : l'application f n'est pas surjective.

méthode

On montre la surjectivité d'une application en déterminant un antécédent à chaque valeur de l'ensemble d'arrivée.

Soit $y \in \mathbb{N}$ une valeur arbitraire de l'ensemble d'arrivée de g . Posons² $k = 2y$. La valeur k appartient à l'ensemble de départ de g et, puisqu'il s'agit d'un nombre pair, on vérifie $g(k) = y$. Ainsi, l'application g est surjective.

1. Cela signifie que la valeur existe et est unique.

2. On peut aussi proposer $k = 2y + 1$.

(d) Soit $k \in \mathbb{N}$. On a $(g \circ f)(k) = g(2k)$. Puisque $2k$ est un entier pair, $g(2k) = k$ et, finalement¹, $g \circ f = \text{Id}_{\mathbb{N}}$.

Soit $k \in \mathbb{N}$. On a $(f \circ g)(k) = f(g(k))$.

méthode

|| Afin d'évaluer $g(k)$ il est nécessaire de discuter selon la parité de k .

Si k est pair, $(f \circ g)(k) = f(k/2) = k$.

Si k est impair, $(f \circ g)(k) = f((k-1)/2) = k-1$.

Les valeurs successives prises par la composée $f \circ g$ sont $0, 0, 2, 2, 4, 4, \dots$

L'application $g \circ f$ est bijective mais l'application $f \circ g$ ne l'est pas, elle n'est même ni injective, ni surjective.

Exercice 13

Soit $s: \mathbb{N} \rightarrow \mathbb{N}^*$ l'application définie par $s(n) = n + 1$. Montrer que l'application s est bijective :

- (a) En constatant la définition d'une application bijective.
- (b) En vérifiant injectivité et surjectivité.
- (c) En déterminant une application susceptible d'être son application réciproque.

Solution

Commençons par souligner que l'application s est bien définie, notamment car ses valeurs sont prises dans \mathbb{N}^* .

(a) méthode

|| Par retour à la définition, on montre qu'une application $f: E \rightarrow F$ est bijective en résolvant, pour chaque valeur de y dans F , l'équation $f(x) = y$ et en observant que celle-ci admet une unique solution x dans E .

Soit $y \in \mathbb{N}^*$. On a immédiatement

$$s(n) = y \iff n = y - 1.$$

L'équation $s(n) = y$ admet donc une unique solution et celle-ci est bien élément de \mathbb{N} : on peut affirmer que s est bijective.

(b) méthode

|| En étudiant séparément² surjectivité et injectivité, on retrouve l'étude de l'existence et de l'unicité d'une solution à l'équation $f(x) = y$.

1. La composée $g \circ f$ est égale à l'identité et, pourtant, ni l'application f , ni l'application g n'est bijective. Pour pouvoir affirmer que deux applications f et g sont bijectives et réciproques l'une de l'autre, il faut que les deux composées $f \circ g$ et $g \circ f$ soient égales à des applications identités.

2. Dans les situations faciles, comme celle en cours, il est assez inefficace de raisonner ainsi. Dans des situations plus délicates, la séparation des deux problématiques peut être pertinente. Dans des résultats futurs (Th. 4 p. 206 et Th. 15 p. 278) la résolution d'une des deux problématiques entraînera la résolution de l'autre ce qui conduira à la bijectivité.

Soit n et n' dans \mathbb{N} . Si $s(n) = s(n')$ alors $n + 1 = n' + 1$, puis en simplifiant, $n = n'$. L'application s est injective.

Soit y dans l'ensemble d'arrivée \mathbb{N}^* . Pour $n = y - 1$, n est élément de l'ensemble de départ et vérifie $s(n) = y$: l'application s est surjective.

(c) Introduisons¹ $p: \mathbb{N}^* \rightarrow \mathbb{N}$ l'application définie par $p(m) = m - 1$. L'application p est bien définie.

méthode

On montre que les applications s et p sont bijectives et réciproques l'une de l'autre en étudiant les deux² composées $s \circ p$ et $p \circ s$.

Pour tout $n \in \mathbb{N}$, on a $(p \circ s)(n) = p(s(n)) = p(n + 1) = n$ et, pour tout $m \in \mathbb{N}^*$, on a aussi $(s \circ p)(m) = s(p(m)) = s(m - 1) = m$. On peut donc affirmer $p \circ s = \text{Id}_{\mathbb{N}}$ et $s \circ p = \text{Id}_{\mathbb{N}^*}$: l'application s est bijective et p est son application réciproque.

1.5.5 Relations binaires

Exercice 14

Soit $f: E \rightarrow \mathbb{R}$ une application injective. On définit sur E une relation binaire \preccurlyeq par

$$x \preccurlyeq y \iff f(x) \leq f(y).$$

- (a) Montrer que \preccurlyeq est une relation d'ordre sur E .
- (b) S'agit-il d'une relation d'ordre totale ?

Solution

(a) méthode

On vérifie que la relation \preccurlyeq est réflexive, antisymétrique et transitive.

Soit $x \in E$. On a $f(x) \leq f(x)$ et donc $x \preccurlyeq x$. La relation est réflexive.

Soit x et y des éléments de E tels que $x \preccurlyeq y$ et $y \preccurlyeq x$. On a $f(x) \leq f(y)$ et $f(y) \leq f(x)$ donc $f(x) = f(y)$. Or f est injective et l'on peut poursuivre en affirmant $x = y$. La relation est antisymétrique.

Enfin, soit x , y et z des éléments de E tels que $x \preccurlyeq y$ et $y \preccurlyeq z$. On a $f(x) \leq f(y)$ et $f(y) \leq f(z)$ donc $f(x) \leq f(z)$ puis $x \preccurlyeq z$. La relation est transitive.

Finalement, \preccurlyeq est une relation d'ordre sur E .

(b) méthode

Une relation d'ordre est totale si, et seulement si, tous les éléments sont deux à deux comparables.

1. Il serait anticipé de noter s^{-1} cette application : on ne sait pas encore que c'est la bijection réciproque de s même si nous allons l'établir.

2. Avoir une seule composée égale à l'identité ne suffit pas : voir sujet 12 p. 24.

Soit x et y deux éléments de E . Les réels $f(x)$ et $f(y)$ sont comparables. Si $f(x) \leq f(y)$ alors $x \preccurlyeq y$, sinon $y \preccurlyeq x$. La relation d'ordre \preccurlyeq est totale.

Exercice 15

On définit une relation¹ binaire \mathcal{R} sur l'ensemble $E = \mathbb{R}_+$ en posant

$$x \mathcal{R} y \iff \exists (k, \ell) \in (\mathbb{N}^*)^2, kx = \ell y.$$

- (a) Montrer que \mathcal{R} définit une relation d'équivalence.
- (b) Décrire la classe d'équivalence d'un réel x .

Solution

(a) méthode

|| On vérifie que la relation \mathcal{R} est réflexive, symétrique et transitive.

Soit x un réel positif. On a $x \mathcal{R} x$ car $kx = \ell x$ pour $(k, \ell) = (1, 1) \in (\mathbb{N}^*)^2$: la relation est réflexive.

Soit x et y deux réels positifs. On suppose $x \mathcal{R} y$. Il existe $(k, \ell) \in (\mathbb{N}^*)^2$ tel que $kx = \ell y$ et l'on peut alors aussi écrire $k'y = \ell'x$ avec $(k', \ell') = (\ell, k) \in (\mathbb{N}^*)^2$. Ainsi, $y \mathcal{R} x$ et l'on peut affirmer que la relation est symétrique.

Enfin, soit x, y et z trois réels positifs. On suppose $x \mathcal{R} y$ et $y \mathcal{R} z$. On peut donc introduire deux couples (k, ℓ) et (k', ℓ') dans $(\mathbb{N}^*)^2$ tels que $kx = \ell y$ et $k'y = \ell'z$. On a alors $kk'x = \ell\ell'z$, c'est-à-dire $k''x = \ell''z$ avec $(k'', \ell'') = (kk', \ell\ell') \in (\mathbb{N}^*)^2$. Ainsi, $x \mathcal{R} z$ et l'on peut affirmer que la relation est transitive : c'est une relation d'équivalence.

(b) méthode

|| Pour déterminer la classe d'équivalence de x , on recherche les y qui sont en relation avec x .

Soit y un réel. On a

$$\begin{aligned} x \mathcal{R} y &\iff \exists (k, \ell) \in (\mathbb{N}^*)^2, kx = \ell y \\ &\iff \exists (k, \ell) \in (\mathbb{N}^*)^2, y = \frac{k}{\ell}x \\ &\iff \exists r \in \mathbb{Q}_+, y = rx. \end{aligned}$$

La classe d'équivalence peut donc s'écrire

$$\text{Cl}(x) = \{rx \mid r \in \mathbb{Q}_+\}.$$

1. C'est la relation de commensurabilité : deux longueurs sont *commensurables* lorsqu'elles sont multiples d'une longueur commune.

1.6 Exercices d'entraînement

1.6.1 Opérations dans $\wp(E)$

Exercice 16 *

Soit A , B et C trois parties d'un ensemble E . Vérifier

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

Solution

méthode

On factorise partiellement le premier membre par

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

On poursuit par distributivité

$$\begin{aligned} (A \cup B) \cap (B \cup C) \cap (C \cup A) &= (B \cup (A \cap C)) \cap (C \cup A) \\ &= (B \cap (C \cup A)) \cup ((A \cap C) \cap (C \cup A)) \end{aligned}$$

Or, puisqu'il y a inclusion, $(A \cap C) \cap (C \cup A) = A \cap C$ et l'on peut poursuivre par distributivité et conclure

$$\begin{aligned} (A \cup B) \cap (B \cup C) \cap (C \cup A) &= (B \cap (C \cup A)) \cup (A \cap C) \\ &= (B \cap C) \cup (B \cap A) \cup (A \cap C). \end{aligned}$$

Exercice 17 ** (Différence symétrique)

On définit la *différence symétrique* de deux parties A et B d'un ensemble E par

$$A \Delta B = (A \cup B) \cap (\overline{A \cap B}).$$

Soit A , B et C trois parties de E .

- (a) Calculer $A \Delta A$, $A \Delta \bar{A}$, $A \Delta E$ et $A \Delta \emptyset$.
- (b) Vérifier la propriété d'associativité

$$A \Delta (B \Delta C) = (A \Delta B) \Delta C.$$

- (c) Établir

$$A \Delta B = A \Delta C \implies B = C.$$

Solution

(a) Directement

$$A \Delta A = (A \cup A) \cap (\overline{A \cap A}) = A \cap \emptyset = \emptyset$$

$$A \Delta \overline{A} = (A \cup \overline{A}) \cap (\overline{A \cap \overline{A}}) = E \cap E = E$$

$$A \Delta E = (A \cup E) \cap (\overline{A \cap E}) = E \cap \overline{A} = \overline{A}$$

$$A \Delta \emptyset = (A \cup \emptyset) \cap (\overline{A \cap \emptyset}) = A \cap E = A.$$

(b) méthode

On dresse un tableau traitant toutes les possibilités d'appartenance d'un élément x de E .

On exprime par V et F l'appartenance ou non de x à l'ensemble précisé en tête de colonne.

A	B	C	$B \Delta C$	$A \Delta B$	$A \Delta (B \Delta C)$	$(A \Delta B) \Delta C$
V	V	V	F	F	V	V
V	V	F	V	F	F	F
V	F	V	V	V	F	F
V	F	F	F	V	V	V
F	V	V	F	V	F	F
F	V	F	V	V	V	V
F	F	V	V	F	V	V
F	F	F	F	F	F	F

Les deux dernières colonnes sont identiques, les ensembles $A \Delta (B \Delta C)$ et $(A \Delta B) \Delta C$ sont donc constitués des mêmes éléments, à savoir ceux qui sont communs aux trois ensembles A, B, C ainsi que ceux qui appartiennent à l'un des trois sans figurer dans l'un des deux autres.

(c) méthode

Δ est une opération pour laquelle \emptyset joue le rôle d'un élément neutre et pour laquelle toute partie est son propre symétrique : on simplifie l'équation en composant par A .

Supposons $A \Delta B = A \Delta C$. En composant par A , on peut affirmer par associativité $(A \Delta A) \Delta B = (A \Delta A) \Delta C$ ce qui donne $\emptyset \Delta B = \emptyset \Delta C$ puis $B = C$.

1.6.2 Injection, surjection, bijection

Exercice 18 *

Soit f une fonction réelle définie au départ d'un intervalle I .

Montrer que si f est strictement monotone alors f est injective.

Solution**méthode**

On peut montrer qu'une application f est injective en vérifiant

$$x \neq y \implies f(x) \neq f(y).$$

Quitte à considérer $-f$, ce qui ne change pas la nature du problème, on peut supposer la fonction f strictement croissante. On sait alors

$$\forall (x, y) \in I^2, \quad x < y \implies f(x) < f(y).$$

Soit x et y deux éléments distincts de I . Quitte à les échanger, on peut supposer $x < y$ auquel cas l'implication au-dessus donne $f(x) < f(y)$ et donc $f(x) \neq f(y)$.

L'application f est injective.

Exercice 19 *

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

Que dire de la restriction de f au départ de E et à valeurs dans $\text{Im}(f)$? Que dire de plus si l'on sait f injective?

Solution**méthode**

Une application est surjective lorsqu'elle prend toutes les valeurs de l'ensemble qui exprime son ensemble d'arrivée¹.

Notons $f': E \rightarrow \text{Im}(f)$ la restriction considérée. Celle-ci est parfaitement définie car elle prend bien ses valeurs dans $\text{Im}(f)$. Au surplus, cette restriction est surjective car, pour tout y dans $\text{Im}(f)$, il existe x dans E tel que $f(x) = y$ auquel cas $f'(x) = y$.

Si de plus l'application f est injective, la restriction f' l'est aussi² car, pour tous x et y éléments de E ,

$$\begin{aligned} f'(x) = f'(y) &\implies f(x) = f(y) \\ &\implies x = y. \end{aligned}$$

On peut alors affirmer que f' est bijective.

En substance, on retiendra que la restriction d'une injection à son image induit une bijection.

1. Il est important de savoir distinguer l'ensemble image d'une application de l'ensemble d'arrivée. Ce dernier est introduit en même temps que l'ensemble de départ lorsque l'on présente une application.

2. De façon générale, la restriction d'une injection est une injection.

Exercice 20 *

Soit a, b et c trois réels tels que $c \neq 0$ et $a^2 + bc \neq 0$. On introduit $E = \mathbb{R} \setminus \{\pm\}$.

On considère la fonction $f: E \rightarrow E$ définie par

$$f(x) = \frac{ax+b}{cx-a}.$$

(a) Justifier que l'application f est bien définie.

(b) Calculer $f \circ f$. En déduire que f est une bijection dont on déterminera l'application réciproque.

Solution(a) **méthode**

On vérifie non seulement que l'on peut calculer $f(x)$ mais aussi que cette valeur est bien élément de E .

Soit $x \in E$. On peut calculer le réel $y = \frac{ax+b}{cx-a}$ car le dénominateur ne s'annule pas puisque $x \neq a/c$. Au surplus

$$\begin{aligned} \frac{ax+b}{cx-a} = \frac{a}{c} &\iff (ax+b)c = a(cx-a) \\ &\iff a^2 + bc = 0. \end{aligned}$$

Or on a supposé $a^2 + bc \neq 0$ et l'on peut donc affirmer que y appartient à E . La fonction f est donc bien définie au départ de E et à valeurs dans E .

(b) Pour $x \in E$

$$(f \circ f)(x) = f(f(x)) = \frac{af(x)+b}{cf(x)-a} = \frac{a\frac{ax+b}{cx-a}+b}{c\frac{ax+b}{cx-a}-a}$$

Après réduction au même dénominateur puis simplification

$$(f \circ f)(x) = \frac{a^2x + ab + bcx - ab}{acx + bc - ac + a^2} = x \quad \text{car} \quad a^2 + bc \neq 0.$$

Ainsi, $f \circ f = \text{Id}_E$. L'application f est alors bijective d'application réciproque égale à elle-même¹.

Exercice 21 **

Soit $f: \mathbb{C} \rightarrow \mathbb{C}$ l'application définie par $f(z) = z^2$.

(a) Montrer que l'application f est surjective mais non injective.

On note $\Omega = \{z \in \mathbb{C} \mid \text{Re}(z) > 0\}$.

(b) Montrer que la restriction f' de f au départ de Ω et à valeurs dans $\mathbb{C} \setminus \mathbb{R}_-$ est bijective.

1. On dit que f est une *involution*.

Solution

(a) L'application f n'est pas injective puisque $f(1) = f(-1)$.

Soit $Z \in \mathbb{C}$. Déterminons $z \in \mathbb{C}$ tel que $f(z) = Z$.

méthode

|| On écrit Z sous forme trigonométrique : $Z = |Z| e^{i\theta}$ avec θ un argument de Z .

En posant $z = \sqrt{|Z|} e^{i\theta/2} \in \mathbb{C}$, on obtient $f(z) = Z$. Toute valeur de l'ensemble d'arrivée \mathbb{C} possède au moins un antécédent : la fonction f est surjective.

(b) méthode

|| On commence par vérifier la bonne définition de la restriction f' .

Soit $z \in \Omega$. On peut écrire $z = a + ib$ avec $a > 0$ et l'on a alors $f(z) = a^2 - b^2 + 2iab$. Si $b \neq 0$, $f(z)$ n'est pas réel et a fortiori n'appartient pas à \mathbb{R}_- . Si $b = 0$, $f(z) = a^2 > 0$ n'est pas non plus élément de \mathbb{R}_- . Ainsi,

$$\forall z \in \Omega, \quad f(z) \in \mathbb{C} \setminus \mathbb{R}_-.$$

La restriction f' de f au départ de Ω et à valeurs dans $\mathbb{C} \setminus \mathbb{R}_-$ est donc bien définie.

Vérifions qu'elle est injective. Soit z et z' dans \mathbb{C} tels que $f'(z) = f'(z')$, c'est-à-dire tels que $z^2 = z'^2$. Après différence des membres, on obtient $(z - z')(z + z') = 0$. On a alors $z = z'$ ou $z = -z'$. Cependant, les parties réelles de z et z' sont toutes deux strictement positives et l'alternative $z = -z'$ est impossible. Il reste $z = z'$ et l'on peut affirmer que la restriction f' est injective.

Vérifions qu'elle est aussi surjective. Soit Z un élément de l'ensemble d'arrivée $\mathbb{C} \setminus \mathbb{R}_-$. Comme au-dessus, on écrit $Z = |Z| e^{i\theta}$ avec θ un argument de Z que l'on choisit dans l'intervalle $]-\pi; \pi[$. En posant $z = \sqrt{|Z|} e^{i\theta/2}$, on définit un complexe dont le carré vaut Z et dont la partie réelle est strictement positive. En effet,

$$\operatorname{Re}(z) = \sqrt{|Z|} \cos\left(\frac{\theta}{2}\right) > 0 \quad \text{car} \quad |Z| \neq 0 \text{ et } \frac{\theta}{2} \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$$

Ainsi, $z \in \Omega$ et $f'(z) = Z$. La fonction f' est surjective et, finalement, bijective.

Exercice 22 **

On considère l'application $f: \mathbb{N} \rightarrow \mathbb{Z}$ définie par

$$f(n) = \begin{cases} \frac{n}{2} & \text{si } n \text{ est pair} \\ -\frac{n+1}{2} & \text{sinon.} \end{cases}$$

Montrer que l'application f est bijective et exprimer sa bijection réciproque.

Solution

Afin de nous familiariser avec la fonction f , calculons ses premières valeurs :

n	0	1	2	3	4	...
$f(n)$	0	-1	1	-2	2	...

On vérifie que l'application f est bien définie au départ de \mathbb{N} et à valeurs dans \mathbb{Z} . En effet, si n est pair, on peut écrire $n = 2k$ avec $k \in \mathbb{N}$ et alors $f(n) = k \in \mathbb{N} \subset \mathbb{Z}$. Aussi, si n est impair, on peut écrire $n = 2k + 1$ avec $k \in \mathbb{N}$ et alors $f(n) = -(k + 1) \in \mathbb{Z}$. Dans les deux cas, les valeurs prises par f sont bien des entiers relatifs.

Montrons maintenant que l'application f est bijective.

Soit $p \in \mathbb{Z}$. Résolvons l'équation $f(n) = p$ d'inconnue $n \in \mathbb{N}$.

méthode

On observe que les entiers pairs sont envoyés sur les entiers positifs tandis que les entiers impairs sont envoyés sur les entiers strictement négatifs : on détermine un antécédent en discutant selon son signe.

Si p est positif, l'équation $f(n) = p$ ne possède pas de solutions parmi les entiers impairs. Pour n pair, l'équation $f(n) = p$ se résout en $n = 2p \in \mathbb{N}$ ce qui détermine une unique solution à l'équation $f(n) = p$.

De façon semblable, si p est un entier strictement négatif, les solutions de l'équation $f(n) = p$ sont à rechercher parmi les entiers impairs. Après résolution, on obtient que $n = -(2p + 1) \in \mathbb{N}$ est la seule solution.

Finalement, pour chaque $p \in \mathbb{Z}$, l'équation $f(n) = p$ possède une unique solution n dans \mathbb{N} . L'application f est bijective et sa bijection réciproque $f^{-1}: \mathbb{Z} \rightarrow \mathbb{N}$ est déterminée par les résolutions qui précèdent

$$f^{-1}(p) = \begin{cases} 2p & \text{si } p \geq 0 \\ -(2p + 1) & \text{sinon.} \end{cases}$$

Exercice 23 **

Soit $f: E \rightarrow F$ et $g: F \rightarrow G$ deux applications. Établir :

- (a) $g \circ f$ injective $\implies f$ injective.
- (b) $g \circ f$ surjective $\implies g$ surjective.
- (c) $g \circ f$ injective et f surjective $\implies g$ injective.
- (d) $g \circ f$ surjective et g injective $\implies f$ surjective.

Solution

Notons que la composée $g \circ f$ définit une application de E vers G .

méthode

Pour ne pas s'égarer, il peut être utile de convenir de noter x, x' les éléments choisis dans E , de noter y, y' ceux choisis dans F et z, z' ceux choisis dans G .

(a) Supposons $g \circ f$ injective. Soit x et x' dans E . Si $f(x) = f(x')$, il vient en composant par g l'égalité $g(f(x)) = g(f(x'))$, c'est-à-dire $(g \circ f)(x) = (g \circ f)(x')$. Or la fonction $g \circ f$ est injective et donc $x = x'$. Ainsi, la fonction f est injective.

(b) Supposons $g \circ f$ surjective. Soit $z \in G$. Il existe $x \in E$ tel que $z = g(f(x))$. En posant $y = f(x) \in F$, on obtient $g(y) = z$. Ainsi, la fonction g est surjective.

(c) Supposons $g \circ f$ injective et f surjective.

méthode

|| On observe¹ que f est bijective et l'on introduit f^{-1} .

Puisque $g \circ f$ est injective, l'étude du (a) assure que f est injective et donc bijective (Th. 7 p. 11). On peut alors introduire sa bijection réciproque et écrire

$$g = (g \circ f) \circ f^{-1}.$$

L'application g est donc injective par composition d'injections (Th. 8 p. 11).

(d) Supposons $g \circ f$ surjective et g injective. Par l'étude du (b), on peut affirmer que g est surjective donc bijective. On conclut alors que f est surjective par la composition² de surjections

$$f = g^{-1} \circ (g \circ f).$$

Exercice 24 ***

Soit E un ensemble et $f: E \rightarrow E$ une application telle que $f \circ f \circ f = f$.

Montrer que f est injective si, et seulement si, f est surjective.

Solution

On raisonne par double implication.

(\implies) Supposons f injective. Soit $y \in E$. On écrit

$$f(y) = (f \circ f \circ f)(y) = f((f \circ f)(y)).$$

La fonction f étant injective, on obtient $y = (f \circ f)(y)$. Ceci suffit à déterminer un antécédent de y puisque, pour $x = f(y) \in E$, on a $f(x) = f(f(y)) = y$.

Ainsi, la fonction f est surjective.

(\impliedby) Supposons f surjective.

Soit $x, x' \in E$ tels que $f(x) = f(x')$.

méthode

|| Par la surjectivité de f , on écrit x et x' de sorte de faire apparaître $f \circ f \circ f$ afin d'exploiter l'hypothèse $f \circ f \circ f = f$.

1. Une démonstration de l'implication « $g(y) = g(y') \implies y = y'$ » est aussi possible en introduisant x et x' antécédents de y et y' par f .

2. La détermination d'un antécédent par f à un élément $y \in F$ quelconque est aussi possible en introduisant $z = g(y)$ qui possède un antécédent par $g \circ f$.

Puisque f est surjective, la composée $f \circ f$ l'est aussi et l'on peut introduire a et a' dans E tels que $x = (f \circ f)(a)$ et $x' = (f \circ f)(a')$. L'égalité $f(x) = f(x')$ se relit alors $(f \circ f \circ f)(a) = (f \circ f \circ f)(a')$, c'est-à-dire $f(a) = f(a')$. On en déduit $f(f(a)) = f(f(a'))$ donc $x = x'$.

Ainsi, la fonction f est injective¹.

1.6.3 Images directes et images réciproques

Exercice 25 *

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

(a) Soit A_1 et A_2 deux parties de E . Montrer

$$f(A_1 \cup A_2) = f(A_1) \cup f(A_2) \quad \text{et} \quad f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2).$$

(b) Soit B_1 et B_2 deux parties de F . Montrer

$$f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2) \quad \text{et} \quad f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2).$$

Solution

(a) **méthode**

|| On raisonne par inclusions.

Montrons l'égalité $f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$ par double inclusion.

Soit $y \in f(A_1 \cup A_2)$. Par définition, y est la valeur prise par f sur un certain élément x de $A_1 \cup A_2$. Si x est élément de A_1 alors y appartient à $f(A_1)$. Sinon, x est élément de A_2 et y appartient à $f(A_2)$. Dans les deux cas y est élément de l'union $f(A_1) \cup f(A_2)$ ce qui fournit une première inclusion : $f(A_1 \cup A_2) \subset f(A_1) \cup f(A_2)$.

Soit $y \in f(A_1) \cup f(A_2)$. Si y appartient à $f(A_1)$, il existe x dans A_1 tel que $y = f(x)$. Or A_1 est inclus dans $A_1 \cup A_2$ et donc y peut se comprendre comme l'image d'un élément de $A_1 \cup A_2$: $y \in f(A_1 \cup A_2)$. De même², si $y \in f(A_2)$, on obtient $y \in f(A_1 \cup A_2)$. Ceci fournit la seconde inclusion : $f(A_1) \cup f(A_2) \subset f(A_1 \cup A_2)$. On peut conclure à l'égalité.

Montrons maintenant l'inclusion $f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2)$.

Soit $y \in f(A_1 \cap A_2)$. Il existe x dans $A_1 \cap A_2$ tel que $y = f(x)$. Puisque $A_1 \cap A_2$ est inclus dans A_1 , x est élément de A_1 et l'on peut écrire $y \in f(A_1)$. De même³, $y \in f(A_2)$ et donc $y \in f(A_1) \cap f(A_2)$.

1. La fonction est alors bijective et vérifie $f \circ f = \text{Id}_E$. Son application réciproque est elle-même, c'est une involution.

2. On pourrait aussi employer le résultat du sujet 11 p. 23 : $A \subset A' \implies f(A) \subset f(A')$ avec $A = A_1$ (ou $A = A_2$) et $A' = A_1 \cup A_2$.

3. De nouveau, on peut exploiter le résultat du sujet 11 p. 23.

(b) **méthode**

|| On raisonne par équivalences.

Soit $x \in E$.

$$\begin{aligned} x \in f^{-1}(B_1 \cup B_2) &\iff f(x) \in B_1 \cup B_2 \\ &\iff f(x) \in B_1 \text{ ou } f(x) \in B_2 \\ &\iff x \in f^{-1}(B_1) \text{ ou } x \in f^{-1}(B_2) \\ &\iff x \in f^{-1}(B_1) \cup f^{-1}(B_2). \end{aligned}$$

On en déduit la première égalité : $f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$.

On obtient la seconde égalité par la même démonstration en changeant union en intersection et « et » en « ou ».

Exercice 26 **

Soit $f: E \rightarrow F$ une application. À quelle condition sur f peut-on affirmer

$$\forall (A_1, A_2) \in \wp(E)^2, \quad f(A_1 \cap A_2) = f(A_1) \cap f(A_2) ?$$

Solution**méthode**

|| On montre que l'égalité est toujours vraie si, et seulement si, f est injective.

Supposons f injective. Soit $A_1, A_2 \in \wp(E)$. Par l'étude du sujet précédent, on sait déjà l'inclusion $f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2)$. Étudions l'inclusion réciproque. Soit y un élément de $f(A_1) \cap f(A_2)$. L'élément y est une valeur prise par f en certain x de A_1 et aussi une valeur prise par f en certain x' de A_2 . Cependant, la fonction f est supposée injective et donc $x = x'$ ce qui détermine un élément commun à A_1 et A_2 . Ainsi, y est élément de $f(A_1 \cap A_2)$.

Finalement, on a par double inclusion l'égalité $f(A_1 \cap A_2) = f(A_1) \cap f(A_2)$.

Inversement, supposons l'égalité vraie pour toutes parties A_1 et A_2 de E et montrons que f est injective. Soit x et x' dans E tels que $f(x) = f(x')$.

méthode

|| On introduit des parties A_1 et A_2 adaptées au contexte.

Considérons $A_1 = \{x\}$ et $A_2 = \{x'\}$. On a

$$f(A_1 \cap A_2) = f(A_1) \cap f(A_2) = \{f(x)\} \neq \emptyset.$$

Nécessairement, $A_1 \cap A_2$ est non vide et donc $x = x'$. On peut alors conclure que f est injective.

Exercice 27 **

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

(a) Établir

$$\forall A \in \wp(E), A \subset f^{-1}(f(A)) \quad \text{et} \quad \forall B \in \wp(F), f(f^{-1}(B)) \subset B.$$

(b) Montrer

$$f \text{ est injective} \iff \forall A \in \wp(E), A = f^{-1}(f(A)).$$

(c) Montrer

$$f \text{ est surjective} \iff \forall B \in \wp(F), f(f^{-1}(B)) = B.$$

Solution

(a) Soit A une partie de E . Pour $x \in A$, l'élément $y = f(x)$ est évidemment une valeur de $f(A)$ et, puisque x en est un antécédent, on peut écrire $x \in f^{-1}(f(A))$. Ainsi, on obtient l'inclusion $A \subset f^{-1}(f(A))$.

Soit B une partie de F . Soit $y \in f(f^{-1}(B))$. L'élément y est une valeur prise par f sur un certain élément x de $f^{-1}(B)$. Or les éléments de $f^{-1}(B)$ ont tous leur image dans B . En particulier, $y = f(x)$ appartient à B . Ainsi, on obtient $f(f^{-1}(B)) \subset B$.

(b) On raisonne par double implication.

(\implies) Supposons f injective. Soit $A \in \wp(E)$. On sait déjà que $A \subset f^{-1}(f(A))$. Étudions l'inclusion inverse. Pour $x \in f^{-1}(f(A))$, on a $f(x) \in f(A)$ et il existe donc x' dans A tel que $f(x) = f(x')$. Puisque f est injective, on a nécessairement $x = x'$ et x est élément de A . Ainsi, $f^{-1}(f(A)) \subset A$ puis l'égalité.

(\impliedby) Supposons $A = f^{-1}(f(A))$ pour toute partie A de E et montrons que f est injective. Soit x et x' deux éléments de A tels que $f(x) = f(x')$.

méthode

|| On introduit une partie A permettant d'exploiter l'hypothèse.

Considérons $A = \{x\} \subset E$. On a $f(A) = \{f(x)\}$ et donc $x' \in f^{-1}(f(A)) = A = \{x\}$. On en déduit $x = x'$. Ainsi, la fonction f injective.

(c) On raisonne à nouveau par double implication.

(\implies) Supposons f surjective. Soit $B \in \wp(F)$. On sait déjà $f(f^{-1}(B)) \subset B$. Étudions l'inclusion réciproque. Soit $y \in B$. Puisque f est surjective, il existe un antécédent $x \in E$ à la valeur y . Puisque $f(x) = y \in B$, cet antécédent x est élément de $f^{-1}(B)$ et donc $y = f(x)$ est une valeur prise par f sur $f^{-1}(B)$, autrement dit $y \in f(f^{-1}(B))$. Ainsi, on obtient $B \subset f(f^{-1}(B))$ puis l'égalité.

(\Leftarrow) Supposons $f(f^{-1}(B)) = B$ pour toute partie B de F . Soit $y \in F$. Montrons que cet élément y admet un antécédent par f .

méthode

|| On introduit une partie B permettant d'exploiter l'hypothèse.

Considérons $B = \{y\} \subset F$. On a $f(f^{-1}(\{y\})) = \{y\}$ et donc $f^{-1}(\{y\}) \neq \emptyset$. Il existe donc au moins un antécédent à l'élément y , l'application f est surjective.

Exercice 28 ***

Soit $f: E \rightarrow F$ une application. Montrer

$$f \text{ est bijective} \iff \forall A \in \wp(E), f(\complement_E A) = \complement_F f(A)$$

Solution

On raisonne par double implication.

(\Rightarrow) On suppose f bijective.

méthode

|| Soit A et A' deux parties de E . Puisque f est injective, on vérifie :

$$f(A' \setminus A) = f(A') \setminus f(A).$$

En effet, un élément de $f(A' \setminus A)$ est une valeur prise par f sur A' mais ne peut être une valeur prise par f sur A car f est injective. Inversement, un élément de $f(A') \setminus f(A)$ est une valeur prise par f sur un élément de A' qui ne peut pas être élément de A .

En choisissant $A' = E$, la surjectivité de f donne $f(E) = F$ et donc

$$f(\complement_E A) = f(E \setminus A) = f(E) \setminus f(A) = F \setminus f(A) = \complement_F f(A).$$

(\Leftarrow) Montrons que f est injective. Soit x et x' deux éléments distincts de E et considérons la partie $A = \{x\}$. L'élément x' appartient au complémentaire de A et donc

$$f(x') \in f(\complement_E A) = \complement_F f(A) = \complement_F \{f(x)\}.$$

On en déduit $f(x) \neq f(x')$. L'application f est injective.

Montrons que f est aussi surjective. Considérons $A = E$.

$$\complement_F \text{Im}(f) = \complement_F f(E) = f(\complement_E E) = f(\emptyset) = \emptyset.$$

On en déduit $\text{Im}(f) = F$. L'application f est surjective et, finalement, bijective.

1.6.4 Relations binaires

Exercice 29 * (Ordre lexicographique)

Sur $E = \mathbb{R}^2$, on définit une relation binaire \preccurlyeq par

$$(x, y) \preccurlyeq (x', y') \iff x < x' \text{ ou } (x = x' \text{ et } y \leq y').$$

(a) Vérifier que \preccurlyeq définit une relation d'ordre sur E .

(b) S'agit-il d'une relation d'ordre totale ?

Solution

(a) Soit $(x, y) \in E$. On a $x = x$ et $y \leq y$ donc $(x, y) \preccurlyeq (x, y)$. La relation est réflexive.

Soit (x, y) et (x', y') dans E tels que $(x, y) \preccurlyeq (x', y')$ et $(x', y') \preccurlyeq (x, y)$

méthode

On remarque

$$(x, y) \preccurlyeq (x', y') \implies x \leq x'$$

ainsi que

$$(x, y) \preccurlyeq (x', y') \text{ et } x = x' \implies y \leq y'.$$

On a simultanément $x \leq x'$ et $x' \leq x$ donc $x = x'$. La comparaison $(x, y) \preccurlyeq (x', y')$ donne alors $y \leq y'$. On obtient de même $y' \leq y$ donc $y = y'$. On peut alors conclure $(x, y) = (x', y')$ et affirmer que la relation est antisymétrique.

Soit (x, y) , (x', y') et (x'', y'') dans E tels que $(x, y) \preccurlyeq (x', y')$ et $(x', y') \preccurlyeq (x'', y'')$. On a $x \leq x'$ et $x' \leq x''$ donc $x \leq x''$. Poursuivons par disjonction de cas.

Cas : $x < x''$. On peut conclure immédiatement $(x, y) \preccurlyeq (x'', y'')$.

Cas : $x = x''$. On a $x = x' = x''$ et nécessairement $y \leq y'$ et $y' \leq y''$ donc $y \leq y''$. On peut à nouveau affirmer $(x, y) \preccurlyeq (x'', y'')$. La relation est transitive.

Finalement, \preccurlyeq est bien une relation d'ordre sur E .

(b) Vérifions que la relation d'ordre \preccurlyeq est totale. Soit (x, y) et (x', y') deux éléments de E .

Si $x \neq x'$ alors $x < x'$ ou $x' < x$ et donc $(x, y) \preccurlyeq (x', y')$ ou $(x', y') \preccurlyeq (x, y)$.

Si $x = x'$, on compare y et y' . Si $y \leq y'$ alors $(x, y) \preccurlyeq (x', y')$, sinon $(x', y') \preccurlyeq (x, y)$.

Dans tous les cas, on peut comparer les couples (x, y) et (x', y') .

Exercice 30 **

Soit \preccurlyeq la relation binaire définie sur le demi-plan $E = \{(a, b) \in \mathbb{R}^2 \mid a \leq b\}$ par

$$(a, b) \preccurlyeq (a', b') \iff (a, b) = (a', b') \text{ ou } b \leq a'.$$

(a) Montrer que \preccurlyeq est une relation d'ordre sur E .

(b) S'agit-il d'une relation d'ordre totale ?

Solution

(a) La relation \preccurlyeq est évidemment réflexive.

Soit (a, b) , (a', b') et (a'', b'') trois éléments du demi-plan E vérifiant $(a, b) \preccurlyeq (a', b')$ et $(a', b') \preccurlyeq (a'', b'')$. Si deux éléments sont égaux, on a immédiatement $(a, b) \preccurlyeq (a'', b'')$. Sinon, on a $b \leqslant a'$ et $b' \leqslant a''$ donc $b \leqslant a''$ car on sait aussi $a' \leqslant b'$. On trouve alors $(a, b) \preccurlyeq (a'', b'')$ et l'on peut affirmer que la relation est transitive.

Soit (a, b) et (a', b') des éléments du demi-plan E . On suppose $(a, b) \preccurlyeq (a', b')$ et $(a', b') \preccurlyeq (a, b)$. On a alors

$$((a, b) = (a', b') \text{ ou } b \leqslant a') \text{ et } ((a', b') = (a, b) \text{ ou } b' \leqslant a).$$

En factorisant par distributivité, on obtient

$$(a, b) = (a', b') \text{ ou } (b \leqslant a' \text{ et } b' \leqslant a). \quad (*)$$

Cependant, on a aussi $a \leqslant b$ et $a' \leqslant b'$ de sorte que $b \leqslant a'$ et $b' \leqslant a$ entraînent

$$a \leqslant b \leqslant a' \leqslant b' \leqslant a$$

et donc $(a, b) = (a, a) = (a', b')$. L'assertion $(*)$ se résume alors en $(a, b) = (a', b')$ et la relation est antisymétrique.

Finalement, \preccurlyeq est une relation d'ordre.

(b) Vérifions que la relation d'ordre \preccurlyeq n'est pas totale.

méthode

On établit qu'une relation d'ordre n'est pas totale en exhibant deux éléments qui ne sont pas comparables.

Les couples $(1, 3)$ et $(2, 4)$ sont éléments de E mais ne sont pas comparables :

$$(1, 3) \not\preccurlyeq (2, 4) \quad \text{et} \quad (2, 4) \not\preccurlyeq (1, 3).$$

La relation d'ordre¹ n'est que partielle.

Exercice 31 *

Soit $f: E \rightarrow F$ une application. On définit une relation binaire sur E par :

$$x \mathcal{R} y \iff f(x) = f(y).$$

(a) Montrer que \mathcal{R} définit une relation d'équivalence sur E .

(b) Exprimer la classe d'équivalence de x élément quelconque de E .

1. Si a et b se comprennent comme les extrémités d'un segment $[a; b]$, la relation $(a, b) \preccurlyeq (a', b')$ signifie que les segments $[a; b]$ sont confondus ou qu'ils « se suivent ».

Solution

(a) Soit x dans E . On a $f(x) = f(x)$ et donc $x \mathcal{R} x$. La relation est réflexive.

Soit x et y dans E . Si $x \mathcal{R} y$ alors $f(x) = f(y)$ et donc $f(y) = f(x)$ ce qui permet d'écrire $y \mathcal{R} x$. La relation est symétrique.

Soit x , y et z dans E . Si $x \mathcal{R} y$ et $y \mathcal{R} z$ alors $f(x) = f(y) = f(z)$ et donc $x \mathcal{R} z$. La relation est transitive.

Finalement, \mathcal{R} est une relation d'équivalence.

(b) méthode

|| On vérifie que les éléments en relation avec x sont les antécédents de $f(x)$.

Soit $y \in E$.

$$\begin{aligned} x \mathcal{R} y &\iff f(x) = f(y) \\ &\iff f(y) \in \{f(x)\} \\ &\iff y \in f^{-1}(\{f(x)\}). \end{aligned}$$

On a donc

$$\text{Cl}(x) = f^{-1}(\{f(x)\}).$$

1.7 Exercices d'approfondissement

Exercice 32 *

Soit A et B des parties de E .

Discuter et résoudre l'équation $A \cup X = B$ d'inconnue $X \in \wp(E)$.

Solution**méthode**

|| On raisonne par analyse-synthèse.

Analyse : Supposons X solution de l'équation $A \cup X = B$. On a nécessairement A inclus dans B et X inclus dans B . Cependant, ceci ne suffit pas, il faut aussi que les éléments de B qui ne sont pas dans A se retrouvent dans X et donc $B \setminus A \subset X$.

Résumons, si l'équation $A \cup X = B$ admet une solution X , nécessairement

$$A \subset B \quad \text{et} \quad B \setminus A \subset X \subset B.$$

Synthèse : Distinguons deux cas :

Cas : $A \not\subset B$. Il n'y a pas de solutions à l'équation $A \cup X = B$.

Cas : $A \subset B$. Pour toute partie X vérifiant $B \setminus A \subset X \subset B$, on vérifie par double inclusion $A \cup X = B$ et donc X est solution. L'ensemble des solutions de l'équation est alors

$$\{X \in \wp(E) \mid B \setminus A \subset X \subset B\}.$$

Exercice 33 **

Soit A et B deux parties d'un ensemble E et

$$\begin{cases} \wp(E) \rightarrow \wp(A) \times \wp(B) \\ f: X \mapsto (X \cap A, X \cap B). \end{cases}$$

- (a) Montrer que f est injective si, et seulement si, $A \cup B = E$.
- (b) À quelle condition la fonction f est-elle surjective ?

Solution

- (a) Supposons f injective.

méthode

|| Lorsqu'une fonction f est injective, il suffit d'établir $f(x) = f(y)$ pour pouvoir affirmer $x = y$.

On a $f(E) = (A, B)$ et $f(A \cup B) = (A, B)$. Par injectivité, on peut affirmer $E = A \cup B$. Inversement, supposons $A \cup B = E$. Soit X et Y dans $\wp(E)$ telles que $f(X) = f(Y)$. On a donc

$$\begin{cases} X \cap A = Y \cap A \\ X \cap B = Y \cap B. \end{cases}$$

En écrivant $X = X \cap E = X \cap (A \cup B)$, il vient par distributivité

$$X = (X \cap A) \cup (X \cap B) = (Y \cap A) \cup (Y \cap B) = Y \cap (A \cup B) = Y.$$

L'application f est donc injective.

(b) méthode

|| La surjectivité est une notion « quelque peu duale » de la notion d'injectivité. En renversant la condition d'injectivité $A \cup B = E$, on peut s'attendre à ce que la condition de surjectivité soit $A \cap B = \emptyset$.

Supposons f surjective. L'élément (A, \emptyset) de $\wp(A) \times \wp(B)$ possède un antécédent X dans $\wp(E)$. Pour celui-ci, on a $X \cap A = A$ et $X \cap B = \emptyset$. Par commutativité et associativité, on obtient alors

$$A \cap B = (X \cap A) \cap B = A \cap (X \cap B) = A \cap \emptyset = \emptyset.$$

Inversement, supposons $A \cap B = \emptyset$. Soit (A', B') un élément arbitraire de $\wp(A) \times \wp(B)$. Pour $X = A' \cup B'$, on a

$$f(X) = ((\underbrace{A' \cap A}_{= \emptyset}) \cup (\underbrace{B' \cap A}_{= \emptyset}), (\underbrace{A' \cap B}_{= \emptyset}) \cup (\underbrace{B' \cap B}_{= \emptyset})) = (A', B').$$

L'application f est donc surjective.

Exercice 34 **

Soit E et F des ensembles non vides. Montrer qu'il existe une injection de E dans F si, et seulement si, il existe une surjection de F sur E .

Solution**méthode**

À partir d'une injection $i: E \rightarrow F$, on construit une surjection $s: F \rightarrow E$ vérifiant $s \circ i = \text{Id}_E$ et inversement.

Supposons qu'il existe une injection $i: E \rightarrow F$. La restriction de celle-ci à son image définit¹ une bijection $j: E \rightarrow \text{Im}(i)$. Choisissons arbitrairement un élément a dans l'ensemble non vide E et considérons l'application $s: F \rightarrow E$ déterminée par

$$s(x) = \begin{cases} j^{-1}(x) & \text{si } x \in \text{Im}(i) \\ a & \text{sinon.} \end{cases}$$

L'application s vérifie $s \circ i = \text{Id}_E$, elle est donc surjective².

Inversement, supposons qu'il existe une surjection $s: F \rightarrow E$. Chaque x appartenant à E possède au moins un antécédent par s . Parmi ceux-ci, choisissons en un et posons $i(x)$ égal à celui-ci. On définit ainsi une application $i: E \rightarrow F$ vérifiant $s \circ i = \text{Id}_E$, l'application i est donc injective³.

Exercice 35 ***

Soit E un ensemble. Montrer qu'il n'existe pas d'applications surjectives de E vers $\wp(E)$.

Solution**méthode**

En s'inspirant du paradoxe du menteur⁴, on introduit une partie de E qui ne peut avoir d'antécédent par une application donnée $f: E \rightarrow \wp(E)$.

Soit $f: E \rightarrow \wp(E)$. Considérons la partie A de E définie par

$$A = \{x \in E \mid x \notin f(x)\}.$$

Montrons par l'absurde que cette partie A ne peut avoir d'antécédent par f . Supposons qu'il existe x dans E tel que $f(x) = A$ et interrogeons-nous sur l'appartenance de x à A .

Si x est élément de A alors x est élément de $f(x)$ et donc x n'est pas élément de A puisque, par définition, les éléments de A sont les x tels que $x \notin f(x)$.

Si x n'est pas élément de A alors x n'est pas élément de $f(x)$ et donc x est élément de A . C'est absurde !

L'application f ne peut pas être surjective.

1. Voir sujet 19 p. 31.

2. Si la composée $g \circ f$ est surjective, l'application g est surjective, voir sujet 23 p. 34.

3. Si la composée $g \circ f$ est injective, l'application f est injective, voir encore le sujet 23 p. 34.

4. Ni un menteur, ni une personne affirmant la vérité, ne peuvent dire « je suis un menteur ».

Exercice 36 ***

Soit $f: \wp(E) \rightarrow \wp(E)$ une application croissante au sens de l'inclusion, c'est-à-dire une application vérifiant

$$\forall (A, B) \in \wp(E)^2, \quad A \subset B \implies f(A) \subset f(B).$$

Montrer qu'il existe une partie A de E vérifiant $f(A) = A$.

Solution**méthode**

|| On introduit la plus grande partie A vérifiant $A \subset f(A)$.

Considérons B la réunion de toutes les parties A de E vérifiant $A \subset f(A)$:

$$B = \bigcup_{A \in \mathcal{S}} A \quad \text{avec} \quad \mathcal{S} = \{A \in \wp(E) \mid A \subset f(A)\}.$$

Par définition de la partie B , on a la propriété

$$A \subset f(A) \implies A \subset B. \tag{*}$$

Pour toute partie A de \mathcal{S} , on a $A \subset B$ et donc $f(A) \subset f(B)$ par croissance de f . Or on a aussi $A \subset f(A)$ et donc $A \subset f(B)$. On en déduit $B \subset f(B)$ car B est la réunion des parties A éléments de \mathcal{S} .

De plus, on a alors par croissance de f l'inclusion $f(B) \subset f(f(B))$ et donc $f(B) \subset B$ en vertu de la propriété (*) utilisée par $A = f(B)$.

Finalement¹, on a obtenu par double inclusion $f(B) = B$: on dit que B est un point fixe² de f .

1. On aurait aussi pu introduire la plus petite partie A vérifiant $A \subset f(A)$, à savoir, l'intersection de toutes les parties ayant cette propriété.

2. On pourra mettre en résonance cet exercice avec le sujet 28 du chapitre 1 de l'ouvrage *Exercices d'analyse MPSI* dans la même collection.

CHAPITRE 2

Calculs algébriques

2.1 Les entiers naturels et le principe de récurrence

2.1.1 Les entiers naturels

On désigne par \mathbb{N} l'ensemble infini des entiers naturels : $0, 1, 2, \dots$

Cet ensemble est muni d'une relation d'ordre totale \leq vérifiant :

Théorème 1

Toute partie non vide de \mathbb{N} possède un plus petit élément¹.

Lorsqu'il existe au moins un naturel vérifiant une propriété donnée, ce résultat permet d'affirmer l'existence d'un plus petit entier naturel vérifiant cette propriété. Aussi, on peut établir que toute partie non vide et majorée de \mathbb{N} possède un plus grand élément².

Théorème 2 (Principe de récurrence)

Si E est une partie de \mathbb{N} contenant 0 et vérifiant

$$\forall p \in \mathbb{N}, \quad p \in E \implies p + 1 \in E$$

alors $E = \mathbb{N}$.

Ce théorème permet d'établir la validité des raisonnements par récurrence qui suivent.

1. On dit que \mathbb{N} muni de \leq est un ensemble *bien ordonné*.

2. Plus généralement, toute partie non vide et minorée de \mathbb{Z} (resp. majorée) admet un plus petit élément (resp. un plus grand élément).

2.1.2 Les raisonnements par récurrence

Soit $n_0 \in \mathbb{N}$ et $\mathcal{P}(n)$ une assertion dépendant d'un paramètre entier $n \geq n_0$.

Théorème 3 (Récurrence simple)

Si

$$\begin{cases} \mathcal{P}(n_0) \text{ est vérifié (initialisation)} \\ \forall n \geq n_0, \mathcal{P}(n) \implies \mathcal{P}(n+1) \text{ (héritéité)} \end{cases}$$

alors $\mathcal{P}(n)$ est vraie pour tout $n \geq n_0$.

Il est quelquefois nécessaire de disposer de l'hypothèse de récurrence sur plusieurs rangs précédents afin d'établir la propriété au rang supérieur. Dans ce cas, on rédige une récurrence forte :

Théorème 4 (Récurrence forte)

Si

$$\begin{cases} \mathcal{P}(n_0) \text{ est vérifiée (initialisation)} \\ \forall n \geq n_0, (\mathcal{P}(n_0), \mathcal{P}(n_0+1), \dots, \mathcal{P}(n)) \implies \mathcal{P}(n+1) \text{ (héritéité forte)} \end{cases}$$

alors $\mathcal{P}(n)$ est vraie pour tout $n \geq n_0$.

On peut aussi énoncer des principes de récurrence multiples, parmi lesquels figure la récurrence double : si

$$\begin{cases} \mathcal{P}(n_0) \text{ et } \mathcal{P}(n_0+1) \text{ est vérifiée (initialisation double)} \\ \forall n \geq n_0, (\mathcal{P}(n) \text{ et } \mathcal{P}(n+1)) \implies \mathcal{P}(n+2) \text{ (héritéité double)} \end{cases}$$

alors $\mathcal{P}(n)$ est vraie pour tout $n \geq n_0$.

2.2 Sommes et produits

2.2.1 Sommes numériques

Soit n un entier naturel non nul et $(a_i)_{1 \leq i \leq n} = (a_1, a_2, \dots, a_n)$ une famille de nombres réels ou complexes.

Définition

On appelle *somme* de la famille $(a_i)_{1 \leq i \leq n}$ le nombre

$$\sum_{i=1}^n a_i \stackrel{\text{déf}}{=} a_1 + a_2 + \dots + a_n.$$

Dans la description de cette somme, l'indice i joue un rôle muet.

Plus généralement, si $(a_i)_{i \in I}$ désigne une famille finie de nombres réels ou complexes, on introduit la somme des éléments de cette famille notée

$$\sum_{i \in I} a_i.$$

Lorsque l'ensemble d'indexation I est vide, on convient que cette somme est nulle.

Théorème 5

Si $(a_i)_{i \in I}$ et $(b_i)_{i \in I}$ sont des familles finies de nombres réels ou complexes alors, pour tout λ réel ou complexe,

$$\sum_{i \in I} \lambda a_i = \lambda \sum_{i \in I} a_i \quad \text{et} \quad \sum_{i \in I} (a_i + b_i) = \sum_{i \in I} a_i + \sum_{i \in I} b_i.$$

Théorème 6

Si $(a_i)_{i \in I}$ est une famille de réels positifs alors

$$\sum_{i \in I} a_i \geq 0.$$

De plus, cette somme n'est nulle que si tous les a_i sont nuls.

2.2.2 Sommes remarquables

On retient les sommes suivantes, pour¹ $n \in \mathbb{N}^*$,

$$\begin{aligned} \sum_{k=1}^n k &= 1 + 2 + \cdots + n = \frac{n(n+1)}{2} \\ \sum_{k=1}^n k^2 &= 1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6} \end{aligned}$$

ainsi que la *somme géométrique*, pour $n \in \mathbb{N}$,

$$\sum_{k=0}^n q^k = 1 + q + q^2 + \cdots + q^n = \frac{1 - q^{n+1}}{1 - q} \quad \text{pour } q \neq 1.$$

Plus généralement, la somme des termes successifs d'une suite géométrique de raison différente de 1 s'obtient par la formule :

$$(\text{premier terme}) \times \frac{1 - \text{raison}^{\text{nombre de termes}}}{1 - \text{raison}}.$$

1. Ces deux formules sont aussi valables pour $n = 0$ en rappelant qu'une « somme vide » est nulle.

On rencontre aussi fréquemment des *sommes télescopiques*¹

$$\sum_{i=0}^n (a_{i+1} - a_i) = a_{n+1} - a_0.$$

2.2.3 Réorganisation de la sommation

Théorème 7 (Changement d'indice)

Soit $(a_j)_{j \in J}$ une famille finie de nombres réels ou complexes. Si $\varphi: I \rightarrow J$ est une application bijective alors

$$\sum_{j \in J} a_j = \sum_{i \in I} a_{\varphi(i)}$$

La transformation d'une somme en l'autre est appelée *changement d'indice* défini par la relation $j = \varphi(i)$. En particulier, lorsque φ est de la forme $i \mapsto i + C^{te}$, on parle de *glissement d'indice*. C'est le cas de la transformation suivante :

$$\sum_{j=0}^n a_j = \sum_{i=1}^{n+1} a_{i-1}.$$

Lorsque φ est de la forme $i \mapsto C^{te} - i$, on parle de *renversement d'indice* comme l'illustre la transformation ci-dessous

$$\sum_{j=0}^n a_j = \sum_{i=0}^n a_{n-i}.$$

On vérifie à chaque fois que les termes sommés sont parfaitement identiques².

Théorème 8 (Regroupement de termes)

Soit $(a_i)_{i \in I}$ une famille finie de nombres réels ou complexes. Si I est la réunion de sous-ensembles I_1, I_2, \dots, I_p , deux à deux disjoints, on a

$$\sum_{i \in I} a_i = \sum_{i \in I_1} a_i + \dots + \sum_{i \in I_p} a_i = \sum_{j=1}^p \left(\sum_{i \in I_j} a_i \right).$$

Il importe que les ensembles I_j soient deux à deux disjoints afin de ne pas adjoindre de nouveaux termes à la somme. En particulier, lorsque l'on découpe une somme en deux, on sera attentif à ne pas dédoubler le terme frontière : pour $p \in [1; n]$,

$$\sum_{i=1}^n a_i = \sum_{i=1}^p a_i + \sum_{i=p+1}^n a_i$$

1. La somme est alors égale à « la différence des extrêmes ».

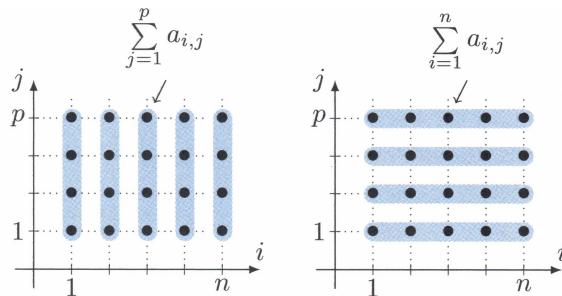
2. Poser $j = 2i$ n'est pas un changement d'indice correct : la somme $\sum_{j=0}^{2n} a_j$ comporte des termes d'indices impairs alors que $\sum_{i=0}^n a_{2i}$ n'en comporte aucun.

2.2.4 Sommes doubles

Lorsque l'ensemble d'indexation d'une somme correspond à un ensemble de couples (i, j) , on dit que l'on a affaire à une *somme double*.

Parmi celles-ci figurent les *sommes rectangulaires* dont le calcul peut être organisé par regroupement de termes

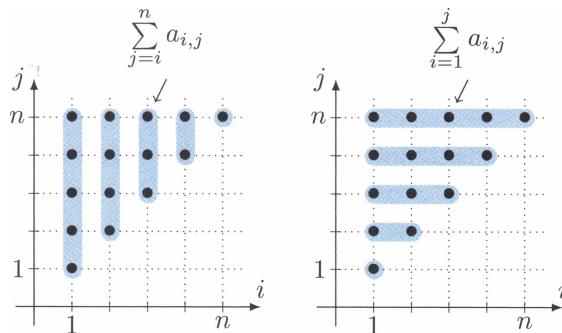
$$\sum_{(i,j) \in \llbracket 1;n \rrbracket \times \llbracket 1;p \rrbracket} a_{i,j} = \sum_{i=1}^n \left(\sum_{j=1}^p a_{i,j} \right) = \sum_{j=1}^p \left(\sum_{i=1}^n a_{i,j} \right).$$



Deux organisations du calcul d'une somme double.

On rencontre aussi fréquemment des *sommes triangulaires* dont le calcul peut aussi être organisé par regroupement de termes¹

$$\sum_{1 \leqslant i \leqslant j \leqslant n} a_{i,j} = \sum_{i=1}^n \left(\sum_{j=i}^n a_{i,j} \right) = \sum_{j=1}^n \left(\sum_{i=1}^j a_{i,j} \right).$$



Deux organisations du calcul d'une somme triangulaire.

¹. Dans les deux écritures proposées, on sera attentif à l'ordre dans lequel les deux sommes sont écrites car la somme contenue dépend de l'indice de la somme contenante.

2.2.5 Produits numériques

Définition

On appelle *produit* d'une famille $(a_i)_{1 \leq i \leq n}$ de nombres réels ou complexes le nombre

$$\prod_{i=1}^n a_i \stackrel{\text{déf}}{=} a_1 \times a_2 \times \cdots \times a_n.$$

Plus généralement, on définit aussi le produit d'une famille finie $(a_i)_{i \in I}$ de nombres réels ou complexes

$$\prod_{i \in I} a_i$$

Lorsque l'ensemble d'indexation I est vide, on convient que le produit vaut 1.

Théorème 9

Si $(a_i)_{i \in I}$ et $(b_i)_{i \in I}$ sont des familles finies de nombres réels ou complexes alors, pour tout λ réel ou complexe,

$$\prod_{i \in I} \lambda a_i = \lambda^n \prod_{i \in I} a_i \quad \text{et} \quad \prod_{i \in I} (a_i b_i) = \left(\prod_{i \in I} a_i \right) \left(\prod_{i \in I} b_i \right)$$

avec n le nombre d'éléments de I .

Changement d'indice et regroupement de termes sont possibles pour les produits. Parmi les produits remarquables, citons les *produits télescopiques* :

$$\prod_{i=1}^n \frac{a_{i+1}}{a_i} = \frac{a_{n+1}}{a_1}.$$

2.3 Formules du binôme et de factorisation

2.3.1 Coefficients binomiaux

Définition

On appelle *factorielle* d'un naturel n le produit des entiers allant de 1 à n :

$$n! \stackrel{\text{def}}{=} \prod_{k=1}^n k = 1 \times 2 \times \cdots \times n.$$

Le tableau ci-dessous figure les valeurs des dix premières factorielles¹ :

n	0	1	2	3	4	5	6	7	8	9
$n!$	1	1	2	6	24	120	720	5 040	40 320	362 880

1. On remarquera que la factorielle de 0 vaut 1 car correspond à un produit vide.

Définition

Soit $p \in \mathbb{Z}$ et $n \in \mathbb{N}$. On définit le *coefficient binomial* p parmi n par

$$\binom{n}{p} \stackrel{\text{déf}}{=} \frac{n!}{p!(n-p)!} \text{ si } p \in [0; n] \text{ et } \binom{n}{p} \stackrel{\text{déf}}{=} 0 \text{ sinon.}$$

Théorème 10 (Symétrie et formule du triangle de Pascal)

Pour tout $p \in \mathbb{Z}$ et tout $n \in \mathbb{N}$, on a

$$\binom{n}{p} = \binom{n}{n-p} \text{ et } \binom{n}{p} + \binom{n}{p+1} = \binom{n+1}{p+1}.$$

On retient les valeurs remarquables

$$\binom{n}{0} = \binom{n}{n} = 1, \quad \binom{n}{1} = \binom{n}{n-1} = n \quad \text{et} \quad \binom{n}{2} = \binom{n}{n-2} = \frac{n(n-1)}{2}.$$

2.3.2 Formule du binôme**Théorème 11 (Formule du binôme de Newton¹)**

Soit $n \in \mathbb{N}$. Pour tous a et b nombres réels ou complexes

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k$$

En particulier, on retrouve l'identité remarquable $(a+b)^2 = a^2 + 2ab + b^2$ et l'on peut généraliser celle-ci :

$$\begin{aligned} (a+b)^3 &= a^3 + 3a^2b + 3ab^2 + b^3, \\ (a+b)^4 &= a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4, \\ (a+b)^5 &= a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5, \dots \end{aligned}$$

En remplaçant b par $-b$, on peut aussi développer $(a-b)^n$.

2.3.3 Formule de factorisation**Théorème 12 (Formule de factorisation géométrique)**

Soit $n \in \mathbb{N}$. Pour tous a et b nombres réels ou complexes

$$a^n - b^n = (a - b) \sum_{k=0}^{n-1} a^{n-1-k} b^k.$$

1. Par symétrie des rôles joués par a et b , on peut aussi énoncer une formule où les exposants de a et b dans la somme sont échangés.

En particulier, on retrouve l'identité remarquable $a^2 - b^2 = (a - b)(a + b)$ et l'on peut généraliser celle-ci :

$$\begin{aligned} a^3 - b^3 &= (a - b)(a^2 + ab + b^2), \\ a^4 - b^4 &= (a - b)(a^3 + a^2b + ab^2 + b^3), \\ a^5 - b^5 &= (a - b)(a^4 + a^3b + a^2b^2 + ab^3 + b^4), \dots \end{aligned}$$

2.4 Systèmes d'équations linéaires

Ici, paramètres et inconnues sont des nombres réels ou complexes.

2.4.1 Présentation

Définition

On appelle *système d'équations linéaires* à n équations et p inconnues tout système de la forme

$$(\Sigma): \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,p}x_p = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,p}x_p = b_2 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,p}x_p = b_n \end{cases}$$

d'inconnues x_1, x_2, \dots, x_p et où les $a_{i,j}$ et les b_i désignent des paramètres.

Théorème 13

Le système est transformé en un système équivalent lorsque :

- on échange les équations d'indices i et j (on note $L_i \leftrightarrow L_j$);
- on multiplie l'équation d'indice i par un nombre α non nul (on note $L_i \leftarrow \alpha L_i$);
- on ajoute à l'équation d'indice i l'équation d'indice j (avec $j \neq i$) multipliée par un facteur λ (on note $L_i \leftarrow L_i + \lambda L_j$).

2.4.2 Algorithme du pivot

Résolvons le système d'équations linéaires

$$(\Sigma): \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,p}x_p = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,p}x_p = b_2 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,p}x_p = b_n. \end{cases}$$

Étape 1 : (Détermination du premier pivot)

Si les coefficients $a_{1,1}, \dots, a_{n,1}$ devant l'inconnue x_1 sont tous nuls, on réordonne les inconnues pour se ramener à une situation où ceci n'a pas lieu. Quitte à échanger deux équations, on peut alors supposer $a_{1,1} = a_{1,1} \neq 0$ (c'est le premier pivot).

Étape 2 : (Annulation des coefficients sous le premier pivot)

Par des opérations du type $L_i \leftarrow L_i + \lambda L_1$ avec λ bien choisi, on annule les coefficients devant x_1 dans les équations d'indices allant de 2 à n .

Par équivalences, on a alors transformé le système étudié en un système de la forme

$$\left\{ \begin{array}{l} p_1 x_1 + a'_{1,2} x_2 + \cdots + a'_{1,p} x_p = b'_1 \\ a'_{2,2} x_2 + \cdots + a'_{2,p} x_p = b'_2 \\ \vdots \\ a'_{n,2} x_2 + \cdots + a'_{n,p} x_p = b'_n. \end{array} \right.$$

Etapes suivantes :

On reprend les étapes précédentes en opérant avec les équations allant de 2 à n , pour déterminer un deuxième pivot, puis les équations 3 à n , etc.

Après avoir éventuellement renommé les inconnues, on parvient à un système de la forme :

$$\left\{ \begin{array}{l} p_1 x_1 + \cdots + a''_{1,r} x_r + a''_{1,r+1} x_{r+1} + \cdots + a''_{1,p} x_p = b''_1 \\ \vdots \\ p_r x_r + a''_{r,r+1} x_{r+1} + \cdots + a''_{r,p} x_p = b''_r \\ 0 = b''_{r+1} \\ \vdots \\ 0 = b''_n \end{array} \right.$$

avec p_1, \dots, p_r des nombres non nuls appelés *pivots*.

Définition

|| Les r premières équations sont les *équations principales* et les suivantes sont les *équations de compatibilité*.

Si l'une des équations de compatibilité est fausse, le système n'a pas de solutions.

Si toutes les équations de compatibilité sont vérifiées, le système (Σ) est équivalent au *système triangulaire*

$$\left\{ \begin{array}{l} p_1 x_1 + \cdots + a''_{1,r} x_r = b''_1 - (a''_{1,r+1} x_{r+1} + \cdots + a''_{1,p} x_p) \\ \vdots \\ p_r x_r = b''_r - (a''_{r,r+1} x_{r+1} + \cdots + a''_{r,p} x_p). \end{array} \right.$$

Ce système se résout en cascade en partant de la dernière équation pour remonter jusqu'à la première. On exprime alors séparément les *inconnues principales* x_1, \dots, x_r en fonction des *inconnues paramètres* x_{r+1}, \dots, x_p :

$$\left\{ \begin{array}{l} x_1 = \beta_1 + \alpha_{1,r+1} x_{r+1} + \cdots + \alpha_{1,p} x_p \\ \vdots \\ x_r = \beta_r + \alpha_{r,r+1} x_{r+1} + \cdots + \alpha_{r,p} x_p \end{array} \right.$$

et ceci suffit à décrire l'ensemble des solutions du système.

2.5 Exercices d'apprentissage

2.5.1 Principe de récurrence

Exercice 1

Exprimer simplement le terme général de la suite (u_n) déterminée par :

- (a) $u_0 = 0$ et $\forall n \in \mathbb{N}$, $u_{n+1} = u_n + 2n + 1$.
- (b) $u_0 = 1$, $u_1 = 1$ et $\forall n \in \mathbb{N}$, $u_{n+2} - (n+1)(u_{n+1} + u_n) = 0$.
- (c) $u_0 = 1$ et $\forall n \in \mathbb{N}$, $u_{n+1} = u_0 + u_1 + \dots + u_n$.

Solution

méthode

On calcule les premiers termes de chaque suite afin de proposer une formule « crédible ». On valide ensuite celle-ci par une récurrence adaptée.

- (a) Les premiers termes de la suite (u_n) sont

n	0	1	2	3	4	5
u_n	0	1	4	9	16	25

Il semble que le terme u_n soit égal à n^2 . On le vérifie par une récurrence simple.

Pour $n = 0$, on a bien $u_0 = 0^2$.

Supposons l'égalité $u_n = n^2$ vraie à un certain¹ rang $n \geq 0$ et vérifions que celle-ci a lieu au rang suivant :

$$u_{n+1} = u_n + 2n + 1 = n^2 + 2n + 1 = (n+1)^2.$$

La récurrence est établie.

- (b) Les premiers termes de la suite (u_n) sont

n	0	1	2	3	4	5
u_n	1	1	2	6	24	120

Il semble que u_n soit égal à $n!$.

méthode

|| Puisque les termes de la suite se calculent à partir des deux rangs qui précèdent, on raisonne par récurrence double².

1. On rédige « à un certain rang » et non « pour tout rang ».

2. Pour que le mécanisme de récurrence progresse correctement, il importe de supposer la propriété vraie aux rangs n et $n+1$ pour l'établir au rang $n+2$ suivant. Supposer cette propriété aux rangs n et $n+2$ pour l'établir au rang intermédiaire $n+1$ est incorrect.

On commence par une initialisation double : L'égalité $u_n = n!$ est vérifiée aux rangs initiaux $n = 0$ et $n = 1$.

On poursuit avec une hypothèse de récurrence double : on suppose l'identité $u_n = n!$ vraie aux rangs n et $n + 1$ (avec $n \geq 0$). On vérifie ensuite qu'elle est vraie au rang $n + 2$:

$$u_{n+2} = (n+1)(u_{n+1} + u_n) = (n+1)\underbrace{((n+1)! + n!)}_{=(n+2).n!} = (n+2)!$$

La récurrence est établie.

(c) Les premiers termes de la suite (u_n) sont

n	0	1	2	3	4	5
u_n	1	1	2	4	8	16

Il semble que u_n soit égal à 2^{n-1} à partir du rang 1.

méthode

Puisque le terme u_n se calcule à partir de tous les termes précédents, on raisonne par récurrence forte.

L'égalité $u_n = 2^{n-1}$ est vérifiée pour $n = 1$.

Supposons l'égalité $u_k = 2^{k-1}$ vraie pour tous les rangs k allant de 1 à n (avec $n \geq 1$). Au rang suivant, on obtient par sommation géométrique

$$u_{n+1} = u_0 + u_1 + \cdots + u_n = 1 + \sum_{k=1}^n 2^{k-1} = 1 + \frac{1 - 2^n}{1 - 2} = 2^n.$$

La récurrence est établie.

2.5.2 Sommes et produits

Exercice 2

Soit $n \in \mathbb{N}^*$. Calculer les sommes suivantes :

- | | |
|---|---|
| (a) $A_n = 1 \times 2 + 2 \times 3 + \cdots + n \times (n+1)$ | (b) $B_n = 1 \times n + 2 \times (n-1) + \cdots + n \times 1$ |
| (c) $C_n = 1^2 + 3^2 + \cdots + (2n+1)^2$ | (d) $D_n = \frac{1}{1 \times 2} + \frac{1}{2 \times 3} + \cdots + \frac{1}{n(n+1)}$ |

Solution

méthode

On exprime les sommes étudiées avec le symbole \sum et l'on exploite si besoin les égalités

$$\sum_{k=1}^n k = \frac{n(n+1)}{2} \quad \text{et} \quad \sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}.$$

(a) Les termes sommés sont les produits de deux entiers successifs : ce sont les $k(k+1)$ pour k allant de 1 à n . On peut alors écrire

$$A_n = 1 \times 2 + 2 \times 3 + \cdots + n \times (n+1) = \sum_{k=1}^n k(k+1) = \sum_{k=1}^n (k^2 + k).$$

En séparant la somme en deux, il vient

$$A_n = \sum_{k=1}^n k^2 + \sum_{k=1}^n k = \frac{n(n+1)}{2} + \frac{n(n+1)(2n+1)}{6} = \frac{n(n+1)(n+2)}{3}$$

(b) Les termes sommés sont des produits d'entiers de somme égale à $n+1$: ce sont les $k(n+1-k)$ pour k allant de 1 à n . On peut alors écrire

$$B_n = 1 \times n + 2 \times (n-1) + \cdots + n \times 1 = \sum_{k=1}^n k(n+1-k) = \sum_{k=1}^n ((n+1)k - k^2).$$

En séparant la somme en deux

$$B_n = \sum_{k=1}^n \underbrace{(n+1)}_{\text{ne dépend pas de } k} k - \sum_{k=1}^n k^2 = (n+1) \sum_{k=1}^n k - \sum_{k=1}^n k^2 = \frac{n(n+1)(n+2)}{6}$$

(c) Les termes sommés sont les carrés des entiers impairs : ce sont les $(2k+1)^2$ pour k allant de 0 jusqu'à n .

$$C_n = 1^2 + 3^2 + \cdots + (2n+1)^2 = \sum_{k=0}^n (2k+1)^2 = \sum_{k=0}^n (4k^2 + 4k + 1).$$

En séparant la somme en trois

$$C_n = 4 \sum_{k=0}^n k^2 + 4 \sum_{k=0}^n k + \underbrace{\sum_{k=0}^n 1}_{n+1 \text{ termes}} = 4 \sum_{k=0}^n k^2 + 4 \sum_{k=0}^n k + n + 1 = \frac{(n+1)(2n+1)(2n+3)}{3}.$$

(d) Les termes sommés sont les inverses des produits de deux entiers successifs : ce sont les $\frac{1}{k(k+1)}$ pour k allant de 1 à n .

$$D_n = \frac{1}{1 \times 2} + \frac{1}{2 \times 3} + \cdots + \frac{1}{n(n+1)} = \sum_{k=1}^n \frac{1}{k(k+1)}.$$

méthode

|| On fait apparaître une somme télescopique en écrivant $1 = (k+1) - k$.

On transforme l'écriture du terme sommé

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

On peut alors achever le calcul en simplifiant les termes intermédiaires

$$\begin{aligned} D_n &= \underbrace{\left(\frac{1}{1} - \frac{1}{2} \right)}_{k=1} + \underbrace{\left(\frac{1}{2} - \frac{1}{3} \right)}_{k=2} + \cdots + \underbrace{\left(\frac{1}{n} - \frac{1}{n+1} \right)}_{k=n} \\ &= \frac{1}{1} + \left(-\frac{1}{2} + \frac{1}{2} \right) + \left(-\frac{1}{3} + \frac{1}{3} \right) + \cdots + \left(\frac{1}{n} - \frac{1}{n} \right) - \frac{1}{n+1} \\ &= 1 - \frac{1}{n+1} = \frac{n}{n+1}. \end{aligned}$$

Exercice 3

Soit $n \in \mathbb{N}$. Calculer les sommes suivantes

(a) $A_n = \sum_{k=0}^n (-1)^k x^{2k}$	(b) $B_n = \sum_{k=1}^{2n} (-1)^k k^3$
(c) $C_n = \sum_{1 \leq i, j \leq n} (i+j)$	(d) $D_n = \sum_{1 \leq i \leq j \leq n} (i+j)$

Solution

(a) méthode

|| On reconnaît une somme géométrique.

On écrit

$$A_n = \sum_{k=0}^n (-1)^k x^{2k} = \sum_{k=0}^n (-x^2)^k.$$

Il s'agit d'une somme géométrique de raison $-x^2 \neq 1$:

$$\begin{aligned} A_n &= (\text{premier terme}) \times \frac{1 - \text{raison}^{\text{nombre de termes}}}{1 - \text{raison}} \\ &= 1 \times \frac{1 - (-x^2)^{n+1}}{1 + x^2} = \frac{1 + (-1)^n x^{2n+2}}{1 + x^2}. \end{aligned}$$

(b) méthode

|| On scinde la somme en deux (Th. 8 p. 50) selon la parité de k afin de résoudre la puissance de (-1) .

$$B_n = \sum_{k=1}^{2n} (-1)^k k^3 = \sum_{\substack{k=1 \\ k \text{ pair}}}^{2n} (-1)^k k^3 + \sum_{\substack{k=1 \\ k \text{ impair}}}^{2n} (-1)^k k^3.$$

Les indices k pairs de la première somme peuvent s'écrire $2p$ pour p allant de 1 à n alors que les indices k impairs de la seconde somme s'écrivent $2p - 1$ pour les mêmes valeurs de p . On obtient alors

$$B_n = \sum_{p=1}^n \underbrace{(-1)^{2p}(2p)^3}_{=1} + \sum_{p=1}^n \underbrace{(-1)^{2p-1}(2p-1)^3}_{=-1}.$$

On peut ensuite combiner les deux sommes en une seule puisque la plage d'indexation est identique

$$\begin{aligned} B_n &= \sum_{p=1}^n (2p)^3 - \sum_{p=1}^n (2p-1)^3 = \sum_{p=1}^n ((2p)^3 - (2p-1)^3) \\ &= \sum_{p=1}^n (12p^2 - 6p + 1) = n^2(4n + 3). \end{aligned}$$

(c) méthode

Il s'agit d'une somme double rectangulaire, on l'exprime comme deux sommes emboîtées.

$$C_n = \sum_{1 \leq i \leq j \leq n} (i+j) = \sum_{i=1}^n \left(\sum_{j=1}^n (i+j) \right) = \underbrace{\sum_{i=1}^n \left(\sum_{j=1}^n i \right)}_{=S_1} + \underbrace{\sum_{i=1}^n \left(\sum_{j=1}^n j \right)}_{=S_2}.$$

Par symétrie, les sommes S_1 et S_2 sont égales : on se contente de calculer S_1 . Le terme de la somme en l'indice j étant constant, il est facile de déterminer celle-ci et de poursuivre le calcul

$$\sum_{j=1}^n i = ni \quad \text{donc} \quad S_1 = \sum_{i=1}^n ni = n \sum_{i=1}^n i = \frac{n^2(n+1)}{2}.$$

Finalement,

$$C_n = 2S_1 = n^2(n+1).$$

(d) méthode

Il s'agit d'une somme triangulaire, on l'exprime comme deux sommes emboîtées : au choix une somme sur i d'une somme sur j supérieur à i ou une somme sur j d'une somme sur i inférieur à j .

On choisit la deuxième description car un peu plus simple :

$$D_n = \sum_{1 \leq i \leq j \leq n} (i+j) = \sum_{j=1}^n \left(\sum_{i=1}^j (i+j) \right)$$

Calculons à part la somme en l'indice i . On sépare celle-ci en deux

$$\sum_{i=1}^j (i + j) = \sum_{i=1}^j i + \underbrace{\sum_{i=1}^j j}_{j \text{ termes}} = \frac{j(j+1)}{2} + j \times j = \frac{3}{2}j^2 + \frac{1}{2}j.$$

On peut alors reprendre et terminer le calcul initial

$$D_n = \frac{3}{2} \sum_{j=1}^n j^2 + \frac{1}{2} \sum_{j=1}^n j = \frac{n(n+1)^2}{2}.$$

Exercice 4

Soit $n \in \mathbb{N}^*$. Calculer les produits suivants :

$$(a) \prod_{k=1}^n q^k \quad \text{avec } q \in \mathbb{C}$$

$$(b) \prod_{k=1}^n \left(1 + \frac{1}{k}\right)$$

Solution

(a) Multiplier les puissances de q conduit à sommer les exposants :

$$\prod_{k=1}^n q^k = q \times q^2 \times \cdots \times q^n = q^{1+2+\cdots+n} = q^{\frac{n(n+1)}{2}}.$$

(b) méthode

|| On exprime un produit télescopique.

En détaillant les facteurs du produit, on fait apparaître des simplifications

$$\prod_{k=1}^n \left(1 + \frac{1}{k}\right) = \prod_{k=1}^n \left(\frac{k+1}{k}\right) = \underbrace{\frac{2}{1}}_{k=1} \times \underbrace{\frac{3}{2}}_{k=2} \times \cdots \times \underbrace{\frac{n+1}{n}}_{k=n} = \frac{n+1}{1} = n+1.$$

Exercice 5

Soit $n \in \mathbb{N}$. Exprimer à l'aide de nombres factoriels les produits suivants

$$(a) 2 \times 4 \times 6 \times \cdots \times (2n)$$

$$(b) 1 \times 3 \times 5 \times \cdots \times (2n+1).$$

Solution

(a) méthode

|| On regroupe les 2 de chaque facteur pair avant de reconnaître un nombre factoriel.

$$2 \times 4 \times 6 \times \cdots \times (2n) = \prod_{k=1}^n (2k) = 2^n \left(\prod_{k=1}^n k \right) = 2^n n!$$

Lors de ce calcul 2 apparaît avec une puissance n car il figure dans chacun des n facteurs constituant le produit.

(b) méthode

- || On introduit les facteurs pairs intermédiaires pour faire apparaître un nombre factoriel.

$$1 \times 3 \times \cdots \times (2n+1) = \frac{1 \times 2 \times 3 \times 4 \times \cdots \times (2n) \times (2n+1)}{2 \times 4 \times \cdots \times (2n)}.$$

Au numérateur figure le produit de tous les entiers allant de 1 à $2n+1$ et au dénominateur le produit des entiers pairs calculé au-dessus. On conclut

$$1 \times 3 \times \cdots \times (2n+1) = \frac{(2n+1)!}{2^n n!}.$$

2.5.3 Coefficients binomiaux

Exercice 6

Soit $n \in \mathbb{N}$. Calculer

$$\sum_{k=0}^n 2^k \binom{n}{k}.$$

Solution
méthode

- || On reconnaît le développement $(1+2)^n$ par la formule du binôme de Newton (Th. 11 p. 53).

$$\sum_{k=0}^n 2^k \binom{n}{k} = \sum_{k=0}^n \binom{n}{k} 1^{n-k} \times 2^k = (1+2)^n = 3^n.$$

Exercice 7

Soit $n \in \mathbb{N}^*$.

- (a) Calculer

$$\sum_{k=0}^n \binom{n}{k} \quad \text{et} \quad \sum_{k=0}^n (-1)^k \binom{n}{k}.$$

- (b) En déduire les valeurs de

$$A = \sum_{p=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2p} \quad \text{et} \quad B = \sum_{p=0}^{\lfloor \frac{(n-1)}{2} \rfloor} \binom{n}{2p+1}.$$

Solution

(a) On reconnaît les développements de $(1+1)^n$ et de $(1-1)^n$.

$$\sum_{k=0}^n \binom{n}{k} = \sum_{k=0}^n \binom{n}{k} 1^{n-k} \times 1^k = (1+1)^n = 2^n$$

et

$$\sum_{k=0}^n (-1)^k \binom{n}{k} = \sum_{k=0}^n \binom{n}{k} 1^{n-k} \times (-1)^k = (1+(-1))^n = 0.$$

Notons que, lorsque n est nul, cette dernière somme vaut $0^0 = 1$.

(b) Les sommes définissant A et B correspondent respectivement aux sommes des termes d'indices pairs et des termes d'indices impairs de la somme des k parmi n .

méthode

|| On forme un système dont les deux quantités sont solutions.

On observe

$$A+B = \sum_{k=0}^n \binom{n}{k} = 2^n \quad \text{et} \quad A-B = \sum_{k=0}^n (-1)^k \binom{n}{k} = 0.$$

On en déduit que les deux sommes sont égales à 2^{n-1} .

Exercice 8

Vérifier que, pour tout $n \in \mathbb{N}^*$ et tout $p \in \llbracket 1 ; n \rrbracket$,

$$\binom{n}{p} = \frac{n}{p} \binom{n-1}{p-1}.$$

En déduire la valeur de

$$\sum_{p=1}^n p \binom{n}{p}.$$

Solution**méthode**

|| On écrit les coefficients binomiaux à l'aide de nombres factoriels.

On vérifie :

$$\binom{n}{p} = \frac{n!}{p!(n-p)!} = \frac{n \times (n-1)!}{p \times (p-1)!(n-p)!} = \frac{n}{p} \cdot \frac{(n-1)!}{(p-1)!(\underbrace{(n-p)}_{(n-1)-(p-1)})!} = \frac{n}{p} \binom{n-1}{p-1}$$

On exploite la formule précédente pour exprimer autrement le terme sommé

$$\sum_{p=1}^n p \binom{n}{p} = \sum_{p=1}^n n \binom{n-1}{p-1}.$$

Dans la somme en second membre, le facteur n ne dépend pas de l'indice de sommation ce qui permet de le factoriser

$$\sum_{p=1}^n p \binom{n}{p} = n \sum_{p=1}^n \binom{n-1}{p-1}.$$

Un glissement d'indice permet de terminer le calcul¹

$$\sum_{p=1}^n p \binom{n}{p} = n \sum_{k=0}^{n-1} \binom{n-1}{k} = n(1+1)^{n-1} = n2^{n-1}.$$

2.5.4 Systèmes d'équations linéaires

Exercice 9

Résoudre dans \mathbb{R} les systèmes d'équations linéaires suivants en discutant selon la valeur du paramètre m :

$$(a) \begin{cases} x + y + z + t = 1 \\ x + y + 2z = 0 \\ x + y + 2t = m \end{cases}$$

$$(b) \begin{cases} mx + y + z = 1 \\ x + my + z = m \\ x + y + mz = m^2 \end{cases}$$

Solution

(a) méthode

On applique l'algorithme du pivot de Gauss. Il sera commode de figurer les inconnues respectives les unes en dessous des autres.

$$\begin{array}{l} \left\{ \begin{array}{l} x + y + z + t = 1 \\ x + y + 2z = 0 \\ x + y + 2t = m \end{array} \right. \xrightarrow[L_3 \leftarrow L_3 - L_1]{L_2 \leftarrow L_2 - L_1} \left\{ \begin{array}{l} x + y + z + t = 1 \\ z - t = -1 \\ -z + t = m - 1 \end{array} \right. \\ \xrightarrow[L_3 \leftarrow L_3 + L_2]{L_1 \leftarrow L_1 + L_2} \left\{ \begin{array}{l} x + y + z + t = 1 \\ z - t = -1 \\ 0 = m - 2 \end{array} \right. \end{array}$$

La troisième équation est une équation de compatibilité.

Cas : $m \neq 2$. Le système est incompatible, il n'y a pas de solutions.

1. On verra une démarche alternative dans le sujet 19 p. 74.

Cas : $m = 2$. On poursuit la résolution en simplifiant l'équation de compatibilité et en considérant les inconnues x et z comme principales¹

$$\begin{cases} x + y + z + t = 1 \\ z - t = -1 \end{cases} \iff \begin{cases} x + z = 1 - y - t \\ z = -1 + t \end{cases} \iff \begin{cases} x = 2 - y - 2t \\ z = -1 + t. \end{cases}$$

L'ensemble des solutions est alors l'ensemble des quadruplets²

$$(2 - y - 2t, y, -1 + t, t) \quad \text{avec } (y, t) \in \mathbb{R}^2.$$

(b) méthode

On applique à nouveau l'algorithme du pivot en prenant soin de réduire au minimum le nombre de discussions selon les valeurs du paramètre m .

$$\begin{cases} mx + y + z = 1 \\ x + my + z = m \\ x + y + mz = m^2 \end{cases} \xleftarrow[L_1 \leftrightarrow L_3]{\quad} \begin{cases} x + y + mz = m^2 \\ x + my + z = m \\ mx + y + z = 1 \end{cases} \xleftarrow[L_2 \leftarrow L_2 - L_1]{\quad} \begin{cases} x + y + mz = m^2 \\ (m-1)y + (1-m)z = m - m^2 \\ (1-m)y + (1-m^2)z = 1 - m^3 \end{cases} \xleftarrow[L_3 \leftarrow L_3 - mL_1]{\quad} \begin{cases} x + y + mz = m^2 \\ (m-1)y + (1-m)z = m - m^2 \\ (1-m)(2+m)z = (1-m)(m+1)^2. \end{cases}$$

méthode

On poursuit la résolution dans le cas où les « pivots » facteurs de y et z sont non nuls puis on traite les cas particuliers.

Cas : $m \neq 1$ et $m \neq -2$. Le système présente un unique triplet (x, y, z) solution à savoir

$$\left(-\frac{m+1}{m+2}, \frac{1}{m+2}, \frac{(m+1)^2}{m+2} \right)$$

Cas : $m = 1$. Le système se résume à l'équation $x + y + z = 1$ et, en considérant x comme inconnue principale, l'ensemble des solutions est constitué des triplets³

$$(1 - y - z, y, z) \quad \text{avec } (y, z) \in \mathbb{R}^2.$$

Cas : $m = -2$. La dernière équation du système se relit $0 = 3$. Le système n'a pas de solutions.

1. On pourrait aussi considérer les inconnues x et t , y et z ou encore y et t comme principales, mais pas x et y ou z et t . Le choix de tel ou tel couple d'inconnues principales modifiera la description de l'ensemble des solutions.

2. On peut aussi proposer la description $(2, 0, -1, 0) + y(-1, 1, 0, 0) + t(-2, 0, 1, 1)$ avec $y, t \in \mathbb{R}$ qui fait apparaître la structure affine de l'ensemble des solutions (voir p. 320).

3. On peut aussi proposer la description affine $(1, 0, 0) + y(-1, 1, 0) + z(-1, 0, 1)$ avec $y, z \in \mathbb{R}$.

Exercice 10

Soit a, b et θ des réels. Résoudre le système suivant d'inconnue $(x, y) \in \mathbb{R}^2$:

$$(\Sigma) : \begin{cases} \cos(\theta)x - \sin(\theta)y = a & (1) \\ \sin(\theta)x + \cos(\theta)y = b & (2). \end{cases}$$

Solution

Il n'est pas possible d'appliquer simplement l'algorithme du pivot sans discuter selon les valeurs du paramètre θ et les éventuelles annulations de $\cos \theta$ ou de $\sin \theta$. On peut cependant résoudre le système sans discussion en raisonnant par combinaison d'équations :

méthode

|| On isole les inconnues par combinaison d'équations sachant $\cos^2 \theta + \sin^2 \theta = 1$.

Soit (x, y) un couple solution du système (Σ) . La combinaison $\cos \theta \times (1) + \sin \theta \times (2)$ isole x et simplifie le terme en y . Parallèlement, la combinaison $-\sin \theta \times (1) + \cos \theta \times (2)$ détermine y . On obtient alors

$$\begin{cases} x = \cos(\theta)a + \sin(\theta)b \\ y = -\sin(\theta)a + \cos(\theta)b. \end{cases}$$

Inversement¹, on vérifie par le calcul que les valeurs x et y proposées ci-dessus déterminent un couple solution :

$$\begin{cases} \cos(\theta)(\cos(\theta)a + \sin(\theta)b) - \sin(\theta)(-\sin(\theta)a + \cos(\theta)b) = a \\ \sin(\theta)(\cos(\theta)a + \sin(\theta)b) + \cos(\theta)(-\sin(\theta)a + \cos(\theta)b) = b. \end{cases}$$

2.6 Exercices d'entraînement

2.6.1 Principe de récurrence

Exercice 11 *

Etablir que tout entier naturel non nul n s'écrit

$$n = 2^k(2p+1) \quad \text{avec} \quad (p, k) \in \mathbb{N}^2$$

en procédant de deux manières :

- (a) En introduisant le plus grand entier k tel que 2^k divise n .
- (b) En raisonnant par récurrence.

1. La résolution du système étant conduite par implication, une vérification de la solution obtenue est nécessaire.

Solution**(a) méthode**

On montre l'existence d'un plus grand entier vérifiant une propriété en observant que l'ensemble des nombres concernés est une partie de \mathbb{N} non vide et majorée.

Introduisons l'ensemble $A = \{\ell \in \mathbb{N} \mid 2^\ell \text{ divise } n\}$. C'est une partie de \mathbb{N} et celle-ci est non vide car 0 en est élément. De plus, elle est majorée car¹

$$\begin{aligned} 2^m \mid n &\implies 2^m \leq n \\ &\implies m \leq n. \end{aligned}$$

La partie A possède donc un plus grand élément k .

Puisque k est élément de A , 2^k divise n ce qui permet d'écrire $n = 2^k m$ avec m entier. Puisque $k+1$ n'est pas élément de A , 2 ne divise pas m qui est donc un nombre impair de la forme $2p+1$ avec $p \in \mathbb{N}$. L'entier n s'écrit alors $2^k(2p+1)$ comme voulu.

(b) Raisonnons par récurrence forte sur $n \in \mathbb{N}^*$.

Pour $n = 1$: k et p égaux à 0 conviennent.

Supposons la propriété établie jusqu'au rang $n \geq 1$ et étudions l'entier $n+1$. Procédons par disjonction de cas :

Cas : $n+1$ est impair. L'écriture est directement² obtenue avec $k=0$ et $n+1=2p+1$.

Cas : $n+1$ est pair. On peut écrire $n+1=2m$ avec $1 \leq m \leq n$. Par l'hypothèse de récurrence forte, il est possible d'écrire $m=2^k(2p+1)$ puis $n+1=2^{k+1}(2p+1)$ avec p et k entiers.

La récurrence est établie.

Exercice 12 *

Montrer par récurrence que, pour tout entier naturel non nul n ,

$$\prod_{k=1}^n (4k-2) = \prod_{k=1}^n (n+k).$$

Solution

Pour $n=1$, les deux produits comportent un seul facteur égal à 2 : l'égalité est vérifiée. Supposons la propriété établie au rang $n \geq 1$.

méthode

Ne pas oublier de changer n en $n+1$ dans le facteur du second produit lorsque celui-ci est exprimé au rang $n+1$.

1. La deuxième implication est justifiée par la comparaison $m \leq 2^m$ que l'on établit par récurrence.
2. On ne fait alors pas usage de l'hypothèse de récurrence.

D'une part, on isole un facteur au premier produit

$$\prod_{k=1}^{n+1} (4k - 2) = \left(\prod_{k=1}^n (4k - 2) \right) \times (4n + 2).$$

D'autre part, en exprimant les facteurs du second produit

$$\prod_{k=1}^{n+1} (n + 1 + k) = (\underbrace{n + 2}_{k=1}) \times (\underbrace{n + 3}_{k=2}) \times \cdots \times (\underbrace{2n}_{k=n-1}) \times (\underbrace{2n + 1}_{k=n}) \times (\underbrace{2n + 2}_{k=n+1})$$

on observe

$$\prod_{k=1}^{n+1} (n + 1 + k) = \left(\prod_{k=1}^n (n + k) \right) \times \frac{(2n + 1) \times (2n + 2)}{(n + 1)}$$

En simplifiant le quotient final, on obtient

$$\prod_{k=1}^{n+1} (n + 1 + k) = \left(\prod_{k=1}^n (n + k) \right) \times (4n + 2).$$

L'hypothèse de récurrence permet alors de conclure

$$\prod_{k=1}^{n+1} (n + 1 + k) = \left(\prod_{k=1}^n (4k - 2) \right) \times (4n + 2) = \prod_{k=1}^{n+1} (4k - 2).$$

La récurrence est établie¹.

Exercice 13 **

Soit x un réel non nul tel que $x + \frac{1}{x}$ soit entier.

(a) Montrer que, pour tout $n \in \mathbb{N}$,

$$x^n + \frac{1}{x^n} \in \mathbb{Z}.$$

(b) Donner un exemple de réel x non trivial ayant la propriété qui précède.

Solution

(a) **méthode**

On exprime $x^{n+1} + \frac{1}{x^{n+1}}$ en fonction de $x^n + \frac{1}{x^n}$ et $x^{n-1} + \frac{1}{x^{n-1}}$.

En développant le produit

$$\left(x^n + \frac{1}{x^n} \right) \left(x + \frac{1}{x} \right) = \left(x^{n+1} + \frac{1}{x^{n+1}} \right) + \left(x^{n-1} + \frac{1}{x^{n-1}} \right). \quad (*)$$

1. On peut aussi proposer une démonstration directe en observant les deux produits égaux à $\frac{1}{x^n}$.

Montrons alors la propriété voulue en raisonnant par récurrence double sur $n \in \mathbb{N}$.

Pour $n = 0$, $x^0 + \frac{1}{x^0} = 2$ est un entier. Pour $n = 1$, $x + \frac{1}{x}$ est un entier par hypothèse.

Supposons la propriété vérifiée aux rangs n et $n - 1$ (avec $n \geq 1$). Par l'identité (*), il vient

$$x^{n+1} + \frac{1}{x^{n+1}} = \underbrace{\left(x^n + \frac{1}{x^n}\right)}_{\in \mathbb{Z}} \underbrace{\left(x + \frac{1}{x}\right)}_{\in \mathbb{Z}} - \underbrace{\left(x^{n-1} + \frac{1}{x^{n-1}}\right)}_{\in \mathbb{Z}} \in \mathbb{Z}.$$

La récurrence est établie.

(b) méthode

On choisit $p \in \mathbb{Z}$ et l'on résout l'équation $x + \frac{1}{x} = p$.

Pour $x \in \mathbb{R}^*$

$$x + \frac{1}{x} = p \iff x^2 - px + 1 = 0.$$

L'équation du second degré est de discriminant $\Delta = p^2 - 4$: elle détermine une solution réelle non triviale dès que $p^2 - 4 > 0$. Pour $p = 3$, on peut proposer

$$x = \frac{3 + \sqrt{5}}{2}.$$

2.6.2 Sommes numériques

Exercice 14 *

Pour x un réel différent de 1 et n un entier naturel, on pose

$$S_n = \sum_{k=0}^n kx^k.$$

(a) Déterminer la valeur de S_n en calculant $xS_n - S_n$.

(b) Retrouver la valeur de S_n en dérivant la fonction $x \mapsto 1 + x + \cdots + x^n$.

Solution

(a) méthode

On simplifie $xS_n - S_n$ en s'aider d'un glissement d'indice.

On développe le x dans la somme

$$xS_n - S_n = \sum_{k=0}^n kx^{k+1} - \sum_{k=0}^n kx^k.$$

On réalise le glissement d'indice $\ell = k + 1$ dans la première somme avant de renommer k l'indice de celle-ci

$$xS_n - S_n = \sum_{\ell=1}^{n+1} (\ell - 1)x^\ell - \sum_{k=0}^n kx^k = \sum_{k=1}^{n+1} ((k-1)x^k) - \sum_{k=0}^n kx^k.$$

En isolant le terme d'indice $n + 1$ de la première somme et le terme d'indice 0 de la seconde, on peut combiner les deux sommes sur leur portion commune

$$xS_n - S_n = \underbrace{nx^{n+1}}_{k=n+1} + \sum_{k=1}^n ((k-1)x^k - kx^k) - \underbrace{0.x^0}_{k=0} = nx^{n+1} - \sum_{k=1}^n x^k.$$

Enfin, la dernière somme est géométrique de raison x (avec $x \neq 1$) et

$$xS_n - S_n = nx^{n+1} - x \frac{x^{n+1} - 1}{x - 1}.$$

Après réduction au même dénominateur et division par $x - 1$, on conclut

$$S_n = \frac{nx^{n+2} - (n+1)x^{n+1} + x}{(x-1)^2}.$$

(b) Pour $x \neq 1$, l'identité géométrique s'écrit

$$1 + x + x^2 + \cdots + x^n = \sum_{k=0}^n x^k = \frac{x^{n+1} - 1}{x - 1}.$$

On dérive en la variable x ce qui fait disparaître le terme constant correspondant à l'indice $k = 0$

$$1 + 2x + \cdots + nx^{n-1} = \sum_{k=1}^n kx^{k-1} = \frac{nx^{n+1} - (n+1)x^n + 1}{(x-1)^2}.$$

En multipliant cette relation par x et en ajoutant un terme nul correspondant à l'indice $k = 0$, on retrouve

$$S_n = \frac{nx^{n+2} - (n+1)x^{n+1} + x}{(x-1)^2}.$$

Exercice 15 **

Soit $n \in \mathbb{N}^*$. Calculer

$$\sum_{k=1}^n k.k! \quad \text{et} \quad \sum_{k=1}^n \frac{k}{(k+1)!}.$$

Solution**méthode**

|| On fait apparaître une somme télescopique en écrivant $k = (k+1) - 1$.

On écrit

$$\sum_{k=1}^n k \cdot k! = \sum_{k=1}^n ((k+1)-1) \cdot k! = \sum_{k=1}^n ((k+1)! - k!).$$

On exprime les termes sommés afin de voir les simplifications

$$\sum_{k=1}^n ((k+1)! - k!) = (2! - 1!) + (3! - 2!) + \cdots + ((n+1)! - n!) = (n+1)! - 1.$$

De manière semblable, on obtient

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

Exercice 16 ***

Soit $n \in \mathbb{N}^*$. Calculer

$$\sum_{1 \leq i, j \leq n} \min(i, j).$$

Solution**méthode**

|| Par regroupement de termes (Th. 8 p. 50), on sépare la somme en plusieurs sommes pour chacune desquelles on sait déterminer $\min(i, j)$.

On écrit

$$\sum_{1 \leq i, j \leq n} \min(i, j) = \sum_{1 \leq i < j \leq n} \underbrace{\min(i, j)}_{=i} + \sum_{1 \leq i = j \leq n} \underbrace{\min(i, j)}_{=i} + \sum_{1 \leq j < i \leq n} \underbrace{\min(i, j)}_{=j}.$$

La somme intermédiaire est une somme simple :

$$\sum_{1 \leq i = j \leq n} i = \sum_{i=1}^n i = \frac{n(n+1)}{2}.$$

Les deux sommes extrêmes sont triangulaires et elles sont égales par un argument de symétrie. On exprime la somme double comme deux sommes emboîtées en étant attentif à la nature triangulaire du domaine de sommation

$$\sum_{1 \leq i < j \leq n} i = \sum_{i=1}^{n-1} \left(\sum_{\substack{j=i+1 \\ n-i \text{ termes}}}^n i \right) = \sum_{i=1}^{n-1} i(n-i) = \frac{n(n-1)(n+1)}{6}.$$

Finalement,

$$\sum_{1 \leq i, j \leq n} \min(i, j) = 2 \times \frac{n(n-1)(n+1)}{6} + \frac{n(n+1)}{2} = \frac{n(n+1)(2n+1)}{6}.$$

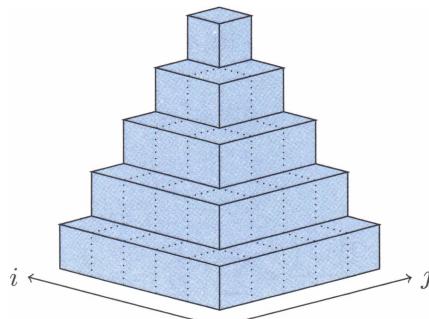


Figure illustrant $\sum_{1 \leq i, j \leq n} \min(i, j) = \sum_{k=1}^n k^2$.

2.6.3 Produits numériques

Exercice 17 *

Soit $a \in \mathbb{R}$. Pour $n \in \mathbb{N}$, on pose

$$P_n = \prod_{k=0}^n (1 + a^{2^k}).$$

- (a) Calculer P_n lorsque $a = 1$.
- (b) On suppose $a \neq 1$. Donner la valeur de P_n en calculant $(1 - a)P_n$.

Solution

- (a) Si $a = 1$, un calcul direct est possible (en remarquant que le produit comporte $n+1$ facteurs)

$$P_n = \prod_{k=0}^n 2 = 2^{n+1}.$$

(b) méthode

|| On exploite l'identité remarquable $(a - b)(a + b) = a^2 - b^2$.

En regroupant successivement les deux premiers facteurs¹

$$\begin{aligned}(1-a)P_n &= \underbrace{(1-a)(1+a)}_{=(1-a^2)}(1+a^2)\dots(1+a^{2^n}) \\ &= \underbrace{(1-a^2)(1+a^2)}_{=1-a^4}\dots(1+a^{2^n}) = \dots \\ &= (1-a^{2^n})(1+a^{2^n}) = 1-a^{2^{n+1}}.\end{aligned}$$

On peut alors conclure²

$$P_n = \frac{1-a^{2^{n+1}}}{1-a}.$$

Exercice 18 **

Soit $n \in \mathbb{N}^*$. Exprimer à l'aide de nombres factoriels le produit

$$\prod_{k=1}^n \left(1 - \frac{1}{4k^2}\right)$$

Solution

méthode

|| Après réduction au même dénominateur, on exprime un produit d'entiers pairs et d'entiers impairs.

On écrit

$$\prod_{k=1}^n \left(1 - \frac{1}{4k^2}\right) = \prod_{k=1}^n \frac{4k^2 - 1}{4k^2} = \prod_{k=1}^n \frac{(2k-1)(2k+1)}{(2k)^2}$$

ce qui conduit au calcul des trois produits

$$\prod_{k=1}^n (2k-1), \quad \prod_{k=1}^n (2k+1) \quad \text{et} \quad \prod_{k=1}^n (2k).$$

D'une part, on a déjà vu³

$$\prod_{k=1}^n (2k) = 2^n n! \quad \text{et} \quad \prod_{k=1}^n (2k+1) = \frac{(2n+1)!}{2^n n!}.$$

1. Lors des calculs, on précise l'étape finale afin de proposer la bonne expression. On sera en particulier attentif à ce que les exposants de a sont des puissances de 2 et non simplement des nombres pairs.

2. La formule obtenue correspond aussi à la somme géométrique $1 + a + a^2 + a^3 + \dots + a^{2^n-1}$: on peut vérifier cette propriété en développant le produit initial et en observant que ce développement fait apparaître, une fois et une seule, tous les termes de la somme précédente.

3. Voir sujet 5 p. 61.

D'autre part, en introduisant les facteurs pairs intermédiaires y compris un dernier facteur $2n$ pour simplifier l'expression finale du résultat

$$\prod_{k=1}^n (2k-1) = 1 \times 3 \times \cdots \times (2n-1) = \frac{1 \times 2 \times 3 \times 4 \times \cdots \times (2n-1) \times (2n)}{2 \times 4 \times \cdots \times (2n)} = \frac{(2n)!}{2^n n!}.$$

Finalement,

$$\prod_{k=1}^n \left(1 - \frac{1}{4k^2}\right) = (2n+1) \frac{((2n)!)^2}{(2^n n!)^4} = \frac{2n+1}{16^n} \binom{2n}{n}^2.$$

2.6.4 Coefficients binomiaux

Exercice 19 *

Soit $n \in \mathbb{N}$. En considérant la fonction $f: x \mapsto (1+x)^n$, calculer

$$\sum_{k=1}^n k \binom{n}{k} \quad \text{et} \quad \sum_{k=1}^n k^2 \binom{n}{k}.$$

Solution

Par la formule du binôme, on a l'expression développée

$$f(x) = \sum_{k=0}^n \binom{n}{k} x^k \quad \text{pour tout } x \in \mathbb{R}.$$

méthode

|| On dérive la fonction f avant d'évaluer en 1.

Par dérivation¹, on a pour tout réel x

$$f'(x) = n(1+x)^{n-1} = \sum_{k=1}^n k \binom{n}{k} x^{k-1}.$$

En évaluant en $x = 1$, on obtient²

$$\sum_{k=1}^n k \binom{n}{k} = n^{n-1}.$$

En multipliant $f'(x)$ par x (afin de faire apparaître x^k au lieu de x^{k-1}) avant de dériver à nouveau, on obtient

$$xf''(x) + f'(x) = n(1+nx)(1+x)^{n-2} = \sum_{k=1}^n k^2 \binom{n}{k} x^{k-1}.$$

1. Le terme constant correspondant à l'indice $k = 0$ disparaît lors de la dérivation.

2. On retrouve le résultat déjà obtenu dans le sujet 8 p. 63.

En évaluant en $x = 1$, il vient

$$\sum_{k=1}^n k^2 \binom{n}{k} = n(n+1)2^{n-2}.$$

Exercice 20 **

Pour tous n, p et q entiers naturels, calculer la somme

$$\sum_{k=p+1}^q \binom{n+k}{k}.$$

Solution
méthode

On exprime une somme télescopique par la formule du triangle de Pascal (Th. 10 p. 53).

Pour tout $k \in \mathbb{N}^*$, on écrit

$$\binom{n+k}{k} = \binom{n+k+1}{k} - \binom{n+k}{k-1}$$

et l'on obtient par télescopage

$$\begin{aligned} \sum_{k=p+1}^q \binom{n+k}{k} &= \sum_{k=p+1}^q \left(\binom{n+k+1}{k} - \binom{n+k}{k-1} \right) \\ &= \binom{n+p+2}{p+1} - \binom{n+p+1}{p} + \binom{n+p+3}{p+2} - \binom{n+p+2}{p+1} \\ &\quad + \cdots + \binom{n+q+1}{q} - \binom{n+q}{q-1} \\ &= \binom{n+q+1}{q} - \binom{n+p+1}{p}. \end{aligned}$$

Par la symétrie des coefficients binomiaux, on peut aussi remarquer

$$\sum_{k=p+1}^q \binom{n+k}{n} = \sum_{k=p+1}^q \binom{n+k}{k} = \binom{n+q+1}{q} - \binom{n+p+1}{p}$$

Exercice 21 **

Calculer

$$S_n = \sum_{k=0}^n (-1)^k \binom{2n+1}{k}$$

Solution**méthode**

|| On décompose le coefficient binomial par la formule du triangle de Pascal.

Pour $k \geq 1$, on écrit

$$\binom{2n+1}{k} = \binom{2n}{k-1} + \binom{2n}{k}.$$

On isole le terme d'indice 0 de la somme et l'on peut percevoir la somme étudiée sous forme télescopique

$$\begin{aligned} S_n &= \binom{2n+1}{0} + \sum_{k=1}^n \left((-1)^k \binom{2n}{k-1} - (-1)^{k+1} \binom{2n}{k} \right) \\ &= 1 - \underbrace{\left(\binom{2n}{0} + \binom{2n}{1} \right)}_{k=1} + \underbrace{\left(\binom{2n}{1} + \binom{2n}{2} \right)}_{k=2} + \cdots + \underbrace{\left(\binom{2n}{n-1} + \binom{2n}{n} \right)}_{k=n} \\ &= 1 - \binom{2n}{0} + (-1)^n \binom{2n}{n} = (-1)^n \binom{2n}{n}. \end{aligned}$$

Exercice 22 ** (Formule de Chu-Vandermonde)

Soit n , p et q trois entiers naturels vérifiant $n \leq p + q$. En développant de deux manières $(1+x)^{p+q}$, établir

$$\sum_{k=0}^n \binom{p}{k} \binom{q}{n-k} = \binom{p+q}{n}.$$

Solution**méthode**

|| On étudie le coefficient de x^n dans $(1+x)^{p+q} = (1+x)^p \times (1+x)^q$.

D'une part,

$$(1+x)^p \times (1+x)^q = (1+x)^{p+q} = \sum_{k=0}^{p+q} \binom{p+q}{k} x^k.$$

Le coefficient de x^n dans $(1+x)^p \times (1+x)^q$ est donc $\binom{p+q}{n}$.

D'autre part,

$$(1+x)^p \times (1+x)^q = \left(\sum_{i=0}^p \binom{p}{i} x^i \right) \left(\sum_{j=0}^q \binom{q}{j} x^j \right)$$

et le coefficient de x^n dans le développement de ce produit est la somme des produits $\binom{p}{i} \binom{q}{j}$ pour les indices i et j vérifiant $i + j = n$.

L'identification¹ de ces deux coefficients donne

$$\binom{p+q}{n} = \sum_{i+j=n} \binom{p}{i} \binom{q}{j} = \sum_{k=0}^n \binom{p}{k} \binom{q}{n-k}$$

quitte à adjoindre quelques coefficients nuls² à la dernière somme.

Exercice 23 **

Soit $n \in \mathbb{N}$. Calculer

$$\sum_{k=0}^n \binom{n}{k}^2 \quad \text{et} \quad \sum_{k=0}^n (-1)^k \binom{n}{k}^2.$$

Solution

méthode

On reprend l'idée du sujet précédent avec la formule de symétrie (Th. 10 p. 53)

$$\binom{n}{k} = \binom{n}{n-k}.$$

Le coefficient de x^n dans $(1+x)^{2n}$ est $\binom{2n}{n}$. On peut retrouver ce coefficient en considérant $(1+x)^{2n} = (1+x)^n \times (1+x)^n$ ce qui donne l'identité

$$\binom{2n}{n} = \sum_{k=0}^n \underbrace{\binom{n}{k}}_{\substack{\text{coeffient de } x^k \\ \text{dans } (1+x)^n}} \times \underbrace{\binom{n}{n-k}}_{\substack{\text{coeffient de } x^{n-k} \\ \text{dans } (1+x)^n}} = \sum_{k=0}^n \binom{n}{k}^2.$$

méthode

Pour faire apparaître $(-1)^k$, on considère le développement de $(1-x)^n$.

Le coefficient de x^n dans $(1-x)^n(1+x)^n$ est

$$\sum_{k=0}^n \underbrace{(-1)^k \binom{n}{k}}_{\substack{\text{coeffient de } x^k \\ \text{dans } (1-x)^n}} \times \underbrace{\binom{n}{n-k}}_{\substack{\text{coeffient de } x^{n-k} \\ \text{dans } (1+x)^n}} = \sum_{k=0}^n (-1)^k \binom{n}{k}^2.$$

Or ce coefficient est aussi celui de x^n dans $(1-x^2)^n$ et donc

$$\sum_{k=0}^n (-1)^k \binom{n}{k}^2 = \begin{cases} 0 & \text{si } n \text{ est impair} \\ (-1)^{n/2} \binom{n}{\frac{n}{2}} & \text{si } n \text{ est pair.} \end{cases}$$

1. Lorsque deux fonctions polynomiales sont égales sur \mathbb{R} , elles sont nécessairement écrites avec les mêmes coefficients (Th. 9 p. 154).

2. Rappelons que le coefficient $\binom{n}{p}$ est nul lorsque p n'est pas compris entre 0 et n .

Exercice 24 **

Soit $n \in \mathbb{N}$ avec $n \geq 2$. Calculer

$$A = \sum_{p=0}^{\lfloor \frac{n}{3} \rfloor} \binom{n}{3p}, \quad B = \sum_{p=0}^{\lfloor \frac{n-1}{3} \rfloor} \binom{n}{3p+1} \quad \text{et} \quad C = \sum_{p=0}^{\lfloor \frac{n-2}{3} \rfloor} \binom{n}{3p+2}.$$

Solution**méthode**

En exploitant la formule du binôme, on forme un système de trois équations vérifiées par A , B et C .

D'une part,

$$A + B + C = \sum_{k=0}^n \binom{n}{k} = (1+1)^n = 2^n. \quad (1)$$

D'autre part, en exploitant ¹ $j^{3p} = 1$, $j^{3p+1} = j$ et $j^{3p+2} = j^2$, on obtient

$$A + Bj + Cj^2 = \sum_{k=0}^n \binom{n}{k} j^k = (1+j)^n = (-j^2)^n. \quad (2)$$

Par conjugaison, on a aussi

$$A + Bj^2 + Cj = (-j)^n. \quad (3)$$

La combinaison (1) + (2) + (3) détermine A car $1 + j + j^2 = 0$

$$A = \frac{1}{3}(2^n + (-j)^n + (-j^2)^n).$$

On simplifie cette expression en écrivant

$$(-j)^n + (-j^2)^n = (-j)^n + \overline{(-j)^n} = 2 \operatorname{Re}((-j)^n) = 2 \operatorname{Re}(e^{-in\pi/3}) = 2 \cos\left(\frac{n\pi}{3}\right)$$

et donc

$$A = \frac{1}{3} \left(2^n + 2 \cos\left(\frac{n\pi}{3}\right) \right)$$

La combinaison (1) + $j^2(2) + j(3)$ détermine B :

$$B = \frac{1}{3}(2^n + j^2(-j^2)^n + j(-j)^n) = \frac{1}{3} \left(2^n + 2 \cos\left(\frac{(n-2)\pi}{3}\right) \right)$$

1. Le nombre j désigne la racine troisième de l'unité $e^{2i\pi/3}$, elle vérifie $j^3 = 1$, $1 + j + j^2 = 0$ et $j = j^2$.

et de même $(1) + j(2) + j^2(3)$ détermine C :

$$C = \frac{1}{3} \left(2^n + 2 \cos\left(\frac{(n+2)\pi}{3}\right) \right).$$

Exercice 25 **

Soit $n \in \mathbb{N}^*$.

(a) Vérifier que, pour tout $k \in \llbracket 1 ; n \rrbracket$,

$$\binom{n}{k} = \frac{n-k+1}{k} \binom{n}{k-1}.$$

(b) En déduire que

$$\binom{n}{k-1} < \binom{n}{k} \quad \text{si } 1 \leq k \leq \frac{n}{2}$$

$$\binom{n}{k+1} < \binom{n}{k} \quad \text{si } \frac{n}{2} \leq k \leq n-1.$$

(c) Application : Montrer

$$\binom{2n}{n} \geq \frac{4^n}{2n+1}.$$

Solution

(a) En exploitant l'écriture factorielle des coefficients binomiaux

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{(n-k+1)}{k} \cdot \frac{n!}{(k-1)!(n-k+1)!} = \frac{(n-k+1)}{k} \binom{n}{k-1}.$$

(b) Si $1 \leq k \leq n/2$ alors $2k < n+1$ et donc $n-k+1 > k$ puis

$$\binom{n}{k} = \frac{n-k+1}{k} \binom{n}{k-1} > \binom{n}{k-1}.$$

méthode

|| La deuxième inégalité se déduit de la formule de symétrie : $\binom{n}{k} = \binom{n}{n-k}$

Si $n/2 \leq k \leq n-1$ alors $\ell = n-k$ vérifie $1 \leq \ell \leq n/2$ donc

$$\binom{n}{\ell-1} < \binom{n}{\ell} \quad \text{c'est-à-dire} \quad \binom{n}{k+1} < \binom{n}{k}.$$

(c) méthode

|| Par ce qui précède, on peut affirmer que, dans la formule du binôme, c'est le ¹
 « coefficient binomial du milieu » qui est le plus grand.

On a

$$\sum_{k=0}^{2n} \binom{2n}{k} = (1+1)^{2n} = 2^{2n} \quad \text{et} \quad \binom{2n}{k} \leq \binom{2n}{n} \text{ pour tout } k \in [0; 2n].$$

On en déduit

$$2^{2n} \leq \underbrace{(2n+1)}_{\substack{\text{nombre de termes} \\ \text{dans la somme}}} \binom{2n}{n}$$

puis l'inégalité proposée.

2.7 Exercices d'approfondissement

Exercice 26 *

Soit $n \in \mathbb{N}^*$ et a_1, \dots, a_n des réels positifs et $s_n = a_1 + \dots + a_n$. Vérifier

$$\prod_{k=1}^n (1+a_k) \leq \sum_{k=0}^n \frac{s_n^k}{k!}.$$

Solution

On vérifie l'inégalité proposée par récurrence.

Pour $n = 1$, l'inégalité est vérifiée, il s'agit même d'une égalité.

Supposons la propriété vérifiée au rang n avec $n \geq 1$. Soit a_1, \dots, a_n, a_{n+1} des réels positifs. Considérons $s_n = a_1 + \dots + a_n$ et $s_{n+1} = s_n + a_{n+1}$. Par hypothèse de récurrence on peut écrire

$$\prod_{k=1}^{n+1} (1+a_k) \leq \left(\sum_{k=0}^n \frac{s_n^k}{k!} \right) (1+a_{n+1}).$$

En développant puis en opérant un glissement d'indice

$$\left(\sum_{k=0}^n \frac{s_n^k}{k!} \right) (1+a_{n+1}) = \sum_{k=0}^n \frac{s_n^k}{k!} + \sum_{k=0}^n \frac{s_n^k a_{n+1}}{k!} = \sum_{k=0}^n \frac{s_n^k}{k!} + \sum_{k=1}^{n+1} \frac{s_n^{k-1} a_{n+1}}{(k-1)!}$$

On combine ensuite les deux sommes sur leur portion commune

$$\left(\sum_{k=0}^n \frac{s_n^k}{k!} \right) (1+a_{n+1}) = 1 + \sum_{k=1}^n \frac{s_n^k + ks_n^{k-1} a_{n+1}}{k!} + \frac{s_n^n a_{n+1}}{n!}$$

méthode

Si x et y sont des réels positifs, on a¹

$$x^k + kx^{k-1}y \leq (x+y)^k \quad \text{pour tout } k \geq 1.$$

1. Plus précisément, si n est pair, ce coefficient est unique et correspond à l'indice $n/2$, si n est impair, il sont deux correspondant aux indices successifs $(n-1)/2$ et $(n+1)/2$.

On a $s_n^k + ks_n^{k-1}a_{n+1} \leq s_{n+1}^k$ pour tout k compris entre 1 et n . On a aussi par un argument semblable $(n+1)s_n^na_{n+1} \leq s_{n+1}^{n+1}$. On en déduit

$$\prod_{k=1}^{n+1} (1+a_k) \leq 1 + \sum_{k=1}^n \frac{s_{n+1}^k}{k!} + \frac{s_{n+1}^{n+1}}{(n+1)!} = \sum_{k=0}^{n+1} \frac{s_{n+1}^k}{k!}.$$

La récurrence est établie.

Exercice 27 **

Soit n un entier naturel.

(a) Montrer l'existence et l'unicité de nombres entiers a_n et b_n vérifiant

$$(1+\sqrt{2})^n = a_n + b_n\sqrt{2}.$$

(b) Calculer $a_n^2 - 2b_n^2$.

(c) Montrer qu'il existe un unique $p \in \mathbb{N}^*$ tel que

$$(1+\sqrt{2})^n = \sqrt{p} + \sqrt{p-1}.$$

Solution

(a) Montrons l'existence par la formule du binôme de Newton²

$$(1+\sqrt{2})^n = \sum_{k=0}^n \binom{n}{k} (\sqrt{2})^k.$$

En séparant la somme en deux selon la parité de l'indice, il vient

$$(1+\sqrt{2})^n = \sum_{p=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2p} \underbrace{(\sqrt{2})^{2p}}_{=2^p} + \sum_{p=0}^{\lfloor \frac{n-1}{2} \rfloor} \binom{n}{2p+1} \underbrace{(\sqrt{2})^{2p+1}}_{2^p\sqrt{2}} = a_n + b_n\sqrt{2}$$

avec les entiers

$$a_n = \sum_{p=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2p} 2^p \quad \text{et} \quad b_n = \sum_{p=0}^{\lfloor \frac{n-1}{2} \rfloor} \binom{n}{2p+1} 2^p.$$

méthode

|| L'unicité de l'écriture provient de l'irrationalité de $\sqrt{2}$.

Supposons

$$(1+\sqrt{2})^n = a + b\sqrt{2} = a' + b'\sqrt{2}$$

1. En effet, x^k et $kx^{k-1}y$ sont les deux premiers termes du développement de $(x+y)^k$, les autres étant tous positifs.

2. On peut aussi raisonner par récurrence en constatant $a_0 = 1$ et $b_0 = 0$ et, pour tout $n \in \mathbb{N}$, $a_{n+1} = a_n + 2b_n$ et $b_{n+1} = a_n + b_n$.

avec a, b, a', b' entiers. Par différence de membres, on a $(b' - b)\sqrt{2} = a - a'$. Par l'absurde, si $b \neq b'$, on peut exprimer $\sqrt{2}$ comme quotient de deux nombres entiers ce qui contredit son irrationalité. On a donc $b = b'$ puis, nécessairement, $a = a'$.

(b) **méthode**

|| Par le même calcul qu'au-dessus, on constate

$$(1 - \sqrt{2})^n = a_n - b_n\sqrt{2}.$$

Dès lors

$$a_n^2 - 2b_n^2 = (a_n - b_n\sqrt{2})(a_n + b_n\sqrt{2}) = (1 + \sqrt{2})^n(1 - \sqrt{2})^n = (-1)^n.$$

(c) La fonction $x \mapsto \sqrt{x} + \sqrt{x-1}$ est strictement croissante donc injective : ceci assure l'unicité de l'écriture. Reste à établir l'existence.

Cas : n est pair. On a $a_n^2 = 1 + 2b_n^2$. Pour $p = a_n^2 \in \mathbb{N}$,

$$(1 + \sqrt{2})^n = a_n + \sqrt{2}b_n = \sqrt{p} + \sqrt{p-1}.$$

Cas : n est impair. On a $2b_n^2 = a_n^2 + 1$. Pour $p = 2b_n^2 \in \mathbb{N}$,

$$(1 + \sqrt{2})^n = \sqrt{2}b_n + a_n = \sqrt{p} + \sqrt{p-1}.$$

Exercice 28 ***

Soit n un entier ≥ 2 . Montrer que

$$H_n = \sum_{k=1}^n \frac{1}{k}$$

n'est pas entier.

Solution

Les premières valeurs de H_n sont données dans le tableau ci-dessous :

n	2	3	4	5	6	7	8
H_n	$\frac{3}{2}$	$\frac{11}{6}$	$\frac{25}{12}$	$\frac{137}{60}$	$\frac{49}{20}$	$\frac{363}{140}$	$\frac{761}{280}$

méthode

|| On montre par récurrence forte que H_n est le quotient d'un entier impair par un entier pair.

Pour $n = 2$, la propriété est vérifiée.

Supposons la propriété vraie jusqu'au rang $n - 1$ (avec $n \geq 3$). Raisonnons par disjonction de cas.

Cas : n impair. On écrit par l'hypothèse de récurrence

$$H_n = H_{n-1} + \frac{1}{n} \quad \text{avec} \quad H_{n-1} = \frac{2p+1}{2q}, \quad (p, q) \in \mathbb{N} \times \mathbb{N}^*.$$

Après réduction au même dénominateur, on obtient l'écriture voulue

$$H_n = \frac{(2p+1)n + 2q}{2pq}$$

car $(2p+1)n$ est impair puisque produit de deux entiers impairs.

Cas : n pair. On peut écrire $n = 2m$ avec $m \geq 2$ puis, en séparant les termes d'indices pairs de ceux d'indices impairs dans la somme, on écrit

$$\begin{aligned} H_n &= \sum_{k=1}^{2m} \frac{1}{k} = \sum_{j=1}^m \frac{1}{2p} + \sum_{j=1}^m \frac{1}{2p-1} \\ &= \frac{1}{2}H_m + 1 + \frac{1}{3} + \cdots + \frac{1}{2m-1}. \end{aligned}$$

Par l'hypothèse de récurrence forte, H_m est le quotient d'un entier impair par un entier pair, *a fortiori*, $\frac{1}{2}H_m$ l'est aussi. Au surplus, comme on l'a vu dans l'étude du cas précédent, l'ajout de l'inverse d'un entier impair conserve cette propriété. On en déduit que H_n est le quotient d'un entier impair par un entier pair.

La récurrence est établie et l'on peut affirmer que, pour tout $n \geq 2$, le nombre H_n n'est pas entier.

Exercice 29 ***

Soit $n \in \mathbb{N}^*$. Vérifier

$$\sum_{k=1}^n \frac{(-1)^{k-1}}{k} \binom{n}{k} = \sum_{k=1}^n \frac{1}{k}.$$

Solution

méthode

|| On écrit $\frac{1}{k}$ comme l'intégrale de 0 à 1 d'une puissance de x .

Pour tout $k \in \mathbb{N}^*$,

$$\int_0^1 x^{k-1} dx = \left[\frac{x^k}{k} \right]_0^1 = \frac{1}{k}.$$

On peut alors écrire par linéarité de l'intégrale

$$\sum_{k=1}^n \frac{(-1)^{k-1}}{k} \binom{n}{k} = \sum_{k=1}^n \left((-1)^{k-1} \binom{n}{k} \int_0^1 x^{k-1} dx \right) = \int_0^1 \left(\sum_{k=1}^n (-1)^{k-1} \binom{n}{k} x^{k-1} \right) dx.$$

1. On peut aussi raisonner par récurrence en exploitant la formule de Pascal.

Soit $x \in [0; 1]$. Par la formule du binôme

$$(1-x)^n = \sum_{k=0}^n (-1)^k \binom{n}{k} x^k.$$

Pour $x \neq 0$, on isole le terme d'indice 0 de la somme et l'on obtient en divisant par x l'égalité

$$\sum_{k=1}^n (-1)^{k-1} \binom{n}{k} x^{k-1} = \frac{1 - (1-x)^n}{x}.$$

Dans le second membre, on reconnaît la somme des termes d'une suite géométrique de raison $1-x$

$$\sum_{k=1}^{n-1} (1-x)^k = \frac{1 - (1-x)^n}{1 - (1-x)} = \frac{1 - (1-x)^n}{x}.$$

On a ainsi obtenu la formule

$$\sum_{k=1}^n (-1)^{k-1} \binom{n}{k} x^{k-1} = \sum_{k=0}^{n-1} (1-x)^k.$$

Celle-ci est encore vraie¹ pour $x = 0$.

On peut alors poursuivre le calcul initial

$$\begin{aligned} \sum_{k=1}^n \frac{(-1)^{k-1}}{k} \binom{n}{k} &= \int_0^1 \left(\sum_{k=0}^{n-1} (1-x)^k \right) dx = \sum_{k=0}^{n-1} \left(\int_0^1 (1-x)^k dx \right) \\ &= \sum_{k=0}^{n-1} \left[-\frac{(1-x)^{k+1}}{k+1} \right]_0^1 = \sum_{k=0}^{n-1} \frac{1}{k+1} = \sum_{k=1}^n \frac{1}{k}. \end{aligned}$$

Exercice 30 **** (Développement factoriel d'un entier)

(a) Montrer que, pour tout $n \in \mathbb{N}^*$, il existe $p \in \mathbb{N}^*$ et $(a_1, a_2, \dots, a_p) \in \mathbb{N}^p$ tels que

$$n = \sum_{k=1}^p a_k \cdot k! \quad \text{avec} \quad 0 \leq a_k \leq k \text{ et } a_p \neq 0.$$

(b) Vérifier l'unicité de cette écriture.

1. On peut vérifier que l'on obtient n dans les deux membres ou utiliser, sans calculs, un argument de continuité : de part et d'autre les fonctions sont continues et elles sont égales au voisinage de 0 donc égales en 0.

Solution(a) **méthode**

On raisonne par récurrence forte en opérant la division euclidienne (Th. 4 p. 88) de n par la plus grande factorielle qui lui est inférieure.

Pour $n = 1$, on peut immédiatement écrire $n = a_1 \cdot 1!$ avec $a_1 = 1$.

Supposons l'existence de l'écriture vraie jusqu'au rang $n - 1$ (avec $n \geq 2$). Introduisons¹ p le plus grand entier tels que $p! \leq n$ et réalisons la division euclidienne de n par $p!$:

$$n = p!q + r \quad \text{avec} \quad 0 \leq r < p!$$

Posons alors $a_p = q \in \mathbb{N}$. D'une part, $a_p \geq 1$ car $n \geq p!$. D'autre part, $a_p < p + 1$ car $n < (p + 1)!$ par définition de p . Il suffit ensuite de décomposer le reste r .

Si r est nul, on pose $a_1 = \dots = a_{p-1} = 0$ et l'on obtient l'écriture attendue de n .

Si r est non nul, on peut exploiter l'hypothèse de récurrence forte pour décomposer r car $r < p! \leq n$. On peut ainsi écrire

$$r = \sum_{k=1}^q b_k \cdot k! \quad \text{avec} \quad 0 \leq b_k \leq k \text{ et } b_q \neq 0.$$

Puisque b_q est au moins égal à 1, on a $q! \leq r$ et donc $q < p$: la décomposition de r est de longueur strictement inférieure à p . En posant $(a_1, \dots, a_q) = (b_1, \dots, b_q)$ et, s'il y a lieu, a_{q+1}, \dots, a_{p-1} égaux à 0, on obtient l'écriture

$$n = \sum_{k=1}^p a_k \cdot k! \quad \text{avec} \quad 0 \leq a_k \leq k \text{ et } a_p \neq 0.$$

La récurrence est établie.

(b) Supposons

$$n = \sum_{k=1}^p a_k \cdot k! = \sum_{k=1}^q b_k \cdot k!$$

avec les conditions requises.

Puisque $a_p \geq 1$, on a $n \geq p!$. Aussi, sachant $a_k \leq k$ pour tout k , on a²

$$n \leq \sum_{k=1}^p k \cdot k! = \sum_{k=1}^p ((k+1)-1) \cdot k! = (p+1)! - 1.$$

On peut alors affirmer l'encadrement

$$p! \leq n < (p+1)!$$

1. Ce plus grand entier existe car l'ensemble des entiers p tels que $p! \leq n$ est une partie non vide et majorée.

2. Voir le calcul de la somme dans le sujet 15 p. 70.

De même, on a aussi l'encadrement

$$q! \leq n < (q+1)!$$

On en déduit $p = q$. On dispose alors de l'égalité suivante avec deux sommes de mêmes longueurs

$$\sum_{k=1}^p a_k.k! = \sum_{k=1}^p b_k.k!$$

Par l'absurde, supposons $(a_1, \dots, a_p) \neq (b_1, \dots, b_p)$ et considérons le plus grand entier r de $\llbracket 1 ; p \rrbracket$ tel que $a_r \neq b_r$. Après simplification, on dispose de l'égalité

$$\sum_{k=1}^r a_k.k! = \sum_{k=1}^r b_k.k! \quad \text{avec} \quad a_r \neq b_r. \quad (*)$$

Quitte à permute les écritures, on peut supposer $a_r < b_r$, c'est-à-dire $a_r + 1 \leq b_r$. On a alors

$$n = \sum_{k=1}^r a_k.k! \leq \underbrace{\sum_{k=1}^{r-1} k.k!}_{< r!} + a_r.r! < r! + a_r.r! \leq b_r.r! \leq \sum_{k=1}^r b_k.k! = n.$$

Cette inégalité stricte est absurde : on a justifié l'unicité de l'écriture.

CHAPITRE 3

Arithmétique des entiers

3.1 Divisibilité

3.1.1 Divisibilité dans \mathbb{Z}

Définition

On dit qu'un entier¹ a divise un entier b , et l'on note $a | b$, lorsqu'il existe un entier k tel que $b = ak$. On dit alors que a est un *diviseur* de b ou encore que b est un *multiple* de a .

Les entiers 1 , a , -1 et $-a$ sont des diviseurs de a , ce sont ses *diviseurs triviaux*.

En dehors du cas $a = 0$ qui est divisible par tout entier, les diviseurs de a sont assurément inférieurs à a en valeur absolue.

Théorème 1 (Double divisibilité)

Soit a et b deux entiers

$$a | b \text{ et } b | a \implies |a| = |b|.$$

En particulier, lorsque a et b sont naturels, on obtient l'antisymétrie de la relation de divisibilité :

$$a | b \text{ et } b | a \implies a = b.$$

1. Lorsque l'on parle d'entier, on signifie : entier relatif.

Théorème 2 (Transitivité)

Soit a, b et c trois entiers.

$$a \mid b \text{ et } b \mid c \implies a \mid c.$$

Théorème 3 (Divisibilité des combinaisons entières)

Soit a, b et c trois entiers.

$$a \mid b \text{ et } a \mid c \implies \forall (u, v) \in \mathbb{Z}^2, a \mid (bu + cv).$$

3.1.2 La division euclidienne**Théorème 4 (Division euclidienne dans \mathbb{Z})**

Pour tout $a \in \mathbb{Z}$ et tout $b \in \mathbb{N}^*$, il existe un unique couple (q, r) d'entiers vérifiant

$$a = bq + r \quad \text{et} \quad 0 \leq r < b.$$

Les entiers q et r sont respectivement nommés *quotient* et *reste* de la *division euclidienne* de a par b .

L'entier b est un diviseur de a si, et seulement si, le reste r est nul.

3.1.3 Calculs en congruence

Soit $n \in \mathbb{N}^*$.

Définition

On dit qu'un entier a est *congru* à un entier b *modulo* n lorsque n divise la différence $b - a$. On note alors

$$a \equiv b \pmod{n}.$$

En particulier, n divise a si, et seulement si, $a \equiv 0 \pmod{n}$.

Théorème 5

La congruence modulo n définit une relation d'équivalence sur \mathbb{Z} , relation pour laquelle tout entier est congru à un unique entier r compris entre 0 et $n - 1$.

Cet entier r correspond au reste de la division euclidienne de a par n .

La relation de congruence modulo n est compatible avec les opérations arithmétiques :

Théorème 6

Soit a, a', b et b' des entiers. Si $a \equiv a' \pmod{n}$ et $b \equiv b' \pmod{n}$ alors

$$a + b \equiv a' + b' \pmod{n} \quad \text{et} \quad ab \equiv a'b' \pmod{n}.$$

En particulier, on obtient par récurrence $a^p \equiv b^p \pmod{n}$ pour tout $p \in \mathbb{N}$.

3.2 PGCD et PPCM

3.2.1 PGCD

Définition

On appelle PGCD de deux entiers a et b le plus grand¹ diviseur commun à a et b . On le note $a \wedge b$.

Les diviseurs communs à a et b étant aussi les diviseurs communs à $|a|$ et $|b|$, il est usuel de travailler avec des entiers naturels lorsque l'on étudie un PGCD.

Théorème 7

Les diviseurs communs à deux entiers a et b sont exactement les diviseurs du PGCD de a et b .

On dispose ainsi de l'équivalence suivante où l'implication directe² est remarquable :

$$\forall d \in \mathbb{Z}, \quad (d | a \text{ et } d | b) \iff d | (a \wedge b).$$

3.2.2 Propriétés calculatoires

Théorème 8

Pour tous a , b et c entiers :

- a) $a \wedge 0 = |a|$;
- b) $a \wedge b = b \wedge a$ (commutativité);
- c) $a \wedge (b \wedge c) = (a \wedge b) \wedge c$ (associativité).

Par la propriété d'associativité, on peut écrire $a \wedge b \wedge c$ sans préciser le parenthésage organisant le calcul. On définit ainsi le PGCD des trois entiers a , b et c . Plus généralement, on peut introduire le PGCD $a_1 \wedge \dots \wedge a_n$ d'une famille (a_1, \dots, a_n) d'entiers. Celui-ci est l'unique entier naturel vérifiant

$$\forall d \in \mathbb{Z}, \quad (d | a_1 \text{ et } \dots \text{ et } d | a_n) \iff d | a_1 \wedge \dots \wedge a_n.$$

3.2.3 L'algorithme d'Euclide

Soit a et b deux entiers. Quitte à considérer leurs valeurs absolues, on suppose a et b positifs.

Si b est nul, on connaît le PGCD de a et b : $a \wedge 0 = a$.

Si b est non nul, le PGCD de a et b est aussi celui de b et de r avec r le reste de la division euclidienne de a par b : $a \wedge b = b \wedge r$.

1. Plus précisément, lorsque a ou b est non nul, l'ensemble des diviseurs communs à a et b est fini non vide et il existe donc un plus grand diviseur commun. Lorsque a et b sont tous deux nuls, on pose le PGCD égal à 0. Dans les deux cas, le PGCD est un plus grand diviseur pour la relation de divisibilité.

2. L'implication réciproque est immédiate par transitivité.

Ces deux propriétés sont à l'origine de l'*algorithme d'Euclide* calculant le PGCD de deux entiers naturels :

Théorème 9 (Algorithme d'Euclide)

Partant de la division euclidienne de a par b , lorsque l'on opère une succession de divisions euclidiennes du diviseur par le reste tant que le reste obtenu est non nul, le PGCD de a et b correspond au dernier reste non nul.

Par cette succession de divisions euclidiennes il est possible d'exprimer le PGCD de deux entiers comme une combinaison entière de ces deux entiers :

Théorème 10 (Relation de Bézout)

Si d désigne le PGCD de deux entiers a et b , on peut écrire

$$d = au + bv \quad \text{avec} \quad (u, v) \in \mathbb{Z}^2.$$

Plus généralement, si d est le PGCD d'une famille (a_1, \dots, a_n) d'entiers, on peut écrire

$$d = a_1u_1 + \dots + a_nu_n \quad \text{avec} \quad (u_1, \dots, u_n) \in \mathbb{Z}^n.$$

3.2.4 PPCM

Définition

On appelle PPCM de deux entiers a et b le plus petit entier naturel multiple commun à a et b . On le note $a \vee b$.

Par la propriété $a \vee b = |a| \vee |b|$, il est usuel d'étudier des PPCM d'entiers naturels.

Théorème 11

Les multiples communs à deux entiers a et b sont exactement les multiples du PPCM de a et b .

On dispose ainsi de l'équivalence suivante où l'implication directe est remarquable :

$$\forall m \in \mathbb{Z}, \quad (a \mid m \text{ et } b \mid m) \iff (a \vee b) \mid m.$$

On retrouve pour le PPCM les propriétés de commutativité et d'associativité déjà vues pour le PGCD. Par cette dernière, on peut introduire le PPCM d'une famille finie d'entiers.

3.2.5 Relation entre le PGCD et le PPCM

Les PGCD et PPCM de deux¹ entiers sont liés par la formule :

Théorème 12

Pour tous a et b entiers,

$$(a \wedge b)(a \vee b) = |ab|.$$

1. La relation n'est pas valable si l'on considère le PGCD et le PPCM d'une famille de trois nombres entiers ou plus.

3.3 Entiers premiers entre eux

3.3.1 Définition

Définition

Deux entiers a et b sont dits *premiers entre eux* lorsque leurs seuls diviseurs communs sont 1 et -1 .

Il revient au même de dire que leur PGCD vaut 1, c'est-à-dire d'écrire $a \wedge b = 1$.

Des diviseurs de deux nombres premiers entre eux sont eux-mêmes premiers entre eux.

Théorème 13 (Factorisation du PGCD)

Si a et b sont deux entiers de PGCD d , on peut écrire

$$a = da' \text{ et } b = db' \quad \text{avec} \quad a' \text{ et } b' \text{ des entiers premiers entre eux.}$$

C'est par factorisation du PGCD du numérateur et du dénominateur d'un nombre rationnel que l'on forme son représentant irréductible.

3.3.2 Le théorème de Bézout

Théorème 14 (Théorème de Bézout)

Deux entiers a et b sont premiers entre eux si, et seulement si, il existe un couple (u, v) formé d'entiers vérifiant $au + bv = 1$.

En application, si un entier a est premier avec les entiers b et c , il l'est aussi avec le produit bc . Par récurrence, on obtient que a est premier avec b^p pour tout $p \in \mathbb{N}$.

3.3.3 Le lemme de Gauss

Lorsqu'un entier a divise un produit bc , il ne divise pas nécessairement l'un des facteurs b ou c . Cependant, on a le résultat suivant :

Théorème 15 (Lemme de Gauss)

Soit a, b et c trois entiers.

$$a \mid bc \text{ et } a \wedge b = 1 \implies a \mid c.$$

En application, si deux entiers a et b premiers entre eux divisent un entier c , leur produit ab divise aussi c .

3.4 Nombres premiers

3.4.1 Définition

Définition

Soit p un entier au moins égal à 2. On dit que p est un *nombre premier* si ses seuls diviseurs positifs sont 1 et p . Sinon, l'entier p est dit *composé*.

Les douze premiers nombres¹ premiers sont : 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37.

Théorème 16 (Théorème d'Euclide sur les nombres premiers)

L'ensemble \mathcal{P} des nombres premiers est infini.

3.4.2 Le lemme d'Euclide

Le caractère atomique des nombres premiers se traduit par le résultat suivant :

Théorème 17 (Lemme d'Euclide)

Si un nombre premier p divise un produit ab de deux entiers, il divise au moins l'un des deux facteurs a ou b .

3.4.3 Le théorème fondamental de l'arithmétique

Théorème 18 (Théorème fondamental de l'arithmétique)

Tout entier naturel n non nul s'écrit de façon unique à l'ordre près des facteurs

$$n = \prod_{k=1}^r p_k^{\alpha_k}$$

avec $r \in \mathbb{N}$, p_1, \dots, p_r des nombres premiers deux à deux distincts et $\alpha_1, \dots, \alpha_r$ des entiers naturels non nuls.

Définition

|| L'écriture de l'entier n ci-dessus se nomme sa *décomposition en facteurs premiers*².

3.4.4 Valuation p -adique

Soit p un nombre premier.

1. On ne considère pas que 1 soit un nombre premier. On ne considère pas non plus qu'il s'agit d'un nombre composé : c'est un élément inversible dans l'anneau \mathbb{Z} (voir chapitre suivant).

2. On parle aussi de *décomposition primaire*. Lorsque $n = 1$, cette écriture est réalisée avec un produit vide.

Définition

On appelle *valuation p-adique* d'un entier non nul n l'exposant de la plus grande puissance de p qui divise n . Cette valuation est notée $v_p(n)$.

La valuation p -adique de n correspond à l'exposant¹ de p dans sa décomposition en facteurs premiers. On peut alors écrire²

$$n = \pm \prod_{p \in \mathcal{P}} p^{v_p(n)}.$$

Théorème 19

Si a et b sont deux entiers non nuls

- a) $v_p(ab) = v_p(a) + v_p(b)$;
- b) $a | b \iff \forall p \in \mathcal{P}, v_p(a) \leq v_p(b)$;
- c) $a \wedge b = 1 \iff \forall p \in \mathcal{P}, v_p(a) = 0$ ou $v_p(b) = 0$.

Si un entier n s'écrit

$$n = \prod_{k=1}^r p_k^{\alpha_k} \quad \text{avec } p_1, \dots, p_r \in \mathcal{P} \text{ deux à deux distincts}$$

les diviseurs positifs de n sont les

$$d = \prod_{k=1}^r p_k^{\beta_k} \quad \text{avec } \forall k \in [1; r], 0 \leq \beta_k \leq \alpha_k.$$

Si a et b sont des entiers non nuls

$$a \wedge b = \prod_{p \in \mathcal{P}} p^{\min(v_p(a), v_p(b))} \quad \text{et} \quad a \vee b = \prod_{p \in \mathcal{P}} p^{\max(v_p(a), v_p(b))}.$$

Le PGCD de a et b regroupe ce qui est « commun » aux entiers a et b .

3.4.5 Le petit théorème de Fermat**Théorème 20 (Petit théorème de Fermat)**

Si p est un nombre premier alors, pour tout entier a ,

$$p \nmid a \implies a^{p-1} \equiv 1 \pmod{p}.$$

En conséquence $a^p \equiv a \pmod{p}$ que p divise a ou non.

1. Cet exposant est nul si p n'est pas un facteur premier de n .

2. Soulignons que ce produit ne contient qu'un nombre fini de facteurs différents de 1 ce qui permet de lui attribuer un sens bien qu'il porte sur un ensemble d'indexation infini.

3.5 Exercices d'apprentissage

3.5.1 Divisibilité

Exercice 1

Soit $n \in \mathbb{N}$. Établir les divisibilités suivantes :

$$(a) 6 \mid 5n^3 + n$$

$$(b) 5 \mid 2^{2n+1} + 3^{2n+1}$$

$$(c) 9 \mid 4^n - 1 + 6n.$$

Solution

(a) méthode

|| On vérifie¹ que $5n^3 + n$ est nul modulo 6.

En étudiant les six valeurs possibles de n modulo 6, on observe $n^3 \equiv n \pmod{6}$. On a donc $5n^3 + n \equiv 6n^3 \equiv 0 \pmod{6}$. Ceci assure la divisibilité affirmée.

(b) méthode

|| On vérifie que $2^{2n+1} + 3^{2n+1}$ est nul modulo 5.

On a $2^{2n} = 4^n$ et $3^{2n} = 9^n \equiv 4^n \pmod{5}$ donc

$$2^{2n+1} + 3^{2n+1} \equiv 2 \times 4^n + 3 \times 4^n \equiv 5 \times 4^n \equiv 0 \pmod{5}.$$

(c) méthode

|| On factorise² $4^n - 1 + 6n$ afin de faire apparaître des divisibilités par 3.

Par la formule de factorisation géométrique (Th. 12 p. 53)

$$4^n - 1 = (4 - 1)(1 + 4 + \cdots + 4^{n-1})$$

et donc

$$4^n - 1 + 6n = 3 \times (1 + 4 + \cdots + 4^{n-1} + 2n).$$

Or

$$1 + 4 + \cdots + 4^{n-1} + 2n \equiv \underbrace{1 + 1 + \cdots + 1}_{n \text{ termes}} + 2n \equiv 3n \equiv 0 \pmod{3}.$$

On peut donc conclure que 9 divise $4^n - 1 + 6n$.

Exercice 2 (Équations diophantiennes)

Déterminer les couples $(x, y) \in \mathbb{Z}^2$ vérifiant

$$(a) xy = 3x + y + 2$$

$$(b) x^2 - 6x - y^2 - 2y = 4.$$

1. On peut aussi raisonner par récurrence.
2. On peut encore raisonner par récurrence.

Solution**(a) méthode**

On forme une équation équivalente où le problème est transformé en celui de la recherche des diviseurs d'un entier.

En passant les inconnues x et y en premier membre, on factorise l'expression en x et y à une constante additive près

$$\begin{aligned} xy = 3x + y + 2 &\iff xy - 3x - y = 2 \\ &\iff (x-1)(y-3) = 5. \end{aligned}$$

Un couple (x, y) est solution si, et seulement si, le couple $(x-1, y-3)$ est constitué de deux diviseurs associés de 5. Les couples correspondants étant $(5, 1)$, $(1, 5)$, $(-5, -1)$ et $(-1, -5)$, les solutions respectives sont

$$(6, 4), (2, 8), (-4, 2) \text{ et } (0, -2).$$

(b) méthode

On écrit sous forme canonique les deux expressions du second degré en x et y avant de factoriser une différence de deux carrés.

$$\begin{aligned} x^2 - 6x - y^2 - 2y = 4 &\iff (x-3)^2 - (y+1)^2 = 12 \\ &\iff ((x-3) - (y+1))((x-3) + (y+1)) = 12 \\ &\iff (x-y-4)(x+y-2) = 12. \end{aligned}$$

Les couples (x, y) solutions se déduisent des factorisations $12 = ab$ avec a, b entiers.

$$\left\{ \begin{array}{l} x-y-4=a \\ x+y-2=b \end{array} \right. \iff \left\{ \begin{array}{l} x=\frac{a+b}{2}+3 \\ y=\frac{b-a}{2}-1 \end{array} \right.$$

En éliminant les couples qui ne sont pas constitués d'entiers, les solutions de l'équation étudiée sont les couples formés à partir de $(a, b) = (6, 2), (2, 6), (-6, -2)$ et $(-2, -6)$ à savoir respectivement

$$(7, -3), (7, 1), (-1, 1) \text{ et } (-1, -3).$$

3.5.2 PGCD et PPCM**Exercice 3**

Calculer le PGCD d des entiers $a = 33$ et $b = 24$ ainsi qu'un couple (u, v) de coefficients entiers exprimant une relation de Bézout $d = au + bv$.

Solution**méthode**

On calcule le PGCD de a par b par l'algorithme d'Euclide (Th. 9 p. 90) en exprimant parallèlement chaque reste comme combinaison entière de a et b .

$$\begin{array}{ll} 33 = 24 \times 1 + 9 & 9 = 33 - 24 \times 1 = a - b \\ 24 = 9 \times 2 + 6 & 6 = 24 - 9 \times 2 = b - (a - b) \times 2 = 3b - 2a \\ 9 = 6 \times 1 + 3 & 3 = 9 - 6 \times 1 = (a - b) - (3b - 2a) = 3a - 4b \\ 6 = 3 \times 2 + 0. & \end{array}$$

Le PGCD de 33 et 24 vaut 3 et l'on peut écrire $3 = au + bv$ avec¹ $(u, v) = (3, -4)$.

Exercice 4

Soit a et b deux entiers naturels avec b non nul.

(a) Montrer que si r est le reste de la division euclidienne de a par b alors $2^r - 1$ est le reste de la division euclidienne de $2^a - 1$ par $2^b - 1$.

(b) En déduire

$$(2^a - 1) \wedge (2^b - 1) = 2^{a \wedge b} - 1.$$

Solution

(a) La division euclidienne de a par b s'écrit $a = bq + r$ avec $0 \leq r < b$.

méthode

On vérifie que $2^b - 1$ divise la différence de $2^a - 1$ et $2^r - 1$.

$$(2^a - 1) - (2^r - 1) = 2^{bq+r} - 2^r = 2^r(2^{bq} - 1).$$

Par la formule de factorisation géométrique (Th. 12 p. 53)

$$2^{bq} - 1 = (2^b - 1)(1 + 2^b + \dots + 2^{b(q-1)})$$

et donc $2^b - 1$ divise $(2^a - 1) - (2^r - 1)$. Ceci permet d'écrire

$$2^a - 1 = (2^b - 1)Q + 2^r - 1 \quad \text{avec} \quad Q \in \mathbb{Z}.$$

méthode

Cette identité ne suffit pas pour identifier le reste de la division euclidienne, il faut aussi vérifier un encadrement (Th. 4 p. 88).

Puisque $0 \leq r < b$, on a $0 \leq 2^r - 1 < 2^b - 1$ et donc $2^r - 1$ est bien le reste de la division euclidienne de $2^a - 1$ par $2^b - 1$.

¹. Ce couple n'est pas unique, les autres couples solutions sont les $(3+8k, -4-11k)$ avec k parcourant \mathbb{Z} (voir sujet 13 p. 102).

(b) méthode

Par l'algorithme d'Euclide, on calcule le PGCD de $2^a - 1$ et $2^b - 1$ parallèlement au calcul du PGCD de a et b .

Posons $a_0 = a$, $a_1 = b$, \dots le reste de la division euclidienne de a_0 par a_1 et, plus généralement, a_{k+1} le reste de la division euclidienne de a_{k-1} par a_k tant que a_k est non nul. En notant $m+1$ le rang pour lequel le reste a_{m+1} est nul, on sait que a_m désigne le PGCD de a et b .

Pour tout k compris entre 1 et m , on a

$$(2^{a_{k-1}} - 1) \wedge (2^{a_k} - 1) = (2^{a_k} - 1) \wedge (2^{a_{k+1}} - 1)$$

car $2^{a_{k+1}} - 1$ est le reste de la division euclidienne de $2^{a_{k-1}} - 1$ par $2^{a_k} - 1$. On en déduit

$$\begin{aligned} (2^a - 1) \wedge (2^b - 1) &= (2^{a_0} - 1) \wedge (2^{a_1} - 1) = \dots \\ &= \underbrace{(2^{a_m} - 1)}_{2^{a \wedge b} - 1} \wedge \underbrace{(2^{a_{m+1}} - 1)}_{=0} = 2^{a \wedge b} - 1. \end{aligned}$$

Exercice 5

Soit a, b et k des entiers. Etablir

$$a \wedge b = (a + kb) \wedge b.$$

Solution**méthode**

En arithmétique, on établit fréquemment l'égalité de deux entiers naturels par double divisibilité (Th. 1 p. 87).

Posons d le PGCD de a et b et δ celui de $a + kb$ et b . Le PGCD d divise la combinaison entière $a + kb$ (Th. 3 p. 88) car c'est un diviseur commun à a et b . L'entier d est donc un diviseur commun à $a + kb$ et b et c'est donc un diviseur de δ (Th. 7 p. 89). Inversement, le PGCD δ divise $a + kb$ et b , il divise donc aussi $a = a + kb - kb$ et c'est alors un diviseur de d . Par double divisibilité, on peut conclure que d et δ sont égaux¹.

3.5.3 Entiers premiers entre eux**Exercice 6**

Montrer que, pour tout $n \in \mathbb{Z}$, les entiers n , $n + 1$ et $2n + 1$ sont premiers entre eux deux à deux. Que dire des entiers $n^2 + n$ et $2n + 1$?

1. Lorsque b est non nul, si $a \equiv a' \pmod{b}$ alors $a \wedge b = a' \wedge b$.

Solution**méthode**

On peut vérifier que deux entiers sont premiers entre eux en calculant leur PGCD ou en exploitant le théorème de Bézout (Th. 14 p. 91).

Si d désigne le PGCD de n et $n+1$ alors d divise n , d divise $n+1$ et donc divise la différence $n+1-n=1$. On en déduit $d=1$ et l'on peut affirmer que les entiers n et $n+1$ sont premiers entre eux. Plus rapidement, on peut dire que n et $n+1$ sont premiers entre eux en vertu de l'égalité de Bézout

$$n \times (-1) + (n+1) \times 1 = 1.$$

On affirme de même que n et $2n+1$ sont premiers entre eux d'une part, et aussi $n+1$ et $2n+1$ d'autre part, par les égalités

$$n \times (-2) + (2n+1) \times 1 = 1 \quad \text{et} \quad (n+1) \times 2 + (2n+1) \times (-1) = 1.$$

Enfin¹, $2n+1$ étant premier avec les facteurs n et $n+1$, il l'est aussi avec leur produit $n(n+1)=n^2+n$.

Exercice 7

Soit a , b et c trois entiers avec a et c premiers entre eux.

- (a) Montrer $a \wedge (bc) = a \wedge b$.
- (b) Que dire de $a \vee (bc)$?

Solution

(a) Posons d le PGCD de a et b et δ celui de a et bc . Montrons que d et δ sont égaux par double divisibilité.

L'entier d divise a , il divise aussi b et *a fortiori* bc , il divise donc δ .

Inversement, l'entier δ divise a et le produit bc .

méthode

Sachant que δ divise le produit bc , il suffit de vérifier qu'il est premier² avec c pour affirmer qu'il divise b par le lemme de Gauss (Th. 15 p. 91).

Puisque δ divise a et que ce dernier est premier avec c , δ est aussi premier³ avec c . Par le lemme de Gauss, on peut alors affirmer que δ divise b et, finalement, δ divise d car il est diviseur commun à a et b .

Par double divisibilité, on peut conclure que d et δ sont égaux.

1. On pourrait aussi constater une égalité de Bézout : $(n^2+n) \times (-2) + (2n+1) \times (1-n) = 1$.

2. Affirmer que δ ne divise pas c n'est pas suffisant !

3. Les diviseurs communs à δ et c sont *a fortiori* des diviseurs communs à a et b .

(b) méthode

|| PGCD et PPCM sont liés par la relation $(a \wedge b)(a \vee b) = |ab|$ (Th. 12 p. 90).

On a donc

$$\underbrace{(a \wedge (bc))}_{=a \wedge b} (a \vee (bc)) = |ab| |c| = (a \wedge b)(a \vee b) |c|.$$

Cas : $a \wedge b$ est non nul. On peut simplifier par $a \wedge b$ et affirmer $a \vee (bc) = (a \vee b) |c|$.

Cas : $a \wedge b$ est nul. Les entiers a et b qui en sont multiples sont aussi nuls et la relation ci-dessus est encore vraie car elle correspond à l'égalité de deux 0.

Exercice 8

Soit a et b deux entiers relatifs. Montrer

$$a \mid b \iff a^2 \mid b^2.$$

Solution

L'implication directe est immédiate : si $b = ak$ avec k entier alors $b^2 = a^2\ell$ avec $\ell = k^2$ entier. Étudions l'implication réciproque.

méthode

|| On écrit¹ a et b en factorisant leur PGCD (Th. 13 p. 91).

Soit $d = a \wedge b$. On peut écrire $a = da'$ et $b = db'$ avec a' et b' entiers premiers entre eux. Supposons que a^2 divise b^2 c'est-à-dire $d^2a'^2$ divise $d^2b'^2$. Pour pouvoir simplifier par d^2 , on traite séparément le cas où $d = 0$.

Cas : $d = 0$. On a a et b nuls car multiples de d et la propriété voulue est immédiate.

Cas : $d \neq 0$. On peut simplifier par d^2 puis affirmer que a' divise le produit $b'^2 \times 1$. Or a' est premier avec b' donc aussi avec b'^2 . Par le lemme de Gauss (Th. 15 p. 91) on peut affirmer que a' divise le facteur 1 et donc $a' = \pm 1$. On en déduit $a = \pm d$ et a est alors un diviseur de b .

3.5.4 Nombres premiers**Exercice 9**

Soit p un nombre premier et a un entier. Montrer

$$p \mid a \text{ ou } p \text{ et } a \text{ sont premiers entre eux.}$$

1. Dans le cas où a et b sont non nuls, on peut aussi exploiter $a^2 \mid b^2 \iff v_p(a^2) \leq v_p(b^2)$ pour tout p nombre premier (Th. 19 p. 93).

Solution**méthode**

|| On calcule le PGCD de p et a .

Soit $d = p \wedge a$. L'entier d est diviseur positif du nombre premier p , il vaut donc 1 ou p .

Cas : $d = 1$. Les entiers a et p sont premiers entre eux.

Cas : $d = p$. L'entier p divise a car le PGCD d est un diviseur de a .

Exercice 10 (Morphisme de Frobenius)

Soit p un nombre premier.

(a) Pour tout $1 \leq k < p$, montrer que le coefficient binomial $\binom{p}{k}$ est un multiple de p .

(b) En déduire que, pour tous entiers a et b , $(a + b)^p \equiv a^p + b^p \pmod{p}$.

Solution

(a) Soit $k \in \llbracket 1 ; p - 1 \rrbracket$.

méthode

|| On exploite la formule¹

$$\binom{p}{k} = \frac{p(p-1)}{k(k-1)}.$$

On réécrit² la formule proposée

$$p \binom{p-1}{k-1} = k \binom{p}{k}.$$

Les coefficients binomiaux étant des nombres entiers, on peut affirmer que p divise le produit $k \binom{p}{k}$.

méthode

|| Lorsqu'un entier divise un produit et que l'on souhaite établir qu'il divise l'un des facteurs, on peut penser au lemme de Gauss (Th. 15 p. 91).

Puisque le nombre premier p ne divise pas k , p et k sont premiers entre eux³ et donc p divise le coefficient binomial $\binom{p}{k}$.

(b) **méthode**

|| On développe $(a + b)^p$ par la formule du binôme (Th. 11 p. 53).

1. Voir sujet 8 p. 63.

2. Il n'est pas possible de faire de l'arithmétique avec des nombres rationnels : il faut réécrire la formule de sorte de ne manipuler que des nombres entiers.

3. Voir le sujet précédent.

On isole les termes d'indices 0 et p de la somme et l'on simplifie par congruence les termes intermédiaires

$$(a+b)^p = \sum_{k=0}^p \binom{p}{k} a^{p-k} b^k = a^p + \underbrace{\sum_{k=1}^{p-1} \binom{p}{k} a^{p-k} b^k}_{\equiv 0 [p]} + b^p \equiv a^p + b^p [p].$$

Notons qu'en raisonnant par récurrence sur n , on peut alors aisément établir $n^p \equiv n$ [p] : on retrouve ainsi un énoncé possible du petit théorème de Fermat (Th. 20 p. 93).

3.6 Exercices d'entraînement

3.6.1 Études arithmétiques

Exercice 11 *

Déterminer les $x \in \mathbb{Z}$ tels que

$$(a) (x-2) \mid (x+2) \quad (b) (x-1) \mid (x^2+x+1).$$

Solution

(a) méthode

On détermine les valeurs de x pour lesquelles le quotient est un entier en décomposant celui-ci en somme.

Soit $x \in \mathbb{Z}$. L'entier $x = 2$ n'est pas solution, on suppose dans la suite $x \neq 2$.

$$\begin{aligned} (x-2) \mid (x+2) &\iff \frac{x+2}{x-2} = 1 + \frac{4}{x-2} \in \mathbb{Z} \\ &\iff (x-2) \mid 4. \end{aligned}$$

A partir des diviseurs de 4, on détermine les valeurs de x solutions : $-2, 0, 1, 3, 4$ et 6 .

(b) On reprend la même démarche. Soit $x \in \mathbb{Z}$. L'entier $x = 1$ n'est pas solution, on suppose dans la suite $x \neq 1$.

$$\begin{aligned} (x-1) \mid (x^2+x+1) &\iff \frac{x^2+x+1}{x-1} = x+2 + \frac{3}{x-1} \in \mathbb{Z} \\ &\iff (x-1) \mid 3. \end{aligned}$$

Les valeurs de x solutions sont : $-2, 0, 2, 4$.

Exercice 12 *

Soit x et y deux entiers. Montrer

$$7 \mid x \text{ et } 7 \mid y \iff 7 \mid x^2 + y^2.$$

Solution**méthode**

|| On calcule les valeurs possibles des carrés modulo 7.

On obtient le tableau suivant¹ :

x	0	1	2	3	4	5	6
x^2	0	1	4	2	2	4	1

En sommant deux entiers parmi 0, 1, 2 et 4, la seule possibilité pour obtenir 0 modulo 7 est de sommer deux fois 0. On en déduit que, pour qu'une somme de deux carrés soit nulle modulo 7, il faut et il suffit que les deux carrés le soient et donc

$$\begin{aligned} 7 \mid x^2 + y^2 &\iff x^2 + y^2 \equiv 0 \pmod{7} \\ &\iff x \equiv 0 \pmod{7} \text{ et } y \equiv 0 \pmod{7} \\ &\iff 7 \mid x \text{ et } 7 \mid y. \end{aligned}$$

Exercice 13 ** (Coefficients de l'égalité de Bézout)

Soit a et b deux entiers non nuls de PGCD d . Notre but est de déterminer tous les couples $(u, v) \in \mathbb{Z}^2$ tels que $au + bv = d$.

(a) Justifier l'existence d'un couple solution (u_0, v_0) .

(b) Exprimer à partir de (u_0, v_0) tous les couples solutions.

Solution

(a) Une relation de Bézout (Th. 10 p. 90) assure immédiatement l'existence d'un couple (u_0, v_0) solution.

(b) **méthode**

|| On exploite la solution particulière (u_0, v_0) pour retraduire l'équation étudiée et exploiter les outils d'arithmétique.

Soit $(u, v) \in \mathbb{Z}^2$.

$$\begin{aligned} au + bv = d &\iff au + bv = au_0 + bv_0 \\ &\iff a(u - u_0) = b(v_0 - v). \end{aligned}$$

Le PGCD d divise les deux membres de cette égalité.

méthode

|| On factorise a et b par leur PGCD (Th. 13 p. 91) afin de pouvoir simplifier par celui-ci.

1. Ce tableau comporte une symétrie que l'on peut aisément expliquer : $(7 - k)^2 \equiv k^2 \pmod{7}$.

On peut écrire $a = da'$ et $b = db'$ avec a' et b' premiers entre eux. On poursuit alors la chaîne d'équivalence précédente en simplifiant par d qui est non nul

$$au + bv = d \iff a'(u - u_0) = b'(v_0 - v).$$

Poursuivons la résolution en raisonnant par double implication.

Soit (u, v) un couple solution. L'entier a' divise le produit $b'(v_0 - v)$, or il est premier avec le facteur b' , le lemme de Gauss assure donc que a' divise $v_0 - v$. Il existe alors un entier k permettant d'écrire $v_0 - v = ka'$ soit encore $v = v_0 - ka'$. En injectant cette écriture dans l'égalité $a'(u - u_0) = b'(v_0 - v)$, il vient $a'(u - u_0) = a'b'k$. En simplifiant par a' (qui est non nul), on obtient $u = u_0 + kb'$.

La réciproque étant immédiate, on peut affirmer que les couples (u, v) solutions sont :

$$(u_0 + kb', v_0 - kd') \quad \text{avec } k \text{ parcourant } \mathbb{Z}.$$

Exercice 14 **

Soit m et n deux entiers naturels non nuls premiers entre eux.

(a) Soit x un nombre rationnel. On suppose que x^n est entier, montrer que x est entier.

(b) Soit a et b deux entiers non nuls tels que $a^n = b^m$. Montrer qu'il existe $c \in \mathbb{N}^*$ tel que $a = c^m$ et $b = c^n$.

Solution

(a) méthode

On écrit x sous forme irréductible et l'on traduit la propriété $x^n \in \mathbb{Z}$ comme une égalité entre nombres entiers.

On écrit le nombre rationnel x sous forme irréductible p/q avec $p \in \mathbb{Z}$, $q \in \mathbb{N}^*$, p et q premiers entre eux. En posant k la valeur de x^n , on a l'identité $p^n = kq^n$. On en déduit que q divise le produit $p^n = p^n \times 1$. Or p et q sont premiers entre eux et donc p^n et q le sont aussi. Par le lemme de Gauss (Th. 15 p. 91), on peut affirmer que q divise le facteur 1 et donc $q = 1$. On en déduit $x = p \in \mathbb{Z}$.

(b) méthode

Les entiers m et n étant premiers entre eux, on peut introduire $(u, v) \in \mathbb{Z}^2$ tel que $mu + nv = 1$ (Th. 14 p. 91).

Analyse : Supposons que c soit une valeur solution. On peut calculer c en écrivant

$$c = c^1 = c^{mu+nv} = (c^m)^u (c^n)^v = b^u a^v.$$

Synthèse : Posons $c = b^u a^v$. On vérifie :

$$c^n = b^{nu} a^{nv} = (b^n)^u a^{nv} = a^{mu} a^{nv} = a^{mu+nv} = a.$$

De même, on obtient $c^m = b$. Cependant, c est *a priori* un nombre rationnel car les exposants u et v sont des entiers relatifs. Puisque c^n est un nombre entier, l'étude de la question précédente assure que c est nombre entier, au surplus, non nul.

Exercice 15 ** (Triplets pythagoriciens)

On étudie l'équation (E): $a^2 + b^2 = c^2$ d'inconnue $(a, b, c) \in \mathbb{Z}^3$.

Soit (a, b, c) une solution non nulle. Si d désigne le PGCD des trois entiers a , b et c , on peut simplifier l'équation par d^2 et se ramener à la situation où $d = 1$: nous supposons par la suite être dans cette situation.

- (a) Montrer que a , b et c sont deux à deux premiers entre eux.
- (b) Quelles sont les congruences possibles d'un carré modulo 4 ? En déduire que l'une des deux valeurs a ou b est paire et l'autre impaire.
- On suppose que b est pair.
- (c) Montrer qu'il existe x et y entiers tels que $c + a = 2x^2$ et $c - a = 2y^2$.
- (d) Quels sont les triplets $(a, b, c) \in \mathbb{Z}^3$ solutions de l'équation $a^2 + b^2 = c^2$?

Solution

(a) Si k est diviseur commun à a et b alors k^2 divise c^2 et donc¹ k divise c . Or le PGCD de a , b et c est supposé égal à 1 et donc k divise 1. Ainsi, a et b sont premiers entre eux. On montre de même que a et c d'une part, b et c d'autre part, sont aussi premiers entre eux.

(b) Selon que x est pair ou impair son carré est congru à 0 ou 1 modulo 4. Si a et b sont tous deux impairs, $c^2 = a^2 + b^2$ est congru à 2 modulo 4 ce qui est impossible. Si a et b sont tous deux pairs, ils ne sont pas premiers entre eux ce qui est aussi exclu. L'une des valeurs est donc paire et l'autre impaire. Notons que c est un nombre impair.

(c) méthode

|| On factorise le second membre de l'égalité $b^2 = c^2 - a^2$.

On écrit $b = 2p$ avec p entier et l'égalité $a^2 + b^2 = c^2$ donne

$$p^2 = k\ell \quad \text{avec} \quad k = \frac{c+a}{2} \quad \text{et} \quad \ell = \frac{c-a}{2}.$$

Les nombres a et c étant impairs, k et ℓ sont des entiers.

méthode

|| On montre que k et ℓ sont des carrés en vérifiant qu'ils sont premiers entre eux.

Un diviseur commun à k et ℓ divise $a = k - \ell$ et $c = k + \ell$. Les entiers a et c étant premiers entre eux, il en est de même de k et ℓ . Le produit $k\ell$ étant un carré et les entiers k et ℓ n'ayant aucun facteurs premiers en commun, ce sont chacun des carrés². On en déduit les écritures proposées de $c + a$ et de $c - a$.

1. Voir sujet 8 p. 99.

2. Plus précisément, $k = x^2$ avec $x = p \wedge k$ et $\ell = y^2$ avec $y = p \wedge \ell$.

(d) On a alors

$$a = x^2 - y^2, \quad b = 2xy \quad \text{et} \quad c = x^2 + y^2 \quad \text{avec} \quad (x, y) \in \mathbb{Z}^2.$$

On vérifie par un simple calcul qu'un tel triplet est solution. Enfin, on retrouve la généralité des solutions, en multipliant a , b et c par un même facteur d entier et en offrant la possibilité d'échanger les expressions de a et b .

Exercice 16 **

On note $\text{Div}(n)$ l'ensemble des diviseurs positifs d'un entier $n \in \mathbb{Z}$.

Soit a, b deux entiers premiers entre eux. Établir la bijectivité de l'application

$$\varphi : \begin{cases} \text{Div}(a) \times \text{Div}(b) \rightarrow \text{Div}(ab) \\ (k, \ell) \mapsto k\ell \end{cases}$$

Solution

On vérifie facilement que φ est bien définie à valeurs dans $\text{Div}(ab)$: si k divise a et ℓ divise b alors $k\ell$ divise ab . On montre ensuite que φ réalise une bijection en observant qu'elle est injective et surjective.

Soit (k, ℓ) et (k', ℓ') deux couples éléments de $\text{Div}(a) \times \text{Div}(b)$ tels que $\varphi(k, \ell) = \varphi(k', \ell')$, c'est-à-dire $k\ell = k'\ell'$.

méthode

Les diviseurs de deux entiers premiers entre eux sont eux-mêmes premiers entre eux.

Les entiers a et b étant premiers entre eux, les entiers k et ℓ sont premiers entre eux. Or k divise le produit $k'\ell'$, il divise donc le facteur k' en vertu du lemme de Gauss. Par un raisonnement symétrique, on obtient aussi que k' divise k et l'on en déduit que les naturels k et k' sont égaux. Un raisonnement analogue¹ donne $\ell = \ell'$ et l'on peut affirmer que l'application φ est injective.

Montrons maintenant que φ est surjective. Soit $d \in \text{Div}(ab)$.

méthode

On définit un antécédent $(k, \ell) \in \text{Div}(a) \times \text{Div}(b)$ au diviseur d en choisissant pour k ce qui est « commun » à a et d et en procédant de même pour ℓ avec b et d .

Posons² $k = a \wedge d$ et $\ell = b \wedge d$. On a évidemment $(k, \ell) \in \text{Div}(a) \times \text{Div}(b)$. Vérifions par double divisibilité que le produit $k\ell$ correspond à d .

D'une part, k et ℓ sont premiers entre et divisent chacun d . Leur produit divise donc aussi d .

1. Plutôt qu'une simplification par k , cet argument évite d'avoir à traiter séparément le cas $k = 0$, cas que l'on peut rencontrer lorsque $a = 0$ et $b = 1$.

2. En d'autres termes, on regroupe dans la décomposition primaire de d les facteurs premiers qui sont en commun à a pour former k et les autres pour former ℓ .

D'autre part, on peut écrire les relations de Bézout $k = au + dv$ et $\ell = bu' + dv'$ avec u, v, u', v' des entiers. On en déduit $k\ell = abU + dV$ avec U, V entiers. Puisque d est diviseur de ab , on obtient que d divise aussi $k\ell$.

Finalement, on a $d = k\ell$ et l'on peut affirmer que l'application \square est surjective.

Exercice 17 ** (Théorème RSA)

Soit p et q deux nombres premiers distincts, $n = pq$ et e un entier naturel premier avec le produit $(p - 1)(q - 1)$.

- Justifier qu'il existe un entier $d \geq 0$ tel que $ed \equiv 1 \pmod{(p-1)(q-1)}$.
- Montrer que $x^{ed} \equiv x \pmod{n}$ pour tout entier x .

Solution

- (a) Par le théorème de Bézout, il existe des entiers u et v tels que

$$eu + (p - 1)(q - 1)v = 1.$$

On a donc

$$eu \equiv 1 \pmod{(p-1)(q-1)}.$$

Cependant, on ignore si l'entier u est positif. Considérons alors le reste d de la division euclidienne de u par $(p - 1)(q - 1)$. Celui-ci convient car $d \geq 0$ et $u \equiv d \pmod{(p-1)(q-1)}$ donc

$$ed \equiv 1 \pmod{(p-1)(q-1)}.$$

- (b) Soit $x \in \mathbb{Z}$.

méthode

On vérifie par le petit théorème de Fermat (Th. 20 p. 93) que p et q divisent le nombre $x^{ed} - x$.

On introduit¹ $k \in \mathbb{N}$ tel que $ed = 1 + k(p - 1)(q - 1)$.

Cas : p divise x . L'entier p divise x^{ed} car $ed \geq 1$ et donc divise aussi $x^{ed} - x$.

Cas : p ne divise pas x . Le petit théorème de Fermat donne $x^{p-1} \equiv 1 \pmod{p}$. On en déduit

$$x^{ed} = x \times (x^{p-1})^{k(q-1)} \equiv x \times 1^{k(q-1)} \equiv x \pmod{p}$$

et donc à nouveau p divise $x^{ed} - x$.

On montre de même que q divise $x^{ed} - x$. Sachant que p et q sont deux nombres premiers distincts, on peut affirmer que $n = pq$ divise $x^{ed} - x$.

1. L'entier k est nécessairement positif car e et d le sont.

Exercice 18 * (Suite de Fibonacci)**

On considère la suite (φ_n) déterminée par

$$\varphi_0 = 0, \varphi_1 = 1 \quad \text{et} \quad \forall n \in \mathbb{N}, \varphi_{n+2} = \varphi_{n+1} + \varphi_n.$$

- (a) Vérifier que, pour tout $n \in \mathbb{N}$, φ_n et φ_{n+1} sont des entiers premiers entre eux.
 (b) Soit $k \in \mathbb{N}^*$. Montrer

$$\varphi_{k+n} = \varphi_k \varphi_{n+1} + \varphi_{k-1} \varphi_n \quad \text{pour tout } n \in \mathbb{N}.$$

Soit $a \in \mathbb{N}$ et $b \in \mathbb{N}^*$.

- (c) Établir

$$\varphi_{a+b} \wedge \varphi_b = \varphi_a \wedge \varphi_b \quad \text{puis} \quad \varphi_a \wedge \varphi_b = \varphi_b \wedge \varphi_r$$

où r est le reste de la division euclidienne de a par b .

- (d) Conclure

$$\varphi_a \wedge \varphi_b = \varphi_{a \wedge b}.$$

Solution

(a) On vérifie par récurrence double que φ_n est un nombre entier : la propriété est vraie aux rangs 0 et 1 et l'est aussi au rang $n+2$ lorsqu'elle l'est aux rangs n et $n+1$.

méthode

|| On sait¹ $(a+kb) \wedge b = a \wedge b$ pour tous entiers a, b et k .

On a donc $\varphi_{n+2} \wedge \varphi_{n+1} = (\varphi_{n+1} + \varphi_n) \wedge \varphi_n = \varphi_{n+1} \wedge \varphi_n$. Le PGCD de deux termes consécutifs de la suite est constant égal à $\varphi_1 \wedge \varphi_0 = 1$. Les entiers φ_n et φ_{n+1} sont donc premiers entre eux pour tout $n \in \mathbb{N}$.

- (b) Soit $k \in \mathbb{N}^*$.

méthode

|| On vérifie que les deux suites $(\varphi_{k+n})_{n \in \mathbb{N}}$ et $(\varphi_k \varphi_{n+1} + \varphi_{k-1} \varphi_n)_{n \in \mathbb{N}}$ satisfont les mêmes conditions de récurrence.

Posons $u_n = \varphi_{k+n}$ et $v_n = \varphi_k \varphi_{n+1} + \varphi_{k-1} \varphi_n$ pour tout $n \in \mathbb{N}$.

On a $u_0 = \varphi_k = \varphi_k \varphi_1 + \varphi_{k-1} \varphi_0 = v_0$.

On a aussi $u_1 = \varphi_{k+1} = \varphi_k + \varphi_{k-1} = \varphi_k \varphi_2 + \varphi_{k-1} \varphi_1 = v_1$ car $\varphi_2 = 1$.

Enfin, on vérifie aisément les relations $u_{n+2} = u_{n+1} + u_n$ et $v_{n+2} = v_{n+1} + v_n$ pour tout naturel n . Les suites (u_n) et (v_n) satisfont la même relation de récurrence double ainsi que les deux mêmes conditions initiales : elles sont égales.

1. Voir sujet 5 p. 97.

(c) Par la propriété $(a + kb) \wedge b = a \wedge b$,

$$\varphi_{a+b} \wedge \varphi_b = (\varphi_a \varphi_{b-1} + \varphi_{a+1} \varphi_b) \wedge \varphi_b \underset{\in \mathbb{Z}}{=} (\varphi_a \varphi_{b-1}) \wedge \varphi_b.$$

Aussi, les entiers φ_b et φ_{b-1} étant premiers entre eux, on peut poursuivre¹ et affirmer $\varphi_{a+b} \wedge \varphi_b = \varphi_a \wedge \varphi_b$.

Par récurrence, on obtient $\varphi_a \wedge \varphi_b = \varphi_{a+qb} \wedge \varphi_b$ pour tout $q \in \mathbb{N}$. En écrivant $a = bq + r$ la division euclidienne de a par b et en exploitant la propriété précédente avec r au lieu de a , on obtient $\varphi_r \wedge \varphi_b = \varphi_{r+bq} \wedge \varphi_b = \varphi_a \wedge \varphi_b$.

(d) Comme dans le sujet 4 p. 96, on calcule le PGCD de φ_a et φ_b parallèlement au calcul du PGCD de a et b par l'algorithme d'Euclide. En reprenant les notations de ce sujet, on a pour tout k compris entre 1 et m

$$\varphi_{a_{k-1}} \wedge \varphi_{a_k} = \varphi_{a_k} \wedge \varphi_{a_{k+1}}$$

et donc

$$\varphi_a \wedge \varphi_b = \varphi_{a_0} \wedge \varphi_{a_1} = \cdots = \varphi_{a_m} \wedge \varphi_0 = \varphi_a \wedge \varphi_b.$$

3.6.2 Nombres premiers

Exercice 19 *

Soit p un nombre premier supérieur à 5. Montrer que $p^2 - 1$ est divisible par 24.

Solution

méthode

|| On vérifie que 3 et 8 divisent $p^2 - 1 = (p - 1)(p + 1)$.

Le nombre premier p est impair et donc les nombres $p - 1$ et $p + 1$ sont tous deux des entiers pairs. De plus, ce sont des entiers pairs consécutifs et donc l'un d'eux est divisible par 4. Ainsi, 8 divise $p^2 - 1$.

Les entiers $p - 1$, p et $p + 1$ sont trois entiers consécutifs, l'un d'eux est divisible par 3. Ce ne peut pas être l'entier p car celui-ci est premier au moins égal à 5. Ainsi, 3 divise $p^2 - 1$.

Enfin, 3 et 8 sont premiers entre eux et leur produit divise aussi $p^2 - 1$.

Exercice 20 **

Soit a et p deux entiers supérieurs à 2. Montrer que, si $a^p - 1$ est un nombre premier, alors $a = 2$ et p est premier.

1. Voir sujet 7 p. 98.

Solution**méthode**

On exploite la factorisation géométrique (Th. 12 p. 53) :

$$a^p - 1 = (a - 1)(1 + a + \cdots + a^{p-1}).$$

Supposons $a^p - 1$ premier. Par l'identité ci-dessus, $a - 1$ divise $a^p - 1$ et donc $a - 1 = 1$ ou $a - 1 = a^p - 1$ car les seuls diviseurs positifs d'un nombre premier sont 1 et lui-même. Le cas $a - 1 = a^p - 1$ est à exclure car a et p sont supérieurs à 2. Il reste le cas $a - 1 = 1$, c'est-à-dire¹ $a = 2$.

Montrons maintenant que p est un nombre premier. Soit d un diviseur de p strictement inférieur à p . On peut écrire $p = cd$ avec $c > 1$ puis

$$2^p - 1 = (2^d)^c - 1 = (2^d - 1) \left(1 + 2^d + \cdots + (2^d)^{c-1} \right).$$

Comme au-dessus, on peut affirmer que $2^d - 1$ divise le nombre premier $2^p - 1$ et, puisque $2^d - 1 < 2^p - 1$, on obtient $2^d - 1 = 1$ puis $d = 1$. Ainsi, les seuls diviseurs de p sont 1 et lui-même et l'on peut affirmer que p est un nombre premier.

3.6.3 L'infiniété des nombres premiers**Exercice 21 ** (Nombres de Fermat)**

Pour $n \in \mathbb{N}$, on introduit $F_n = 2^{2^n} + 1$.

- (a) On suppose n et m sont deux entiers naturels distincts. Montrer $F_n \wedge F_m = 1$.
- (b) Exploiter cette propriété afin de retrouver le théorème d'Euclide affirmant l'existence d'une infinité de nombres premiers (Th. 16 p. 92).

Solution

- (a) Quitte à échanger, supposons $n < m$.

méthode

On remarque que $F_m - 1 = (F_n - 1)^{2^{m-n}}$.

En développant la relation ci-dessus par la formule du binôme, on obtient l'écriture

$$F_m - 1 = \sum_{k=0}^{2^{m-n}} \binom{2^{m-n}}{k} (-1)^{2^{m-n}-k} F_n^k = \underbrace{1}_{k=0} + v \underbrace{F_n}_{k \geq 1}$$

avec $v \in \mathbb{Z}$ car les coefficients binomiaux sont des entiers.

1. Les $2^p - 1$ sont les *nombres de Mersenne*.

En réorganisant les termes, il vient $F_m - vF_n = 2$ et l'on en déduit que le PGCD de F_n et F_m divise 2 : il vaut donc 1 ou 2. Cependant, F_n et F_m ne sont pas des nombres pairs, ils sont donc premiers entre eux.

(b) Les entiers F_n sont en nombre infini et possèdent chacun des facteurs premiers distincts, il existe donc une infinité de nombres premiers.

Exercice 22 **

Montrer qu'il existe une infinité de nombres premiers de la forme $4n - 1$ avec n entier.

Solution

méthode

En raisonnant par l'absurde, on introduit $4N - 1$ avec N le produit de tous les nombres premiers de la forme $4n - 1$.

Par l'absurde, supposons qu'il n'existe qu'un nombre fini de nombres premiers de la forme $4n - 1$. On peut introduire le nombre N égal au produit de ceux-ci et considérer l'entier $4N - 1$.

2 n'est pas un facteur premier de $4N - 1$ car c'est un nombre impair. Aucun facteur premier de $4N - 1$ ne peut s'écrire de la forme $4n - 1$ car un tel facteur divise N mais ne divise pas 1. Tous les facteurs premiers de $4N - 1$ sont donc des entiers impairs de la forme $4n + 1$ avec n entier. Ceux-ci sont tous congrus à 1 modulo 4 et, par produit, $4N - 1$ l'est aussi. Cependant, $4N - 1$ est congru à -1 modulo 4. C'est absurde !

Exercice 23 ***

On désire établir qu'il existe une infinité de nombres premiers de la forme $4n + 1$. Pour cela on raisonne par l'absurde et l'on suppose que ceux-ci sont en nombre fini. On pose N le produit de ceux-ci et l'on introduit

$$M = (2N)^2 + 1.$$

(a) On suppose qu'il existe un facteur premier q de M de la forme $4n + 3$. Établir

$$(2N)^{q-1} \equiv -1 \pmod{q}.$$

(b) Conclure en exploitant le petit théorème de Fermat.

Solution

(a) Puisque q divise M , on a $(2N)^2 \equiv -1 \pmod{q}$ et alors

$$(2N)^{q-1} \equiv ((2N)^2)^{\frac{q-1}{2}} \equiv (-1)^{\frac{q-1}{2}} \equiv -1 \pmod{q}.$$

(b) Puisque le nombre premier q ne divise pas $2N$, le petit théorème de Fermat donne

$$(2N)^{q-1} \equiv 1 \pmod{q}.$$

Sachant que 1 et -1 ne sont pas congrus modulo q , on obtient une absurdité : aucun facteur premier de M n'est de la forme $4n + 3$.

méthode

|| On observe que M ne possède pas de facteurs premiers.

Les facteurs premiers de M ne peuvent donc qu'être 2 et de la forme $4n + 1$. Cependant, ceux-ci divisent $2N$ mais ne divisent pas $1 = M - 4N^2$. C'est absurde.

Finalement, il existe une infinité de nombres premiers de la forme¹ $4n + 1$.

3.6.4 Décomposition en facteurs premiers

Exercice 24 **

Soit $n \in \mathbb{N}$ au moins égal à 2. Montrer que n est le produit de ses diviseurs non triviaux² si, et seulement si, $n = p^3$ avec p nombre premier ou $n = pq$ avec p et q des nombres premiers distincts.

Solution

Raisonnons par double implication.

(\Rightarrow) Si $n = p^3$ avec p nombre premier, les diviseurs non triviaux de n sont p et p^2 . Si $n = pq$ avec p, q nombres premiers distincts, les diviseurs non triviaux de n sont p et q . Dans les deux cas, n est le produit de ses diviseurs non triviaux.

(\Leftarrow) Supposons n égal au produit de ses diviseurs non triviaux.

méthode

|| On introduit un facteur premier p de n et l'on étudie $d = n/p$.

L'entier n est divisible par un nombre premier p et ne peut lui être égal. On peut donc écrire $n = dp$ avec $1 < d < n$. Distinguons deux cas :

Cas : d possède un facteur premier q distinct de p . Les trois entiers d , p et q sont des diviseurs non triviaux de n . Si d est différent de q , ces diviseurs sont distincts ce qui est absurde car $dqp > dp = n$. On a donc $d = q$ et l'on obtient l'écriture $n = pq$ avec p et q nombres premiers distincts.

Cas : p est le seul diviseur premier de d . On peut écrire $d = p^\alpha$ avec $\alpha \in \mathbb{N}^*$ et les diviseurs non triviaux de n sont les p, p^2, \dots, p^α . Si $\alpha = 1$ ou $\alpha \geq 3$, l'entier n n'est pas le produit de ses diviseurs non triviaux. Il reste $\alpha = 2$ ce qui donne $n = p^3$.

Finalement, seuls les entiers de la forme p^3 et pq (avec p et q nombres premiers distincts) sont égaux au produit de leurs diviseurs non triviaux.

1. Plus généralement, le théorème de la progression arithmétique de Dirichlet assure que, pour deux entiers naturels a et b premiers entre eux donnés, il existe une infinité de nombres premiers de la forme $an + b$.

2. Dans ce sujet, on se limite aux diviseurs positifs. Les diviseurs triviaux de n sont 1 et n .

Exercice 25 **

Soit $n \in \mathbb{N} \setminus \{0, 1\}$ dont la décomposition en facteurs premiers s'écrit

$$n = \prod_{k=1}^r p_k^{\alpha_k} \quad \text{avec } p_1, \dots, p_r \text{ nombres premiers deux à deux distincts.}$$

Montrer que la somme $\sigma(n)$ des diviseurs positifs de n vaut

$$\sigma(n) = \prod_{k=1}^r \frac{p_k^{\alpha_k+1} - 1}{p_k - 1}.$$

Solution

Les diviseurs positifs de n s'écrivent (Th. 19 p. 93) :

$$\prod_{k=1}^r p_k^{\beta_k} \quad \text{avec } \forall k \in \llbracket 1 ; r \rrbracket, 0 \leq \beta_k \leq \alpha_k.$$

Par l'unicité de la décomposition en facteurs premiers d'un entier, à chaque choix des β_1, \dots, β_r correspond un diviseur unique. On a donc

$$\begin{aligned} \sigma(n) &= \sum_{(\beta_1, \dots, \beta_r)} \left(\prod_{k=1}^r p_k^{\beta_k} \right) \\ &= \sum_{\beta_1=0}^{\alpha_1} \sum_{\beta_2=0}^{\alpha_2} \cdots \sum_{\beta_r=0}^{\alpha_r} p_1^{\beta_1} p_2^{\beta_2} \cdots p_r^{\beta_r} \\ &= \left(\sum_{\beta_1=0}^{\alpha_1} p_1^{\beta_1} \right) \left(\sum_{\beta_2=0}^{\alpha_2} p_2^{\beta_2} \right) \cdots \left(\sum_{\beta_r=0}^{\alpha_r} p_r^{\beta_r} \right). \end{aligned}$$

Par sommes géométriques de raisons différentes de 1, on conclut :

$$\sigma(n) = \prod_{k=1}^r \frac{p_k^{\alpha_k+1} - 1}{p_k - 1}.$$

3.7 Exercices d'approfondissement

Exercice 26 *

Soit A un ensemble de $n + 1 \geq 2$ entiers distincts tous inférieurs ou égaux à $2n$. Montrer qu'il existe deux éléments de A tels que l'un divise l'autre.

Solution**méthode**

Par le principe des tiroirs, on détermine dans A deux éléments s'écrivant $2^k(2p+1)$ et $2^\ell(2p+1)$ avec la même valeur de p .

Tout entier m compris entre 1 et $2n$ s'écrit de façon unique $m = 2^k(2p+1)$ avec p compris entre 0 et $n-1$.

Il y a exactement n valeurs de p possibles, et donc, parmi les $n+1$ éléments de A , il en existe au moins deux pour lesquels les valeurs de p sont les mêmes. Ces éléments s'écrivent $2^k(2p+1)$ et $2^\ell(2p+1)$: le plus petit des deux divise l'autre.

Exercice 27 **

On note $d(n)$ le nombre de diviseurs positifs de $n \in \mathbb{N}^*$. Montrer

$$\frac{1}{n} \sum_{k=1}^n d(k) = \sum_{k=1}^n \frac{1}{k} + b_n$$

avec (b_n) une suite bornée.

Solution**méthode**

On exprime la somme des $d(k)$ comme une somme double et l'on réorganise la sommation.

Notons $\text{Div}(n)$ l'ensemble des diviseurs positifs de n . On peut écrire

$$\sum_{k=1}^n d(k) = \sum_{k=1}^n \sum_{d \in \text{Div}(k)} 1.$$

En échangeant les deux sommes

$$\sum_{k=1}^n d(k) = \sum_{d=1}^n \sum_{k \in M_d(n)} 1$$

avec $M_d(n)$ l'ensemble des multiples de d inférieurs à n : $d, 2d, \dots, pd$ avec $p = \lfloor \frac{n}{d} \rfloor$.

On peut alors poursuivre le calcul en écrivant

$$\sum_{k=1}^n d(k) = \sum_{d=1}^n \left\lfloor \frac{n}{d} \right\rfloor$$

Sachant $x - 1 < \lfloor x \rfloor \leq x$, pour tout réel x , on obtient l'encadrement

$$n \left(\sum_{d=1}^n \frac{1}{d} - 1 \right) \leq \sum_{k=1}^n d(k) \leq n \sum_{d=1}^n \frac{1}{d}$$

puis

$$-1 < \frac{1}{n} \sum_{k=1}^n d(k) - \sum_{d=1}^n \frac{1}{d} < 0.$$

Le terme encadré définit la suite (b_n) qui est effectivement bornée.

Exercice 28 *** (Formule d'inversion de Möbius)

Pour $n \in \mathbb{N}^*$ on pose

$$\mu(n) = \begin{cases} 0 & \text{si } n \text{ est divisible par le carré d'un nombre premier} \\ (-1)^r & \text{si } n \text{ est le produit de } r \text{ nombres premiers deux à deux distincts.} \end{cases}$$

En particulier, $\mu(1) = 1$. Enfin, pour $n \in \mathbb{N}^*$, on pose¹

$$s(n) = \sum_{d|n} \mu(d).$$

(a) Soit p un nombre premier et $\alpha \in \mathbb{N}^*$. Calculer $s(p^\alpha)$.

(b) Soit m et $n \in \mathbb{N}^*$ premiers entre eux, vérifier $s(mn) = s(m)s(n)$.

(c) En déduire la valeur de $s(n)$ pour tout $n \in \mathbb{N}^*$.

On considère une fonction $u: \mathbb{N}^* \rightarrow \mathbb{C}$ et la fonction $v: \mathbb{N}^* \rightarrow \mathbb{C}$ définie par

$$v(n) = \sum_{d|n} u(d).$$

(d) Soit $n \in \mathbb{N}^*$. Vérifier la formule d'inversion

$$u(n) = \sum_{d|n} \mu\left(\frac{n}{d}\right) v(d).$$

Solution

(a) Les diviseurs de p^α sont les p^β avec $0 \leq \beta \leq \alpha$ et donc

$$s(p^\alpha) = \sum_{\beta=0}^{\alpha} \mu(p^\beta) = \underbrace{1}_{\beta=0} + \underbrace{(-1)}_{\beta=1} + \underbrace{0}_{\beta \geq 2} = 0.$$

(b) méthode

|| Puisque m et n sont premiers entre eux, les diviseurs de mn sont² les produits $k\ell$ avec k diviseur de m et ℓ diviseur de n .

1. La somme introduite porte sur les entiers d diviseurs positifs de n .
 2. Voir sujet 16 p. 105.

On peut alors organiser le calcul de la somme définissant $s(mn)$

$$s(mn) = \sum_{d|mn} \mu(d) = \sum_{k|m \text{ et } \ell|n} \mu(k\ell) = \sum_{k|m} \left(\sum_{\ell|n} \mu(k\ell) \right).$$

Les entiers k et ℓ parcourant les sommes étant deux à deux premiers entre eux, on a $\mu(k\ell) = \mu(k)\mu(\ell)$. En effet, si l'un des deux entiers est divisible par le carré d'un nombre premier, les deux membres de l'égalité sont nuls. Sinon, ils s'écrivent respectivement avec q et r nombres premiers tous distincts et l'égalité se relit $(-1)^{q+r} = (-1)^q(-1)^r$. On peut alors poursuivre le calcul en factorisant à l'intérieur des sommes

$$\begin{aligned} s(mn) &= \sum_{k|m} \left(\sum_{\substack{\ell|n \\ \text{indépendant de } k}} \underbrace{\mu(k)\mu(\ell)}_{\text{indépendant de } \ell} \right) = \sum_{k|m} \left(\mu(k) \underbrace{\sum_{\substack{\ell|n \\ \text{indépendant de } k}} \mu(\ell)}_{\text{indépendant de } k} \right) \\ &= \left(\sum_{k|m} \mu(k) \right) \left(\sum_{\ell|n} \mu(\ell) \right) = s(m)s(n). \end{aligned}$$

(c) Si n est supérieur à 2, la décomposition de n en produit $p_1^{\alpha_1} \cdots p_r^{\alpha_r}$ de facteurs premiers donne

$$s(n) = s(p_1^{\alpha_1}) \times \cdots \times s(p_r^{\alpha_r}) = 0.$$

Si n vaut 1, $s(n) = 1$ car la somme se limite à un terme égal à 1.

(d) Soit $n \in \mathbb{N}^*$. Simplifions

$$\sum_{d|n} \mu\left(\frac{n}{d}\right) v(d) = \sum_{d|n} \left(\mu\left(\frac{n}{d}\right) \sum_{c|d} u(c) \right) = \sum_{d|n} \left(\sum_{c|d} \mu\left(\frac{n}{d}\right) u(c) \right).$$

méthode

|| On réorganise le calcul en échangeant les deux sommes.

La double somme considérée porte sur les couples (c, d) avec $c | d$ et $d | n$. Par transitivité, c divise n . L'entier d étant divisible par c , il peut s'écrire ck . La condition $d | n$ devient alors $k | \frac{n}{c}$. On obtient donc

$$\sum_{d|n} \mu\left(\frac{n}{d}\right) v(d) = \sum_{c|n} \left(\sum_{k|\frac{n}{c}} \underbrace{\mu\left(\frac{n}{ck}\right) u(c)}_{\text{indépendant de } k} \right) = \sum_{c|n} \left(\sum_{k|\frac{n}{c}} \mu\left(\frac{n}{ck}\right) u(c) \right).$$

Lorsque k parcourt les diviseurs de $\frac{n}{c}$, l'entier $\ell = \frac{n}{ck}$ parcourt aussi les diviseurs de $\frac{n}{c}$ et donc

$$\sum_{d|n} \mu\left(\frac{n}{d}\right) v(d) = \sum_{c|n} \left(\sum_{\ell|\frac{n}{c}} \mu(\ell) \right) u(c) - \sum_{c|n} s\left(\frac{n}{c}\right) u(c) = u(n).$$

Exercice 29 * (Formule de Legendre)**

(a) Soit p un nombre premier et $n \in \mathbb{N}$. Établir

$$v_p(n!) = \sum_{i=1}^r \left\lfloor \frac{n}{p^i} \right\rfloor \quad \text{avec} \quad r = \left\lfloor \frac{\ln n}{\ln p} \right\rfloor.$$

(b) Application : Soit $a, b \in \mathbb{N}$, montrer

$$\frac{(2a)!(2b)!}{a!b!(a+b)!} \in \mathbb{N}.$$

Solution(a) **méthode**

On simplifie dans le produit décrivant $n!$ les facteurs qui ne sont pas multiples de p .

Dans le produit $n! = 1 \times 2 \times \cdots \times n$ figurent exactement k multiples de p avec k le plus grand entier tel que $pk \leq n$, c'est-à-dire $k = \left\lfloor \frac{n}{p} \right\rfloor$. On peut donc écrire

$$n! = 1 \times \cdots \times p \times \cdots \times 2p \times \cdots \times kp \times \cdots \times n.$$

Dans ce produit, les facteurs autres que les ip avec $1 \leq i \leq k$ n'influent pas sur la valeur de la valuation p -adique de $n!$: on peut les simplifier et écrire

$$v_p(n!) = v_p(p \times 2p \times \cdots \times kp) = v_p(p^k) + v_p(1 \times 2 \times \cdots \times k) = k + v_p(k!).$$

Autrement dit,

$$v_p(n!) = \left\lfloor \frac{n}{p} \right\rfloor + v_p\left(\left\lfloor \frac{n}{p} \right\rfloor!\right).$$

Par le même calcul, on obtient encore

$$v_p(n!) = \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{1}{p} \left\lfloor \frac{n}{p} \right\rfloor \right\rfloor + v_p\left(\left\lfloor \frac{1}{p} \left\lfloor \frac{n}{p} \right\rfloor \right\rfloor!\right). \quad (*)$$

Or¹

$$\left\lfloor \frac{1}{p} \left\lfloor \frac{n}{p} \right\rfloor \right\rfloor = \left\lfloor \frac{n}{p^2} \right\rfloor$$

et donc (*) se relit

$$v_p(n!) = \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + v_p\left(\left\lfloor \frac{n}{p^2} \right\rfloor!\right).$$

1. En effet, les deux inégalités $\frac{1}{p} \left\lfloor \frac{n}{p} \right\rfloor \leq \frac{n}{p^2}$ et $\left\lfloor \frac{n}{p^2} \right\rfloor \leq \frac{1}{p} \left\lfloor \frac{n}{p} \right\rfloor$ donnent l'égalité voulue par croissance de la fonction partie entière.

On répète ce calcul jusqu'au rang r introduit dans l'énoncé pour lequel $p^r \leq n < p^{r+1}$ et l'on obtient¹

$$v_p(n!) = \left\lfloor \frac{n}{p} \right\rfloor + \left\lfloor \frac{n}{p^2} \right\rfloor + \dots + \underbrace{\left\lfloor \frac{n}{p^r} \right\rfloor}_{=0} + v_p\left(\left\lfloor \frac{n}{p^r} \right\rfloor!\right) = \sum_{i=1}^r \left\lfloor \frac{n}{p^i} \right\rfloor.$$

(b) méthode

On compare les valuations p -adiques du numérateur et du dénominateur pour tout nombre premier p (Th. 19 p. 93).

Soit x et y deux réels. En discutant selon l'appartenance de x à l'un ou l'autre des intervalles $\lfloor x \rfloor ; [x] + 1/2[$ et $\lfloor x \rfloor + 1/2 ; [x] + 1[$ et une discussion analogue en y , on vérifie l'inégalité

$$\lfloor x \rfloor + \lfloor y \rfloor + \lfloor x + y \rfloor \leq \lfloor 2x \rfloor + \lfloor 2y \rfloor.$$

On a donc, pour tout nombre premier p et tout entier $i \geq 1$,

$$\left\lfloor \frac{a}{p^i} \right\rfloor + \left\lfloor \frac{b}{p^i} \right\rfloor + \left\lfloor \frac{a+b}{p^i} \right\rfloor \leq \left\lfloor \frac{2a}{p^i} \right\rfloor + \left\lfloor \frac{2b}{p^i} \right\rfloor.$$

En sommant ces inégalités jusqu'à une valeur de i suffisamment grande (ce qui a pour effet d'adoindre des 0 aux sommes étudiées), on obtient

$$v_p(a!) + v_p(b!) + v_p((a+b)!) \leq v_p((2a)!) + v_p((2b)!).$$

c'est-à-dire

$$v_p(a!b!(a+b)!) \leq v_p((2a)!(2b)!).$$

On peut donc affirmer que le quotient étudié est entier.

Exercice 30 *** (Théorème d'Aubry)

Soit N un entier strictement positif et \mathcal{C} le cercle d'équation $x^2 + y^2 = N$.

On suppose que le cercle \mathcal{C} possède un point (x_0, y_0) à coordonnées rationnelles. On introduit (x_0, y_0) un point à coordonnées entières obtenues par arrondis des coordonnées de (x_0, y_0) . En étudiant, lorsque cela a un sens, l'intersection du cercle \mathcal{C} avec la droite joignant (x_0, y_0) et (x_0, y_0) , montrer que ce cercle contient un point à coordonnées entières.

Solution

méthode

On vérifie que les coordonnées du point intersection proposé sont des nombres rationnels s'écrivant avec un dénominateur commun inférieur au dénominateur commun des coordonnées du point de départ.

¹. La formule peut aussi être comprise ainsi : $v_p(n!)$ est la somme du nombre de multiples de p , de p^2 , de p^3 , etc. inférieurs à n .

Si le couple (x_0, y_0) est à coordonnées entières la conclusion est entendue.

Sinon, on peut écrire en choisissant un dénominateur commun

$$x_0 = \frac{p_0}{d_0} \text{ et } y_0 = \frac{q_0}{d_0} \quad \text{avec } p_0, q_0 \in \mathbb{Z} \text{ et } d_0 \in \mathbb{N} \setminus \{0, 1\}.$$

Considérons alors un couple (x_0, y_0) à coordonnées entières obtenues par arrondis des coordonnées (x_0, y_0) . En notant D la distance entre les deux points considérés, on a

$$0 < D^2 = (x_0 - x'_0)^2 + (y_0 - y'_0)^2 \leq \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

En développant les carrés, on obtient

$$D^2 = N^2 - 2(x_0 x'_0 + y_0 y'_0) + (x'_0)^2 + (y'_0)^2.$$

Par réduction au même dénominateur, on observe

$$0 < D^2 = \frac{d_1}{d_0} \leq \frac{1}{2}$$

avec $d_1 \in \mathbb{N}^*$ et $d_1 < d_0$.

La droite joignant les deux points distincts (x_0, y_0) et (x'_0, y'_0) est formée des points de coordonnées

$$\begin{cases} x = x'_0 + t(x_0 - x'_0) \\ y = y'_0 + t(y_0 - y'_0) \end{cases} \quad \text{avec } t \in \mathbb{R}.$$

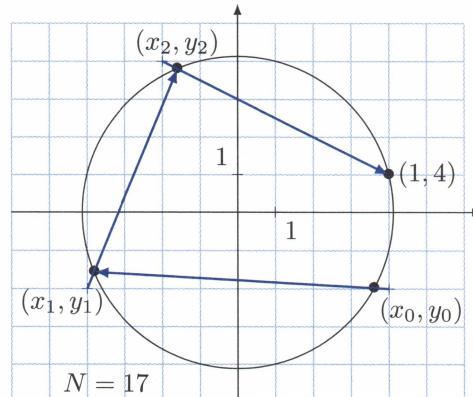
Cette droite coupe le cercle en (x_0, y_0) pour $t = 1$ et recoupe celui-ci en (x_1, y_1) obtenu¹ pour

$$t = \frac{(x'_0)^2 + (y'_0)^2 - N^2}{D^2}.$$

Ce nombre t est donc de la forme $\frac{d_0}{k}$ avec k entier. Les coordonnées (x_1, y_1) sont alors de la forme

$$x_1 = \frac{p_1}{d_1} \text{ et } y_1 = \frac{q_1}{d_1} \quad \text{avec } p_1, q_1 \in \mathbb{Z} \text{ et } d_1 \in \mathbb{N}^*, d_1 < d_0.$$

Si $d_1 = 1$, le processus s'arrête : on a obtenu un point à coordonnées entières. Sinon, il suffit de répéter la manipulation. Le processus va nécessairement s'arrêter car les dénominateurs communs aux deux coordonnées décroissent strictement à chaque itération.



1. L'appartenance de (x, y) au cercle C se traduit par une équation du second degré en t , connaître une des racines suffit à déterminer l'autre car le produit des racines se lit sur les coefficients de l'équation.

CHAPITRE 4

Structures algébriques usuelles

4.1 Loi de composition interne

4.1.1 Définition

Définition

On appelle *loi de composition interne* (ou *opération*) sur un ensemble E toute application de $E \times E$ vers E . Lorsque l'on convient de noter \star une loi de composition interne, on note $x \star y$ l'image du couple (x, y) par l'application \star . L'élément $x \star y$ est appelé *composé* de x par y via la loi \star .

L'addition et la multiplication définissent des lois de composition internes sur \mathbb{N} , \mathbb{Z} , \mathbb{R} , etc. L'union et l'intersection définissent des lois de composition internes sur $\wp(E)$. La composition des applications définit une loi de composition interne sur $\mathcal{F}(E, E)$.

4.1.2 Propriétés

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

Définition

On dit que la loi \star est *commutative* lorsque $a \star b = b \star a$ pour tous a et b dans E .

Définition

On dit que la loi \star est *associative* lorsque $a \star (b \star c) = (a \star b) \star c$ pour tous a , b et c dans E .

Dans \mathbb{R} , l'addition et la multiplication sont commutatives et associatives. Dans $\wp(E)$, il en est de même pour l'union et l'intersection. Dans $\mathcal{F}(E, E)$, la composition des applications est associative mais n'est pas commutative dès que E possède au moins deux éléments.

Lorsqu'une loi \star est associative, on peut écrire $a \star b \star c$ sans préciser de parenthèses organisant le calcul. De manière générale, on définit par récurrence

$$\underset{i=1}{\overset{n}{\star}} a_i \stackrel{\text{déf}}{=} a_1 \star a_2 \star \cdots \star a_n \quad \text{pour tous } a_1, \dots, a_n \in E.$$

Soit de plus \top une autre loi de composition interne sur E .

Définition

On dit que la loi \star est *distributive* sur la loi \top lorsque $a \star (b \top c) = (a \star b) \top (a \star c)$ et $(b \top c) \star a = (b \star a) \top (c \star a)$ pour tous a, b et c dans E .

Dans \mathbb{R} , la multiplication est distributive sur l'addition. Dans $\wp(E)$, l'intersection est distributive sur l'union et inversement.

4.1.3 Élément neutre

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

Définition

On dit qu'un élément e de E est *neutre* pour la loi \star si $e \star x = x \star e = x$ pour tout x de E .

Lorsqu'un tel élément existe, il est unique.

Dans \mathbb{R} , 0 est le neutre pour l'addition et 1 le neutre pour la multiplication. Dans $\wp(E)$, l'ensemble vide \emptyset est le neutre pour l'union et E est le neutre pour l'intersection. Dans $\mathcal{F}(E, E)$, l'application Id_E est le neutre pour la composition des applications.

4.1.4 Éléments symétrisables

Soit \star une loi associative sur un ensemble E possédant un neutre e .

Définition

On dit qu'un élément x de E est *symétrisable*¹ pour la loi \star s'il existe un élément x' dans E vérifiant $x \star x' = x' \star x = e$. Cet élément x' est alors unique, on l'appelle le *symétrique* de x et on le note $\text{sym}(x)$.

L'élément neutre e est symétrisable et égal à son symétrique : $\text{sym}(e) = e$.

Dans \mathbb{R} , tout nombre est symétrisable pour l'addition mais seul les réels non nuls sont symétrisables pour la multiplication. Dans $\mathcal{F}(E, E)$, les éléments symétrisables pour la composition des applications sont les bijections.

Théorème 1

Si x est un élément symétrisable de E , $\text{sym}(x)$ l'est aussi et $\text{sym}(\text{sym}(x)) = x$.

Si x et y sont des éléments symétrisables de E , le composé $x \star y$ l'est aussi et² $\text{sym}(x \star y) = \text{sym}(y) \star \text{sym}(x)$.

1. On dit aussi que l'élément est *inversible*.

2. On sera attentif au renversement de l'ordre de x et y dans cette formule.

4.1.5 Itérés d'un élément

Soit \star une loi associative sur un ensemble E possédant un neutre e .

Pour $x \in E$ et $n \in \mathbb{N}^*$, on pose

$$x^n \stackrel{\text{déf}}{=} \underbrace{x \star x \star \cdots \star x}_{n \text{ termes}} \quad \text{et} \quad x^0 \stackrel{\text{déf}}{=} e.$$

Définition

|| L'élément x^n est appelé *itéré* d'ordre n de l'élément x .

Théorème 2

Soit $x \in E$. Pour tous p et $q \in \mathbb{N}$, $x^p \star x^q = x^{p+q}$ et $(x^p)^q = x^{pq}$.

Si x et y sont deux éléments de E , $(x \star y)^n$ se comprend $(x \star y) \star \cdots \star (x \star y)$. Lorsque les éléments x et y commutent (c'est-à-dire si $x \star y = y \star x$), on peut réorganiser ce calcul en $x^n \star y^n$.

Si un élément x est inversible, x^n l'est aussi et l'on définit

$$x^{-n} \stackrel{\text{déf}}{=} \text{sym}(x^n) = \underbrace{\text{sym}(x) \star \text{sym}(x) \star \cdots \star \text{sym}(x)}_{n \text{ termes}}.$$

En particulier, x^{-1} désigne le symétrique de x et l'on a la formule $(x \star y)^{-1} = y^{-1} \star x^{-1}$ pour x et y symétrisables.

Lorsque x est inversible, les formules du théorème 2 sont valables pour $p, q \in \mathbb{Z}$.

4.1.6 Partie stable

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

Définition

|| On dit qu'une partie A de E est *stable* pour la loi \star lorsque $x \star y \in A$ pour tous x et y dans A .

Dans \mathbb{R} , l'ensemble \mathbb{N} est une partie stable pour l'addition et la multiplication.

Lorsqu'une partie A est stable, on peut définir une restriction de la loi \star opérant sur A . Celle-ci s'appelle la *loi induite* par \star sur la partie stable A et est communément encore notée \star .

Si \star est commutative (resp. associative) sur E , la loi induite sur une partie stable A l'est aussi. Si E admet un élément neutre et qu'il appartient à A , il est évidemment aussi élément neutre pour la loi induite. Si de plus la loi est associative, un élément de A est symétrisable si, et seulement si, il est symétrisable dans E et que son symétrique appartient à A .

4.2 Structure de groupe

4.2.1 Définition

Définition

On appelle *groupe*¹ tout couple (G, \star) formé d'un ensemble G et d'une loi de composition interne \star sur G vérifiant :

- 1) la loi \star est associative ;
- 2) la loi \star possède un élément neutre dans G ;
- 3) tout élément de G est symétrisable pour la loi \star .

Si de plus la loi \star est commutative, on dit que le groupe (G, \star) est *commutatif* (ou encore *abélien*).

On dit qu'un groupe est noté *additivement* lorsque sa loi est notée $+$. Dans ce cas, on note $x + y$ le composé de x par y , on note 0 l'élément neutre et $-x$ le symétrique² de x que l'on appelle *opposé*. L'itéré d'ordre $n \in \mathbb{Z}$ de x est noté nx .

L'usage veut que l'on réserve l'utilisation de la notation additive aux groupes commutatifs.

On dit qu'un groupe est noté *multiplicativement* lorsque sa loi est notée \times ou (\cdot) . Dans ce cas, on note xy le composé de x par y et 1 l'élément neutre. Le symétrique de x est noté³ x^{-1} et s'appelle l'*inverse* de x . Il est fréquent de manipuler des multiplications non commutatives (comme le produit matriciel présenté dans le chapitre 9).

$(\mathbb{C}, +)$ et (\mathbb{C}^*, \times) sont des groupes abéliens usuels⁴.

4.2.2 Groupe des permutations d'un ensemble

Définition

On appelle *permutation* d'un ensemble E toute bijection de E vers lui-même. L'ensemble des permutations de E est noté \mathcal{S}_E ou $\mathfrak{S}(E)$.

Théorème 3

(\mathcal{S}_E, \circ) est un groupe de neutre la permutation identité Id_E .

Ce groupe n'est pas commutatif dès que E possède au moins 3 éléments. Il est fini et possède $n!$ éléments lorsque E est fini à n éléments.

4.2.3 Sous-groupe

Soit (G, \star) un groupe de neutre e .

1. Abusivement et s'il n'y a pas d'ambiguïté sur la loi \star , on dit simplement que G est un groupe.
2. Ceci permet aussi de définir l'opération de soustraction : $x - y$ devant se comprendre $x \star \text{sym}(y)$.
3. On n'emploie pas la notation $\frac{1}{x}$ lorsque la multiplication n'est pas commutative.
4. (\mathbb{R}, \times) et (\mathbb{C}, \times) ne sont pas des groupes car 0 n'y est pas inversible.

Définition

On appelle *sous-groupe* de (G, \star) toute partie H de G contenant le neutre e , stable par passage au symétrique et stable par composition, c'est-à-dire vérifiant pour tous x et y de H

$$e \in H, \quad \text{sym}(x) \in H \quad \text{et} \quad x \star y \in H.$$

On peut aussi affirmer rapidement qu'une partie H de G est un sous-groupe de (G, \star) en constatant qu'il s'agit d'une partie non vide stable par composition avec le symétrique¹ :

$$H \neq \emptyset \quad \text{et} \quad \forall (x, y) \in H^2, \quad x \star \text{sym}(y) \in H.$$

Les parties $\{e\}$ et G sont des sous-groupes de (G, \star) , on dit que ce sont ses *sous-groupes triviaux*.

Théorème 4

Si H est un sous-groupe d'un groupe (G, \star) alors (H, \star) est un groupe² de même neutre que G .

\mathbb{Z} , \mathbb{Q} et \mathbb{R} sont des sous-groupes de $(\mathbb{C}, +)$: ce sont des groupes additifs.

\mathbb{R}^* , \mathbb{R}_+^* et \mathbb{U} sont des sous-groupes de (\mathbb{C}^*, \times) : ce sont des groupes multiplicatifs.

4.3 Structure d'anneau

4.3.1 Définition

Définition

On appelle *anneau*³ tout triplet $(A, +, \times)$ formé d'un ensemble A et de deux lois de composition internes $+$ et \times sur A vérifiant :

- 1) $(A, +)$ est un groupe abélien de neutre noté 0 (ou 0_A) ;
- 2) \times est associative et possède un neutre dans A noté 1 (ou 1_A) ;
- 3) \times est distributive sur $+$.

Si de plus la loi \times est commutative, on dit que l'anneau est *commutatif*.

$(\mathbb{Z}, +, \times)$, $(\mathbb{Q}, +, \times)$, $(\mathbb{R}, +, \times)$ et $(\mathbb{C}, +, \times)$ sont des anneaux commutatifs.

Dans le chapitre 9, on évoquera l'anneau des matrices carrées qui n'est généralement pas commutatif.

L'ensemble $A = \{0\}$ muni des lois $+$ et \times déterminées par $0 + 0 = 0$ et $0 \times 0 = 0$ est un anneau appelé *anneau nul*. C'est le seul dans lequel $0_A = 1_A$.

1. En notation additive, on lit $x - y \in H$ et l'on parle de stabilité par différence. En notation multiplicative, on lit $xy^{-1} \in H$ et l'on parle de stabilité par composition avec l'inverse.

2. La loi \star sur H est la loi induite sur la partie stable H par la loi \star de G .

3. En l'absence d'ambiguïté, on dit simplement que A est un anneau.

4.3.2 Calculs dans un anneau

Soit $(A, +, \times)$ un anneau.

Théorème 5

Pour tout $a \in A$, on a $0_A \times a = a \times 0_A = 0_A$.

On en déduit $(-a) \times b = a \times (-b) = -(ab)$ pour tous $a, b \in A$ et, plus généralement, on a

$$(na)b = n(ab) = a(nb) \quad \text{pour tout } n \in \mathbb{Z}.$$

Définition

|| On dit que deux éléments a et b d'un anneau commutent lorsque $ab = ba$.

Théorème 6 (Formule du binôme)

Si a et b sont deux éléments de l'anneau A qui commutent, alors

$$(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k}b^k \quad \text{pour tout } n \in \mathbb{N}.$$

Lorsque a et b ne commutent pas, on a seulement $(a+b)^2 = a^2 + ab + ba + b^2$.

Théorème 7 (Formule de factorisation géométrique)

Si a et b sont deux éléments de l'anneau A qui commutent, alors

$$a^n - b^n = (a-b) \sum_{k=0}^{n-1} a^{n-1-k}b^k \quad \text{pour tout } n \in \mathbb{N}^*.$$

4.3.3 Éléments inversibles

Soit $(A, +, \times)$ un anneau.

Définition

|| Un élément $a \in A$ est dit *inversible* s'il existe $b \in A$ tel que $ab = ba = 1_A$. Cet élément b est alors unique, on l'appelle l'*inverse de a* et on le note a^{-1} .

L'élément 1_A est inversible et $1_A^{-1} = 1_A$. L'élément 0_A n'est pas inversible sauf dans l'anneau nul.

Théorème 8

L'ensemble A^\times des éléments inversibles de l'anneau A est un groupe multiplicatif¹.

1. C'est-à-dire un groupe pour la multiplication définie sur l'anneau A .

4.3.4 Corps

Définition

- || On appelle corps tout anneau commutatif $(K, +, \times)$, non réduit à $\{0_K\}$ et dont tous les éléments, sauf le nul, sont inversibles.
- $(\mathbb{Q}, +, \times)$, $(\mathbb{R}, +, \times)$ et $(\mathbb{C}, +, \times)$ sont des corps fameux.

4.4 Exercices d'apprentissage

4.4.1 Loi de composition interne

Exercice 1

Soit E un ensemble muni d'une loi \star . On dit qu'un élément e est *neutre à gauche* (resp. *à droite*) lorsque $e \star x = x$ (resp. $x \star e = x$) pour tout $x \in E$.

Montrer que si la loi \star possède un neutre à gauche et un neutre à droite, elle possède un élément neutre.

Solution

On suppose que la loi \star possède un neutre à gauche e et un neutre à droite e' .

méthode

- || On calcule de deux façons $e \star e'$.

Puisque e est neutre à gauche $e \star e' = e'$. Puisque e' est neutre à droite, on a aussi $e \star e' = e$. On en déduit $e = e'$. Cet élément est alors neutre pour la loi \star .

Exercice 2

Soit E un ensemble muni d'une loi \star associative possédant un neutre e . Montrer que si x et y sont deux éléments de E tels que les composés $x \star y$ et $y \star x$ sont symétrisables alors x et y sont symétrisables.

Solution

Soit x et y deux éléments de E tels que $x \star y$ et $y \star x$ sont symétrisables de symétriques respectifs z et t .

méthode

- || En raisonnant par analyse-synthèse, on peut proposer des candidats pour les inverses de x et y .

Analyse : Supposons x et y symétrisables. Par la formule $(x \star y)^{-1} = y^{-1} \star x^{-1}$ on a

$$x^{-1} = y \star (x \star y)^{-1} = y \star z \quad \text{et aussi} \quad x^{-1} = (y \star x)^{-1} \star y = t \star y.$$

Synthèse : Posons $x' = y \star z$ et $x'' = t \star y$.

Par associativité, on a¹ $x * x' = x * (y * z) = (x * y) * z = e$ et, par un calcul analogue, on vérifie $x'' * x = e$. Pour affirmer que x est inversible, il reste à vérifier² $x' = x''$ ce qui se fait en calculant de deux façons $x'' * x * x'$ par associativité :

$$x'' * (x * x') = x'' * e = x'' \quad \text{et} \quad (x'' * x) * x' = e * x' = x'.$$

Ainsi, x est inversible d'inverse $x' = x''$. On établit que y est inversible par un raisonnement analogue ou simplement par composition d'éléments inversibles : $y = (y * x) * x^{-1}$.

Exercice 3

On note $E = [0 ; 1]$ et, pour $x, y \in E$, on pose

$$x * y = x + y - xy.$$

- (a) Vérifier que $*$ définit une loi de composition interne sur E .
- (b) Étudier la commutativité et l'associativité de la loi $*$.
- (c) Existe-t-il un élément neutre ? Quels sont les éléments symétrisables ?
- (d) Soit $\alpha \in [0 ; 1]$. Vérifier que $A = [\alpha ; 1]$ est une partie stable.

Solution

(a) **méthode**

Vérifier qu'une loi de composition interne sur E est bien définie consiste non seulement à vérifier l'existence du composé $x * y$ mais aussi à vérifier son appartenance à E .

Pour $x, y \in E$, le réel $x * y$ est parfaitement défini. Il s'agit alors de vérifier qu'il appartient à E . D'une part,

$$x * y = \underbrace{x}_{\geq 0} + \underbrace{y(1-x)}_{\geq 0} \geq 0.$$

D'autre part,

$$1 - x * y = (1 - x)(1 - y) \geq 0.$$

On peut donc affirmer que $x * y$ est bien élément de E .

(b) Pour tous x et y de E , on vérifie $x * y = y * x$ grâce la commutativité des opérations $+$ et \times . La loi $*$ est donc commutative.

Soit $x, y, z \in E$. Par les propriétés calculatoires usuelles sur les réels

$$(x * y) * z = (x + y - xy) + z - (x + y - xy)z = x + y + z - (xy + xz + yz) + xyz.$$

1. On dit alors que x est *inversible à droite* et que x' est un *inverse à droite* de x . De façon similaire, on définit l'inversibilité à gauche.

2. De façon générale, si un élément est inversible à droite et à gauche, il est inversible car ses inverses à droite et à gauche sont nécessairement égaux.

Le calcul de $x * (y * z)$ conduit à une expression identique et l'on peut affirmer que la loi $*$ est associative.

(c) méthode

|| Pour établir l'existence d'un élément neutre on mène souvent une analyse afin de déterminer celui-ci. Cependant, la situation en cours est assez évidente...

Pour $y = 0$, on observe $x * 0 = x = 0 * x$ pour tout $x \in E$. On peut donc affirmer que 0 est élément neutre.

méthode

|| Pour déterminer les éléments symétrisables, on raisonne par analyse-synthèse.

Analyse : Si $x \in E$ est symétrisable, il existe $x' \in E$ vérifiant $x * x' = 0$, c'est-à-dire $x'(x - 1) = x$. Cette équation d'inconnue x' n'a pas de solution lorsque $x = 1$ et une solution qui est $x/(x - 1)$ lorsque $x \neq 1$. Si $x > 0$, cette solution est strictement négative et ne détermine donc pas un élément de E . Le seul cas restant est $x = 0$.

Synthèse : $x = 0$ est évidemment symétrisable puisque c'est le neutre de la loi $*$.

En résumé, seul 0 est symétrisable.

(d) Soit $x, y \in E$. On sait déjà $x * y \leq 1$ et l'on observe

$$x * y = x + \underbrace{y(1-x)}_{\geq 0} \geq x.$$

On en déduit que si x et y sont éléments de A , $x * y \geq x \geq \alpha$ et donc $x * y \in A$.

4.4.2 Groupes

Exercice 4

On note $G =]-1; 1[$ et, pour $x, y \in G$, on pose

$$x * y = \frac{x+y}{1+xy}.$$

Montrer que la loi $*$ munit G d'une structure de groupe abélien.

Solution

méthode

|| On vérifie que $*$ définit bien une loi de composition interne sur G .

Pour $x, y \in G$, le dénominateur $1 + xy$ est strictement positif car $|xy| = |x||y|$ est strictement inférieur à 1 : on peut définir le réel $x * y$. Vérifions ensuite que $x * y$ est élément de G . On a la chaîne d'équivalence

$$\begin{aligned} -1 < x * y < 1 &\iff -1 - xy < x + y < 1 + xy \\ &\iff (1+x)(1+y) < 0 < (1-x)(1-y). \end{aligned}$$

Sachant $1+x > 0$, $1-x > 0$ et les mêmes propriétés en y , on peut affirmer que $x \star y$ est bien élément de G .

méthode

On vérifie ensuite que la loi est associative, possède un neutre et que tout élément est symétrisable.

Commençons par souligner que la loi \star est commutative : pour tous x et y de G , on vérifie $x \star y = y \star x$ grâce la commutativité des opérations $+$ et \times .

Soit $x, y, z \in G$. Par les propriétés calculatoires usuelles sur les réels

$$(x \star y) \star z = \frac{x+y}{1+xy} \star z = \frac{\frac{x+y}{1+xy} + z}{1 + \frac{x+y}{1+xy} z} = \frac{\frac{x+y+z+xyz}{1+xy} + z}{1 + \frac{x+y+z+xyz}{1+xy}} = \frac{x+y+z+xyz}{1+xy+yz+zx}.$$

Le calcul de $x \star (y \star z)$ conduit à une expression identique et l'on peut affirmer que la loi \star est associative.

On observe $x \star 0 = x = 0 \star x$ pour tout $x \in G$: on peut affirmer que 0 est élément neutre. Enfin, pour $x \in G$, on a $x \star (-x) = 0 = (-x) \star x$ avec $-x \in G$: tout élément de G est symétrisable. On peut conclure que (G, \star) est un groupe abélien¹.

Exercice 5

Soit a et b deux réels. Montrer que $a\mathbb{Z} + b\mathbb{Z} = \{ak + bl \mid k, l \in \mathbb{Z}\}$ est un sous-groupe de $(\mathbb{R}, +)$.

Solution

méthode

On vérifie que $a\mathbb{Z} + b\mathbb{Z}$ est une partie de \mathbb{R} , non vide et stable par composition avec le symétrique².

$a\mathbb{Z} + b\mathbb{Z}$ est évidemment une partie de \mathbb{R} et elle est non vide car 0 lui appartient puisque l'on peut écrire $0 = ak + bl$ avec $k = l = 0 \in \mathbb{Z}$.

Soit x et y deux éléments de $a\mathbb{Z} + b\mathbb{Z}$. Étudions si $x - y$ en est aussi élément. On peut écrire $x = ak + bl$ et $y = ak' + bl'$ avec $k, l, k', l' \in \mathbb{Z}$. On a alors $x - y = ak'' + bl''$ avec $k'' = k - k' \in \mathbb{Z}$ et $l'' = l - l' \in \mathbb{Z}$. On a donc $x - y \in a\mathbb{Z} + b\mathbb{Z}$.

Finalement, on peut affirmer que $a\mathbb{Z} + b\mathbb{Z}$ est un sous-groupe de $(\mathbb{R}, +)$.

Exercice 6

(a) Soit $n \in \mathbb{N}^*$. On considère $\mathbb{U}_n = \{z \in \mathbb{C} \mid z^n = 1\}$ l'ensemble des racines n -ièmes de l'unité. Montrer que \mathbb{U}_n muni de la multiplication des nombres complexes est un groupe.

(b) Soit a un élément d'un ensemble E . On considère $H = \{f \in \mathcal{S}_E \mid f(a) = a\}$ l'ensemble des permutations de E fixant a . Montrer que H muni du produit de composition des applications est un groupe.

1. C'est-à-dire un groupe commutatif.

2. Ici, la loi étant additive, on parle plutôt de stabilité par différence.

Solution**méthode**

|| Lorsque la loi est connue, il est fréquent d'établir qu'une structure est un groupe en observant que c'est un sous-groupe d'un groupe déjà référencé (Th. 4 p. 123).

(a) Vérifions que \mathbb{U}_n est un sous-groupe du groupe¹ (\mathbb{C}^*, \times) . Les racines n -ièmes de l'unité sont des complexes de module 1, elles sont *a fortiori* non nulles et donc $\mathbb{U}_n \subset \mathbb{C}^*$. Il y a exactement n racines n -ièmes de l'unité et donc \mathbb{U}_n est non vide. Reste à vérifier la stabilité par composition avec l'inverse. Pour $z, z' \in \mathbb{U}_n$, on a $z \times z'^{-1} \in \mathbb{U}_n$ car

$$(zz'^{-1})^n = \left(z \times \frac{1}{z'}\right)^n = \frac{z^n}{z'^n} = \frac{1}{1} = 1$$

Finalement, \mathbb{U}_n est un sous-groupe de (\mathbb{C}^*, \times) et donc (\mathbb{U}_n, \times) est un groupe : c'est le *groupe des racines n -ièmes de l'unité*.

(b) Vérifions que H est un sous-groupe du groupe des permutations $(\mathcal{S}(E), \circ)$ (Th. 3 p. 122). H est clairement une partie de \mathcal{S}_E . H est non vide² puisque le neutre Id_E en est élément car $\text{Id}_E(a) = a$. Vérifions la stabilité par composition et par passage au symétrique. Soit $f, g \in H$. D'une part, $(f \circ g)(a) = f(g(a)) = f(a) = a$ et donc $f \circ g \in H$. D'autre part, $f^{-1}(a) = f^{-1}(f(a)) = a$ et donc $f^{-1} \in H$.

Finalement, H est un sous-groupe de (\mathcal{S}_E, \circ) et donc (H, \circ) est un groupe.

Exercice 7

Soit E un ensemble. On définit la différence symétrique³ $A \Delta B$ de deux parties A et B de E par la relation $A \Delta B = (A \cup B) \cap (\overline{A \cap B})$.

Montrer que $(\wp(E), \Delta, \cap)$ est un anneau commutatif.

Solution**méthode**

|| On vérifie les axiomes de définition d'un anneau.

Dans le sujet 17 p. 29, il a déjà été établi que l'opération Δ est associative, possède un neutre \emptyset et que toute partie A est symétrisable pour Δ de symétrique elle-même. Il est immédiat que l'opération Δ est commutative car \cup et \cap le sont. On peut donc affirmer que $(\wp(E), \Delta)$ est un groupe abélien.

L'opération d'intersection est associative et possède un neutre : E . Elle est aussi commutative.

1. Soulignons que (\mathbb{C}, \times) n'est pas un groupe car 0 n'est pas inversible.

2. Pour vérifier qu'une partie est non vide lorsque l'on veut établir que c'est un sous-groupe d'un groupe donné, il est usuel d'étudier l'appartenance du neutre car celui-ci appartient à tous les sous-groupes.

3. Voir sujet 17 p. 29.

Il reste à établir la distributivité de \cap sur Δ en vérifiant¹ $A \cap (B \Delta C) = (A \cap B) \Delta (A \cap C)$ pour tous $A, B, C \in \wp(E)$.

Par définition de la différence symétrique

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

On écrit $(A \cap B) \cup (A \cap C) = A \cap (B \cup C)$ et $\overline{A \cap B \cap C} = \overline{A} \cup \overline{B \cap C}$ et alors

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

Par distributivité de l'intersection sur la deuxième union

$$(A \cap B) \Delta (A \cap C) = \underbrace{(A \cap (B \cup C) \cap \overline{A})}_{=\emptyset} \cup \underbrace{(A \cap (B \cup C) \cap \overline{B \cap C})}_{=A \cap (B \Delta C)}$$

ce qui fournit l'identité de distributivité.

Finalement, $(\wp(E), \Delta, \cap)$ est un anneau commutatif de neutres \emptyset et E .

4.5 Exercices d'entraînement

4.5.1 Loi de composition interne

Exercice 8 *

Soit E un ensemble muni d'une loi \star associative possédant un neutre e . Montrer qu'un élément a est symétrisable si, et seulement si, l'application $f: E \rightarrow E$ définie par $f(x) = a \star x$ est bijective.

Solution

Raisonnons par double implication.

(\Rightarrow) Si l'élément a est symétrisable, on peut introduire l'application $g: E \rightarrow E$ définie par $g(x) = a^{-1} \star x$. On vérifie par associativité les égalités² $f \circ g = \text{Id}_E$ et $g \circ f = \text{Id}_E$. On peut donc affirmer que f est bijective (et g est sa bijection réciproque).

(\Leftarrow) Supposons f bijective et introduisons a' l'antécédent du neutre e par f . On a par définition $a \star a' = e$, mais il faut aussi vérifier $a' \star a = e$ pour pouvoir affirmer que a est symétrisable de symétrique a' .

méthode

|| On compare $f(a' \star a)$ et $f(e)$.

Par associativité

$$f(a' \star a) = a \star (a' \star a) = (a \star a') \star a = e \star a = a = f(e).$$

L'application f étant injective, on peut conclure $a' \star a = e$ et l'élément a est donc symétrisable.

1. Par commutativité de l'intersection, l'étude de cette seule identité suffit : il n'est pas nécessaire de considérer $(B \Delta C) \cap A = (B \cap A) \Delta (C \cap A)$.

2. On constate l'égalité $f \circ g = \text{Id}_E$ en montrant $f(g(x)) = x$ pour tout $x \in E$.

Exercice 9 **

Soit E un ensemble muni d'une loi \star associative. On suppose qu'il existe $a \in E$ telle que, pour tout $x \in E$, il est possible d'écrire $x = a \star y = z \star a$ avec $(y, z) \in E^2$.

Montrer que (E, \star) possède un neutre et que a est symétrisable.

Solution

Pour $x = a$, l'hypothèse permet d'introduire e et e' tels que $a = a \star e = e' \star a$.

méthode

|| On vérifie que e et e' sont neutres à droite et à gauche pour la loi \star .

Soit $x \in E$. On peut écrire $x = a \star y = z \star a$ avec $y, z \in E$ et alors

$$x \star e = (z \star a) \star e = z \star (a \star e) = z \star a = x.$$

De même, on vérifie $e' \star x = x$. Comme déjà vu dans le sujet 1 p. 125, le calcul de $e \star e'$ assure que les deux éléments e et e' sont égaux et l'on peut affirmer que \star possède un neutre.

Enfin, pour $x = e$, on peut introduire y et z tels que $e = a \star y = z \star a$. En calculant de deux façons $z \star a \star y$, on obtient $y = z$ et l'on peut affirmer que a est symétrisable et y est son symétrique.

Exercice 10 ***

Soit E un ensemble fini non vide muni d'une loi de composition interne associative \star . Montrer qu'il existe un élément e dans E vérifiant $e \star e = e$.

Solution**méthode**

|| Pour $x \in E$, la suite des x^{2^n} (avec $n \in \mathbb{N}$) comporte des répétitions.

Soit $x \in E$. Puisque l'ensemble E est fini, la suite x^{2^n} ne peut être formée d'éléments deux à deux distincts et il existe donc $p < q$ tels que

$$x^{2^p} = x^{2^q}.$$

Posons alors $a = x^{2^p}$ et $n = q - p \in \mathbb{N}^*$ de sorte que

$$a^{2^n} = x^{2^p \times 2^n} = x^{2^q} = a.$$

Si $n = 1$, l'élément $e = a$ convient. Si $n > 1$, considérons¹ $e = a^{2^{n-1}} \in E$ et vérifions que celui-ci convient :

$$e \star e = a^{2^{n+1}-2} = a^{2^n} \star a^{2^n-2} = a \star a^{2^n-2} = a^{2^n-1} = e.$$

1. Ce choix peut être motivé : si E est un groupe, son neutre vérifie $e \star e = e$ et correspond à l'élément introduit puisque $a \star a^{2^{n-1}} = a$.

4.5.2 Groupes

Exercice 11 *

Soit (G, \star) un groupe. On suppose $x^2 = e$ pour tout $x \in G$. Montrer que le groupe G est commutatif.

Solution

méthode

|| On exploite la formule d'inversion $(x \star y)^{-1} = y^{-1} \star x^{-1}$.

Pour tout $x \in G$, l'égalité $x^2 = e$ donne $x = x^{-1}$: dans le groupe G chaque élément est égal à son inverse. Pour tous x et y de G , le composé $x \star y$ est élément de G et donc

$$x \star y = (x \star y)^{-1} = y^{-1} \star x^{-1} = y \star x.$$

Ainsi, les éléments du groupe G commutent deux à deux.

Exercice 12 *

Soit a et b deux éléments d'un groupe G noté multiplicativement. Montrer

$$(ab)^n = 1 \implies (ba)^n = 1.$$

Solution

Rappelons que $(ab)^n$ doit être compris $(ab)(ab)\dots(ab)$ produit à n facteurs.

méthode

|| Par associativité, $a(ba)^n = aba\dots ba = (ab)^n a$.

L'élément a étant inversible, on peut écrire $(ba)^n = a^{-1}(ab)^n a$. Si $(ab)^n = 1$, on simplifie et l'on conclut $(ba)^n = a^{-1}1a = 1$.

Exercice 13 ** (Transport de loi)

Soit (G, \star) un groupe et $\varphi: G \rightarrow E$ une application bijective au départ de G et à valeurs dans un ensemble E . On définit une loi de composition interne \top sur E en posant

$$x \top y = \varphi(\varphi^{-1}(x) \star \varphi^{-1}(y)).$$

Montrer que (E, \top) est un groupe.

Solution

Commençons par souligner que la loi \top est bien définie car on peut composer par \star les éléments $\varphi^{-1}(x)$ et $\varphi^{-1}(y)$ ce qui détermine un élément de G dont l'image par φ est dans E .

Soit x, y et z dans E . On a

$$x \top (y \top z) = x \top \varphi(\varphi^{-1}(y) \star \varphi^{-1}(z)) = \varphi\left(\varphi^{-1}(x) \star (\varphi^{-1}(y) \star \varphi^{-1}(z))\right).$$

La loi \star étant associative, on poursuit

$$x \top (y \top z) = \varphi((\varphi^{-1}(x) \star \varphi^{-1}(y)) \star \varphi^{-1}(z)) = \varphi(\varphi^{-1}(x) \star \varphi^{-1}(y)) \top z = (x \top y) \top z.$$

Ainsi, la loi \top est associative.

méthode

|| Un peu d'intuition, ou une petite analyse, suffit à déterminer le neutre pour la loi \top .

Notons e le neutre de G et introduisons¹ $e' = \varphi(e)$. Pour tout $x \in E$

$$x \top e' = \varphi(\varphi^{-1}(x) \star e) = \varphi(\varphi^{-1}(x)) = x$$

et l'on vérifie de même $e' \top x = x$. L'élément e' est donc neutre pour la loi \top .

Reste à établir que tout élément de E est inversible pour la loi \top . Soit x un élément arbitraire de E . On introduit² $x' = \varphi(a^{-1})$ avec $a = \varphi^{-1}(x)$ et l'on vérifie que x' est l'inverse de x :

$$x \top x' = \varphi(u \star a^{-1}) = \varphi(e) = e' \quad \text{et} \quad x' \top x = \varphi(a^{-1} \star a) = e'.$$

Ainsi, x est inversible et x' est son inverse.

Finalement, (E, \top) est un groupe.

Exercice 14 **

Soit $G = \mathbb{R}^* \times \mathbb{R}$ et \star la loi de composition interne définie sur G par

$$(x, y) \star (x', y') = (xx', xy' + y).$$

- (a) Observer que la loi \star n'est pas commutative.
- (b) Montrer que (G, \star) est un groupe dont on précisera le neutre.
- (c) Vérifier que $\mathbb{R}_+^* \times \mathbb{R}$ est un sous-groupe de (G, \star) .

Solution

Commençons par souligner que la loi \star est bien définie sur G car $xx' \neq 0$ quand $x \neq 0$ et $x' \neq 0$.

(a) méthode

|| Exhiber deux éléments qui ne commutent pas suffit pour affirmer que la loi \star n'est pas commutative.

On a $(1, 2) \star (3, 4) = (3, 6)$ et $(3, 4) \star (1, 2) = (3, 10)$. La loi \star n'est donc pas commutative.

1. Une petite analyse suffit à déterminer e' : si la loi \top admet un neutre e' , l'égalité $x \top e' = x$ donne $\varphi^{-1}(x) \star \varphi^{-1}(e') = \varphi^{-1}(x)$ et donc $\varphi^{-1}(e')$ est neutre pour la loi \star .

2. Si l'intuition fait défaut, une petite analyse suffit encore à révéler x' . Notons que ce dernier est noté x' et non x^{-1} car on ignore à ce stade si x est inversible.

(b) Commençons par étudier l'associativité de la loi \star . Soit (x, y) , (x', y') et (x'', y'') trois éléments de G . D'une part,

$$((x, y) \star (x', y')) \star (x'', y'') = (xx', xy' + y) \star (x'', y'') = (xx'x'', xx'y'' + xy' + y).$$

D'autre part,

$$(x, y) \star ((x', y') \star (x'', y'')) = (x, y) \star (x'x'', x'y'' + y') = (xx'x'', xx'y'' + xy' + y).$$

La loi \star est donc associative.

méthode

|| La loi n'étant pas commutative, neutralité et inversibilité s'étudient en calculant les compositions dans les deux sens.

Une petite recherche suffit à révéler un candidat pour le neutre de la loi \star : on veut $e = (a, b) \in G$ vérifiant $e \star (x, y) = (x, y)$ et $(x, y) \star e = (x, y)$ pour tout $(x, y) \in G$. Ceci revient à écrire

$$(ax, ay + b) = (x, y) \quad \text{et} \quad (xa, xb + y) = (x, y).$$

Prendre $a = 1$ et $b = 0$ convient : $e = (1, 0) \in G$ est élément neutre.

Il reste à vérifier que tout élément $(x, y) \in G$ est inversible. On cherche alors (x', y') vérifiant $(x, y) \star (x', y') = (1, 0)$ et $(x', y') \star (x, y) = (1, 0)$. Après résolution, $x' = 1/x$ et $y' = -y/x$ déterminent un couple $(x', y') \in G$ convenable.

Finalement, (G, \star) est un groupe.

(c) $H = \mathbb{R}_+^* \times \mathbb{R}$ est une partie de G non vide. Soit (x, y) et (x', y') deux éléments de H . On a

$$(x, y) \star (x', y')^{-1} = (x, y) \star \left(\frac{1}{x'}, -\frac{y'}{x'} \right) = \left(\frac{x}{x'}, -\frac{xy'}{x'} + y \right) \in H \quad \text{car} \quad \frac{x}{x'} > 0.$$

Ainsi, H est un sous-groupe de (G, \star) .

4.5.3 Sous-groupes

Exercice 15 *

Montrer que $\{x + y\sqrt{3} \mid (x, y) \in \mathbb{Z}^2 \text{ et } x^2 - 3y^2 = 1\}$ est un sous-groupe de (\mathbb{R}^*, \times) .

Solution

Nommons H la partie étudiée. Celle-ci est incluse dans \mathbb{R}^* car

$$x^2 - 3y^2 = (x + y\sqrt{3})(x - y\sqrt{3}) = 1 \implies x + y\sqrt{3} \neq 0$$

H est non vide car $1 \in H$ puisque l'on peut écrire $1 = 1 + 0\sqrt{3}$ avec $1^2 - 3 \cdot 0^2 = 1$.

Soit a et b deux éléments de H . On écrit $a = x + y\sqrt{3}$ et $b = x' + y'\sqrt{3}$ avec les conditions entendues sur x, y, x' et y' . On a alors

$$ab = \underbrace{xx' + 3yy'}_{=x''} + \underbrace{(xy' + x'y)\sqrt{3}}_{=y''} \quad \text{avec } x'', y'' \in \mathbb{Z}.$$

On vérifie par développement

$$(x'')^2 - 3(y'')^2 = (xx' + 3yy')^2 - 3(xy' + x'y)^2 = (x^2 - 3y^2)(x'^2 - 3y'^2) = 1.$$

Ainsi, ab est élément de H . Enfin, en multipliant par la quantité conjuguée $x - y\sqrt{3}$ qui est non nulle

$$a^{-1} = \frac{1}{x + y\sqrt{3}} = \frac{x - y\sqrt{3}}{x^2 - 3y^2} = x - y\sqrt{3} \in H.$$

Finalement, H est un sous-groupe de (\mathbb{R}^*, \times) .

Exercice 16 **

Montrer que

$$V = \{z \in \mathbb{C} \mid \exists n \in \mathbb{N}^*, z^n = 1\}$$

est un groupe multiplicatif.

Solution

Montrons que V est un sous-groupe¹ du groupe (\mathbb{C}^*, \times) . La partie V est incluse dans \mathbb{C}^* et est non vide. Soit z et z' deux éléments de V .

méthode

- || Les exposants pour lesquels les puissances de z et z' sont égales à 1 ne sont pas forcément identiques.

Il existe n et $m \in \mathbb{N}^*$ tels que $z^n = z'^m = 1$. Considérons l'exposant $nm \in \mathbb{N}^*$ et constatons

$$(zz'^{-1})^{nm} = \frac{z^{nm}}{z'^{nm}} = \frac{(z^n)^m}{(z'^m)^n} = \frac{1^m}{1^n} = 1.$$

On a donc $zz'^{-1} \in V$.

Finalement, V est un sous-groupe de (\mathbb{C}^*, \times) et donc (V, \times) est un groupe.

Exercice 17 **

On appelle *centre* d'un groupe (G, \star) l'ensemble

$$Z(G) = \{x \in G \mid \forall y \in G, x \star y = y \star x\}.$$

Montrer que $Z(G)$ est un sous-groupe de (G, \star) .

1. V est la réunion des groupes \mathbb{U}_n des racines n -ièmes de l'unité. V ne se confond pas avec le groupe \mathbb{U} des complexes de module 1 car il existe des complexes de module 1 qui ne sont pas racines de l'unité.

Solution**méthode**

|| $Z(G)$ se comprend comme l'ensemble formé des éléments de G qui commutent avec tous les éléments de G .

$Z(G)$ est une partie non vide de G car e en est élément puisque $e \star y = y = y \star e$ pour tout y de G . Soit x et x' deux éléments de $Z(G)$. Pour tout $y \in G$

$$\begin{aligned} (x \star x') \star y &= x \star (x' \star y) = x \star (y \star x') \quad \text{car } x' \in Z(G) \\ &= (x \star y) \star x' = (y \star x) \star x' \quad \text{car } x \in Z(G) \\ &= y \star (x \star x'). \end{aligned}$$

Ainsi, $x \star x' \in Z(G)$. Reste à vérifier la stabilité par passage au symétrique.

Soit $x \in Z(G)$.

méthode

|| On exploite la commutation de x avec y^{-1} .

Soit $y \in G$. L'élément x commute avec tout élément de G et donc notamment avec y^{-1} . On peut alors écrire $x \star y^{-1} = y^{-1} \star x$. En passant cette relation au symétrique, on obtient

$$y \star x^{-1} = (x \star y^{-1})^{-1} = (y^{-1} \star x)^{-1} = x^{-1} \star y$$

et donc $x^{-1} \in Z(G)$.

Exercice 18 **

Soit H une partie finie non vide d'un groupe (G, \star) .

On suppose que H est stable pour la loi \star . Montrer que H est un sous-groupe de G .

Solution

La partie H étant par hypothèse non vide et stable par \star , il suffit de vérifier qu'elle est aussi stable par passage au symétrique. Soit $x \in H$.

méthode

|| Les itérés de x ne peuvent être deux à deux distincts.

Par stabilité de H , les itérés x^n avec $n \in \mathbb{N}^*$ sont tous éléments de H . Or H est une partie finie, les itérés qui précèdent étant en nombre infini, ils comportent des répétitions. On peut alors introduire $m, n \in \mathbb{N}^*$ tels que $x^{n+m} = x^n$. En simplifiant par x^n (ce qui est possible dans le groupe G car tous les éléments y sont inversibles), il vient $x^m = e$.

Cas : $m > 1$. Le symétrique de x est x^{m-1} avec $m-1 \in \mathbb{N}^*$, c'est donc un élément de H .

Cas : $m = 1$. On a $x = e$ et le symétrique de x est simplement x donc élément de H .

Exercice 19 **

Soit H et K deux sous-groupes d'un groupe G noté multiplicativement. On forme

$$HK = \{xy \mid x \in H \text{ et } y \in K\} \quad \text{et} \quad KH = \{yx \mid y \in K \text{ et } x \in H\}.$$

Etablir que HK est un sous-groupe de G si, et seulement si, $KH \subset HK$ et qu'alors $HK = KH$.

Solution

Raisonnons par double implication.

(\Leftarrow) Supposons $KH \subset HK$. Les parties H et K étant non vides, la partie HK est aussi non vide. Étudions la stabilité de HK par produit et par passage à l'inverse.

Soit a et b deux éléments de HK . On peut écrire $a = xy$ et $b = x'y'$ avec $x, x' \in H$ et $y, y' \in K$. On a alors $ab = xyx'y'$.

méthode

|| L'hypothèse $KH \subset HK$ permet, en modifiant les éléments, de transformer le produit d'un élément de K par un élément de H en celui d'un élément de H par un élément de K .

Le facteur yx' est élément de KH , on peut donc l'écrire $x''y''$ avec $x'' \in H$ et $y'' \in K$. On a alors

$$ab = x(x''y'')y' = (xx'')(y''y') = \bar{x}\bar{y}$$

avec $\bar{x} = xx'' \in H$ et $\bar{y} = y''y' \in K$ car H et K sont des sous-groupes de G . Ainsi, le produit ab appartient à HK .

Enfin, $a^{-1} = y^{-1}x^{-1} \in KH \subset HK$. La partie HK est donc un sous-groupe de G .

(\Rightarrow) Supposons que HK soit un sous-groupe de G . Soit $a = yx$ avec $y \in K$ et $x \in H$ un élément de KH .

méthode

|| Par passage à l'inverse, on échange les positions des éléments de H et K .

On peut écrire $a = (x^{-1}y^{-1})^{-1}$ avec $x^{-1} \in H$ et $y^{-1} \in K$. L'élément a est alors l'inverse d'un élément du sous-groupe HK et c'est donc aussi un élément de HK . Ainsi, $KH \subset HK$.

Aussi un élément de a de HK appartient à KH . En effet, a^{-1} appartient à HK ce qui permet d'écrire $a^{-1} = xy$ avec $x \in H$ et $y \in K$ et alors $a = y^{-1}x^{-1} \in KH$. Par double inclusion, on conclut $HK = KH$.

4.5.4 Groupes finis**Exercice 20 ****

Soit G un groupe possédant 4 éléments. Montrer que G est commutatif.

Solution

Adoptons une notation multiplicative pour la loi de G et notons $1, a, b, c$ les quatre éléments de G . Il est entendu que 1 commute avec les trois autres éléments. Il reste à vérifier que ces derniers commutent entre eux et, compte tenu de la généralité de l'étude, il suffit de vérifier que a et b commutent.

méthode

|| On montre $ab = 1$ ou $ab = c$ par exclusion des autres possibilités.

Le cas $ab = a$ est impossible car, en composant par a^{-1} à gauche, on obtient $b = 1$. Le cas $ab = b$ est de même impossible. Il reste donc $ab = 1$ ou $ab = c$. De même, on a aussi $ba = 1$ ou $ba = c$.

Si $ab = 1$ alors b est l'élément inverse de a et donc $ba = a^{-1}a = 1$. De même, si $ba = 1$, on a $ab = 1$. Si $ab \neq 1$ et $ba \neq 1$, il reste $ab = c = ba$.

Dans tous les cas les éléments a et b commutent¹.

Exercice 21 **

Soit G un groupe noté multiplicativement possédant un nombre pair d'éléments.

Montrer qu'il existe $x \in G$ tel que $x^2 = 1$ et $x \neq 1$.

Solution**méthode**

|| On regroupe les éléments de G avec leur inverse.

Pour chaque $x \in G$, on introduit la partie $A_x = \{x, x^{-1}\}$. Cet ensemble se confond avec $A_{x^{-1}}$ et il est constitué de deux éléments sauf si $x = x^{-1}$, c'est-à-dire, sauf si $x^2 = 1$. Au surplus, les parties A_x sont disjointes ou confondues. En effet, si $A_x \cap A_y \neq \emptyset$, les éléments x et y sont égaux ou inverses l'un de l'autre mais, dans les deux cas, $A_x = A_y$.

L'ensemble G est formé d'un nombre pair d'éléments et est la réunion des A_x . Sachant que A_1 ne possède qu'un élément, il existe au moins un élément x dans G , différent du neutre, tel que A_x soit un singleton. Pour cet élément, on a $x^2 = 1$.

4.5.5 Anneaux et corps**Exercice 22 * (Anneau de Boole)**

Soit $(A, +, \times)$ un *anneau de Boole*², c'est-à-dire une anneau dans lequel $x^2 = x$ pour tout $x \in A$.

(a) Montrer que $2x = 0_A$ pour tout $x \in A$. En déduire que A est un anneau commutatif.

(b) Montrer que l'on définit une relation d'ordre \leqslant sur A en posant

$$x \leqslant y \iff xy = x.$$

1. Le plus petit groupe fini non commutatif a 6 éléments : c'est le groupe des permutations de $\{1, 2, 3\}$.
2. L'anneau $(\wp(E), \Delta, \cap)$ étudié dans le sujet 7 p. 129 est un exemple non trivial d'anneau de Boole.

Solution(a) Soit $x \in A$.**méthode**|| On développe $(1_A + x)^2$.

L'égalité $(1_A + x)^2 = 1_A + x$ donne $1_A + 2x + x^2 = 1_A + x$ car $x^2 = x$. Après simplification, on a directement $2x = 0_A$.

Soit $x, y \in A$. En développant¹ l'égalité $(x + y)^2 = x + y$, on obtient

$$x^2 + xy + yx + y^2 = x + y.$$

En simplifiant, on parvient à l'identité $xy + yx = 0_A$ car $x^2 = x$ et $y^2 = y$. Or on a aussi par l'étude précédente $2xy = 0_A$ et l'on en déduit $xy = yx$.

(b) Soit $x \in A$. Puisque $x^2 = x$, la relation \preccurlyeq est réflexive.

Soit $x, y \in A$. Si $x \preccurlyeq y$ et $y \preccurlyeq x$ alors $xy = x$ et $yx = y$. Par commutativité, on conclut que x et y sont égaux.

Soit $x, y, z \in A$. Si $x \preccurlyeq y$ et $y \preccurlyeq z$ alors $xy = x$ et $yz = y$. On en déduit $xz = (xy)z$ puis $xz = x(yz) = xy = x$ c'est-à-dire $x \preccurlyeq z$.

Finalement, \preccurlyeq est une relation d'ordre² sur A .

Exercice 23 **

Soit $(A, +, \times)$ un anneau vérifiant $xyx = x^2y$ pour tous x et y dans A . Montrer que l'anneau est commutatif.

Solution**méthode**

|| Le sujet est posé dans un anneau : ceci invite à raisonner en introduisant l'opération d'addition.

Soit x et y deux éléments de A . On exploite l'hypothèse avec les éléments $1_A + x$ et y . En développant l'égalité $(1_A + x)y(1_A + x) = (1_A + x)^2y$, on obtient

$$y + xy + yx + xyx = y + 2xy + x^2y.$$

En simplifiant, il reste $yx = xy$.

Exercice 24 ** (Nilpotence)

On dit qu'un élément x d'un anneau $(A, +, \times)$ est *nilpotent* lorsqu'il existe $n \in \mathbb{N}^*$ vérifiant $x^n = 0_A$. Soit x et y deux éléments de l'anneau $(A, +, \times)$.

- (a) Montrer que si x est nilpotent et que x et y commutent, alors xy est nilpotent.
- (b) Montrer que si xy est nilpotent, alors yx l'est aussi.
- (c) Montrer que si x et y sont nilpotents et commutent, alors $x + y$ est nilpotent.
- (d) Montrer que si x est nilpotent alors $1_A - x$ est inversible et préciser $(1_A - x)^{-1}$.

1. La formule $(x + y)^2 = x^2 + 2xy + y^2$ n'est pas adaptée car on ignore si x et y commutent.
 2. Dans l'anneau $(\wp(E), \Delta, \cap)$, la relation \preccurlyeq se confond avec \subset .

Solution

(a) Soit $n \in \mathbb{N}^*$ tel que $x^n = 0_A$. Puisque x et y commutent, on peut réordonner les facteurs du produit

$$(xy)^n = \underbrace{(xy)(xy)\dots(xy)}_{n \text{ facteurs}} = x^n y^n = 0_A y^n = 0_A.$$

Ainsi, l'élément xy est nilpotent.

(b) Soit $n \in \mathbb{N}^*$ tel que $(xy)^n = 0_A$. En multipliant à gauche par y et à droite par x

$$(yx)^{n+1} = y(xy)^n x = y0_A x = 0_A.$$

L'élément yx est donc nilpotent.

(c) Soit $n, m \in \mathbb{N}^*$ tels que $x^n = y^m = 0_A$.

méthode

- || On développe $(x+y)^p$ par la formule du binôme (Th. 6 p. 124) en choisissant un exposant suffisamment grand pour être assuré de ne sommer que des 0_A .

Puisque x et y commutent, on peut exploiter la formule du binôme et écrire pour $p \in \mathbb{N}$

$$(x+y)^p = \sum_{k=0}^p \binom{p}{k} x^k y^{p-k}.$$

Lorsque $k \geq n$, on a¹ $x^k = x^n x^{k-n} = 0_A x^{k-n} = 0_A$ et le terme d'indice k de la somme est nul.

Lorsque $p - k \geq m$, on obtient comme au-dessus $y^{p-k} = 0_A$ et encore une fois le terme d'indice k de la somme est nul.

En choisissant $p = n + m - 1 \in \mathbb{N}^*$ (ou une valeur supérieure), on est assuré que, pour chaque k compris entre 0 et p , l'une des deux conditions $k \geq n$ ou $p - k \geq m$ est satisfaite. On obtient alors $(x+y)^p = 0_A$ car somme de termes tous nuls.

(d) Soit $n \in \mathbb{N}^*$ tel que $x^n = 0_A$.

méthode

- || On exploite la formule de factorisation géométrique (Th. 7 p. 124).

Puisque x et 1_A commutent, on peut écrire les factorisations

$$\begin{aligned} 1_A &= 1_A - x^n = (1_A - x)(1_A + x + \dots + x^{n-1}) \text{ et} \\ 1_A &= 1_A - x^n = (1_A + x + \dots + x^{n-1})(1_A - x). \end{aligned}$$

On en déduit que $1_A - x$ est inversible et $1_A + x + \dots + x^{n-1}$ est son inverse.

1. L'écriture $x^k = x^n x^{k-n}$ est possible car l'exposant $k - n$ est positif.

4.6 Exercices d'approfondissement

Les sujets 25 à 28 nécessitent des résultats présentés dans le chapitre 6 relatif au dénombrement.

Exercice 25 **

On dit qu'un élément a d'un ensemble E muni d'une loi \star est *régulier* lorsque, pour tout $(x, y) \in E^2$,

$$\begin{aligned} a \star x = a \star y &\implies x = y \quad (\text{régularité à gauche}) \\ x \star a = y \star a &\implies x = y \quad (\text{régularité à droite}). \end{aligned}$$

- (a) Montrer que tous les éléments d'un groupe sont réguliers.
- (b) Soit E un ensemble fini muni d'une loi associative pour laquelle il existe un élément régulier. Montrer que la loi \star possède un élément neutre.
- (c) Soit G un ensemble fini non vide muni d'une loi associative pour laquelle tous les éléments sont réguliers. Montrer que G est un groupe.

Solution

- (a) Soit a, x et y des éléments d'un groupe (G, \star) .

Si $a \star x = a \star y$, en composant avec a^{-1} à gauche, il vient $a^{-1} \star (a \star x) = a^{-1} \star (a \star y)$, puis, par associativité, $e \star x = e \star y$ ce qui donne $x = y$. Ainsi, a est régulier à droite et l'on montre de même que a est régulier à gauche.

- (b) Soit $a \in E$.

méthode

|| On montre par cardinalité que l'application $x \mapsto a \star x$ est bijective.

La régularité à gauche de a signifie que l'application $x \mapsto a \star x$ est injective. Or cette application va de l'ensemble fini E vers lui-même, elle est donc bijective (Th. 4 p. 206). Par surjectivité, on peut introduire un élément $e \in E$ pour lequel $a \star e = a$. Vérifions alors que e élément neutre pour la loi \star . Soit $x \in E$ arbitraire. On a par associativité

$$a \star (e \star x) = (a \star e) \star x = a \star x.$$

Par régularité à gauche de a , on peut simplifier et affirmer $e \star x = x$. En particulier, on peut écrire $e \star a = a$ et aussi

$$(x \star e) \star a = x \star (e \star a) = x \star a.$$

Par régularité à droite de a , on obtient $x \star e = x$.

Finalement, e est élément neutre pour la loi \star .

(c) L'ensemble G étant non vide on peut exploiter le résultat précédent et affirmer l'existence d'un neutre e . Reste à vérifier que tout élément de G est symétrisable.

Soit $a \in G$. Comme au-dessus on peut affirmer la bijectivité de l'application $x \mapsto a \star x$ et introduire un élément a' pour lequel $a \star a' = e$. De plus, par associativité,

$$a \star (a' \star a) = (a \star a') \star a = e \star a = a = a \star e.$$

Par régularité de a , on obtient $a' \star a = e$. Ainsi, a est symétrisable et l'on peut conclure que G est un groupe.

Exercice 26 *

Soit (G, \star) un groupe possédant $2n$ éléments avec $n > 2$.

On suppose qu'il existe deux sous-groupes H et K possédant chacun n éléments et vérifiant $H \cap K = \{e\}$. Montrer $n = 2$.

Solution

On a

$$\text{Card}(H \cup K) = \text{Card}(H) + \text{Card}(K) - \text{Card}(H \cap K) = 2n - 1.$$

Il existe donc un unique élément g dans G qui n'appartient ni à H ni à K .

méthode

|| On étudie la composition d'un élément de H par un élément de K .

Soit $x \in H$ et $y \in K$ deux éléments distincts du neutre e (ce qui est possible car n est supérieur à 2). Par l'absurde, si le composé $x \star y$ appartient à H , alors, par opérations dans le sous-groupe H , l'élément $y = x^{-1} \star (x \star y)$ appartient aussi à H . Ceci est exclu car y est un élément de K distinct de e . Ainsi, $x \star y$ n'appartient pas à H et l'on montre de même qu'il n'appartient pas à K : il est nécessairement égal à g .

Si l'on considère alors $x, x' \in H$ et $y \in K$ tous distincts de e , on peut affirmer $x \star y = g$ et $x' \star y = g$ ce qui entraîne $x = g \star y^{-1} = x'$: la partie H ne possède qu'un élément en plus de e et donc $n = 2$.

Exercice 27 **

Soit (G, \star) un groupe fini dans lequel $x^2 = e$ pour tout $x \in E$.

(a) Soit H un sous-groupe strict¹ de (G, \star) et a un élément de $G \setminus H$.

Montrer que $K = H \cup H'$ avec $H' = \{a \star x \mid x \in H\}$ est un sous-groupe de G .

(b) Avec les notations qui précèdent, donner le cardinal de K en fonction de celui de H .

(c) En déduire que le cardinal de G est une puissance de 2.

Solution

Commençons par souligner que le groupe G est commutatif (voir sujet 11 p. 132).

(a) K est une partie non vide de G . Elle est évidemment stable par passage au symétrique car, dans G , tout élément est égal à son symétrique. Il reste à vérifier la stabilité

1. Un sous-groupe strict est un sous-groupe différent du groupe.

de K par composition. Soit x et y deux éléments de K . On discute selon l'appartenance des éléments à H ou H' .

Si x et y appartiennent au sous-groupe H , $x * y$ est élément de H .

Si x et y sont éléments de H' (qui n'est pas un sous-groupe), on peut écrire $x = a * x'$ et $y = a * y'$ avec $x', y' \in H$. Sachant la loi commutative et $a^2 = e$, on simplifie

$$x * y = (a * x') * (a * y') = a^2 * x' * y' = x' * y' \in H.$$

Si x est élément de H et y élément de H' (le cas symétrique est analogue), on écrit $y = a * y'$ avec $y' \in H$ et par commutativité de la loi

$$x * y = x * (a * y') = a * (x * y') \in H'.$$

Dans tous les cas, $x * y$ est élément de K .

(b) méthode

|| On vérifie que H et H' sont deux parties disjointes de mêmes cardinaux.

L'application $x \mapsto a * x$ est injective et transforme H en H' . On en déduit que H et H' ont le même nombre d'éléments. Aussi, les parties H et H' sont disjointes car a est choisi dans $G \setminus H$. En effet, si $H \cap H' \neq \emptyset$, il existe un élément x de H qui s'écrit $a * x'$ avec $x' \in H$ et alors $a = x * x'^{-1} \in H$ ce qui est exclu.

On en déduit $\text{Card}(K) = \text{Card}(H) + \text{Card}(H') = 2 \text{Card}(H)$.

(c) méthode

|| Partant de $\{e\}$, on construit des sous-groupes emboîtés jusqu'à parvenir à G .

Soit $H_0 = \{e\}$. Pour $i \in \mathbb{N}$, si $H = H_i$ est distinct de G , on choisit un élément a dans $G \setminus H$ et l'on définit $H_{i+1} = K$ comme ci-dessus. On construit ainsi une suite de sous-groupes H_0, H_1, H_2, \dots chacun de cardinal double du précédent. Le processus de construction s'arrête nécessairement car le groupe G est fini. Si i désigne l'indice pour lequel $H_i = G$, on a $\text{Card}(G) = 2^i$.

Exercice 28 *** (Conjugaison dans un groupe)

Soit G un groupe fini noté multiplicativement.

(a) On appelle *normalisateur* de $x \in G$ l'ensemble $N(x) = \{g \in G \mid gxg^{-1} = x\}$. Montrer que $N(x)$ est un sous-groupe de G .

(b) Montrer que l'on définit une relation d'équivalence \mathcal{R} sur G en posant

$$x \mathcal{R} y \iff \exists g \in G, y = gxg^{-1}.$$

(c) On note $\text{Cl}(x)$ la classe d'équivalence d'un élément x pour la relation \mathcal{R} . Montrer

$$\text{Card}(G) = \text{Card}(\text{Cl}(x)) \times \text{Card}(N(x)).$$

(d) Application : On suppose que G est de cardinal p^α avec p premier et $\alpha \in \mathbb{N}^*$. Montrer que son centre¹ $Z(G)$ n'est pas réduit à $\{1\}$.

Solution

(a) $N(x)$ est une partie non vide de G car le neutre 1 en est élément puisque $1x1 = x$. Soit g et h deux éléments de $N(x)$. On a $gxg^{-1} = x$ et $hxh^{-1} = x$ donc

$$(gh^{-1})x(gh^{-1})^{-1} = gh^{-1}xhg = gh^{-1}(hxh^{-1})hg = gxg^{-1} = x.$$

Ainsi, gh^{-1} est élément de $N(x)$ et l'on peut affirmer que $N(x)$ est un sous-groupe² de G .

(b) Soit $x \in G$. On peut écrire $x = gxg^{-1}$ avec $g = 1$ et donc $x \mathcal{R} x$.

Soit $x, y \in G$ tels que $x \mathcal{R} y$. Il existe $g \in G$ tel que $y = gxg^{-1}$ et alors $x = hyh^{-1}$ avec $h = g^{-1} \in G$. On a donc $y \mathcal{R} x$.

Enfin, soit $x, y, z \in G$ tels que $x \mathcal{R} y$ et $y \mathcal{R} z$. Il existe $g, h \in G$ tels que $y = gxg^{-1}$ et $z = hyh^{-1}$. On a alors $z = kxk^{-1}$ avec $k = hg \in G$ et donc $x \mathcal{R} z$.

La relation \mathcal{R} est réflexive, symétrique et transitive, c'est une relation d'équivalence.

(c) méthode

On introduit l'application $\varphi: G \rightarrow \text{Cl}(x)$ définie par $\varphi(g) = gxg^{-1}$ et l'on étudie les antécédents des éléments de $\text{Cl}(x)$.

Les valeurs prises par φ sont bien en relation avec x : l'application φ introduite est bien définie à valeurs dans $\text{Cl}(x)$. De plus, la classe d'équivalence de x réunit les éléments en relation avec x et l'application φ est surjective.

Soit $y \in \text{Cl}(x)$. Étudions l'ensemble des antécédents de y . Par surjectivité, il existe g dans G tel que $\varphi(g) = y$. Pour $h \in G$, on a alors

$$\begin{aligned}\varphi(h) = y &\iff hxh^{-1} = gxg^{-1} \\ &\iff g^{-1}hx(g^{-1}h)^{-1} = x \\ &\iff g^{-1}h \in N(x).\end{aligned}$$

L'ensemble des antécédents de y est donc l'ensemble des gk avec k parcourant $N(x)$. L'application $k \mapsto gk$ étant injective, on peut affirmer qu'il y a exactement $\text{Card}(N(x))$ antécédents à chaque y de $\text{Cl}(x)$. En dénombrant les éléments de G en regroupant entre eux ceux qui prennent la même valeur par φ , on obtient

$$\text{Card}(G) = \underbrace{\text{Card}(\text{Cl}(x))}_{\substack{\text{nombre de} \\ \text{valeurs possibles}}} \times \underbrace{\text{Card}(N(x))}_{\substack{\text{nombre de fois qu'une} \\ \text{valeur donnée est prise}}},$$

(d) méthode

Les classes d'équivalence réalisent une partition de G : le cardinal de G est la somme des cardinaux des classes d'équivalence de G .

1. Voir sujet 17 p. 135.

2. $N(x)$ est simplement le sous-groupe des éléments qui commutent avec x .

Les cardinaux des classes d'équivalence de G divisent le cardinal de G , ils sont donc chacun de la forme p^β pour $\beta \in \mathbb{N}$.

Soit $x \in G$ et $\beta \in \mathbb{N}$ tel que $\text{Card } \text{Cl}(x) = p^\beta$. Si $\beta = 0$, la classe d'équivalence de x est un singleton ce qui signifie que x commute avec tout élément de G : il appartient au centre de G . Sinon, le cardinal de $\text{Cl}(x)$ est un multiple de p .

En dénombrant G selon ses classes d'équivalence, on obtient

$$\begin{aligned} p^\alpha &= \text{Card}(G) = \sum_{\substack{C \text{ classe d'équivalence} \\ \text{Card}(C)=1}} 1 + \sum_{\substack{C \text{ classe d'équivalence} \\ \text{Card}(C)>1}} \text{Card}(C) \\ &= \text{Card}(Z(G)) + \sum_{\substack{C \text{ classe d'équivalence} \\ \text{Card}(C)>1}} \text{Card}(C). \end{aligned}$$

Les termes de la somme étant tous multiples de p , il en est de même du cardinal de $Z(G)$ et donc $Z(G) \neq \{1\}$.

Exercice 29 **

Soit a et b deux éléments d'un anneau $(A, +, \times)$. Montrer que si $1_A - ab$ est inversible alors $1_A - ba$ l'est aussi.

Solution

méthode

Lorsque ab est nilpotent¹, ba l'est aussi et l'on sait exprimer les inverses de $1_A - ab$ et de $1_A - ba$. En déterminant une relation entre ces deux inverses, on peut espérer révéler une relation vraie en situation générale.

Si $(ab)^n = 0_A$, on a $(ba)^{n+1} = 0_A$ et

$$(1_A - ab)^{-1} = 1_A + ab + \cdots + (ab)^{n-1} \quad \text{et} \quad (1_A - ba)^{-1} = 1_A + ba + \cdots + (ba)^n.$$

On observe alors $(1_A - ba)^{-1} = 1_A + b(1_A - ab)^{-1}a$.

Revenons au cas général. Supposons $1_A - ab$ inversible et notons x son inverse. On a $(1_A - ab)x = x(1_A - ab) = 1_A$ ce qui donne $x = 1_A + abx = 1_A + xab$. Introduisons ensuite $y = 1 + bra$. On obtient en remplaçant x dans les termes bxa

$$\begin{aligned} (1_A - ba)y &= 1_A + bxa - ba - babxa = 1_A + b(1_A + abx)a - ba - babxa = 1_A \text{ et} \\ y(1_A - ba) &= 1_A + bxa - ba - bxaba = 1_A + b(1_A + xab)a - ba - bxaba = 1_A. \end{aligned}$$

On en déduit que $1_A - ba$ est inversible et que y est son inverse.

1. Voir sujet 24 p. 139.

Exercice 30 ** (Équation de Pell-Fermat)

On s'intéresse à l'équation (E): $x^2 - 2y^2 = 1$ d'inconnue $(x, y) \in \mathbb{Z}^2$.

Pour la résoudre, on étudie l'ensemble

$$G = \{(x, y) \in \mathbb{N}^* \times \mathbb{Z} \mid x^2 - 2y^2 = 1\}.$$

(a) Pour (x, y) et (x', y') dans G , on pose

$$(x, y) * (x', y') = (xx' + 2yy', xy' + x'y).$$

Montrer que $*$ munit G d'une structure de groupe dont on précisera le neutre e .

Pour $(x, y) \in G$, on pose $\varphi(x, y) = \ln(x + \sqrt{2}y)$.

(b) On introduit $a = (3, 2) \in G$. Montrer

$$\forall (x, y) \in G, \quad 0 \leq \varphi(x, y) < \varphi(a) \implies (x, y) = e.$$

(c) Vérifier que, pour tout $(x, y) \in G$ et tout $(x', y') \in G$,

$$\varphi((x, y) * (x', y')) = \varphi(x, y) + \varphi(x', y'). \quad (*)$$

(d) En déduire que les éléments de G sont les a^n avec $n \in \mathbb{Z}$.

Solution

(a) Commençons par vérifier que la loi $*$ est bien définie sur G . Pour (x, y) et (x', y') deux éléments de G , on a

$$xx' + 2yy' \in \mathbb{Z} \quad \text{et} \quad xx' + 2yy' = \sqrt{1+2y^2}\sqrt{1+2y'^2} + 2yy' > 2|y||y'| + 2yy' \geq 0$$

et donc $xx' + 2yy' \in \mathbb{N}^*$. On a aussi évidemment $xy' + x'y \in \mathbb{Z}$ et enfin on vérifie par développement

$$(xx' + 2yy')^2 - 2(xy' + x'y)^2 = (x^2 - 2y^2)(x'^2 - 2y'^2) = 1 \times 1 = 1.$$

La loi $*$ est donc bien définie sur G et à valeurs dans G .

On vérifie que la loi est associative en observant avec des notations entendues :

$$\begin{aligned} (x, y) * ((x', y') * (x'', y'')) \\ = (xx''x' + 2(xy'y'' + x'y''y + x''yy'), xx'y'' + x'x''y + x''xy' + 2yy'y'') \\ = ((x, y) * (x', y')) * (x'', y''). \end{aligned}$$

Enfin, on vérifie que $e = (1, 0)$ est élément neutre et que tout $(x, y) \in G$ est symétrisable de symétrique $(x, -y) \in G$.

(b) Commençons par souligner que l'application φ est bien définie sur G car, pour tout $(x, y) \in G$,

$$x + \sqrt{2}y = \sqrt{1+2y^2} + \sqrt{2}y > \sqrt{2}|y| + \sqrt{2}y \geq 0.$$

Soit $(x, y) \in G$.

Cas : $y < 0$. On a $x - \sqrt{2}y > 1$ car $x \geq 1$ puis

$$\varphi(x, y) = \ln(x + \sqrt{2}y) = -\ln(x - \sqrt{2}y) < 0.$$

Cas : $y \geq 2$. On a $x = \sqrt{1+2y^2} \geq 3$ donc

$$\varphi(x, y) = \ln(x + \sqrt{2}y) \geq \varphi(a).$$

On en déduit que l'encadrement $0 \leq \varphi(x, y) < \varphi(a)$ n'est possible que si $y = 0$ ou $y = 1$. Dans le cas $y = 0$, on obtient $(x, y) = e$ comme voulu. Le cas restant $y = 1$ est à exclure car $x = \sqrt{3}$ n'est pas un entier.

(c) On remarque par simple développement

$$(x + \sqrt{2}y)(x' + \sqrt{2}y') = (xx' + 2yy') + \sqrt{2}(xy' + x'y).$$

En passant au logarithme, on obtient l'identité (*).

(d) Il est entendu que les itérés a^n avec $n \in \mathbb{Z}$ sont tous éléments de G . Vérifions ensuite que tout élément $(x, y) \in G$ est de cette forme.

méthode

On introduit n la partie entière du quotient $\varphi(x, y)/\varphi(a)$ de sorte que

$$n\varphi(a) \leq \varphi(x, y) < (n+1)\varphi(a).$$

En raisonnant par récurrence, on vérifie¹ grâce à l'identité (*) que $\varphi(a^p) = p\varphi(a)$ d'abord pour $p \in \mathbb{N}$ puis, plus généralement, pour $p \in \mathbb{Z}$. On a alors

$$0 \leq \varphi\left(\underbrace{(x, y) * a^{-n}}_{\in G}\right) = \varphi(x, y) - n\varphi(a) < \varphi(a)$$

et donc $(x, y) * a^{-n} = e$ puis $(x, y) = a^n$.

On peut alors générer les solutions de l'équation $x^2 - 2y^2 = 1$. Les premières valeurs sont :

$$(1, 0), (3, 2), (17, 12), (99, 70), (577, 408), \dots$$

En approfondissant le sujet, on pourrait observer que ces couples (x_n, y_n) correspondent aux écritures $(3 + 2\sqrt{2})^n = x_n + \sqrt{2}y_n$.

1. L'application φ transforme la loi $*$ sur G en l'addition sur \mathbb{R} , elle transforme donc l'itéré de composition d'ordre n dans G en l'itéré additif d'ordre n dans \mathbb{R} . En seconde année, on dira que φ est un morphisme du groupe $(G, *)$ vers $(\mathbb{R}, +)$.

CHAPITRE 5

Polynômes et fractions rationnelles

\mathbb{K} désigne \mathbb{R} ou \mathbb{C} .

5.1 L'anneau des polynômes

5.1.1 Polynômes

Définition

Un polynôme à coefficients dans \mathbb{K} est une expression de la forme

$$P = a_0 + a_1 X + \cdots + a_n X^n$$

avec $n \in \mathbb{N}$ et a_0, \dots, a_n des éléments de \mathbb{K} .

Les nombres a_k sont appelés les *coefficients* du polynôme P et la lettre X servant à exprimer le polynôme, s'appelle une *indéterminée*.

L'ensemble des polynômes à coefficients dans \mathbb{K} est noté $\mathbb{K}[X]$.

Deux polynômes

$$P = a_0 + a_1 X + \cdots + a_n X^n \quad \text{et} \quad Q = b_0 + b_1 X + \cdots + b_m X^m$$

sont égaux si, et seulement si, $a_k = b_k$ pour tout indice k pour lequel les deux coefficients ont un sens, les autres coefficients, s'il y en a, étant nuls¹.

1. Autrement dit, lorsque $n < m$, les deux polynômes seront dits égaux si $a_0 = b_0, a_1 = b_1, \dots, a_n = b_n$ et $b_{n+1} = \dots = b_m = 0$.

Un polynôme de la forme $P = a_0$ avec $a_0 \in \mathbb{K}$ est dit *constant*. Si de plus $a_0 = 0$, on dit que c'est le *polynôme nul*. Un polynôme de la forme $P = a_n X^n$ est appelé un *monôme*.

5.1.2 Degré d'un polynôme

Définition

On appelle *degré* d'un polynôme $P = a_0 + a_1 X + \cdots + a_n X^n$ non nul le plus grand indice k tel que a_k est non nul. Celui-ci est noté $\deg(P)$.

On convient de poser le degré du polynôme nul égal à $-\infty$.

Un polynôme de degré exactement n s'écrit

$$P = a_0 + a_1 X + \cdots + a_n X^n \quad \text{avec} \quad a_n \neq 0.$$

Un polynôme de degré inférieur à n s'écrit

$$P = a_0 + a_1 X + \cdots + a_n X^n$$

sans conditions particulières sur les coefficients. On note $\mathbb{K}_n[X]$ l'ensemble des polynômes de degrés inférieurs ou égaux à n .

Définition

Si P est un polynôme non nul, on appelle *coefficient dominant* de P le coefficient d'indice n égal au degré P .

Lorsque le coefficient dominant est égal à 1, on dit que le polynôme est *unitaire*.

5.1.3 Somme de deux polynômes

Soit P et Q deux polynômes de $\mathbb{K}[X]$

$$P = a_0 + a_1 X + \cdots + a_n X^n \quad \text{et} \quad Q = b_0 + b_1 X + \cdots + b_m X^m$$

Quitte à adjoindre des coefficients nuls à l'un des polynômes, on peut supposer $m = n$.

Définition

On définit le *polynôme somme* $P + Q$ par l'expression

$$P + Q \stackrel{\text{def}}{=} (a_0 + b_0) + (a_1 + b_1)X + \cdots + (a_n + b_n)X^n.$$

Théorème 1

Si P et Q sont deux polynômes de $\mathbb{K}[X]$, on a

$$\deg(P + Q) \leq \max(\deg(P), \deg(Q)).$$

De plus, si les deux polynômes sont de degrés distincts, il y a égalité.

5.1.4 Produit de deux polynômes

Soit P et Q deux polynômes de $\mathbb{K}[X]$

$$P = a_0 + a_1 X + \cdots + a_n X^n \quad \text{et} \quad Q = b_0 + b_1 X + \cdots + b_m X^m.$$

Définition

On définit le *polynôme produit* $P \times Q$ par l'expression

$$P \times Q \stackrel{\text{def}}{=} c_0 + c_1 X + \cdots + c_{n+m} X^{n+m}$$

avec¹

$$c_k = \sum_{i+j=k} a_i b_j \quad \text{pour tout } k \in \llbracket 0 ; n+m \rrbracket.$$

Théorème 2

Si P et Q sont deux polynômes de $\mathbb{K}[X]$, on a

$$\deg(PQ) = \deg(P) + \deg(Q).$$

En particulier, le produit de deux polynômes non nuls est un polynôme non nul.

Les opérations d'addition et de multiplication sont compatibles avec les notations employées² pour exprimer un polynôme et munissent $\mathbb{K}[X]$ d'une structure d'anneau :

Théorème 3

$(\mathbb{K}[X], +, \times)$ est un anneau commutatif d'élément nul le polynôme nul et d'élément unité le polynôme constant égal à 1.

Enfin, en multipliant par un polynôme constant égal à λ , on définit la *multiplication par un scalaire* :

$$\lambda P \stackrel{\text{def}}{=} \lambda a_0 + \lambda a_1 X + \cdots + \lambda a_n X^n.$$

5.1.5 Composition

Définition

Soit $P = a_0 + a_1 X + \cdots + a_n X^n$ un polynôme de $\mathbb{K}[X]$ et $Q \in \mathbb{K}[X]$. On définit le *polynôme composé* $P \circ Q$ par

$$P \circ Q = \sum_{k=0}^n a_k Q^k = a_0 + a_1 Q + \cdots + a_n Q^n.$$

Ce polynôme est aussi noté $P(Q)$. En particulier, $P(X)$ est une autre écriture possible pour désigner le polynôme P .

1. Dans cette écriture, la somme porte sur les indices $i \in \llbracket 0 ; n \rrbracket$ et $j \in \llbracket 0 ; m \rrbracket$ vérifiant $i + j = k$.
2. En particulier, on a $X^p \times X^q = X^{p+q}$ pour tous p et $q \in \mathbb{N}$.

5.1.6 Divisibilité

Définition

Soit A et B deux polynômes de $\mathbb{K}[X]$. On dit que le polynôme A *divise* le polynôme B , et l'on note $A | B$, s'il existe $P \in \mathbb{K}[X]$ tel que $B = AP$. On dit alors que A est un *diviseur* de B ou encore que B est un *multiple* de A .

Les polynômes constants non nuls divisent tous les polynômes.

Le polynôme nul est divisible par n'importe quel polynôme. Celui-ci mis à part, les diviseurs d'un polynôme non nul sont de degrés inférieurs au degré de ce polynôme.

La relation de divisibilité est réflexive et transitive mais pas antisymétrique :

Théorème 4

Si A et B sont deux polynômes de $\mathbb{K}[X]$

$$A | B \text{ et } B | A \iff \exists \lambda \in \mathbb{K}^*, A = \lambda B.$$

On dit alors que les polynômes A et B sont *associés*.

Si A divise B et si $\deg(A) = \deg(B)$ alors A et B sont associés.

5.1.7 Division euclidienne

Théorème 5 (Division euclidienne dans $\mathbb{K}[X]$)

Pour tous A et $B \in \mathbb{K}[X]$ avec B non nul, il existe un unique couple (Q, R) de polynômes de $\mathbb{K}[X]$ vérifiant

$$A = BQ + R \quad \text{et} \quad \deg(R) < \deg(B).$$

Les polynômes Q et R sont respectivement nommés *quotient* et *reste* de la *division euclidienne* de A par B .

Le polynôme B divise A si, et seulement si, le reste R est le polynôme nul.

5.2 Racines d'un polynôme

5.2.1 Fonction polynomiale

Soit $P = a_0 + a_1X + \cdots + a_nX^n$ un polynôme de $\mathbb{K}[X]$.

Définition

On appelle *valeur* de P en $x \in \mathbb{K}$ le nombre

$$P(x) = \sum_{k=0}^n a_k x^k = a_0 + a_1 x + \cdots + a_n x^n.$$

Lorsque P et Q désignent deux polynômes de $\mathbb{K}[X]$ et x un élément de \mathbb{K} , on observe la compatibilité du calcul polynomial avec le calcul sur les valeurs :

$$(P + Q)(x) = P(x) + Q(x), \quad (PQ)(x) = P(x)Q(x) \quad \text{et} \quad (P \circ Q)(x) = P(Q(x)).$$

Définition

Si \mathcal{D} désigne une partie de \mathbb{K} , on appelle *fonction polynomiale* associée à P définie sur \mathcal{D} l'application

$$\tilde{P}: \begin{cases} \mathcal{D} \rightarrow \mathbb{K} \\ x \mapsto P(x). \end{cases}$$

5.2.2 Racines

Définition

On appelle *racine*¹ d'un polynôme $P \in \mathbb{K}[X]$ tout $\lambda \in \mathbb{K}$ tel que $P(\lambda) = 0$.

Dans le cadre complexe, l'existence de racines est acquise par le résultat suivant :

Théorème 6 (Théorème de d'Alembert-Gauss)

Tout polynôme complexe non constant possède au moins une racine.

Dans le cadre réel, l'existence de racines réelles n'est pas certaine. Cependant, tout polynôme réel peut aussi se comprendre comme un polynôme complexe.

Définition

On appelle *racine complexe* d'un polynôme réel P toute racine de P compris comme un polynôme de $\mathbb{C}[X]$.

Les racines complexes d'un polynôme réel sont deux à deux conjuguées.

5.2.3 Racines, divisibilité, degré

En observant que le reste de la division euclidienne d'un polynôme P par $X - \lambda$ est le polynôme constant égal à $P(\lambda)$, on obtient le résultat :

Théorème 7

Une valeur $\lambda \in \mathbb{K}$ est racine de $P \in \mathbb{K}[X]$ si, et seulement si, $X - \lambda$ divise P .

En raisonnant par récurrence, on en déduit que des nombres deux à deux distincts $\lambda_1, \dots, \lambda_n$ sont racines de P si, et seulement si, $(X - \lambda_1) \dots (X - \lambda_n)$ divise P . Une conséquence remarquable est la suivante :

Théorème 8

Le nombre de racines d'un polynôme non nul est inférieur à son degré.

1. On parle aussi de *zéro* d'un polynôme.

Seul le polynôme nul possède plus de racines que son degré. Ce résultat permet d'identifier polynôme et fonction polynomiale définie sur une partie infinie :

Théorème 9

Si deux fonctions polynomiales sont égales sur une partie infinie, elles sont issues du même polynôme¹.

5.2.4 Multiplicité d'une racine

Définition

Soit P un polynôme non nul de $\mathbb{K}[X]$ et $\lambda \in \mathbb{K}$. On appelle *multiplicité* de λ en tant que racine de P le plus grand $\alpha \in \mathbb{N}$ tel que $(X - \lambda)^\alpha$ divise P .

La multiplicité est nulle lorsque λ n'est pas racine de P . Sinon, on parle de racine *simple*, de racine *double*, etc. selon que $\alpha = 1$, $\alpha = 2$, etc.

Théorème 10

Soit P un polynôme de $\mathbb{K}[X]$, $\lambda_1, \dots, \lambda_m \in \mathbb{K}$ deux à deux distincts et $\alpha_1, \dots, \alpha_m \in \mathbb{N}$.

Les $\lambda_1, \dots, \lambda_m$ sont des racines de P de multiplicités respectives au moins $\alpha_1, \dots, \alpha_m$ si, et seulement si, $(X - \lambda_1)^{\alpha_1} \times \dots \times (X - \lambda_m)^{\alpha_m}$ divise P .

On sait alors comparer les multiplicités et le degré d'un polynôme non nul :

Théorème 11

La somme des multiplicités des racines d'un polynôme non nul est inférieure à son degré.

Dans le cadre des polynômes complexes, il y a même égalité en vertu du théorème de d'Alembert-Gauss (Th. 6 p. 153).

Dans le cadre des polynômes réels, il y a aussi égalité sous réserve de décompter les racines complexes et non seulement les racines réelles. Notons que les racines complexes conjuguées d'un polynôme réel ont la même multiplicité.

5.2.5 Polynôme scindé

Définition

Un polynôme non nul P de $\mathbb{K}[X]$ est dit *scindé* sur \mathbb{K} lorsqu'il est possible de le factoriser sous la forme

$$P = a \prod_{i=1}^n (X - \lambda_i) \quad \text{avec} \quad a \in \mathbb{K}^*, \quad n \in \mathbb{N} \text{ et } \lambda_1, \dots, \lambda_n \in \mathbb{K}.$$

¹. Dit autrement : il est possible d'identifier les coefficients qui définissent ces fonctions polynomiales.

Un tel polynôme est de degré n , de coefficient dominant a et les $\lambda_1, \dots, \lambda_n$ en sont les racines *comptées avec multiplicité*¹.

Théorème 12

Un polynôme non nul P de $\mathbb{K}[X]$ est scindé sur \mathbb{K} si, et seulement si, son degré est égal à la somme des multiplicités de ses racines dans \mathbb{K} .

En particulier, tout polynôme non nul de $\mathbb{C}[X]$ est scindé sur \mathbb{C} .

Le polynôme $X^2 + 1$ n'est pas scindé sur \mathbb{R} mais est scindé sur \mathbb{C} : $X^2 + 1 = (X - i)(X + i)$. La notion de polynôme scindé dépend du corps d'étude.

5.2.6 Relations coefficients-racines d'un polynôme scindé

Lorsque l'on identifie l'écriture développée et l'écriture factorisée d'un polynôme scindé

$$a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0 = a_n (X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_n)$$

on peut exprimer les coefficients du polynôme en fonctions de ses racines. On introduit pour cela les *expressions symétriques élémentaires* en les racines $\lambda_1, \dots, \lambda_n$:

$$\sigma_k := \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} \lambda_{i_1} \lambda_{i_2} \cdots \lambda_{i_k} \quad \text{pour tout } k \in \llbracket 1 ; n \rrbracket.$$

La quantité σ_k est la somme de tous les k -produits possibles d'éléments de $\lambda_1, \dots, \lambda_n$. En particulier,

$$\sigma_1 = \lambda_1 + \cdots + \lambda_n \quad \text{et} \quad \sigma_n = \lambda_1 \times \cdots \times \lambda_n.$$

Théorème 13 (Relations coefficients-racines)

Si $P = a_n X^n + \cdots + a_1 X + a_0$ est un polynôme scindé de degré n et si $\lambda_1, \dots, \lambda_n$ désignent ses racines comptées avec multiplicité alors

$$\sigma_k = (-1)^k \frac{a_{n-k}}{a_n} \quad \text{pour tout } k \in \llbracket 1 ; n \rrbracket.$$

En particulier,

$$\begin{aligned} (X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_n) &= X^n - \sigma_1 X^{n-1} + \sigma_2 X^{n-2} + \cdots + (-1)^n \sigma_n \\ &= X^n + \sum_{k=1}^n (-1)^k \sigma_k X^{n-k}. \end{aligned}$$

1. Une racine simple apparaît une fois dans la liste, une racine double apparaît deux fois, etc.

5.3 Dérivation

5.3.1 Polynôme dérivé

Définition

Le *polynôme dérivé* P' d'un polynôme $P = a_0 + a_1X + a_2X^2 + \cdots + a_nX^n$ est

$$P' = a_1 + 2a_2X + \cdots + na_nX^{n-1} = \sum_{k=0}^{n-1} (k+1)a_{k+1}X^k.$$

Ce polynôme s'exprime encore¹

$$P' = \sum_{k=1}^n ka_k X^{k-1}.$$

Le polynôme P' est nul si, et seulement si, P est constant. Sinon², $\deg(P') = \deg(P) - 1$.

Théorème 14

Si P et Q sont des polynômes de $\mathbb{K}[X]$

$$(P+Q)' = P' + Q', \quad (PQ)' = P'Q + PQ' \quad \text{et} \quad (P \circ Q) = Q' \times P' \circ Q.$$

5.3.2 Dérivées d'ordres supérieurs

Définition

On dit que P est le *polynôme dérivé d'ordre 0* de P . Pour $n \in \mathbb{N}$, on appelle *polynôme dérivé d'ordre $n+1$* de P le polynôme dérivé du polynôme dérivé d'ordre n de P . On note $P^{(n)}$ le polynôme dérivé d'ordre n de P .

Théorème 15 (Formule de Leibniz)

Si P et Q sont des polynômes de $\mathbb{K}[X]$ et n un entier naturel

$$(PQ)^{(n)} = \sum_{k=0}^n \binom{n}{k} P^{(n-k)} Q^{(k)}$$

Théorème 16 (Formule de Taylor polynomiale)

Si P est un polynôme de $\mathbb{K}[X]$ de degré au plus n et si $\lambda \in \mathbb{K}$ alors

$$P = \sum_{k=0}^n \frac{P^{(k)}(\lambda)}{k!} (X-\lambda)^k$$

1. Dans cette formule de dérivation, X^k est simplement transformé en kX^{k-1} . On notera la disparition du terme constant correspondant à l'indice $k = 0$.

2. Dans les deux cas, on peut affirmer $\deg(P') \leq \deg(P) - 1$.

En particulier, les coefficients a_k d'un polynôme P sont donnés par la formule

$$a_k = \frac{P^{(k)}(0)}{k!} \quad \text{pour tout } k \in \mathbb{N}.$$

5.3.3 Multiplicité d'une racine et polynômes dérivés successifs

Théorème 17

Soit P un polynôme de $\mathbb{K}[X]$, $\lambda \in \mathbb{K}$ et $\alpha \in \mathbb{N}$. On a équivalence entre :

- (i) λ est racine de multiplicité α de P ;
- (ii) $P(\lambda) = P'(\lambda) = \dots = P^{(\alpha-1)}(\lambda) = 0$ et $P^{(\alpha)}(\lambda) \neq 0$.

Si λ est racine de multiplicité $\alpha \geq 1$ de P alors λ est racine de multiplicité $\alpha - 1$ de P' . Les racines multiples de P sont les racines communes à P et P' .

5.4 Arithmétique des polynômes

5.4.1 PGCD

Soit A et B deux polynômes de $\mathbb{K}[X]$ non tous deux nuls.

Définition

Tout polynôme de degré maximal parmi les diviseurs communs à A et B est appelé un PGCD de A et B .

Théorème 18

Les diviseurs communs aux polynômes A et B sont exactement les diviseurs d'un de leurs PGCD.

Tous les PGCD de A et B sont alors associés : un seul est unitaire, on le note $A \wedge B$. Lorsque A et B sont tous deux nuls, on pose $A \wedge B = 0$ et le théorème précédent reste valable avec cette convention.

5.4.2 Algorithme d'Euclide

Soit A et B deux polynômes de $\mathbb{K}[X]$.

Si B est nul, A est un PGCD de A et B .

Si B est non nul, on peut introduire le reste R de la division euclidienne de A par B et tout PGCD de B et de R est aussi PGCD de A et de B .

Théorème 19 (Algorithme d'Euclide)

Partant de la division euclidienne de A par B , lorsque l'on opère une succession de divisions euclidiennes du diviseur par le reste tant que le reste obtenu est non nul, le dernier reste non nul correspond à un PGCD de A et B .

Par cette succession de divisions euclidiennes il est possible d'exprimer un PGCD de deux polynômes comme une combinaison polynomiale de ceux-ci :

Théorème 20 (Relation de Bézout)

Si D est un PGCD de deux polynômes A et B de $\mathbb{K}[X]$, on peut écrire

$$D = AU + BV \quad \text{avec} \quad (U, V) \in \mathbb{K}[X]^2.$$

5.4.3 PPCM

Soit A et B deux polynômes de $\mathbb{K}[X]$ tous deux non nuls.

Définition

Tout polynôme de degré minimal parmi les polynômes non nuls multiples communs à A et B est appelé un PPCM de A et B .

Théorème 21

Les multiples communs aux polynômes A et B sont exactement les multiples d'un de leurs PPCM.

Tous les PPCM de A et B sont alors associés : un seul est unitaire, on le note $A \vee B$. Lorsque l'un des polynômes A ou B est nul, on pose $A \vee B = 0$ et le théorème précédent reste valable avec cette convention.

Théorème 22

Pour tous A et $B \in \mathbb{K}[X]$, les polynômes $(A \wedge B)(A \vee B)$ et AB sont associés.

5.4.4 Couples de polynômes premiers entre eux

Définition

Deux polynômes A et B de $\mathbb{K}[X]$ sont dits *premiers entre eux* lorsque leurs seuls diviseurs communs sont les polynômes constants non nuls. Il revient au même de dire que le polynôme constant égal à 1 est un PGCD de A et B et l'on note alors $A \wedge B = 1$.

Théorème 23 (Théorème de Bézout)

Deux polynômes A et B sont premiers entre eux si, et seulement si, il existe un couple (U, V) formé de deux polynômes de $\mathbb{K}[X]$ vérifiant $AU + BV = 1$.

Théorème 24 (Lemme de Gauss)

Soit A, B et C trois polynômes de $\mathbb{K}[X]$.

$$A \mid BC \text{ et } A \wedge B = 1 \implies A \mid C.$$

5.4.5 Polynômes irréductibles**Définition**

Un polynôme P de $\mathbb{K}[X]$ non constant est dit *irréductible* sur \mathbb{K} si ses seuls diviseurs positifs sont les polynômes constants non nuls et ses polynômes associés.

Les polynômes irréductibles de $\mathbb{C}[X]$ sont les polynômes de degré 1 et l'on dispose du théorème de factorisation affirmant à nouveau que les polynômes complexes non nuls sont scindés sur \mathbb{C} :

Théorème 25 (Décomposition en facteurs irréductibles dans $\mathbb{C}[X]$)

Tout polynôme complexe P non nul s'écrit de façon unique à l'ordre près des facteurs

$$P = a \prod_{k=1}^m (X - \lambda_k)^{\alpha_k}$$

avec $a \in \mathbb{C}^*$, $m \in \mathbb{N}$, $\lambda_1, \dots, \lambda_m \in \mathbb{C}$ deux à deux distincts et $\alpha_1, \dots, \alpha_m \in \mathbb{N}^*$.

Le nombre a est alors le coefficient dominant de P , m le nombre de racines distinctes, celles-ci sont les $\lambda_1, \dots, \lambda_m$ de multiplicités respectives $\alpha_1, \dots, \alpha_m$.

Les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré 1 ainsi que les polynômes de degré 2 sans racines réelles.

Théorème 26 (Décomposition en facteurs irréductibles dans $\mathbb{R}[X]$)

Tout polynôme réel P non nul s'écrit de façon unique à l'ordre près des facteurs

$$P = a \prod_{k=1}^m (X - \lambda_k)^{\alpha_k} \prod_{j=1}^p (X^2 + \mu_j X + \nu_j)^{\beta_j}$$

avec $a \in \mathbb{R}^*$, $m, p \in \mathbb{N}$, $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ deux à deux distincts, $(\mu_1, \nu_1), \dots, (\mu_p, \nu_p)$ des couples deux à deux distincts de réels vérifiant la condition $\Delta_j = \mu_j^2 - 4\nu_j < 0$ et $\alpha_1, \dots, \alpha_m, \beta_1, \dots, \beta_p \in \mathbb{N}^*$.

Le nombre a est alors le coefficient dominant de P , les $\lambda_1, \dots, \lambda_m$ sont ses racines réelles de multiplicités respectives $\alpha_1, \dots, \alpha_m$ et les facteurs $(X^2 + \mu_j X + \nu_j)$ sont associés aux racines complexes conjuguées de P .

5.4.6 Arithmétique et racines

Théorème 27

Dans $\mathbb{C}[X]$, un polynôme A divise un polynôme B si, et seulement si, les racines de A sont racines de B de multiplicités au moins égales.

Théorème 28

Dans $\mathbb{C}[X]$, deux polynômes sont premiers entre eux si, et seulement si, ils ne possèdent pas de racines en commun.

En particulier, les polynômes P et P' sont premiers entre eux si, et seulement si, les racines de P sont toutes simples.

Dans $\mathbb{R}[X]$, ces résultats sont encore valables à condition de considérer les racines complexes des polynômes et non seulement les racines réelles.

5.5 Le corps des fractions rationnelles

5.5.1 Les fractions rationnelles

Définition

Une *fraction rationnelle* à coefficients dans \mathbb{K} en l'indéterminée X est un élément F représenté par un quotient

$$\frac{A}{B} \quad \text{avec } A, B \in \mathbb{K}[X] \text{ et } B \neq 0.$$

On note $\mathbb{K}(X)$ l'ensemble de ces fractions rationnelles.

Soit $A, B, C, D \in \mathbb{K}[X]$ avec B et D non nuls. On définit l'égalité sur les fractions rationnelles de sorte que

$$\frac{A}{B} = \frac{C}{D} \iff AD = BC.$$

Il n'y a donc pas unicité de la façon de figurer une fraction rationnelle. Cependant :

Théorème 29

Si F est une fraction de $\mathbb{K}(X)$, il existe un unique couple $(P, Q) \in \mathbb{K}[X]^2$ vérifiant

$$F = \frac{P}{Q}, \quad P \wedge Q = 1 \quad \text{et} \quad Q \text{ est unitaire.}$$

Définition

Le quotient P/Q est alors appelé *représentant irréductible* de la fraction rationnelle F . Un polynôme P de $\mathbb{K}[X]$ s'identifie à la fraction rationnelle irréductible $P/1$.

5.5.2 Opérations

On définit une *addition* et une *multiplication* sur $\mathbb{K}(X)$ prolongeant celles sur $\mathbb{K}[X]$ en posant

$$\frac{A}{B} + \frac{C}{D} \stackrel{\text{def}}{=} \frac{AD + BC}{BD} \quad \text{et} \quad \frac{A}{B} \times \frac{C}{D} \stackrel{\text{def}}{=} \frac{AC}{BD}.$$

Toute fraction non nulle est inversible et donc

Théorème 30

$(\mathbb{K}(X), +, \times)$ est un corps.

L'inverse d'une fraction $F = A/B$ non nulle est $1/F = B/A$.

5.5.3 Degré

Définition

On appelle *degré* d'une fraction rationnelle $F = A/B$ le nombre

$$\deg(F) \stackrel{\text{def}}{=} \deg(A) - \deg(B) \in \mathbb{Z} \cup \{-\infty\}.$$

Théorème 31

Pour F et G deux fractions de $\mathbb{K}(X)$,

$$\deg(F + G) \leq \max(\deg(F), \deg(G)) \quad \text{et} \quad \deg(FG) = \deg(F) + \deg(G).$$

5.5.4 Racines et pôles

Soit F une fraction de $\mathbb{K}(X)$ de représentant irréductible P/Q .

Définition

On appelle *racine* de F toute racine de P . On appelle *pôle* de F toute racine de Q .

On définit la *multiplicité* de ces racines et pôles à partir des multiplicités respectives relatives aux polynômes P et Q .

Supposons la fraction rationnelle F représentée par le quotient A/B et soit $\lambda \in \mathbb{K}$. Si λ est racine de A de multiplicité $\alpha \in \mathbb{N}$ et racine de B de multiplicité $\beta \in \mathbb{N}$ alors :

- si $\alpha > \beta$, λ est racine de F de multiplicité $\alpha - \beta$;
- si $\alpha < \beta$, λ est pôle de F de multiplicité $\beta - \alpha$;
- si $\alpha = \beta$, λ n'est ni racine, ni pôle, de F .

Une fraction rationnelle n'admet qu'un nombre fini de pôles et seule la fraction rationnelle nulle admet une infinité de racines.

5.5.5 Fonction rationnelle associée

Soit F une fraction de $\mathbb{K}(X)$ de représentant irréductible P/Q .

Définition

Si $\lambda \in \mathbb{K}$ n'est pas pôle de F , on dit que F est *définie* en λ et l'on appelle *valeur* de F en λ le nombre :

$$F(\lambda) = \frac{P(\lambda)}{Q(\lambda)}.$$

Si la fraction rationnelle F est représentée par le quotient A/B et si λ n'est pas racine de B alors F est définie en λ et $F(\lambda) = A(\lambda)/B(\lambda)$.

Définition

Si \mathcal{D} est une partie de \mathbb{K} ne contenant pas de pôles de F , on appelle *fonction rationnelle* associée à F définie sur \mathcal{D} l'application

$$\bar{F}: \begin{cases} \mathcal{D} \rightarrow \mathbb{K} \\ \lambda \mapsto F(\lambda). \end{cases}$$

Si deux fonctions rationnelles sont définies et égales sur une partie \mathcal{D} infinie, elles sont issues de la même fraction rationnelle.

5.5.6 Décompositions en éléments simples

Théorème 32

Pour toute fraction F de $\mathbb{K}(X)$, il existe un unique couple $(E, G) \in \mathbb{K}[X] \times \mathbb{K}(X)$ vérifiant $F = E + G$ et $\deg(G) < 0$.

Définition

Le polynôme E ainsi introduit s'appelle *la partie entière* de la fraction F .

Si la fraction F s'écrit A/B avec $A, B \in \mathbb{K}[X]$, $B \neq 0$, la partie entière de F est le quotient de la division euclidienne de A par B .

Théorème 33 (Décomposition en éléments simples dans $\mathbb{C}(X)$)

Soit F une fraction de $\mathbb{C}(X)$ de représentant irréductible P/Q .

Si la décomposition en facteurs irréductibles du dénominateur Q dans $\mathbb{C}[X]$ s'écrit

$$Q = \prod_{k=1}^m (X - \lambda_k)^{\alpha_k}$$

alors la fraction F s'écrit de manière unique

$$F = E + \sum_{k=1}^m \left(\sum_{j=1}^{\alpha_k} \frac{a_{k,j}}{(X - \lambda_k)^j} \right)$$

avec E sa partie entière et $a_{k,j}$ des nombres complexes.

Définition

|| Cette écriture se nomme la *décomposition en éléments simples* de F dans $\mathbb{C}(X)$.

Théorème 34 (Décomposition en éléments simples dans $\mathbb{R}(X)$)

Soit F une fraction de $\mathbb{R}(X)$ de représentant irréductible P/Q .

Si la décomposition en facteurs irréductibles du dénominateur Q dans $\mathbb{R}[X]$ s'écrit

$$Q = \prod_{k=1}^m (X - \lambda_k)^{\alpha_k} \prod_{k=1}^p (X^2 + \mu_k X + \nu_k)^{\beta_k}$$

alors la fraction F s'écrit de manière unique

$$F = E + \sum_{k=1}^m \left(\sum_{j=1}^{\alpha_k} \frac{a_{k,j}}{(X - \lambda_k)^j} \right) + \sum_{k=1}^p \left(\sum_{j=1}^{\beta_k} \frac{b_{k,j}X + c_{k,j}}{(X^2 + \mu_k X + \nu_k)^j} \right)$$

avec E sa partie entière et $a_{k,j}, b_{k,j}, c_{k,j}$ des nombres réels.

Définition

|| Cette écriture se nomme la *décomposition en éléments simples* de F dans $\mathbb{R}(X)$.

Les techniques de calculs de décompositions en éléments simples sont présentées dans le sujet 11 p. 171.

5.6 Exercices d'apprentissage

5.6.1 Polynômes

Exercice 1

Trouver les $P \in \mathbb{R}[X]$ tels que $P(X^2) = (X^2 - X + 1)P(X)$.

Solution**méthode**

|| On commence par déterminer le degré des polynômes solutions puis on calcule ceux-ci en raisonnant par coefficients inconnus.

Le polynôme nul est solution de l'équation proposée. Soit P un polynôme non nul de degré $n \in \mathbb{N}$. Le polynôme composé $P(X^2)$ est de degré $2n$ tandis que $(X^2 - X + 1)P(X)$ est de degré $n + 2$. Les solutions non nulles de l'équation sont donc à chercher parmi les polynômes de degré 2.

Soit $P = aX^2 + bX + c$ un polynôme de degré inférieur¹ à 2.

$$\begin{aligned} P(X^2) &= (X^2 - X + 1)P(X) \\ &\iff aX^4 + bX^2 + c = aX^4 + (b-a)X^3 + (a-b+c)X^2 + (b-c)X + c \\ &\iff \begin{cases} b-a = 0 \\ a-b+c = b \\ b-c = 0. \end{cases} \end{aligned}$$

Après résolution du système, on peut affirmer que les polynômes solutions sont les polynômes $a(X^2 + X + 1)$ avec $a \in \mathbb{R}$.

Exercice 2

Soit $n \in \mathbb{N}$. En étudiant la dérivée n -ième de X^{2n} , établir

$$\sum_{k=0}^n \binom{n}{k}^2 = \binom{2n}{n}.$$

Solution

méthode

- || On réalise² un premier calcul par une dérivation directe et un second en dérivant le produit $X^n \times X^n$ par la formule de Leibniz (Th. 15 p. 156).

On calcule les premières dérivées de X^{2n} :

$$(X^{2n})' = 2nX^{2n-1}, (X^{2n})'' = 2n(2n-1)X^{2n-2}, \dots$$

À chaque dérivation l'exposant de la puissance de X perd une unité. Au bout de n dérivation, le X^{2n} devient X^n multiplié par les exposants intermédiaires $2n, 2n-1$, etc. jusqu'à $n+1$ obtenu lors de l'ultime dérivation, à savoir celle de X^{n+1} . On en déduit l'expression

$$(X^{2n})^{(2n)} = 2n \times (2n-1) \times \cdots \times (n+1)X^n = \frac{(2n)!}{n!} X^n. \quad (*)$$

Parallèlement, l'application de la formule de Leibniz donne

$$(X^{2n})^{(2n)} = \sum_{k=0}^n \binom{n}{k} (X^n)^{(k)} (X^n)^{(n-k)}. \quad (**)$$

En raisonnant comme au-dessus, on vérifie pour tout $k \in \llbracket 0 ; n \rrbracket$

$$(X^n)^{(k)} = n \times (n-1) \times \cdots \times (n-k+1)X^{n-k} = \frac{n!}{(n-k)!} X^{n-k}.$$

1. Ce polynôme est de degré 2 si, et seulement si, $a \neq 0$.

2. On trouvera une autre résolution dans le sujet 23 p. 77.

En remplaçant k par $n - k$ dans la formule ci-dessus, on obtient aussi

$$(X^n)^{(n-k)} = \frac{n!}{k!} X^k.$$

Les identités (*) et (**) donnent alors

$$\frac{(2n)!}{n!} X^n = \sum_{k=0}^n \binom{n}{k} \frac{n!}{(n-k)! k!} \frac{n!}{k!} X^n.$$

En identifiant les coefficients et en divisant par $n!$ il vient

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

5.6.2 Racines

Exercice 3

Soit P un polynôme de degré $n \in \mathbb{N}$ possédant au moins n racines distinctes. Peut-il y en avoir d'autres ? Quelles sont leurs multiplicités ?

Solution

méthode

Un polynôme non nul possède moins de racines que son degré (Th. 8 p. 153). Plus précisément, la somme des multiplicités de ses racines est inférieure à son degré (Th. 11 p. 154).

Le polynôme P étant de degré n , il possède au plus n racines et donc en possède exactement n : un tel polynôme ne peut avoir d'autres racines que celles proposées. De plus, ces dernières sont toutes des racines simples car la somme des multiplicités des n racines est au moins égale à n mais ne peut excéder n .

Exercice 4

Montrer qu'il existe un unique polynôme $P \in \mathbb{R}[X]$ vérifiant

$$\forall t \in \mathbb{R}, \quad P(\sin t) = \sin(3t).$$

Solution

Par développement trigonométrique

$$\begin{aligned} \sin(3t) &= \sin(2t + t) = \underbrace{\sin(2t) \cos(t)}_{=2\sin(t)\cos(t)} + \underbrace{\sin(t) \cos(2t)}_{=1-2\sin^2(t)} = 3\sin(t) - 4\sin^3(t). \end{aligned}$$

Le polynôme $P = 3X - 4X^3$ définit donc une solution. Vérifions que celle-ci est unique.

méthode

|| On peut montrer que deux polynômes sont égaux en observant que leur différence possède plus de racines que son degré.

Soit P et Q deux polynômes solutions du problème posé. Pour tout t réel, on a $P(\sin t) = Q(\sin t)$ et donc le polynôme $P - Q$ s'annule en tout $x \in [-1; 1]$. Ce polynôme possède alors une infinité de racines : c'est le polynôme nul et l'on conclut $P = Q$.

Exercice 5

Vérifier que, pour tout $n \in \mathbb{N}^*$,

$$(X - 1)^2 \mid (nX^{n+1} - (n + 1)X^n + 1).$$

Déterminer le quotient de cette division.

Solution

Soit $n \in \mathbb{N}^*$ et $P_n = nX^{n+1} - (n + 1)X^n + 1$.

méthode

|| Par dérivation, on vérifie que 1 est racine au moins double de P_n (Th. 10 p. 154).

$$P_n(1) = n - (n + 1) + 1 = 0 \quad \text{et} \quad P'_n(1) = n(n + 1) - (n + 1)n = 0.$$

La valeur 1 est racine au moins double de P_n et donc $(X - 1)^2$ divise P_n .

méthode

|| On forme le quotient¹ en factorisant deux fois P_n par $X - 1$ grâce à l'identité géométrique

$$X^n - 1 = (X - 1)(1 + X + \dots + X^{n-1}).$$

On écrit

$$\begin{aligned} nX^{n+1} - nX^n - (X^n - 1) &= n(X - 1)X^n - (X - 1)(1 + X + \dots + X^{n-1}) \\ &= (X - 1)\left(nX^n - (1 + X + \dots + X^{n-1})\right). \end{aligned}$$

Or

$$\begin{aligned} nX^n - (1 + X + \dots + X^{n-1}) &= (X^n - 1) + (X^n - X) + \dots + (X^n - X^{n-1}) \\ &= (X - 1)\left((1 + X + \dots + X^{n-1}) + X(1 + X + \dots + X^{n-2}) + \dots + X^{n-1}\right) \\ &= (X - 1)(1 + 2X + 3X^2 + \dots + nX^{n-1}). \end{aligned}$$

Finalement,

$$nX^{n+1} - (n + 1)X^n + 1 = (X - 1)^2(1 + 2X + 3X^2 + \dots + nX^{n-1}).$$

1. On démontre ainsi de nouveau que $(X - 1)^2$ divise P_n .

Exercice 6

Déterminer les triplets complexes (x, y, z) tels que :

$$(a) \begin{cases} x + y + z = 2 \\ xy + yz + zx = -5 \\ xyz = -6 \end{cases}$$

$$(b) \begin{cases} x + y + z = 2 \\ x^2 + y^2 + z^2 = 6 \\ x^3 + y^3 + z^3 = 8. \end{cases}$$

Solution**méthode**

On détermine les racines du polynôme $P = (X - x)(X - y)(X - z)$ dont on calcule les coefficients à l'aide du système étudié (Th. 13 p. 155).

(a) Soit (x, y, z) un triplet solution du système proposé. Ce système donne directement les valeurs des expressions symétriques élémentaires :

$$\sigma_1 = x + y + z = 2, \quad \sigma_2 = xy + yz + zx = -5 \quad \text{et} \quad \sigma_3 = xyz = -6.$$

Par développement

$$P = (X - x)(X - y)(X - z) = X^3 - \sigma_1 X^2 + \sigma_2 X - \sigma_3 = X^3 - 2X^2 - 5X + 6.$$

La valeur 1 est racine apparente du polynôme P ce qui permet de factoriser¹ par $(X - 1)$:

$$P = (X - 1)(X^2 - X - 6) = (X - 1)(X + 2)(X - 3).$$

Par conséquent, le triplet (x, y, z) vaut $(1, -2, 3)$ à l'ordre près des valeurs.

Inversement, on vérifie que de tels triplets sont solutions, soit par le calcul, soit en remontant le raisonnement qui ne contient pas véritablement de ruptures d'équivalences.

(b) Soit (x, y, z) un triplet solution. Posons

$$S_1 = x + y + z = 2, \quad S_2 = x^2 + y^2 + z^2 = 6 \quad \text{et} \quad S_3 = x^3 + y^3 + z^3 = 8.$$

méthode

À partir de S_1 , S_2 , S_3 on peut déterminer les valeurs des expressions symétriques élémentaires σ_1 , σ_2 et σ_3 .

On a immédiatement $\sigma_1 = S_1 = 2$.

Par développement d'un carré, on observe

$$(x + y + z)^2 = \underbrace{x^2 + y^2 + z^2}_{=S_2} + 2\underbrace{(xy + yz + zx)}_{=\sigma_2}$$

et donc $S_1^2 = S_2 + 2\sigma_2$ ce qui donne $\sigma_2 = -1$.

1. Le facteur associé peut être déterminé en raisonnant par coefficients inconnus, en posant une division euclidienne ou directement avec un peu d'habileté.

Par développement d'un cube, on fait apparaître σ_3 et un nouveau terme t

$$(x+y+z)^3 = \underbrace{x^3 + y^3 + z^3}_{=S_3} + 3(\underbrace{x^2y + y^2x + y^2z + z^2y + z^2x + x^2z}_{=t}) + 6xyz.$$

On détermine t en développant le produit S_1S_2 :

$$(x+y+z)(x^2 + y^2 + z^2) = \underbrace{x^3 + y^3 + z^3}_{=S_3} + \underbrace{x^2y + y^2x + y^2z + z^2y + z^2x + x^2z}_{=t}.$$

On obtient $t = S_1S_2 - S_3 = 4$ puis $\sigma_3 = -2$.

On a alors

$$(X-x)(X-y)(X-z) = X^3 - 2X^2 - X + 2 = (X-1)(X+1)(X-2).$$

Le triplet (x, y, z) vaut $(1, -1, 2)$ à l'ordre près des valeurs.

Inversement, on vérifie par le calcul qu'un tel triplet est solution.

5.6.3 Arithmétique des polynômes

Exercice 7

Soit $A = 2X^4 + X^3 - X^2 - X - 1$ et $B = X^3 + X^2 + X - 3$

- (a) Calculer le quotient Q et le reste R de la division euclidienne de A par B .
- (b) Calculer un PGCD D des polynômes A et B .
- (c) Déterminer deux polynômes U et V tels que $D = AU + BV$.

Solution

- (a) Quotient et reste se calculent en posant une division euclidienne :

$$\begin{array}{r} 2X^4 + X^3 - X^2 - X - 1 \\ 2X^4 + 2X^3 + 2X^2 - 6X \\ \hline -X^3 - 3X^2 + 5X - 1 \\ -X^3 - X^2 - X + 3 \\ \hline -2X^2 + 6X - 4 \end{array} \left| \begin{array}{l} X^3 + X^2 + X - 3 \\ 2X - 1 \end{array} \right.$$

On obtient $Q = 2X - 1$ et $R = -2X^2 + 6X - 4$.

- (b) On calcule le PGCD de A et B par l'algorithme d'Euclide (Th. 19 p. 157) :

$$2X^4 + X^3 - X^2 - X - 1 = (X^3 + X^2 + X - 3)(2X - 1) - 2X^2 + 6X - 4$$

$$X^3 + X^2 + X - 3 = (-2X^2 + 6X - 4)\left(-\frac{1}{2}X - 2\right) + 11X - 11$$

$$-2X^2 + 6X - 4 = (11X - 11)\left(-\frac{2}{11}X + \frac{4}{11}\right) + 0.$$

Le dernier reste non nul détermine un PGCD : $11X - 11$. En choisissant ce PGCD unitaire, on écrit $A \wedge B = X - 1$.

(c) Par les divisions euclidiennes qui précèdent, on peut exprimer successivement chaque reste sous la forme $AU + BV$ avec U et V des polynômes :

$$\begin{aligned} -2X^2 + 6X - 4 &= A \times 1 - B \times (2X - 1) \\ 11X - 11 &= B - (-2X^2 + 6X - 4) \left(-\frac{1}{2}X - 2 \right) \\ &= A \times \left(\frac{1}{2}X + 2 \right) + B \times \left(-X^2 - \frac{7}{2}X + 3 \right). \end{aligned}$$

Cette écriture n'est pas unique : les couples $(U - BP, V + AP)$ avec P parcourant $\mathbb{K}[X]$ sont aussi solutions.

Exercice 8

Soit $n \in \mathbb{N}^*$.

- (a) Décomposer $X^n - 1$ en facteurs irréductibles dans $\mathbb{C}[X]$.
- (b) Décomposer $X^n - 1$ en facteurs irréductibles dans $\mathbb{R}[X]$.

Solution

(a) méthode

On peut former la décomposition en facteurs irréductibles d'un polynôme dans $\mathbb{C}[X]$ à partir de la détermination de son coefficient dominant et de ses racines comptées avec multiplicité.

Le polynôme $X^n - 1$ est unitaire et ses racines sont les racines n -ièmes de l'unité, à savoir les $\omega_k = e^{2ik\pi/n}$ avec k parcourant $[0 ; n - 1]$. Puisqu'il y a exactement n racines celles-ci sont simples¹ et l'on peut écrire

$$X^n - 1 = \prod_{k=0}^{n-1} (X - e^{2ik\pi/n}).$$

(b) méthode

On déduit la factorisation dans $\mathbb{R}[X]$ de celle dans $\mathbb{C}[X]$ en regroupant ensemble les racines complexes conjuguées pour former les facteurs irréductibles de degré 2.

Cas : n impair. On écrit $n = 2p + 1$.

1. Voir sujet 3 p. 165.

$\omega_0 = 1$ est la seule racine réelle et ω_{n-k} désigne la racine conjuguée de ω_k . En regroupant celles-ci

$$\begin{aligned} X^{2p+1} - 1 &= (X - 1) \prod_{k=1}^p (X - \omega_k)(X - \bar{\omega}_k) = (X - 1) \prod_{k=1}^p (X - \omega_k)(X - \bar{\omega}_k) \\ &= (X - 1) \prod_{k=1}^p \left(X^2 - 2 \cos\left(\frac{2k\pi}{2p+1}\right)X + 1 \right) \end{aligned}$$

Cas : n pair. On écrit $n = 2p$ et les calculs sont analogues sauf qu'il y a deux racines réelles $\omega_0 = 1$ et $\omega_p = -1$. On obtient

$$X^{2p} - 1 = (X - 1)(X + 1) \prod_{k=1}^{p-1} \left(X^2 - 2 \cos\left(\frac{k\pi}{p}\right)X + 1 \right)$$

Exercice 9

- (a) Décomposer $X^4 + 2X^3 + X^2 - 2X - 2$ en facteurs irréductibles dans $\mathbb{R}[X]$.
- (b) Décomposer $X^4 + 1$ en facteurs irréductibles dans $\mathbb{R}[X]$.

Solution

(a) méthode

|| À partir de racines apparentes, il est possible de factoriser un polynôme.

1 et -1 sont racines de $X^4 + 2X^3 + X^2 - 2X - 2$, on peut donc factoriser¹ ce polynôme par $(X - 1)(X + 1)$, on obtient

$$X^4 + 2X^3 + X^2 - 2X - 2 = (X - 1)(X + 1)(X^2 + 2X + 2).$$

Le trinôme qui apparaît en facteur est de discriminant $\Delta < 0$, c'est un facteur irréductible réel et l'on a donc formé la décomposition cherchée.

(b) méthode

|| On factorise l'expression en faisant apparaître une différence de deux carrés.

$$X^4 + 1 = (X^2 + 1)^2 - 2X^2 = (X^2 - \sqrt{2}X + 1)(X^2 + \sqrt{2}X + 1).$$

Les trinômes apparus sont sans racines réelles, ce sont des facteurs irréductibles.

Exercice 10

- (a) Vérifier que $X^2 + X + 1$ divise² $X^{10} + X^5 + 1$.
- (b) Montrer que $X^3 - X^2 + 1$ et $X^2 - 2X + 2$ sont premiers entre eux³.

1. Par exemple, en posant la division euclidienne de $X^4 + 2X^3 + X^2 - 2X - 2$ par $(X - 1)(X + 1)$.
 2. Le sujet ne précise pas si la divisibilité est à entendre dans $\mathbb{R}[X]$ ou dans $\mathbb{C}[X]$ car cela est sans

Solution**(a) méthode**

On vérifie que les racines complexes de $X^2 + X + 1$ sont racines de $X^{10} + X^5 + 1$ de multiplicités au moins égales (Th. 27 p. 160).

Les racines de $X^2 + X + 1$ sont les racines simples $j = e^{2i\pi/3}$ et $\bar{j} = j^2$. On sait $1 + j + j^2 = 0$ et $j^3 = 1$. On en déduit $j^{10} + j^5 + 1 = j + j^2 + 1 = 0$ et donc j est racine de $X^{10} + X^5 + 1$. Les polynômes étudiés étant réels, on est assuré que la racine conjuguée \bar{j} est aussi racine de $X^{10} + X^5 + 1$. Le polynôme $X^2 + X + 1$ divise donc $X^{10} + X^5 + 1$.

(b) méthode

On vérifie⁴ que les deux polynômes n'ont pas de racines complexes en commun (Th. 28 p. 160).

Il n'est pas simple de déterminer les racines de $X^3 - X^2 + 1$ au contraire de celles de $X^2 - 2X + 2$ qui sont $1 + i$ et $1 - i$: on étudie par le calcul si ces dernières sont ou non racines de $X^3 - X^2 + 1$:

$$(1+i)^3 - (1+i)^2 + 1 = 1 + 3i - 3 - i - (1+2i-1) + 1 = -1 \neq 0$$

$1+i$ n'est donc pas racine de $X^3 - X^2 + 1$. Par conjugaison des racines complexes d'un polynôme réel, $1-i$ ne l'est pas non plus : les deux polynômes sont premiers entre eux.

5.6.4 Décompositions en éléments simples**Exercice 11**

Décomposer en éléments simples :

$$(a) \frac{X^3}{X^2 - 3X + 2} \text{ dans } \mathbb{R}[X]$$

$$(b) \frac{X^3 + X^2 + 1}{X^3 + X^2 + X} \text{ dans } \mathbb{C}[X]$$

$$(c) \frac{X - 1}{X^3 - 3X - 2} \text{ dans } \mathbb{R}[X]$$

$$(d) \frac{1}{(X-1)^2(X+1)^2} \text{ dans } \mathbb{R}[X]$$

$$(e) \frac{X}{X^3 - 1} \text{ dans } \mathbb{R}[X]$$

$$(f) \frac{1}{(X^2 + 1)(X^2 + X + 2)} \text{ dans } \mathbb{R}[X].$$

incidence. En effet, les deux polynômes étant réels, quotient et reste de la division euclidienne de l'un par l'autre peuvent se calculer dans $\mathbb{R}[X]$ et ce sont les mêmes que l'on retrouve dans $\mathbb{C}[X]$.

3. Encore une fois il n'est pas nécessaire de préciser si l'étude a lieu dans $\mathbb{R}[X]$ ou dans $\mathbb{C}[X]$ car le PGCD se calcule par une succession de divisions euclidiennes et ce sont les mêmes qui sont réalisées dans $\mathbb{R}[X]$ et dans $\mathbb{C}[X]$. En substance le PGCD unitaire de deux polynômes réels est identique dans $\mathbb{R}[X]$ et dans $\mathbb{C}[X]$.

4. On peut aussi calculer le PGCD par l'algorithme d'Euclide.

Solution**méthode**

Pour décomposer une fraction rationnelle en éléments simples :

- on l'exprime sous forme irréductible ;
- on calcule sa partie entière qui est le quotient de la division euclidienne du numérateur par le dénominateur ;
- on factorise le dénominateur dans le corps spécifié ;
- on exprime la décomposition en éléments simples à l'aide de coefficients inconnus ;
- on détermine ceux-ci.

(a) La fraction étudiée est déjà sous forme irréductible (il n'y a pas de racines complexes communes au numérateur et au dénominateur). La division euclidienne du numérateur par le dénominateur s'écrit¹

$$X^3 = (X^2 - 3X + 2)(X + 3) + 7X - 6.$$

La partie entière de la fraction vaut donc $X + 3$. Enfin, le dénominateur se factorise $X^2 - 3X + 2 = (X - 1)(X - 2)$. La décomposition en éléments simples de la fraction s'exprime alors

$$\frac{X^3}{(X - 1)(X - 2)} = X + 3 + \frac{a}{X - 1} + \frac{b}{X - 2} \quad \text{avec } a, b \in \mathbb{R}. \quad (*)$$

méthode

On détermine a en multipliant (*) par $X - 1$ puis en évaluant² en 1. On procède de façon analogue pour b .

$$a = \left. \frac{X^3}{X - 2} \right|_{X=1} = -1 \quad \text{et} \quad b = \left. \frac{X^3}{(X - 1)} \right|_{X=2} = 8.$$

La décomposition en éléments simples cherchée s'exprime

$$\frac{X^3}{X^2 - 3X + 2} = X + 3 - \frac{1}{X - 1} + \frac{8}{X - 2}.$$

(b) La fraction étudiée est déjà sous forme irréductible et sa partie entière vaut 1 car X^3 est le terme dominant au numérateur et au dénominateur. Dans $\mathbb{C}[X]$, la factorisation du dénominateur s'écrit $X^3 + X^2 + X = X(X - j)(X - j^2)$ avec $j = e^{2i\pi/3}$ racine troisième de l'unité. La décomposition en éléments simples de la fraction s'exprime

$$\frac{X^3 + X^2 + 1}{X^3 + X^2 + X} = 1 + \frac{a}{X} + \frac{b}{X - j} + \frac{c}{X - j^2} \quad \text{avec } a, b, c \in \mathbb{C}.$$

1. Seul le quotient nous intéresse, il n'est pas nécessaire de calculer précisément le reste.

2. Dans le premier membre de (*), on simplifie par $X - 1$ avant d'évaluer en 1. Dans le second membre tous les termes sauf celui contenant a sont multipliés par 0 et disparaissent.

méthode

- || La fraction rationnelle décomposée étant réelle, on doit obtenir par conjugaison une décomposition identique. L'unicité de l'écriture entraîne alors $c = b$.

$$a = \frac{X^3 + X^2 + 1}{X^2 + X + 1} \Big|_{X=0} = 1, \quad b = \frac{X^3 + X^2 + 1}{X(X - j^2)} \Big|_{X=j} = \frac{2 + j^2}{j(j - j^2)} = j \quad \text{et} \quad c = b = j^2.$$

Finalement,

$$\frac{X^3 + X^2 + 1}{X^3 + X^2 + X} = 1 + \frac{1}{X} + \frac{j}{X - j} + \frac{j^2}{X - j^2}.$$

(c) La fraction étudiée est sous forme irréductible et sa partie entière est nulle car elle est de degré strictement négatif. On observe que -1 est racine du dénominateur ce qui permet de le factoriser : $X^3 - 3X - 2 = (X + 1)^2(X - 2)$. La décomposition en éléments simples s'écrit

$$\frac{X - 1}{X^3 - 3X - 2} = \frac{a}{(X + 1)^2} + \frac{b}{X + 1} + \frac{c}{X - 2} \quad \text{avec } a, b, c \in \mathbb{R}. \quad (\Delta)$$

méthode

- || c se détermine comme au-dessus tandis que a s'obtient en multipliant (Δ) par $(X + 1)^2$ avant d'évaluer en -1 .

$$a = \frac{X - 1}{X - 2} \Big|_{X=-1} = \frac{2}{3} \quad \text{et} \quad c = \frac{X - 1}{(X + 1)^2} \Big|_{X=2} = \frac{1}{9}$$

méthode

- || Pour calculer b on peut former une équation liant a, b, c en évaluant (Δ) en un point (par exemple en 0). Plus efficacement, on peut aussi multiplier par X et considérer la limite en $+\infty$.

Par cette dernière démarche, on obtient la relation $0 = a \times 0 + b \times 1 + c \times 1$ et l'on a donc $b = -c$.

Finalement,

$$\frac{X - 1}{X^3 - 3X - 2} = \frac{\frac{2}{3}}{(X + 1)^2} - \frac{\frac{1}{9}}{X + 1} + \frac{\frac{1}{9}}{X - 2}.$$

(d) La fraction étudiée est sous forme irréductible et sa partie entière est nulle car elle est de degré strictement négatif. Son dénominateur est déjà factorisé dans $\mathbb{R}[X]$ et sa décomposition en éléments simples est de la forme

$$\frac{1}{(X - 1)^2(X + 1)^2} = \frac{a}{(X - 1)^2} + \frac{b}{X - 1} + \frac{c}{(X + 1)^2} + \frac{d}{X + 1} \quad \text{avec } a, b, c, d \in \mathbb{R}.$$

On a

$$a = \frac{1}{(X+1)^2} \Big|_{X=1} = \frac{1}{4} \quad \text{et} \quad c = \frac{1}{(X-1)^2} \Big|_{X=-1} = \frac{1}{4}$$

méthode

On calcule b et c en formant des relations liant les coefficients de la décomposition.

En multipliant par X puis en considérant la limite en $+\infty$, on obtient $b+d=0$. En évaluant en 0, on observe $1=a-b+c+d$. On en déduit $d=1/4$ et $b=-1/4$. On conclut¹

$$\frac{1}{(X-1)^2(X+1)^2} = \frac{\frac{1}{4}}{(X-1)^2} - \frac{\frac{1}{4}}{X+1} + \frac{\frac{1}{4}}{(X+1)^2} + \frac{\frac{1}{4}}{X+1}.$$

(e) La fraction étudiée est sous forme irréductible et sa partie entière est nulle. Son dénominateur se factorise $X^3-1=(X-1)(X^2+X+1)$ dans $\mathbb{R}[X]$ et sa décomposition en éléments simples est de la forme

$$\frac{X}{X^3-1} = \frac{a}{X-1} + \frac{bX+c}{X^2+X+1} \quad \text{avec } a, b, c \in \mathbb{R}. \quad (\square)$$

Par des calculs analogues aux précédents on obtient $a=1/3$.

méthode

On calcule simultanément les réels b et c en multipliant (\square) par X^2+X+1 puis en évaluant en j qui en est racine.

On obtient

$$bj+c = \frac{X}{X-1} \Big|_{X=j} = \frac{j}{j-1} = \frac{1-j}{3}.$$

Les nombres b et c étant réels, on peut opérer une identification : $b=-1/3$ et $c=1/3$.

$$\frac{X}{X^3-1} = \frac{\frac{1}{3}}{X-1} + \frac{-\frac{1}{3}X+\frac{1}{3}}{X^2+X+1}.$$

(f) La fraction étudiée est sous forme irréductible et sa partie entière est nulle car elle est de degré strictement négatif. Son dénominateur est déjà factorisé dans $\mathbb{R}[X]$ et sa décomposition en éléments simples est de la forme

$$\frac{1}{(X^2+1)(X^2+X+2)} = \frac{aX+b}{X^2+1} + \frac{cX+d}{X^2+X+2} \quad \text{avec } a, b, c, d \in \mathbb{R}.$$

Les réels a et b peuvent se calculer comme au-dessus

$$ai+b = \frac{1}{X^2+X+2} \Big|_{X=1} = \frac{1}{1+i} = \frac{1}{2}(1-i).$$

1. La fraction décomposée est paire : en changeant X en $-X$ la décomposition en éléments simples ne doit pas être modifiée. Ceci permet de justifier $a=c$ et $d=-b$ avant calculs.

On en déduit $a = -1/2$ et $b = 1/2$.

Les réels c et d peuvent être calculés par un procédé analogue en introduisant¹ une racine ω de $X^2 + X + 2$. On peut cependant être plus rapide en formant des relations sur les coefficients.

En multipliant par X et en étudiant la limite en $+\infty$, on obtient $a+c=0$ donc $c=1/2$. En évaluant en 0, il vient $1/2=b+d/2$ et donc $d=0$.

$$\frac{1}{(X^2+1)(X^2+X+2)} = \frac{-\frac{1}{2}X + \frac{1}{2}}{X^2+1} + \frac{\frac{1}{2}X}{X^2+X+2}.$$

méthode

|| En dernier recours² il est possible de réduire au même dénominateur une décomposition en éléments simples afin d'identifier les coefficients inconnus.

Il est aussi fréquent de former des décompositions en éléments simples par des astuces d'écriture comme la suivante :

$$\frac{1}{(X-\lambda)(X-\mu)} = \frac{1}{(\lambda-\mu)} \cdot \frac{(X-\mu)-(X-\lambda)}{(X-\lambda)(X-\mu)} = \frac{1}{\lambda-\mu} \left(\frac{1}{X-\lambda} - \frac{1}{X-\mu} \right).$$

Exercice 12

Soit P un polynôme complexe non constant.

Exprimer en fonction des racines de P et de leurs multiplicités respectives la décomposition en éléments simples de la fraction P'/P .

Solution

Notons $\lambda_1, \dots, \lambda_m$ les racines de P et $\alpha_1, \dots, \alpha_m$ leurs multiplicités respectives. En introduisant le coefficient dominant a , on peut écrire la factorisation

$$P = a \prod_{k=1}^m (X - \lambda_k)^{\alpha_k}.$$

méthode

|| La dérivée d'un produit est la somme des produits obtenus en ne dérivant qu'un facteur.

$$P' = a \sum_{k=1}^m \underbrace{\left(\alpha_k (X - \lambda_k)^{\alpha_k-1} \times \left(\prod_{\substack{1 \leq j \leq m \\ j \neq k}} (X - \lambda_j)^{\alpha_j} \right) \right)}_{= ((X - \lambda_k)^{\alpha_k})'}$$

En simplifiant, on obtient

$$\frac{P'}{P} = \sum_{k=1}^m \frac{\alpha_k}{X - \lambda_k}.$$

Cette identité correspond à une décomposition en éléments simples. Or il y a unicité de ces décompositions, c'est donc la décomposition en éléments simples de P'/P .

1. Il n'est pas nécessaire de déterminer ω , savoir $\omega^2 = -(\omega + 2)$ suffit à mener les calculs.
2. En pratique, on est rarement aussi désespéré...

5.7 Exercices d'entraînement

5.7.1 Généralités

Exercice 13 *

Déterminer les polynômes P de $\mathbb{K}[X]$ vérifiant $P(X+1) = P(X)$.

Solution

Soit P un polynôme solution.

méthode

|| On étudie les racines du polynôme $Q = P(X) - P(0)$.

Par la propriété $P(X+1) = P(X)$, on obtient $P(k+1) = P(k)$ et donc $P(k) = P(0)$ pour tout $k \in \mathbb{N}$. Le polynôme Q admet donc une infinité de racines (tous les entiers naturels), c'est donc le polynôme nul. On en déduit que le polynôme P est constant. La réciproque est immédiate.

Exercice 14 *

Déterminer les polynômes réels P de degré au plus 3 tels que

$$(X-1)^2 \mid (P-1) \quad \text{et} \quad (X+1)^2 \mid (P+1).$$

Solution

méthode

|| Afin de réduire le nombre de calculs, on commence par déterminer P' .

Analyse : Soit P un polynôme solution. 1 est racine au moins double de $P-1$ et donc aussi racine de $(P-1)' = P'$. Par le même argument, -1 est aussi racine de P' et donc $X^2 - 1 = (X-1)(X+1)$ divise P' (Th. 7 p. 153.). Au surplus, P' est de degré au plus 2 et l'on peut donc écrire $P' = a(X^2 - 1)$ avec $a \in \mathbb{R}$. Par intégration

$$P = \frac{a}{3}X^3 - aX + b \quad \text{avec} \quad (a, b) \in \mathbb{R}^2.$$

Les conditions $P(1) = 1$ et $P(-1) = -1$ permettent de déterminer $a = -3/2$ et $b = 0$:

$$P = -\frac{1}{2}X^3 + \frac{3}{2}X.$$

Synthèse : Le polynôme proposé est bien solution puisque de degré 3 et défini de sorte que 1 et -1 sont racines au moins doubles de respectivement $P-1$ et $P+1$.

Exercice 15 **

Déterminer les polynômes de $\mathbb{K}[X]$ divisibles par leur polynôme dérivé.

Solution

Parmi les polynômes constants, seul le polynôme nul est divisible par son polynôme dérivé. Il reste à déterminer les solutions parmi les polynômes non constants.

Soit P un polynôme non constant et $n \in \mathbb{N}^*$ son degré.

méthode

On exprime P en fonction de P' , puis P' en fonction de P'' et ainsi de suite jusqu'à parvenir au polynôme constant $P^{(n)}$.

Analyse : Si P' divise P , on peut écrire

$$nP = (X - \lambda)P' \quad \text{avec } \lambda \in \mathbb{K} \quad (*)$$

car $\deg(P') = \deg(P) - 1$ et le coefficient dominant de P' vaut n fois celui de P . En dérivant $(*)$, on obtient $nP' = (X - \lambda)P'' + P'$ et donc

$$(n - 1)P' = (X - \lambda)P''.$$

On répète ce calcul jusqu'à écrire

$$P^{(n-1)} = (X - \lambda)P^{(n)}.$$

Or $P^{(n)} = n!a$ avec a le coefficient dominant de P . En remontant les calculs précédents, on obtient $P = a(X - \lambda)^n$.

Synthèse : Les polynômes de la forme $a(X - \lambda)^n$ avec $a, \lambda \in \mathbb{K}$ et $n \in \mathbb{N}^*$ sont divisibles par leur polynôme dérivé.

Exercice 16 ***

Déterminer les polynômes non nuls P de $\mathbb{C}[X]$ vérifiant :

$$(a) P(X^2) = P(X)P(X+1) \qquad (b) P(X^2) = P(X)P(X-1).$$

Solution

(a) Soit P un polynôme non nul vérifiant $P(X^2) = P(X)P(X+1)$.

méthode

On détermine les racines possibles de P en observant que, lorsque λ est racine de P , d'autres racines s'en déduisent.

Si λ est racine de P alors λ^2 l'est aussi car

$$P(\lambda^2) = P(\lambda)P(\lambda+1) = 0.$$

De même, $\lambda^4, \lambda^8, \dots$ sont alors racines de P . Or le polynôme P n'admet qu'un nombre fini de racines. La suite des $\lambda, \lambda^2, \lambda^4, \dots$ doit comporter des répétitions et il existe donc k et ℓ dans \mathbb{N} avec $k < \ell$ tel que $\lambda^{2^\ell} = \lambda^{2^k}$. Ceci entraîne que λ est nul ou bien égal à une racine de l'unité. En particulier, si λ n'est pas nul, il est de module 1.

Aussi, si λ est racine de P , $(\lambda - 1)^2$ est racine de P car

$$P((\lambda - 1)^2) = P(\lambda - 1)P(\lambda) = 0.$$

Comme au-dessus, on peut affirmer $\lambda - 1 = 0$ ou $|\lambda - 1| = 1$.

En dehors de $\lambda = 0$ et $\lambda = 1$, les racines λ possibles figurent à l'intersection des deux cercles déterminés par les conditions $|\lambda| = 1$ et $|\lambda - 1| = 1$: ce sont les complexes $e^{i\pi/3}$ et $e^{-i\pi/3}$. Cependant, ces deux racines sont à exclure car leurs carrés ne sont pas racines de P .

En résumé, seuls 0 et 1 peuvent être racines de P et l'on peut écrire

$$P = aX^\alpha(X - 1)^\beta \quad \text{avec } a \in \mathbb{C}^*, \alpha, \beta \in \mathbb{N}.$$

On vérifie alors qu'un tel polynôme est solution si, et seulement si,

$$aX^{2\alpha}(X^2 - 1)^\beta = a^2X^\alpha(X - 1)^\beta(X + 1)^\alpha X^\beta.$$

Si cette condition est vérifiée, on a $a^2 = a$ donc $a = 1$. Aussi, en identifiant l'exposant des puissances de X , on obtient $2\alpha = \alpha + \beta$ donc $\alpha = \beta$. Inversement, si $a = 1$ et $\alpha = \beta$, l'égalité précédente est vraie car on peut écrire $X^2 - 1 = (X - 1)(X + 1)$.

Finalement, les polynômes non nuls solutions de l'équation sont les

$$(X^2 - X)^\alpha \quad \text{avec } \alpha \in \mathbb{N}.$$

(b) Soit P un polynôme non nul vérifiant $P(X^2) = P(X)P(X - 1)$. En raisonnant comme au-dessus, on obtient que si λ est racine de P alors λ^2 et $(\lambda + 1)^2$ le sont aussi. On en déduit $\lambda = 0$, $\lambda = -1$ ou λ vérifie $|\lambda| = |\lambda + 1| = 1$ ce qui donne $\lambda = j$ ou j^2 avec $j = e^{i\pi/3}$. Le cas $\lambda = 0$ est à exclure car $(0 + 1)^2$ n'est pas racine de P . De même, le cas $\lambda = -1$ peut être exclu car $(-1)^2$ n'est pas racine de P . Ainsi, les seules racines possibles pour P sont j et j^2 ce qui permet d'écrire

$$P = a(X - j)^\alpha(X - j^2)^\beta \quad \text{avec } a \in \mathbb{C}^*, \alpha, \beta \in \mathbb{N}.$$

Un tel polynôme est alors solution si, et seulement si,

$$a(X^2 - j)^\alpha(X^2 - j^2)^\beta = a^2(X - j)^\alpha(X - j^2)^\beta(X - 1 - j)^\alpha(X - 1 - j^2)^\beta.$$

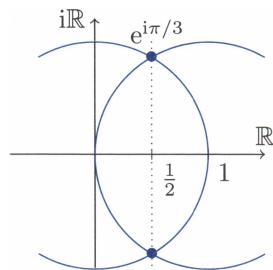
En exploitant $j = j^4$ en premier membre et $1 + j + j^2 = 0$ en second membre, on obtient l'égalité équivalente

$$a(X - j^2)^\alpha(X + j^2)^\alpha(X - j)^\beta(X + j)^\beta = a^2(X - j)^\alpha(X - j^2)^\beta(X + j^2)^\alpha(X + j)^\beta.$$

Cette égalité est vérifiée si, et seulement si, $a = 1$ et $\alpha = \beta$.

Finalement, les polynômes non nuls solutions sont les

$$(X - j)^\alpha(X - j^2)^\alpha \quad \text{c'est-à-dire} \quad (X^2 + X + 1)^\alpha \quad \text{avec } \alpha \in \mathbb{N}.$$



5.7.2 Racines

Exercice 17 *

Soit

$$P = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0$$

un polynôme à coefficients entiers tel que $a_n \neq 0$ et $a_0 \neq 0$.

(a) On suppose que P admet une racine rationnelle $r = p/q$ exprimée sous forme irréductible. Montrer que $p \mid a_0$ et $q \mid a_n$.

(b) Factoriser dans $\mathbb{R}[X]$

$$P = 2X^3 - X^2 - X - 3.$$

Solution

(a) **méthode**

Par réduction au même dénominateur, on transforme l'égalité $P(r) = 0$ en une identité entre entiers permettant d'employer les outils d'arithmétique.

L'égalité $P(r) = 0$ donne

$$a_n p^n + a_{n-1} p^{n-1} q + \cdots + a_1 p q^{n-1} + a_0 q^n = 0$$

où tous les paramètres sont entiers. Puisque p divise la portion $a_n p^n + \cdots + a_1 p q^{n-1}$, il divise aussi $a_0 q^n$. Or p et q sont premiers entre eux et donc p divise a_0 en vertu du lemme de Gauss (Th. 15 p. 91). Un raisonnement analogue donne $q \mid a_n p^n$ donc $q \mid a_n$.

(b) Si P admet un racine rationnelle $r = p/q$ alors p divise 3 et q divise 2. Les racines rationnelles possibles se limitent donc à la liste : $\pm 1, \pm 3, \pm \frac{1}{2}$ et $\pm \frac{3}{2}$. On observe que $\frac{3}{2}$ est racine et l'on peut donc factoriser P par $2X - 3$:

$$P = (2X - 3)(X^2 + X + 1).$$

Le trinôme apparu en second membre étant sans racines réelles, la factorisation dans $\mathbb{R}[X]$ s'arrête à cette écriture.

Exercice 18 *

Soit

$$P = X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0 \in \mathbb{C}[X].$$

Montrer que si un nombre complexe ξ est racine de P alors

$$|\xi| \leq \max(1, |a_0| + |a_1| + \cdots + |a_{n-1}|).$$

Solution

Si ξ est racine de P , l'égalité $P(\xi) = 0$ donne

$$\xi^n = -(a_{n-1}\xi^{n-1} + \cdots + a_1\xi + a_0). \quad (*)$$

méthode

- || On discute selon que $|\xi|$ est inférieur ou supérieur à 1 afin de pouvoir comparer les $|\xi|^k$ entre eux.

Cas : $|\xi| \leq 1$. L'inégalité voulue est vraie.

Cas : $|\xi| > 1$. On a $|\xi|^k \leq |\xi|^{n-1}$ pour tout $k \in [0; n-1]$. L'identité (*) donne alors par l'inégalité triangulaire

$$\begin{aligned} |\xi|^n &\leq |a_{n-1}|\xi^{n-1} + \cdots + |a_1|\xi + |a_0| \\ &\leq |a_{n-1}|\xi^{n-1} + \cdots + |a_1|\xi^{n-1} + |a_0|\xi^{n-1}. \end{aligned}$$

En simplifiant par $|\xi|^{n-1} \neq 0$, on obtient $|\xi| \leq |a_{n-1}| + \cdots + |a_1| + |a_0|$ et l'on peut conclure.

Exercice 19 **

Soit P un polynôme réel non constant.

- On suppose que P est scindé à racines simples. Montrer que le polynôme P' est lui aussi scindé.
- Montrer que le résultat perdure même si les racines de P ne sont pas simples.
- Le polynôme $X^6 - X + 1$ est-il scindé sur \mathbb{R} ?

Solution

(a) Soit $n \in \mathbb{N}^*$ le degré de P . Le polynôme P étant supposé scindé à racines simples, il possède exactement n racines distinctes $\lambda_1, \dots, \lambda_n$.

méthode

- || Par le théorème de Rolle, on montre que P' possède $n-1$ racines distinctes.

Quitte à redéfinir l'indexation des racines, on suppose celles-ci triées en ordre strictement croissant $\lambda_1 < \lambda_2 < \cdots < \lambda_n$ afin que les intervalles $[\lambda_k; \lambda_{k+1}]$ ne se chevauchent pas. Soit $k \in [1; n-1]$. La fonction polynomiale $t \mapsto P(t)$ est continue sur $[\lambda_k; \lambda_{k+1}]$, dérivable sur $[\lambda_k; \lambda_{k+1}]$ et prend les mêmes valeurs en λ_k et λ_{k+1} . Par le théorème de Rolle, il existe $\mu_k \in [\lambda_k; \lambda_{k+1}]$ tel que $P'(\mu_k) = 0$. Ceci détermine $n-1$ racines pour le polynôme P' , toutes distinctes car

$$\lambda_1 < \mu_1 < \lambda_2 < \mu_2 < \cdots < \lambda_{n-1} < \mu_{n-1} < \lambda_n.$$

Le polynôme P' étant de degré $n-1$, il est scindé à racines simples (Th. 12 p. 155).

(b) Soit $n \geq 1$ le degré de P et $m \geq 1$ le nombre de ses racines distinctes $\lambda_1, \dots, \lambda_m$. On introduit aussi leurs multiplicités respectives $\alpha_1, \dots, \alpha_m \in \mathbb{N}^*$. Par dérivation, les réels $\lambda_1, \dots, \lambda_m$ sont racines de P' de multiplicités¹ respectives $\alpha_1 - 1, \dots, \alpha_m - 1$. Comme au-dessus, le théorème de Rolle fournit $m - 1$ racines distinctes des précédentes μ_1, \dots, μ_{m-1} et la somme des multiplicités des racines de P' est au moins égale à

$$\sum_{k=1}^m \underbrace{(\alpha_k - 1)}_{\text{multiplicité de } \lambda_k} + \sum_{k=1}^{m-1} \underbrace{1}_{\leq \text{multiplicité de } \mu_k} = \sum_{k=1}^{m-1} \alpha_k - m + (m - 1) = n - 1.$$

Le polynôme P' étant de degré $n - 1$, il est scindé. On peut ajouter que les μ_k sont des racines simples.

(c) Lorsque P est scindé sur \mathbb{R} , les seules racines multiples de P' sont des racines de P . Ici $P' = 6X^5 - 1$ et $P'' = 30X^4$. 0 est racine multiple de P'' sans être racine de P' . Le polynôme P' n'est donc pas scindé sur \mathbb{R} et, *a fortiori*, P ne l'est pas non plus².

5.7.3 Relations coefficients-racines d'un polynôme scindé

Exercice 20 **

Soit $a \in \mathbb{R}$, $n \in \mathbb{N}^*$ et $P_n = (X + 1)^n - e^{2ina}$

(a) Déterminer les racines du polynôme P_n ainsi que leurs multiplicités.

(b) En déduire la valeur de

$$\prod_{k=0}^{n-1} \sin\left(a + \frac{k\pi}{n}\right).$$

Solution

(a) méthode

On résout l'équation $z^n = Z$ d'inconnue $z \in \mathbb{C}$ en introduisant les racines n -ièmes de l'unité $\omega_k = e^{2ik\pi/n}$ avec $k \in [0; n - 1]$.

Soit $z \in \mathbb{C}$.

$$\begin{aligned} (z+1)^n = e^{2ina} &\iff \left(\frac{z+1}{e^{2ia}}\right)^n = 1 \\ &\iff \exists k \in [0; n - 1], \frac{z+1}{e^{2ia}} = e^{2ik\pi/n} \\ &\iff \exists k \in [0; n - 1], z = \underbrace{e^{2i(a+k\pi/n)}}_{=x_k} - 1. \end{aligned}$$

1. Si λ est racine simple de P alors λ est racine de multiplicité 0 de P' , autrement dit, λ n'est pas racine de P' .

2. En revanche, ce polynôme est scindé sur \mathbb{C} comme le sont tous les polynômes non nuls.

Les racines du polynôme P_n sont les z_k pour $k \in \llbracket 0 ; n - 1 \rrbracket$. Ces valeurs étant deux à deux distinctes et au nombre de $n = \deg(P_n)$, ce sont des racines simples¹.

(b) méthode

|| Le coefficient constant du polynôme P_n est lié au produit des racines.

Le polynôme P_n étant scindé et unitaire, on peut écrire

$$(X + 1)^n - e^{2ina} = \prod_{k=0}^{n-1} (X - z_k).$$

En évaluant² en 0, on obtient

$$1 - e^{2ina} = (-1)^n \prod_{k=0}^{n-1}$$

méthode

|| On factorise par l'exponentielle imaginaire d'angle moitié :

$$e^{i\theta} - 1 = e^{i\frac{\theta}{2}} (e^{i\frac{\theta}{2}} - e^{-i\frac{\theta}{2}}) = 2ie^{i\frac{\theta}{2}} \sin\left(\frac{\theta}{2}\right).$$

On écrit alors

$$\prod_{k=0}^{n-1} z_k = \prod_{k=0}^{n-1} (e^{2i(a + \frac{k\pi}{n})} - 1) = 2^n i^n \left(\prod_{k=0}^{n-1} e^{i(a + \frac{k\pi}{n})} \right) \left(\prod_{k=0}^{n-1} \sin\left(a + \frac{k\pi}{n}\right) \right).$$

Or

$$\prod_{k=0}^{n-1} e^{i(a + \frac{k\pi}{n})} = e^{ina} \exp\left(i \sum_{k=0}^{n-1} \frac{k\pi}{n}\right) = e^{ina} \exp\left(i \frac{n(n-1)\pi}{2n}\right) = i^{n-1} e^{ina}$$

donc

$$\prod_{k=0}^{n-1} z_k = (-1)^{n-1} i 2^n e^{ina} \prod_{k=0}^{n-1} \sin\left(a + \frac{k\pi}{n}\right).$$

On en déduit

$$\prod_{k=0}^{n-1} \sin\left(a + \frac{k\pi}{n}\right) = \frac{i}{2^n} \cdot \frac{1 - e^{2ina}}{e^{ina}} = \frac{\sin(na)}{2^{n-1}}$$

Exercice 21 **

Donner une condition nécessaire et suffisante sur $(p, q) \in \mathbb{C}^2$ pour que le polynôme $X^3 + pX + q$ admette une racine multiple et déterminer celle-ci.

1. Voir sujet 3 p. 165.

2. Ou en identifiant le coefficient constant dans les deux polynômes : cela revient au même.

Solution**méthode**

On simplifie¹ le système exprimant les coefficients du polynôme $X^3 + pX + q$ en fonction de ses racines sachant que deux racines sont identiques.

Notons x, y, z les trois racines comptées avec multiplicité du polynôme $X^3 + pX + q$. On sait (Th. 13 p. 155)

$$\begin{cases} x + y + z = 0 \\ xy + yz + zx = p \\ xyz = -q. \end{cases}$$

Si le polynôme P admet au moins une racine double, on peut supposer $z = x$ quitte à renommer les racines. Dans ce cas le système précédent se réécrit

$$\begin{cases} 2x + y = 0 \\ 2xy + x^2 = p \\ x^2y = -q \end{cases} \text{ puis } \begin{cases} y = -2x \\ -3x^2 = p \\ 2x^3 = q \end{cases} \begin{array}{l} (1) \\ (2). \end{array}$$

Les équations (1) et (2) entraînent alors la condition $4p^3 + 27q^2 = 0$.

Inversement, supposons cette condition remplie.

méthode

On détermine x et y solutions du système précédent en faisant le quotient des équations (1) et (2).

Afin de pouvoir diviser par p , on traite séparément le cas $p = 0$.

Si $p = 0$ alors $q = 0$ et le polynôme X^3 admet au moins une racine double.

Si $p \neq 0$, on peut poser

$$x = -\frac{3q}{2p} \quad \text{et} \quad y = \frac{3q}{p}.$$

On vérifie alors le développement

$$(X - x)^2(X - y) = X^3 + pX + q$$

et l'on peut affirmer que le polynôme $X^3 + pX + q$ admet une racine double.

Exercice 22 ***

Soit x, y, z trois nombres complexes de somme nulle. Vérifier

$$\frac{x^5 + y^5 + z^5}{5} = \left(\frac{x^2 + y^2 + z^2}{2} \right) \left(\frac{x^3 + y^3 + z^3}{3} \right).$$

1. D'autres méthodes sont possibles, voire plus simples, comme rechercher les racines de P' et étudier à quelle condition l'une d'elles est racine de P ou encore calculer le PGCD de P et P' .

Solution

Posons $S_2 = x^2 + y^2 + z^2$, $S_3 = x^3 + y^3 + z^3$ et $S_5 = x^5 + y^5 + z^5$.

méthode

On introduit le polynôme

$$P = (X - x)(X - y)(Z - z) = X^3 + pX + q$$

avec $p = xy + yz + zx$ et $q = -xyz$.

Par développement d'un carré

$$(x + y + z)^2 = x^2 + y^2 + z^2 + 2p$$

et donc $S_2 = -2p$.

Les valeurs x , y et z étant chacune racines de P , on a

$$P(x) + P(y) + P(z) = x^3 + y^3 + z^3 + p(x + y + z) + 3q = 0$$

et donc $S_3 = -3q$.

Enfin, $x^3 = -px - q$ donne $x^5 = -px^3 - qx^2$ puis $x^5 = p^2x + pq - qx^2$. En sommant avec les relations analogues pour y et z , il vient

$$x^5 + y^5 + z^5 = p^2 \times 0 + 3pq - qS_2 = 5pq.$$

La relation proposée est dès lors immédiate.

5.7.4 Arithmétiques des polynômes

Exercice 23 *

Soit λ et μ deux éléments distincts de \mathbb{K} et $P \in \mathbb{K}[X]$.

- (a) Exprimer en fonction de P le reste de la division de P par $(X - \lambda)(X - \mu)$.
- (b) Exprimer en fonction de P le reste de la division de P par $(X - \lambda)^2$.

Solution

- (a) La division euclidienne de P par $(X - \lambda)(X - \mu)$ s'écrit

$$P = (X - \lambda)(X - \mu)Q + R \quad (*)$$

avec Q et R polynômes, $\deg(R) < 2$ (Th. 5 p. 152).

méthode

On exprime le reste à coefficients inconnus et l'on détermine ceux-ci en évaluant en λ et μ .

On écrit $R = aX + b$ avec $a, b \in \mathbb{K}$. En évaluant (*) en λ et μ , on forme un système d'inconnue (a, b)

$$\begin{cases} a\lambda + b = P(\lambda) \\ a\mu + b = P(\mu). \end{cases}$$

Après résolution, on obtient¹

$$R = \frac{P(\mu) - P(\lambda)}{\mu - \lambda} X + \frac{\mu P(\lambda) - \lambda P(\mu)}{\mu - \lambda}$$

(b) La division euclidienne de P par $(X - \lambda)^2$ s'écrit

$$P = (X - \lambda)^2 Q + R \quad (\Delta)$$

avec Q et R polynômes, $R = aX + b$ avec $a, b \in \mathbb{K}$.

méthode

|| On opère comme au-dessus en exploitant au surplus une dérivation.

On forme un système de deux équations d'inconnue (a, b) en évaluant (Δ) en λ puis en dérivant (Δ) avant d'évaluer à nouveau en λ

$$\begin{cases} a\lambda + b = P(\lambda) \\ a = P'(\lambda). \end{cases}$$

Après résolution, on obtient

$$R = P'(\lambda)(X - \lambda) + P(\lambda).$$

Exercice 24 **

Montrer que, pour tous a et $b \in \mathbb{N}^*$,

$$b \mid a \iff X^b - 1 \mid X^a - 1.$$

Solution

méthode

|| On exprime le reste de la division euclidienne de $X^a - 1$ par $X^b - 1$ en fonction de celui de a par b .

La division euclidienne de a par b s'écrit

$$a = bq + r \quad \text{avec} \quad 0 \leq r < b.$$

Vérifions alors que le reste de la division euclidienne de $X^a - 1$ par $X^b - 1$ est $X^r - 1$. On étudie pour cela si $X^b - 1$ divise la différence $(X^a - 1) - (X^r - 1)$.

Par l'identité géométrique

$$\begin{aligned} (X^a - 1) - (X^r - 1) &= X^{bq+r} - X^r = X^r(X^{bq} - 1) \\ &= X^r(X^b - 1)(1 + X^b + \cdots + X^{b(q-1)}). \end{aligned}$$

1. Le polynôme R détermine la droite passant par $(\lambda, P(\lambda))$ et $(\mu, P(\mu))$.

On peut donc écrire la division euclidienne

$$X^a - 1 = (X^b - 1)Q + R$$

avec Q un polynôme et $R = X^r - 1$ vérifiant $\deg(R) < \deg(X^b - 1)$.

Il est alors possible de conclure par une chaîne d'équivalences :

$$\begin{aligned} b \mid a &\iff r = 0 \\ &\iff X^r - 1 = 0 \\ &\iff X^b - 1 \mid X^a - 1. \end{aligned}$$

Exercice 25 **

Soit A et B deux polynômes non constants de $\mathbb{K}[X]$ premiers entre eux.

Montrer qu'il existe un unique couple $(U, V) \in \mathbb{K}[X]^2$ vérifiant

$$AU + BV = 1 \quad \text{et} \quad \begin{cases} \deg(U) < \deg(B) \\ \deg(V) < \deg(A). \end{cases}$$

Solution

Existence : Puisque A et B sont premiers entre eux, le théorème de Bézout (Th. 23 p. 158) assure l'existence d'un premier couple (U, V) de polynômes vérifiant $AU + BV = 1$.

méthode

|| Par division euclidienne, on réduit les degrés des polynômes U et V .

La division euclidienne de U par B s'écrit

$$U = BQ + \hat{U} \quad \text{avec} \quad \deg(\hat{U}) < \deg(B).$$

Posons ensuite $V = V + AQ$. On vérifie $A\hat{U} + BV = AU + BV = 1$ avec la première condition $\deg(\hat{U}) < \deg(B)$. Aussi, on a alors (Th. 1 p. 150)

$$\underbrace{\deg(AU + BV)}_{=1} < \max(\deg(AU), \deg(BV)).$$

Les degrés des polynômes non constants $A\hat{U}$ et BV doivent donc être égaux pour que leurs plus grandes puissances de X se simplifient. On en déduit

$$\deg(V) = \deg(A) + \deg(\hat{U}) - \deg(B) < \deg(A).$$

On a ainsi formé un couple (U, V) tel que voulu.

Unicité : Soit (U, V) et (\hat{U}, \hat{V}) deux couples solutions. On a à la fois $AU + BV = 1$ et $A\hat{U} + B\hat{V} = 1$. Par différence, il vient

$$A(U - \hat{U}) = B(\hat{V} - V).$$

Le polynôme A divise alors $B(V - V)$. Or il est premier avec B et donc, par le lemme de Gauss (Th. 24 p. 159), il divise $V - V$. Cependant,

$$\deg(V - V) \leq \max(\deg(V), \deg(V)) < \deg(A)$$

et $V - V$ est donc le polynôme nul. On en déduit $V = \bar{V}$ puis $U = \bar{U}$.

Exercice 26 **

Soit $P \in \mathbb{K}[X]$. Montrer que $P(X) - X$ divise $P(P(X)) - P(X)$.

Solution
méthode

|| On vérifie d'abord que $P(X) - X$ divise $P(P(X)) - P(X)$.

En introduisant les coefficients de P , on peut écrire

$$P(X) = \sum_{k=0}^n a_k X^k$$

et alors

$$P(P(X)) - P(X) = \sum_{k=0}^n a_k (P(X)^k - X^k).$$

Or, en vertu de l'identité géométrique, $P(X) - X$ divise chaque terme $P(X)^k - X^k$:

$$(P(X)^k - X^k) = (P(X) - X)(P(X)^{k-1} + X P(X)^{k-2} + \cdots + X^{k-1}).$$

On en déduit que $P(X) - X$ divise $P(P(X)) - P(X)$ puis divise

$$P(P(X)) - X = (P(P(X)) - P(X)) + (P(X) - X).$$

Exercice 27 **

Soit $a \in]0 ; \pi[$ et $n \in \mathbb{N}^*$. Décomposer en facteurs irréductibles dans $\mathbb{R}[X]$

$$X^{2n} - 2 \cos(a) X^n + 1.$$

Solution
méthode

|| On factorise le polynôme dans $\mathbb{C}[X]$ en commençant par déterminer ses racines puis on combine entre eux les facteurs correspondant aux racines conjuguées.

Le polynôme $Y^2 - 2 \cos(a)Y + 1$ est de discriminant $\Delta = (2i \sin(a))^2 < 0$ et ses racines sont e^{ia} et e^{-ia} . On a donc la factorisation

$$Y^2 - 2 \cos(a)Y + 1 = (Y - e^{ia})(Y - e^{-ia}).$$

Par substitution

$$X^{2n} - 2 \cos(a)X^n + 1 = (X^n - e^{ia})(X^n - e^{-ia}).$$

Les racines de $X^n - e^{ia}$ sont¹ les $e^{i(a+2k\pi)/n}$ avec $k \in \llbracket 0 ; n-1 \rrbracket$ et celles de $X^n - e^{-ia}$ s'en déduisent par conjugaison. Ainsi,

$$\begin{aligned} X^{2n} - 2 \cos(a)X^n + 1 &= \prod_{k=0}^{n-1} (X - e^{i(a+2k\pi)/n}) \prod_{k=0}^{n-1} (X - e^{i(a+2k\pi)/n}) \\ &= \prod_{k=0}^{n-1} \left(X^2 - 2 \cos\left(\frac{a+2k\pi}{n}\right)X + 1 \right). \end{aligned}$$

Les facteurs du second degré écrits sont assurément irréductibles car le polynôme étudié initialement est sans racines réelles.

5.7.5 Familles de polynômes classiques

Exercice 28 ** (Polynômes de Tchebychev)

Soit $n \in \mathbb{N}$. On pose $f_n : [-1 ; 1] \rightarrow \mathbb{R}$ l'application définie par

$$f_n(x) = \cos(n \arccos x).$$

(a) Soit $x \in [-1 ; 1]$. Simplifier $f_0(x)$, $f_1(x)$ et $f_2(x)$.

Exprimer $f_{n+1}(x) + f_{n-1}(x)$ en fonction de $f_n(x)$ et donner $f_3(x)$.

(b) Établir qu'il existe un unique polynôme T_n de $\mathbb{R}[X]$ dont la fonction polynomiale associée coïncide avec f_n sur $[-1 ; 1]$.

(c) Donner le degré de T_n ainsi que son coefficient dominant.

(d) Montrer que T_n possède n racines distinctes toutes dans $]-1 ; 1[$.

Solution

(a) **méthode**

|| On pose $\theta = \arccos x$ afin d'alléger les écritures qui suivent.

On sait $\cos \theta = x$ et l'on obtient directement

$$f_0(x) = 1, \quad f_1(x) = x \quad \text{et} \quad f_2(x) = \cos(2\theta) = 2 \cos^2 \theta - 1 = 2x^2 - 1.$$

1. Pour déterminer les racines, on résout une équation $z^n = Z$. Pour $z_0 = e^{ia/n}$, on obtient une solution particulière que l'on exploite pour transformer l'équation en $(z/z_0)^n = 1$.

Par la factorisation de $\cos p + \cos q$ ou tout simplement par développement :

$$\begin{aligned} f_{n+1}(x) + f_{n-1}(x) &= \cos((n+1)\theta) + \cos((n-1)\theta) \\ &= 2\cos(\theta)\cos(n\theta) = 2xf_n(x). \end{aligned}$$

On en déduit

$$f_3(x) = 2xf_2(x) - f_1(x) = 4x^3 - 3x.$$

(b) *Unicité* : Soit T_n et \hat{T}_n deux polynômes solutions. Pour tout $x \in [-1; 1]$, on a

$$(T_n - \hat{T}_n)(x) = f_n(x) - f_n(x) = 0.$$

Le polynôme $T_n - T_n$ possède donc une infinité de racines : c'est le polynôme nul.

Existence :

méthode

Inspiré par les calculs de la question précédente, on introduit¹ la suite de polynômes (T_n) déterminée par les conditions

$$T_0 = 1, \quad T_1 = X \quad \text{et} \quad \forall n \in \mathbb{N}^*, \quad T_{n+1} = 2XT_n - T_{n-1}.$$

Vérifions par récurrence double sur $n \in \mathbb{N}$ que $T_n(x) = f_n(x)$ pour tout $x \in [-1; 1]$.

Compte tenu du choix de T_0 et T_1 , la propriété est vraie aux rangs 0 et 1. Supposons la propriété vraie aux rangs n et $n - 1$ avec $n \geq 1$. Pour tout $x \in [-1; 1]$, on vérifie :

$$T_{n+1}(x) = 2XT_n(x) - T_{n-1}(x) = 2xf_n(x) - f_{n-1}(x) = f_{n+1}(x).$$

La récurrence est établie.

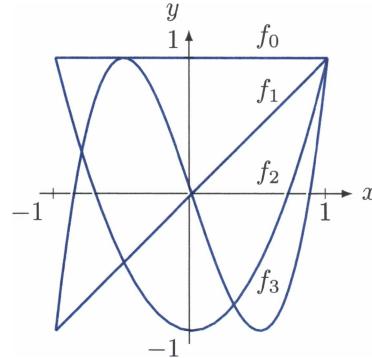
(c) Par récurrence double sur $n \in \mathbb{N}$, on vérifie $\deg(T_n) = n$.

La propriété est vraie pour $n = 0$ et $n = 1$. Supposons la propriété établie aux rangs n et $n - 1$ avec $n \geq 1$. On a $\deg(2XT_n) = n + 1$ et $\deg(T_{n-1}) = n - 1$. Par différence de polynômes de degrés distincts, on peut affirmer $\deg(T_{n+1}) = n + 1$.

La récurrence est établie.

Lors du calcul qui précède, on a vu que la plus grande puissance de X dans T_{n+1} provient du terme $2XT_n$. Le coefficient dominant de T_{n+1} est donc le double de celui de T_n pour $n \geq 1$. Sachant $T_1 = X$, on peut affirmer que le coefficient dominant de T_n est 2^{n-1} . Le coefficient dominant de T_0 se calcule séparément et vaut 1.

1. On trouvera une démarche alternative à celle proposée ici dans le sujet 23 du chapitre 3 de l'ouvrage *Exercices d'analyse MPSI*.



(d) méthode

On résout l'équation $T_n(x) = 0$ dans l'intervalle $[-1; 1]$ sur lequel on sait exprimer simplement T_n .

Soit $x \in [-1; 1]$.

$$\begin{aligned} T_n(x) = 0 &\iff \cos(n \arccos x) = 0 \\ &\iff \exists k \in \mathbb{Z}, n \arccos x = \frac{\pi}{2} + k\pi. \end{aligned}$$

Puisque $n \arccos x$ appartient à $[0; n\pi]$, le paramètre k ci-dessus évolue nécessairement dans $[0; n - 1]$. On reprend alors la chaîne d'équivalences :

$$\begin{aligned} T_n(x) = 0 &\iff \exists k \in [0; n - 1], \arccos x = \frac{(2k + 1)\pi}{2n} \\ &\iff \exists k \in [0; n - 1], x = \cos\left(\frac{(2k + 1)\pi}{2n}\right). \end{aligned}$$

La dernière équivalence étant assurée par la bijectivité de la fonction \cos sur $[0; \pi]$.

Ainsi, on a obtenu que T_n possède n racines dans l'intervalle $[-1; 1]$ à savoir les x_k pour $k \in [0; n - 1]$ avec

$$x_k = \cos\left(\frac{(2k + 1)\pi}{2n}\right).$$

Ces dernières sont deux à deux distinctes (car la fonction \cos est injective sur $[0; \pi]$) et éléments de $]-1; 1[$. Le polynôme T_n étant de degré n , il ne peut pas posséder d'autres racines et celles-ci sont simples.

Exercice 29 ** (Polynômes de Legendre)

Pour tout entier naturel n on pose

$$L_n = \frac{n!}{(2n)!} \left((X^2 - 1)^n \right)^{(n)}.$$

(a) Montrer que L_n est un polynôme unitaire de degré n .

(b) Vérifier que, pour tout polynôme réel Q avec $\deg(Q) < n$, on a

$$\int_{-1}^1 L_n(t) Q(t) dt = 0.$$

(c) En déduire que L_n possède n racines simples toutes dans l'intervalle $]-1; 1[$.

Solution

(a) L_n est le polynôme dérivé d'ordre n du polynôme $U_n = (X^2 - 1)^n$ qui est de degré $2n$. Chaque dérivation abaisse le degré d'une unité et donc $\deg(L_n) = n$.

De plus, le polynôme U_n est unitaire et les dérivations font successivement apparaître les facteurs $2n$, $2n - 1$, ..., $n + 1$ devant la plus grande puissance de X . Le coefficient dominant de L_n est donc

$$\frac{n!}{(2n)!} \times 2n \times (2n - 1) \times \cdots \times (n + 1) = 1.$$

(b) **méthode**

Par intégration par parties, on dérive le polynôme Q jusqu'à ce qu'il disparaîtse.

Les valeurs 1 et -1 sont racines d'ordre de multiplicité n de $U_n = (X - 1)^n(X + 1)^n$. Ces valeurs sont donc racines des polynômes dérivés $U_n^{(k)}$ pour tout $k \in [0; n - 1]$.

On réalise une première intégration par parties avec les fonctions u et v de classe C^1 données par

$$u(t) = U_n^{(n-1)}(t) \quad \text{et} \quad v(t) = Q(t).$$

On obtient

$$\frac{(2n)!}{n!} \int_{-1}^1 L_n(t)Q(t) dt = \underbrace{\left[U_n^{(n-1)}(t)Q(t) \right]_{-1}^1}_{=0} - \int_{-1}^1 U_n^{(n-1)}(t)Q'(t) dt.$$

En renouvelant ce type d'intégration par parties, il vient

$$\begin{aligned} \frac{(2n)!}{n!} \int_{-1}^1 L_n(t)Q(t) dt &= (-1)^2 \int_{-1}^1 U_n^{(n-2)}(t)Q''(t) dt = \cdots \\ &\quad - (-1)^n \int_{-1}^1 U_n(t) \underbrace{Q^{(n)}(t)}_{=0} dt = 0. \end{aligned}$$

(c) **méthode**

On multiplie L_n par un polynôme non nul choisi de sorte que le produit soit de signe constant sur $[-1; 1]$.

Soit a_1, a_2, \dots, a_p les racines d'ordres de multiplicité impairs du polynôme L_n appartenant à l'intervalle $]-1; 1[$. En chacune de celles-ci, la fonction $t \mapsto L_n(t)$ s'annule en changeant de signe. Introduisons alors le polynôme $Q = (X - a_1)(X - a_2) \dots (X - a_p)$ défini de sorte que la fonction $t \mapsto L_n(t)Q(t)$ soit de signe constant sur $[-1; 1]$. Cette fonction est de plus continue et n'est pas la fonction nulle sur $[-1; 1]$, son intégrale ne peut donc être nulle et par conséquent $p = \deg(Q) \geq n$. Sachant que le polynôme L_n est de degré n , il ne peut avoir plus de n racines et l'on peut affirmer qu'il possède exactement n racines simples¹ toutes dans l'intervalle $]-1; 1[$.

1. Une démonstration directe est aussi possible par application répétée du théorème de Rolle à U_n et ses polynômes dérivés : pour $k \in [0; n - 1]$, si $U_n^{(k)}$ s'annule k fois dans $]-1; 1[$, sachant qu'il s'annule aussi en -1 et en 1 , sa dérivée $U_n^{(k+1)}$ s'annule au moins $(k+1)$ fois dans $]-1; 1[$ (en fait exactement $k+1$ fois).

Exercice 30 * (Polynômes de Fibonacci¹)**

On considère la suite de polynômes (F_n) déterminée par

$$F_0 = 0, \quad F_1 = 1 \quad \text{et} \quad \forall n \in \mathbb{N}, \quad F_{n+2} = XF_{n+1} + F_n.$$

(a) Vérifier que, pour tout $n \in \mathbb{N}$, F_n et F_{n+1} sont premiers entre eux.

(b) Soit $k \in \mathbb{N}^*$. Montrer

$$F_{k+n} = F_k F_{n+1} + F_{k-1} F_n \quad \text{pour tout } n \in \mathbb{N}.$$

Soit $a \in \mathbb{N}$ et $b \in \mathbb{N}^*$.

(c) Établir

$$F_{a+b} \wedge F_b = F_a \wedge F_b.$$

(d) Conclure

$$F_a \wedge F_b = F_{a \wedge b}.$$

Solution

(a) On vérifie par récurrence sur $n \in \mathbb{N}$ que F_n et F_{n+1} sont premiers entre eux.

Pour $n = 0$, les polynômes 0 et 1 sont effectivement premiers entre eux.

Supposons la propriété vraie au rang $n \in \mathbb{N}$. Un polynôme diviseur commun à F_{n+1} et F_{n+2} divise aussi $F_n = F_{n+2} - XF_{n+1}$ et c'est donc un polynôme constant car diviseur commun à F_n et F_{n+1} .

La récurrence est établie.

(b) Soit $k \in \mathbb{N}^*$.

méthode

|| On vérifie que les deux suites $(F_{k+n})_{n \in \mathbb{N}}$ et $(F_k F_{n+1} + F_{k-1} F_n)_{n \in \mathbb{N}}$ satisfont les mêmes conditions de récurrence.

Posons $P_n = F_{k+n}$ et $Q_n = F_k F_{n+1} + F_{k-1} F_n$.

On a $P_0 = F_k$ et $Q_0 = F_k F_1 + F_{k-1} F_0 = F_k$.

On a aussi $P_1 = F_{k+1}$ et $Q_1 = F_k F_2 + F_{k-1} F_1 = F_{k+1}$ car $F_2 = X$.

Enfin, on vérifie $P_{n+2} = XP_{n+1} + P_n$ et $Q_{n+2} = XQ_{n+1} + Q_n$ pour tout naturel n . Les deux suites (P_n) et (Q_n) sont donc égales car satisfont la même relation de récurrence double ainsi que les deux mêmes conditions initiales.

(c) Le PGCD de F_a et F_b divise $F_{a+b} = F_b F_{a+1} + F_{b-1} F_a$ et donc divise le PGCD de F_{a+b} et F_b . Inversement, le PGCD de F_{a+b} et F_b divise $F_{b-1} F_a = F_{a+b} - F_b F_{a-1}$. Or il est premier avec F_{b-1} car il divise F_b . Par le lemme de Gauss, on peut affirmer qu'il divise F_a et donc qu'il divise le PGCD de F_a et F_b .

Finalement, les PGCD de F_{a+b} et F_b et de F_a et F_b sont les mêmes.

1. Cet énoncé peut être mis en parallèle avec le sujet 18 p. 107.

(d) Par récurrence sur $q \in \mathbb{N}$, on obtient pour tout $r \in \mathbb{N}$, $F_{bq+r} \wedge F_b = F_b \wedge F_r$ et donc

$$F_a \wedge F_b = F_b \wedge F_r$$

lorsque r désigne le reste de la division euclidienne de a par b .

En suivant l'algorithme d'Euclide calculant le PGCD de a et b et sachant que le dernier reste est nul alors que le précédent est le PGCD de a et b , on obtient

$$F_a \wedge F_b = F_{a \wedge b} \wedge F_0.$$

On en déduit que $F_a \wedge F_b = F_{a \wedge b}$ car $F_{a \wedge b}$ est un polynôme unitaire¹ alors que F_0 est nul.

5.7.6 Les fractions rationnelles

Exercice 31 *

Soit $n \in \mathbb{N}$. Former la décomposition en éléments simples de

$$\frac{n!}{X(X-1)\dots(X-n)}.$$

Solution

La fraction rationnelle est exprimée sous forme irréductible et sa partie entière est nulle. Le dénominateur est déjà factorisé et la décomposition s'écrit

$$\frac{n!}{X(X-1)\dots(X-n)} = \sum_{k=0}^n \frac{a_k}{(X-k)} \quad \text{avec } a_k \in \mathbb{R}.$$

Soit $k \in [0; n]$. On calcule a_k en multipliant par $X - k$ avant d'évaluer en k :

$$\begin{aligned} a_k &= \frac{n!}{X(X-1)\dots(X-(k-1)) \times (X-(k+1))\dots(X-n)} \Big|_{X=k} \\ &= \frac{n!}{k(k-1) \times \dots \times 1 \times (-1) \times \dots \times (k-n)}. \end{aligned}$$

En passant à l'opposé les $n - k$ derniers facteurs du dénominateur

$$a_k = (-1)^{n-k} \frac{n!}{k!(n-k)!} = (-1)^{n-k} \binom{n}{k}.$$

1. On vérifie par récurrence double que F_n est unitaire de degré $n - 1$ pour tout $n \in \mathbb{N}^*$.

Exercice 32 **

(a) Soit $a \in \mathbb{K}$ un pôle simple d'une fraction rationnelle F de $\mathbb{K}(X)$ exprimée sous forme irréductible P/Q . Montrer que la fraction F peut s'écrire

$$\frac{\alpha}{X-a} + G \quad \text{avec} \quad \alpha = \frac{P(a)}{Q'(a)} \quad \text{et} \quad G \in \mathbb{K}(X) \text{ dont } a \text{ n'est pas pôle.}$$

(b) Application : Soit $n \in \mathbb{N}^*$. On pose $\omega_k = e^{2ik\pi/n}$ pour tout $k \in \llbracket 0; n-1 \rrbracket$. Réduire au même dénominateur la fraction complexe

$$F = \frac{1}{X - \omega_0} + \cdots + \frac{1}{X - \omega_{n-1}}.$$

Solution(a) **méthode**

|| On organise les termes de la décomposition en éléments simples de F .

Si a est un pôle simple de F , c'est une racine simple du dénominateur Q . La décomposition en éléments simples de F fait alors apparaître un terme

$$\frac{\alpha}{X-a} \quad \text{avec} \quad \alpha = \left. \frac{(X-a)P}{Q} \right|_{X=a}$$

et d'autres termes qui, une fois regroupés, définissent¹ une fraction rationnelle G .

D'une part, par opérations sur des fractions dont a n'est pas pôle, a n'est pas pôle de G .

D'autre part, a étant racine simple de Q , on peut écrire $Q = (X-a)R$ avec $R \in \mathbb{K}[X]$ vérifiant $R(a) \neq 0$ et alors

$$\alpha = \left. \frac{(X-a)P}{Q} \right|_{X=a} = \left. \frac{P}{R} \right|_{X=a} = \frac{P(a)}{R(a)}$$

Or $Q' = R + (X-a)R'$ et donc $Q'(a) = R(a)$ ce qui donne l'expression voulue² de α .

(b) La réduction au même dénominateur de la fraction F est de la forme

$$F = \frac{P}{Q} \quad \text{avec} \quad P \in \mathbb{C}[X] \text{ et } Q = \prod_{k=0}^{n-1} (X - \omega_k).$$

1. Par exemple, si la décomposition de F s'écrit $E + \frac{\beta}{X-a} + \frac{\gamma}{X-b} + \frac{\gamma}{(X-b)^2} + \frac{\delta}{(X-c)}$ avec a, b, c distincts, la fraction G est la somme de la partie entière E et des termes $\frac{\beta}{X-b}$, $\frac{\gamma}{(X-b)^2}$ et $\frac{\delta}{X-c}$.

2. Celle-ci peut être retenue car utile dans de nombreux sujets. On notera cependant qu'elle n'est valable que pour des pôles simples.

Les complexes ω_k sont deux à deux distincts et correspondent aux n racines n -ièmes de l'unité, on a donc¹ $Q = X^n - 1$. Aussi, la fraction F est de degré strictement négatif et donc $\deg(P) < \deg(Q) = n$.

Par la formule qui précède, on sait

$$\frac{P(\omega_k)}{Q'(\omega_k)} = 1 \quad \text{pour tout } k \in [0; n-1]$$

et l'on a donc $P(\omega_k) = Q'(\omega_k)$. Les polynômes P et $Q' = nX^{n-1}$ sont alors tous deux de degrés inférieurs à $n-1$ et sont égaux en chacune des n valeurs ω_k , ils sont donc égaux².

Finalement, la réduction au même dénominateur de F s'écrit

$$F = \frac{nX^{n-1}}{X^n - 1}$$

Exercice 33 **

Si F est une fraction rationnelle de $\mathbb{K}(X)$ représentée par le quotient A/B (avec A et $B \in \mathbb{K}[X]$, $B \neq 0$), on définit la *fraction dérivée* de F par

$$F' \stackrel{\text{def}}{=} \frac{A'B - AB'}{B^2}.$$

- (a) Montrer que la définition de F' ne dépend pas du quotient A/B représentant F .
- (b) Étudier le degré de F' en fonction de celui de F .
- (c) Montrer qu'il n'existe pas de fractions F de $\mathbb{K}(X)$ vérifiant $F' = 1/X$.

Solution

(a) méthode

Deux quotients A/B et C/D (avec B et D non nuls) représentent la même fraction si, et seulement si, $AD = BC$.

Soit $A, B, C, D \in \mathbb{K}[X]$ avec $B, D \neq 0$. On suppose $AD = BC$ et l'on obtient en dérivant $A'D + AD' = B'C + BC'$. On a alors

$$(A'B - AB')D^2 = (A'D)BD - AB'D^2 = (B'C + BC' - AD')BD - AB'D^2.$$

On développe et l'on exploite l'égalité $AD = BC$ pour poursuivre le calcul

$$(A'B - AB')D^2 = B'CBD + BC'BD - (AD)D'B - AD'B'D = (C'D - CD')B^2.$$

$=BC$ $=BC$

On en déduit que les deux fractions

$$\frac{A'B - AB'}{B^2} \quad \text{et} \quad \frac{C'D - CD'}{B^2}$$

sont égales.

1. Voir sujet 8 p. 169.

2. Leur différence est nulle car celle-ci est un polynôme de degré inférieur à $n-1$ possédant au moins n racines.

(b) Soit A/B avec $A, B \in \mathbb{K}[X]$, $B \neq 0$ un quotient représentant F . Si la fraction F est constante, sa dérivée est nulle donc de degré $-\infty$. Supposons désormais cette situation écartée et introduisons p, q les degrés des polynômes A, B et a, b leurs coefficients dominants. S'il n'est pas nul, le polynôme $A'B$ est de degré $p+q-1$ et de coefficient dominant pab . Aussi, s'il n'est pas nul, le polynôme AB' est de degré $p+q-1$ et de coefficient dominant qab . Dans tous les cas, le polynôme différence $A'B - AB'$ est de degré inférieur à $p+q-1$ et le coefficient de X^{p+q-1} dans celui-ci est $(p-q)ab$. Distinguons alors deux situations :

Cas : $\deg(F) \neq 0$. On a $p \neq q$, le polynôme $A'B - AB'$ est exactement de degré $p+q-1$ donc

$$\deg(F') = \deg(A'B - AB') - \deg(B^2) = p + q - 1 - 2q = p - q - 1 = \deg(F) - 1.$$

Cas : $\deg(F) = 0$. On a $p = q$ et le polynôme $A'B - AB'$ est de degré strictement inférieur à $p+q-1$. On en déduit¹

$$\deg(F') \leq p + q - 2 - 2q = \deg(F) - 2.$$

(c) Par l'étude qui précède, on voit qu'il est impossible que le degré de la dérivée d'une fraction rationnelle soit égal à -1 . Il n'existe donc pas de fractions telles que $F' = 1/X$.

Exercice 34 **

Soit P un polynôme réel scindé sur \mathbb{R} . Montrer que

$$P'(x)^2 - P(x)P''(x) > 0 \quad \text{pour tout } x \in \mathbb{R}.$$

Solution

méthode

|| L'expression en premier membre fait penser à la dérivée d'un quotient.

En considérant la dérivée d'un quotient de deux fonctions polynômes², on peut affirmer que, pour tout x réel qui n'est pas racine de P ,

$$\frac{d}{dx} \left(\frac{P'(x)}{P(x)} \right) = \frac{P(x)P''(x) - P'(x)^2}{P(x)^2}.$$

En introduisant $\lambda_1, \dots, \lambda_m$ les racines de P et $\alpha_1, \dots, \alpha_m \in \mathbb{N}^*$ leurs multiplicités respectives, on sait³ que pour tout x réel différent des $\lambda_1, \dots, \lambda_m$

$$\frac{P'(x)}{P(x)} = \sum_{j=1}^m \frac{\alpha_j}{x - \lambda_j} \quad \text{donc} \quad \frac{d}{dx} \left(\frac{P'(x)}{P(x)} \right) = - \sum_{j=1}^m \frac{\alpha_j}{(x - \lambda_j)^2}.$$

1. On ne peut pas dire mieux : par exemple, pour $F = X^n/(X^n + 1)$ avec $n \in \mathbb{N}^*$, on a $\deg(F) = 0$ et $\deg(F') = -(n+1)$.

2. On peut aussi introduire le concept de dérivée d'une fraction rationnelle voir sujet 33 p. 195.

3. Voir sujet 12 p. 175.

On en déduit

$$P'^2(x) - P(x)P''(x) = \sum_{j=1}^m \alpha_j \left(\frac{P(x)}{x - \lambda_j} \right)^2 \geqslant 0.$$

Enfin, l'inégalité obtenue se généralise aux réels $\lambda_1, \dots, \lambda_m$ par continuité (ou, si l'on préfère, par un calcul direct).

Exercice 35 **

Soit P un polynôme réel unitaire scindé à racines simples x_1, \dots, x_n .

Calculer, pour tout $p \in [0; n - 1]$,

$$\sum_{k=1}^n \frac{x_k^p}{P'(x_k)}$$

Solution

Notons que les racines x_k étant simples, elles ne sont pas racines de P' et la quantité étudiée est bien définie.

méthode

On introduit la décomposition en éléments simples de

$$F = \frac{X^p}{P}.$$

La fraction F est de partie entière nulle. Elle n'est peut-être pas exprimée sous forme irréductible mais, quitte à autoriser la présence d'un terme nul, sa décomposition en éléments simples permet d'écrire

$$\frac{X^p}{P} = \sum_{k=1}^n \frac{a_k}{X - x_k} \quad \text{avec } a_k \in \mathbb{R}.$$

Soit $k \in [1; n]$. On obtient a_k en multipliant par $X - x_k$ puis en évaluant en x_k . Il revient au même de calculer la limite en x_k : ceci permet de faire apparaître¹ un taux d'accroissement

$$a_k = \lim_{x \rightarrow x_k} (x - x_k) \frac{X^p}{P(x)} = x_k^p \times \lim_{x \rightarrow x_k} \frac{x - x_k}{P(x) - P(x_k)} = \frac{x_k^p}{P'(x_k)}.$$

On étudie ensuite la limite de $xF(x)$ quand x tend vers $+\infty$. En comparant les degrés du numérateur et du dénominateur

$$xF(x) = \frac{x^{p+1}}{P(x)} \xrightarrow[x \rightarrow +\infty]{} \begin{cases} 0 & \text{si } p \neq n - 1 \\ 1 & \text{si } p = n - 1 \end{cases} \quad \text{et} \quad xF(x) = \sum_{k=1}^n \frac{a_k x}{x - x_k} \xrightarrow[x \rightarrow +\infty]{} \sum_{k=1}^n a_k.$$

On en déduit

$$\sum_{k=1}^n \frac{x_k^p}{P'(x_k)} = \begin{cases} 0 & \text{si } p \neq n - 1 \\ 1 & \text{si } p = n - 1. \end{cases}$$

¹ Une alternative efficace est d'employer la formule acquise dans le sujet 32 p. 194.

5.8 Exercices d'approfondissement

Exercice 36 * (Irrationalité de π)

On veut montrer que π est un nombre irrationnel. On raisonne par l'absurde et l'on suppose qu'il est possible d'écrire $\pi = a/b$ avec $a, b \in \mathbb{N}^*$. Pour $n \in \mathbb{N}$, on introduit alors

$$P_n = \frac{1}{n!} X^n (bX - a)^n \quad \text{et} \quad I_n = \int_0^\pi P_n(t) \sin t \, dt.$$

- (a) Montrer que P_n et ses polynômes dérivés successifs prennent des valeurs entières en 0.
- (b) Établir la même propriété en $\pi = a/b$.
- (c) Montrer que la suite (I_n) tend vers 0.
- (d) Conclure en observant que I_n est un nombre entier.

Solution

(a) 0 est racine de multiplicité n de P_n et donc $P_n^{(m)}(0) = 0$ pour tout $m < n$. Aussi, le polynôme P_n est de degré $2n$ et donc $P_n^{(m)}(0) = 0$ pour tout $m > 2n$. Reste à résoudre le cas $n \leq m \leq 2n$.

On développe l'expression de P_n par la formule du binôme de Newton

$$P_n = \sum_{k=0}^n \frac{1}{n!} \binom{n}{k} (-a)^{n-k} b^k X^{n+k}.$$

méthode

La valeur de $P^{(m)}(0)$ d'un polynôme P est liée au coefficient de X^m dans celui-ci : $P^{(m)}(0) = m! a_m$.

En considérant le coefficient de X^{n+k} pour $m = n + k$, il vient

$$P_n^{(m)}(0) = m! \times \frac{1}{n!} \binom{n}{m-n} (-a)^{2n-m} b^{m-n}.$$

Cette valeur est entière car a et b sont des entiers, les coefficients binomiaux sont des entiers et le quotient de factorielles $m!/n!$ aussi car $m \geq n$.

(b) On remarque la symétrie $P_n(\pi - X) = P_n(X)$. On obtient donc par dérivation à l'ordre $m \in \mathbb{N}$

$$P_n^{(m)}(\pi) = (-1)^m P_n^{(m)}(0) \in \mathbb{Z}.$$

- (c) Pour $n \in \mathbb{N}$, on obtient par l'inégalité triangulaire et par croissances comparées

de la factorielle et d'un terme géométrique

$$\begin{aligned}|I_n - 0| &= \frac{1}{n!} \int_0^\pi |t|^n |bt - a|^n |\sin t| dt \\&\leq \frac{1}{n!} \int_0^\pi \pi^n \times (b\pi + a)^n \times 1 dt = \pi \frac{(2a\pi)^n}{n!} \xrightarrow[n \rightarrow +\infty]{} 0.\end{aligned}$$

On en déduit que la suite (I_n) est de limite nulle.

(d) méthode

|| On exprime I_n par intégrations par parties dérivant le terme polynomial.

Une première intégration par parties avec les fonctions u et v de classe \mathcal{C}^1 déterminées par

$$u(t) = -\cos t = \sin\left(t - \frac{\pi}{2}\right) \quad \text{et} \quad v(t) = P_n(t)$$

donne

$$I_n = \left[\sin\left(t - \frac{\pi}{2}\right) P_n(t) \right]_0^\pi - \int_0^\pi \sin\left(t - \frac{\pi}{2}\right) P'_n(t) dt.$$

On répète¹ ces intégrations par parties jusqu'à disparition du terme polynomial

$$\begin{aligned}I_n &= \sum_{k=1}^{2n+1} \left[(-1)^{k-1} \sin\left(t - \frac{k\pi}{2}\right) P_n^{(k-1)}(t) \right]_0^\pi \\&\quad + (-1)^{2n+1} \int_0^\pi \sin\left(t - \frac{(2n+1)\pi}{2}\right) \underbrace{P_n^{(2n+1)}(t)}_{=0} dt.\end{aligned}$$

Les valeurs de P_n en 0 et π étant entières, l'intégrale I_n est un nombre entier. Cependant, la suite (I_n) tend vers 0 et donc, à partir d'un certain rang, I_n est constant égal à 0. Ceci est absurde car pour toute valeur de n , la fonction $t \mapsto P_n(t) \sin t$ est continue, positive, sans être la fonction nulle : son intégrale est strictement positive !

Exercice 37 **

Soit $P \in \mathbb{R}[X]$.

(a) On suppose $P(x) \geq 0$ pour tout réel x .

Montrer qu'il existe deux polynômes $A, B \in \mathbb{R}[X]$ tels que $P = A^2 + B^2$.

(b) On suppose $P(x) \geq 0$ pour tout réel $x \neq 0$.

Montrer qu'il existe deux polynômes $A, B \in \mathbb{R}[X]$ tels que $P = A^2 + XB^2$.

1. On intègre les fonctions trigonométriques $t \mapsto \sin(t+a)$ en $t \mapsto \sin(t+a-\pi/2)$.

Solution(a) **méthode**

On vérifie¹ que l'ensemble des polynômes s'écrivant $A^2 + B^2$ est stable par produit.

Pour $A, B, C, D \in \mathbb{R}[X]$, on constate par développement l'identité :

$$(A^2 + B^2)(C^2 + D^2) = (AC - BD)^2 + (AD + BC)^2.$$

Il suffit alors de vérifier que P est un produit de polynômes de la forme $A^2 + B^2$ pour pouvoir conclure. La décomposition en facteurs irréductibles de P dans $\mathbb{R}[X]$ s'écrit de façon générale

$$P = a \prod_{k=1}^m (X - \lambda_k)^{\alpha_k} \prod_{k=1}^p (X^2 + \mu_k X + \nu_k)^{\beta_k}$$

avec les conditions énoncées dans le Th. 26 p. 159, notamment $\Delta_k = \mu_k^2 - 4\nu_k < 0$ pour tout indice k .

Le signe de a détermine la limite de P en $+\infty$, il est nécessairement positif et l'on peut écrire $a = (\sqrt{a})^2 + 0^2$

Chaque exposant α_k est pair car sinon P change de signe en λ_k . On peut donc écrire :

$$(X - \lambda_k)^{\alpha_k} = ((X - \lambda_k)^{p_k})^2 + 0^2 \quad \text{avec } \alpha_k = 2p_k.$$

Enfin, l'écriture canonique de chaque facteur $X^2 + \mu_k X + \nu_k$ est de la forme voulue car ce trinôme est sans racines réelles :

$$X^2 + \mu_k X + \nu_k = \left(X + \frac{\mu_k}{2}\right)^2 + \left(\frac{\delta_k}{2}\right)^2 \quad \text{avec } \delta_k = \sqrt{-\Delta_k}.$$

Par produit de facteurs de la forme $A^2 + B^2$, P est aussi de cette forme.

(b) Pour $A, B, C, D \in \mathbb{R}[X]$, on vérifie par développement l'identité

$$(A^2 + XB^2)(C^2 + XD^2) = (AC - XBD)^2 + X(AD + BC)^2.$$

L'ensemble des polynômes s'écrivant $A^2 + XB^2$ est donc stable par produit. On raisonne alors comme au-dessus sachant que l'hypothèse de travail assure que les racines strictement positives sont d'ordre de multiplicité pairs. Il suffit ensuite d'exploiter les écritures :

$$\begin{aligned} (X - \lambda_k)^2 &= (X - \lambda_k)^2 + X \times 0^2 && \text{pour } \lambda_k > 0 \\ X - \lambda_k &= (\sqrt{-\lambda_k})^2 + X \times 1^2 && \text{pour } \lambda_k \leqslant 0 \end{aligned}$$

1. Une alternative possible est aussi d'organiser une factorisation de P dans $\mathbb{C}[X]$ sous la forme QQ et d'introduire A la partie réelle et B la partie imaginaire de Q .

et

$$X^2 + \mu_k X + \nu_k = (X - \sqrt{\nu_k})^2 + \underbrace{(\mu_k + 2\sqrt{\nu_k})X}_{>0 \text{ car } \Delta_k < 0}$$

Exercice 38 **

Soit P un polynôme réel scindé à racines simples de degré $n \geq 2$.

(a) Montrer que P ne peut pas posséder deux coefficients nuls successifs.

(b) Montrer que les coefficients situés de part et d'autre d'un coefficient nul de P ne peuvent être de même signe.

Solution

On introduit les coefficients $a_0, \dots, a_n \in \mathbb{R}$ de $P : P = a_0 + a_1X + \dots + a_nX^n$.

(a) méthode

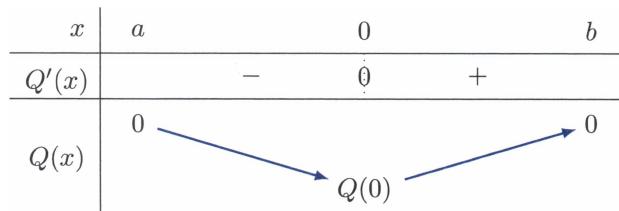
|| Lorsqu'un polynôme réel est scindé à racines simples, ses polynômes dérivés le sont aussi¹.

Par l'absurde, supposons qu'il existe $k < n - 1$ tel que $a_k = a_{k+1} = 0$. Par la formule de Taylor, on a $P^{(k)}(0) = k!a_k = 0$ et de même $P^{(k+1)}(0) = 0$. Le polynôme $P^{(k)}$ admet 0 pour racine double ce qui est absurde puisque ses racines doivent être simples.

(b) méthode

|| Lorsqu'un polynôme réel est scindé à racines simples, ses racines séparent celles de son polynôme dérivé.

Supposons qu'il existe un indice $k \in [1 ; n - 1]$ tel que $a_k = 0$ et $a_{k+1} \neq 0$. Quitte à considérer $-P$, on peut aussi supposer $a_{k+1} > 0$. Considérons alors le polynôme dérivé $Q = P^{(k-1)}$. Par la formule de Taylor, on a $Q'(0) = k!a_k = 0$ et $Q''(0) = (k+1)!a_{k+1} > 0$: le polynôme Q admet un minimum local en 0. Or 0 est racine de Q' et celle-ci est donc encadrée par deux racines de Q sans autres racines de Q' intermédiaires. En notant $a < b$ ces deux racines, on a les variations



On en déduit $Q(0) < 0$ et donc² $a_{k-1} < 0$.

1. Voir sujet 19 p. 180.

2. Plus généralement, on peut montrer $a_{k-1}a_{k+1} < 0$ grâce au sujet 34 p. 196.

Exercice 39 ***

Soit A, B deux polynômes complexes et non constants vérifiant

$$\begin{aligned}\{z \in \mathbb{C} \mid A(z) = 0\} &= \{z \in \mathbb{C} \mid B(z) = 0\} \text{ et} \\ \{z \in \mathbb{C} \mid A(z) = 1\} &= \{z \in \mathbb{C} \mid B(z) = 1\}.\end{aligned}$$

Montrer que $A = B$.

Solution

Quitte à échanger les deux polynômes, on peut supposer $n = \deg(A) \geq \deg(B)$.

méthode

On montre que l'ensemble $\{z \in \mathbb{C} \mid A(z) = 0\} \cup \{z \in \mathbb{C} \mid A(z) = 1\}$ possède au moins $n + 1$ éléments.

Soit p le nombre de racines distinctes de l'équation $A(z) = 0$. Puisque la somme des multiplicités des racines de A vaut n , ces racines sont susceptibles d'être racines du polynôme A' avec une somme de multiplicités égale à $n - p$. Le polynôme A' étant de degré $n - 1$, le polynôme A' possède alors exactement $p - 1$ autres racines comptées avec multiplicité.

Soit q le nombre de racines distinctes de l'équation $A(z) = 1$. Comme au-dessus, la somme des multiplicités de celles-ci en tant que racines du polynôme $A' = (A - 1)'$ vaut $n - q$. Cependant, ces racines ne correspondent pas à celles de A et on peut donc affirmer $n - q \leq p - 1$ ce qui entraîne $p + q \geq n + 1$.

On peut alors aisément conclure, le polynôme $A - B$ est de degré au plus n et s'annule au moins $n + 1$ fois : c'est le polynôme nul.

Exercice 40 * (Équation de Fermat polynomiale)**

(a) Soit P un polynôme complexe non nul. Montrer que le nombre p de ses racines distinctes vérifie :

$$p = \deg(P) - \deg(P \wedge P').$$

(b) Soit P, Q deux polynômes complexes premiers entre eux et vérifiant

$$R = P + Q \text{ est non constant.}$$

On note p, q et r le nombre de racines distinctes des polynômes P, Q et R . En introduisant le polynôme $P'Q - Q'P$, vérifier

$$\deg(R) < p + q + r.$$

(c) Soit $n \in \mathbb{N}$ avec $n \geq 3$. Déterminer les triplets de polynômes complexes (P, Q, R) tels que

$$P^n + Q^n = R^n.$$

Solution**(a) méthode**

|| Les racines du PGCD de P et P' correspondent aux racines multiples de P .

Notons $\lambda_1, \dots, \lambda_p$ les racines de P et $\alpha_1, \dots, \alpha_p$ leurs multiplicités respectives. En introduisant a le coefficient dominant de P , on écrit

$$P = a \prod_{j=1}^p (X - \lambda_j)^{\alpha_j} \quad \text{et} \quad P \wedge P' = \prod_{j=1}^p (X - \lambda_j)^{\alpha_j - 1}.$$

On constate alors

$$\deg(P) - \deg(P \wedge P') = \sum_{j=1}^p \alpha_j - \sum_{j=1}^p (\alpha_j - 1) = p.$$

(b) Commençons par souligner que les polynômes P , Q et R sont deux à deux premiers entre eux puisque, si un polynôme divise deux d'entre eux, il divise le troisième.

Le PGCD de P et P' divise le polynôme $S = P'Q - PQ'$. Il en est de même pour le PGCD de Q et Q' . Aussi, on peut écrire

$$S = P'(R - P) - P(R' - P') = P'R - PR'$$

et affirmer que le PGCD de R et R' divise S . Ces trois PGCD étant deux à deux premiers entre eux¹, on peut écrire

$$(P \wedge P')(Q \wedge Q')(R \wedge R') \mid S. \tag{*}$$

Or le polynôme S n'est pas nul. En effet, si $S = 0$ on a $(P/R)' = 0$ et les polynômes non constants P et R ne sont pas premiers entre eux.

La divisibilité (*) avec $S \neq 0$ donne

$$\deg(P \wedge P') + \deg(Q \wedge Q') + \deg(R \wedge R') \leq \deg(S) < \deg(P) + \deg(Q).$$

On simplifie sachant $\deg(P \wedge P') = \deg(P) - p$ et les relations analogues pour Q et R :

$$\deg(R) < p + q + r.$$

(c) Soit P , Q et R trois polynômes complexes vérifiant $P^n + Q^n = R^n$. Si a est une racine commune aux polynômes P et Q , a est aussi racine de R . De même, a est racine de Q lorsque a est racine commune à P et R , etc. En simplifiant les racines communes, on peut donc se ramener à une situation où les trois polynômes complexes P , Q et R sont deux à deux premiers entre eux. Il en est alors de même de P^n , Q^n et R^n .

1. En effet, ces PGCD divisent respectivement P , Q et R qui sont deux à deux premiers entre eux.

méthode

On montre que les trois polynômes sont alors constants en organisant l'identité $P^n + Q^n = R^n$ pour que polynôme du second membre soit celui de plus haut degré.

Quitte à opérer des passages à l'opposé¹, on peut permute les polynômes P , Q et R de sorte que

$$\max(\deg(P), \deg(Q)) \leq \deg(R) \quad \text{et} \quad P^n + Q^n = R^n.$$

Par l'absurde, si R n'est pas constant, l'étude qui précède donne

$$\deg(R^n) < p + q + r$$

en notant comme au-dessus p, q, r le nombre de racines distinctes des polynômes P, Q, R (et donc des polynômes P^n, Q^n, R^n). Or on sait aussi que le nombre de racines distinctes d'un polynôme est inférieur à son degré et donc

$$n \deg(R) < \deg(P) + \deg(Q) + \deg(R) \leq 3 \deg(R) \quad \text{avec } n \geq 3 \text{ et } \deg(R) > 0.$$

C'est absurde. On peut alors conclure que les trois polynômes P , Q et R sont constants.

Finalement, les solutions (P, Q, R) de l'équation $P^n + Q^n = R^n$ sont les triplets $(\alpha T, \beta T, \gamma T)$ avec $T \in \mathbb{C}[X]$ et $\alpha, \beta, \gamma \in \mathbb{C}$ vérifiant $\alpha^n + \beta^n = \gamma^n$.

1. Un passage à l'opposé peut être intégré à la puissance n en écrivant $-1 = (\mathrm{e}^{i\pi/n})^n$.

CHAPITRE 6

Dénombrément

6.1 Cardinal d'un ensemble fini

6.1.1 Définition

Définition

On dit qu'un ensemble E est *fini* s'il est en bijection avec l'ensemble¹ $\llbracket 1 ; n \rrbracket$ pour une certaine valeur de $n \in \mathbb{N}$. Ce naturel n est alors unique et s'appelle le *cardinal* de E , on le note $\text{Card}(E)$ ou $|E|$.

Lorsqu'un ensemble E n'est pas fini, on le dit *infini* et l'on écrit $\text{Card}(E) = +\infty$.

L'ensemble $\llbracket a ; b \rrbracket$ est fini de cardinal $b - a + 1$.

Si un ensemble E est fini de cardinal n , on peut introduire une bijection $\varphi: \llbracket 1 ; n \rrbracket \rightarrow E$. En posant $x_i = \varphi(i)$ pour tout $i \in \llbracket 1 ; n \rrbracket$, on peut alors écrire

$$E = \{x_1, \dots, x_n\} \quad \text{avec} \quad x_1, \dots, x_n \text{ sans répétition.}$$

Définition

La famille finie (x_1, \dots, x_n) constitue une *énumération* de l'ensemble E .

1. On convient que $\llbracket 1 ; n \rrbracket$ désigne l'ensemble vide lorsque $n = 0$.

6.1.2 Cardinal d'une réunion

Théorème 1

Si A et B sont deux ensembles finis, leur union et leur intersection le sont aussi et

$$\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B) - \text{Card}(A \cap B).$$

Si les ensembles A et B sont disjoints, $\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B)$.

Sans déterminer $A \cap B$, on peut affirmer $\text{Card}(A \cup B) \leq \text{Card}(A) + \text{Card}(B)$. Cette comparaison se généralise par récurrence à une union $A_1 \cup \dots \cup A_p$ d'ensembles finis :

$$\text{Card}\left(\bigcup_{j=1}^p A_j\right) \leq \sum_{j=1}^p \text{Card}(A_j).$$

6.1.3 Cardinal d'une partie

Théorème 2

Toute partie A d'un ensemble fini E est finie et $\text{Card}(A) \leq \text{Card}(E)$.

De plus, $\text{Card}(A) = \text{Card}(E)$ si, et seulement si, $A = E$.

Ajoutons aussi $\text{Card}(\complement_E A) = \text{Card}(E) - \text{Card}(A)$.

6.1.4 Applications entre deux ensembles finis

Théorème 3

Soit $f: E \rightarrow F$ une application opérant entre deux ensembles.

- a) si f est injective et si F est fini alors E est fini et $\text{Card}(E) \leq \text{Card}(F)$;
- b) si f est surjective et si E est fini alors F est fini et $\text{Card}(E) \geq \text{Card}(F)$;
- c) si f est bijective et si l'un des ensembles est fini alors l'autre l'est aussi et $\text{Card}(E) = \text{Card}(F)$;

Par la dernière propriété, il est fréquent de montrer qu'un ensemble est fini tout en calculant son cardinal, en déterminant une bijection entre cet ensemble et un ensemble fini de cardinal connu.

Pour qu'il existe une bijection entre deux ensembles finis, il est nécessaire que ceux-ci aient le même cardinal. Lorsque l'on sait cette condition remplie, on peut caractériser avec efficacité une bijection :

Théorème 4

Si une application entre deux ensembles finis E et F vérifiant $\text{Card}(E) = \text{Card}(F)$ est injective (resp. surjective) alors elle est bijective.

6.2 Cardinaux usuels

6.2.1 Produit cartésien d'ensembles finis

Théorème 5

Si E et F sont deux ensembles finis, le produit cartésien $E \times F$ est fini et

$$\text{Card}(E \times F) = \text{Card}(E) \times \text{Card}(F).$$

Ce résultat s'étend par récurrence à un produit cartésien $E_1 \times \cdots \times E_p$ d'ensembles finis :

$$\text{Card}\left(\prod_{j=1}^p E_j\right) = \prod_{j=1}^p \text{Card}(E_j).$$

6.2.2 Ensemble des parties d'un ensemble fini

Théorème 6

Si E est un ensemble fini, l'ensemble $\wp(E)$ des parties de E est fini et

$$\text{Card}(\wp(E)) = 2^{\text{Card}(E)}.$$

6.2.3 Ensemble des applications d'un ensemble fini vers un autre

Théorème 7

Si E et F sont deux ensembles finis, l'ensemble $\mathcal{F}(E, F) = F^E$ des applications de E vers F est fini et

$$\text{Card}(F^E) = (\text{Card}(F))^{\text{Card}(E)}.$$

Parmi les applications de E vers F , on sait facilement¹ dénombrer les injections :

Théorème 8

Si E et F sont deux ensembles finis avec $p = \text{Card}(E) < n = \text{Card}(F)$, l'ensemble des injections E dans F a pour cardinal

$$n \times (n - 1) \times \cdots \times (n - p + 1) = \frac{n!}{(n - p)!}.$$

Si $\text{Card}(E) > \text{Card}(F)$, rappelons qu'il n'existe pas d'injections de E dans F .

1. Dénombrer les surjections est l'objet du sujet 19 p. 223.

6.2.4 Ensemble des permutations d'un ensemble fini

Si E est un ensemble fini, une injection de E dans E est assurément une permutation :

Théorème 9

Si E est un ensemble fini, l'ensemble S_E des permutations de E est fini et

$$\text{Card}(S_E) = (\text{Card}(E))!$$

6.3 Listes, arrangements, combinaisons

E désigne un ensemble fini à n éléments et p désigne un entier naturel.

6.3.1 Nombre de p -listes

Définition

- || On appelle *liste* de longueur p formée d'éléments de E le choix ordonné d'éléments de E avec possibilité de répétitions de ceux-ci.

Une p -liste s'identifie immédiatement à un p -uplet (x_1, \dots, x_p) élément de E^p .

Théorème 10

Il existe exactement n^p listes de longueur p d'éléments d'un ensemble de cardinal n .

6.3.2 Nombre de p -arrangements

Définition

- || On appelle *arrangement* de longueur p formé d'éléments de E le choix ordonné et sans répétition de p éléments de E .

Un p -arrangement s'identifie à une injection de $\llbracket 1 ; p \rrbracket$ dans E et par conséquent :

Théorème 11

Lorsque $p \leq n$, il existe $n \times (n - 1) \times \cdots \times (n - p + 1)$ arrangements de longueur p d'éléments d'un ensemble de cardinal n .

Si $p > n$, de tels arrangements ne peuvent exister et la formule ci-dessus est encore valable car un facteur nul apparaît dans le produit.

6.3.3 Nombre de p -combinaisons

Définition

- || On appelle *combinaison* de longueur p formée d'éléments de E le choix non ordonné et sans répétition de p éléments de E .

Une p -combinaison s'identifie à une partie à p éléments de E .

Théorème 12

Lorsque $p \leq n$, il existe $\binom{n}{p}$ combinaisons de longueur p d'éléments d'un ensemble à n éléments.

Si $p > n$, de telles combinaisons ne peuvent exister et le résultat ci-dessus demeure valable car le coefficient binomial est nul dans ce cas¹.

6.4 Exercices d'apprentissage

6.4.1 Généralités

Exercice 1

Soit A , B et C trois parties d'un ensemble fini E . Exprimer $\text{Card}(A \cup B \cup C)$ en fonction des cardinaux de A , B , C , $A \cap B$, $B \cap C$, $C \cap A$ et $A \cap B \cap C$.

Solution

méthode

|| On sait exprimer le cardinal d'une union de deux ensembles finis (Th. 1 p. 206).

On considère $A \cup B \cup C$ comme l'union de A et de $B \cup C$:

$$\begin{aligned}\text{Card}(A \cup B \cup C) &= \text{Card}(A \cup (B \cup C)) \\ &= \text{Card}(A) + \text{Card}(B \cup C) - \text{Card}(A \cap (B \cup C))\end{aligned}$$

Par distributivité de l'intersection sur l'union

$$\text{Card}(A \cup B \cup C) = \text{Card}(A) + \text{Card}(B \cup C) - \text{Card}((A \cap B) \cup (A \cap C)).$$

On poursuit avec

$$\text{Card}((A \cap B) \cup (A \cap C)) = \text{Card}(A \cap B) + \text{Card}(A \cap C) - \text{Card}(A \cap B \cap C)$$

et l'on conclut

$$\begin{aligned}\text{Card}(A \cup B \cup C) &= \text{Card}(A) + \text{Card}(B) + \text{Card}(C) \\ &\quad - \text{Card}(B \cap C) - \text{Card}(A \cap B) - \text{Card}(C \cap A) \\ &\quad + \text{Card}(A \cap B \cap C).\end{aligned}$$

Exercice 2

Soit $f: E \rightarrow F$ une application au départ d'un ensemble fini E et à valeurs dans un ensemble F . Montrer

$$f \text{ est injective} \iff \text{Card}(f(E)) = \text{Card}(E).$$

1. Cependant, l'expression factorielle du coefficient binomial $\binom{n}{p}$ ne peut être utilisée quand $p > n$.

Solution**méthode**

La restriction d'une application à l'arrivée dans son image est surjective et même bijective lorsque l'application est injective¹.

(\implies) Si l'application f est injective, elle induit une bijection entre les ensembles E et $\text{Im}(f) = f(E)$. On en déduit que l'ensemble $f(E)$ est fini et de même cardinal que E (Th. 3 p. 206).

(\impliedby) Inversement, supposons $\text{Card}(f(E)) = \text{Card}(E)$. La restriction f' de l'application f au départ de E et à l'arrivée dans $f(E)$ est surjective. Les ensembles E et $f(E)$ étant de cardinaux finis et égaux, c'est une bijection (Th. 4 p. 206). On en déduit que f' est injective et donc f aussi.

6.4.2 Dénombrements**Exercice 3**

Soit $n, p \in \mathbb{N}$. Combien existe-t-il de couples $(x, y) \in \llbracket -p ; n \rrbracket^2$ vérifiant $xy > 0$?

Solution**méthode**

Pour dénombrer un ensemble fini, on peut mettre cet ensemble en bijection avec un ensemble fini de cardinal connu, par exemple en décrivant ses éléments. On peut aussi opérer par réunion ou produit cartésien d'ensembles finis.

Souvent ces démarches se traduisent par la construction algorithmique des éléments de l'ensemble sachant que :

- on multiple les possibilités, lorsque l'on passe d'une étape à l'étape suivante dans la construction ;
- on somme les possibilités, lorsqu'il y a une alternative stricte dans la construction.

La valeur de x peut être nulle, strictement positive ou strictement négative.

Cas : $x = 0$. Toute valeur de y convient ce qui offre $n + p + 1$ choix².

Cas : $x > 0$. Il y a n choix pour la valeur de x et $n + 1$ pour la valeur de y . Au total, cela produit $n(n + 1)$ couples (x, y) solutions vérifiant la condition $x > 0$.

Cas : $x < 0$. C'est analogue et l'on obtient $p(p + 1)$ couples.

Les différentes alternatives s'excluant mutuellement³, le nombre de couples cherché vaut⁴

$$n + p + 1 + n(n + 1) + p(p + 1) = n^2 + p^2 + 2n + 2p + 1 = (n + p + 1)^2 - 2np.$$

1. Voir sujet 19 p. 31.

2. Le cardinal de $\llbracket a ; b \rrbracket$ est $b - a + 1$ et non $b - a$.

3. On peut aisément formaliser : l'ensemble des couples cherché est la réunion disjointe des ensembles $\{0\} \times \llbracket -p ; n \rrbracket$, $\llbracket 1 ; n \rrbracket \times \llbracket 0 ; n \rrbracket$ et $\llbracket -p ; -1 \rrbracket \times \llbracket -p ; 0 \rrbracket$ dont on vient de calculer les cardinaux respectifs.

4. L'expression finale peut aussi être comprise comme étant le nombre de couples (x, y) possibles dont on a retiré ceux formés d'un élément strictement positif et d'un strictement négatif.

Exercice 4

Une urne contient n jetons numérotés de 1 à n (avec $n \geq 2$).

- (a) On tire successivement et avec remise 2 jetons dans l'urne. Combien de tirages sont possibles ? Pour combien d'entre eux le second jeton est-il d'une valeur au moins égale au premier ?
- (b) Mêmes questions pour un tirage sans remise.
- (c) On tire simultanément 2 jetons dans l'urne. Combien de tirages sont possibles ? Pour combien d'entre eux la somme des valeurs vaut n ?

Solution**méthode**

Etudier un tirage, ou un problème qui s'y rapporte, conduit souvent à dénombrer des listes (tirage ordonné avec remise), des arrangements (tirage ordonné sans remise) ou des combinaisons (tirage simultané donc non ordonné).

On note E l'ensemble constitué des jetons. On peut identifier E et $\llbracket 1 ; n \rrbracket$.

- (a) Le nombre de tirages possibles correspond au nombre de listes formées de deux éléments de E : il y en a n^2 .

Si le premier tirage a pour valeur k (avec $k \in \llbracket 1 ; n \rrbracket$), il existe $n - k + 1$ seconds tirages possibles de valeurs supérieures (à savoir $k, k + 1, \dots, n$). Le nombre¹ de tirages cherché est donc

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

- (b) Le nombre de tirages possibles correspond au nombre d'arrangements de deux éléments de E : il y en a $n(n - 1)$.

Si le premier tirage a pour valeur k (avec $k \in \llbracket 1 ; n \rrbracket$), il existe $n - k$ seconds tirages possibles de valeurs supérieures (à savoir $k + 1, \dots, n$). Le nombre de tirages cherché est donc

$$\sum_{k=1}^n (n - k) = n^2 - \sum_{k=1}^n k = \frac{n(n-1)}{2}$$

- (c) Le nombre de tirages possibles correspond au nombre de combinaisons de deux éléments de E : il y en a $\binom{n}{2} = \frac{n(n-1)}{2}$.

Les tirages dont la somme des valeurs vaut n sont constitués d'éléments k et ℓ différents avec $k + \ell = n$. Afin de les distinguer, on peut supposer $k < \ell$ auquel cas $k < n - k$ et donc $1 \leq k < n/2$. La connaissance de k suffit à déterminer la paire $\{k, \ell\}$ avec $k + \ell = n$ et $k < \ell$: le nombre de tirages cherché vaut donc $n/2 - 1$ si n est pair et $(n - 1)/2$ si n est impair.

1. En identifiant l'ensemble des tirages à E^2 , ceux cherchés sont les éléments de la réunion des ensembles deux à deux disjoints $A_k = \{(k, p) \mid k < p \leq n\}$ pour $k \in \llbracket 1 ; n \rrbracket$: cet argument offre une justification rigoureuse du calcul proposé au prix d'un formalisme un peu excessif.

Exercice 5

Cinq cartes d'un jeu de trente deux cartes constituent la main d'un joueur.

- Combien de mains comportent un As ?
- Combien de mains comportent au moins un As ?
- Combien de mains comportent un As et un Cœur ?
- Combien de mains comportent un As ou un Cœur ?
- Combien de mains comportent au moins un As et au moins un Cœur ?

Solution

Une main s'apparente à une combinaison de 5 éléments dans un ensemble à 32 éléments. Il y a $\binom{32}{5} = 201\,376$ mains possibles. Dans les études qui suivent, les pourcentages indiquent la proportion du nombre de mains solutions.

(a) On choisit la couleur de l'As puis on complète la main avec 4 cartes choisies parmi les 28 qui ne sont pas des As. Cela offre $\binom{4}{1} \times \binom{28}{4} = 81\,900$ possibilités (41 %).

(b) méthode

Il est plus commode de déterminer le nombre de mains qui ne comportent pas d'As puis de calculer le cardinal d'un complémentaire.

On forme une main qui ne comporte pas d'As en choisissant 5 cartes parmi les 28 possibles. Il y a donc $\binom{32}{5} - \binom{28}{5} = 103\,096$ mains comportant au moins un As (51 %).

(c) méthode

On distingue les mains comportant l'As de Cœur des autres.

Pour former une main convenable contenant l'As de Cœur on choisit 4 cartes parmi les 21 cartes qui ne sont ni des As, ni des Cœur. Cela donne $\binom{21}{4}$ possibilités.

Pour former une main convenable ne comportant pas l'As de Cœur, on choisit la couleur de l'As parmi les 3 possibles, on choisit la valeur du Cœur parmi les 7 qui ne correspondent pas à l'As de Cœur et l'on complète la main avec 3 cartes choisies parmi les 21 qui ne sont ni des As, ni des Cœur. Cela donne $3 \times 7 \times \binom{21}{3}$.

Au total, il a $\binom{21}{4} + 3 \times 7 \times \binom{21}{3} = 33\,915$ possibilités (17 %).

(d) méthode

On exploite la formule $\text{Card}(A \cup B) = \text{Card}(A) + \text{Card}(B) - \text{Card}(A \cap B)$.

On note A l'ensemble des mains comportant un As et B celui des mains comportant un Cœur. Ci-dessus on a dénombré A et $A \cap B$ et il est aussi facile de dénombrer B : $\text{Card}(B) = 8 \times \binom{24}{4}$. On en déduit $\text{Card}(A \cup B) = 132\,993$ (66 %).

(e) Les mains sans As ni Cœur sont constituées de 5 cartes choisies parmi les 21 qui ne sont ni des As ni des Cœur, il y en a $\binom{21}{5}$. Les mains sans As sont au nombre de $\binom{28}{5}$ et

celle sans Cœur sont au nombre de $\binom{24}{5}$. Il y a donc $\binom{28}{5} + \binom{24}{5} - \binom{21}{5}$ mains ne comportant pas d'As ou ne comportant pas de Cœur. Par passage au complémentaire on obtient

$$\binom{32}{5} - \binom{28}{5} - \binom{24}{5} + \binom{21}{5} = 80941$$

mains comportant au moins un As et au moins un Cœur (40 %).

6.5 Exercices d'entraînement

6.5.1 Démonstrations combinatoires

Exercice 6 *

Soit n et $p \in \mathbb{N}$. Proposer des preuves combinatoires des formules :

$$(a) \sum_{k=0}^n \binom{n}{k} = 2^n \quad (b) \binom{n}{n-p} = \binom{n}{p} \quad (c) \binom{n}{p} + \binom{n}{p+1} = \binom{n+1}{p+1}.$$

Solution

Soit E un ensemble à n éléments.

(a) Il y a 2^n parties dans E (Th. 6 p. 207). Celles-ci ont un cardinal k compris entre 0 et n et il y a exactement $\binom{n}{k}$ parties k éléments dans E . On en tire l'identité

$$\sum_{k=0}^n \binom{n}{k} = 2^n$$

(b) méthode

On raisonne par passage au complémentaire.

L'application qui à une partie A d'un ensemble E associe son complémentaire est bijective¹ et échange les parties à p éléments avec celles à $n - p$ éléments. Il y a donc autant de parties à p éléments que de parties à $n - p$ éléments.

$$\binom{n}{n-p} = \binom{n}{p},$$

(c) On introduit un élément x qui n'appartient pas à E .

méthode

On dénombre les parties à $p+1$ de l'ensemble $E' = E \cup \{x\}$ en discutant selon que x en est ou non élément.

Il y a $\binom{n+1}{p+1}$ parties à $p+1$ dans l'ensemble E' . Parmi celles-ci, celles qui contiennent x sont constituées en outre de p éléments choisis dans E , il y en a $\binom{n}{p}$. Celles qui ne

1. Il s'agit d'une involution : sa bijection réciproque est elle-même.

qui contiennent pas x sont constituées de $p+1$ éléments choisis dans E , il y en a $\binom{n}{p+1}$. On en déduit

$$\binom{n+1}{p+1} = \binom{n}{p} + \binom{n}{p+1}.$$

Exercice 7 **

Soit E un ensemble à n éléments et p un entier avec $1 \leq p \leq n$. En dénombrant les couples (A, x) constitués d'une partie A de E à p éléments et d'un élément x de A , établir l'identité

$$\binom{n}{p} = \frac{n(n-1)}{p(p-1)}.$$

Solution

méthode

On dénombre les couples (A, x) de deux façons, l'une en commençant par choisir A , l'autre en choisissant x .

Il y a $\binom{n}{p}$ parties A à p éléments possibles et pour chacune il y a p choix de l'élément x . Il y a donc un total de $p\binom{n}{p}$ couples (A, x) possibles.

En toute généralité, l'élément x appartient à E , il y a n possibilités pour le choisir. Une fois ce choix fait, on forme la partie A en complétant x par le choix de $p-1$ éléments pris parmi les $n-1$ constituant $E \setminus \{x\}$. Cela offre $\binom{n-1}{p-1}$ possibilités pour compléter x . Il y a donc aussi un total de $n\binom{n-1}{p-1}$ couples (A, x) possibles. On peut alors conclure

$$p\binom{n}{p} = n\binom{n-1}{p-1}.$$

6.5.2 Dénombrements

Exercice 8 *

En écriture binaire, combien de fois utilise-t-on le chiffre « 1 » pour énumérer tous les entiers compris entre 1 et 1 024 ?

Solution

méthode

On forme une relation de récurrence déterminant le nombre u_n de chiffres « 1 » nécessaires à l'écriture des entiers compris entre 0 et $2^n - 1$.

Calculons les premières valeurs pour percevoir le mécanisme.

n	Entiers compris entre 0 et $2^n - 1$	u_n
1	0, 1	1
2	0, 1, 10, 11	4
3	0, 1, 10, 11, 100, 101, 110, 111	12

Soit $n \in \mathbb{N} \setminus \{0, 1\}$. Les écritures binaires des 2^n entiers compris entre 2^n et $2^{n+1} - 1$ s'obtiennent en devançant par 1 les écritures binaires des entiers compris entre 0 et 2^n . On a donc

$$u_{n+1} = \underbrace{\dots}_{\substack{\text{nombre de 1 pour} \\ \text{l'écriture des entiers} \\ \text{allant de } 0 \text{ à } 2^n - 1}} + \underbrace{2^n}_{\substack{\text{nombre de 1 en tête} \\ \text{de l'écriture des entiers} \\ \text{allant de } 2^n \text{ à } 2^{n+1} - 1}} + \underbrace{\dots}_{\substack{\text{nombre de 1 suivant} \\ \text{dans l'écriture des entiers} \\ \text{allant de } 2^n \text{ à } 2^{n+1} - 1}} = 2u_n + 2^n.$$

Cette relation de récurrence invite à introduire (v_n) de terme général $v_n = u_n/2^{n-1}$. On vérifie $v_{n+1} = v_n + 1$ puis $v_n = n$ pour tout $n > 1$.

Finalement, $u_n = n2^{n-1}$ et il ne reste plus qu'à concrétiser pour la valeur $n = 10$ en traitant le nombre 2^{10} séparément. On obtient $10 \cdot 2^9 + 1 = 5121$ occurrences du chiffre « 1 ».

Exercice 9 **

Soit $n \in \mathbb{N}$, $p \in \mathbb{N}^*$ et $E = [[1; n]]$.

- (a) Dénombrer les suites strictement croissantes (x_1, \dots, x_p) d'éléments de E
- (b) Dénombrer les suites croissantes (x_1, \dots, x_p) d'éléments de E .

Solution

(a) Une suite strictement croissante est constituée d'éléments deux à deux distincts. Si $n < p$, une telle suite ne peut exister. On suppose pour poursuivre $p \leq n$.

méthode

Une suite strictement croissante est entièrement déterminée par ses valeurs qu'il suffit d'ordonner.

Une suite strictement croissante $x = (x_1, \dots, x_p)$ d'éléments de E définit une partie $A = \{x_1, \dots, x_p\}$ à p éléments dans E . Inversement, si l'on se donne une partie A de E à p éléments, il existe une et une seule façon d'ordonner ceux-ci pour constituer une suite strictement croissante $x = (x_1, \dots, x_p)$ associée :

$$x_1 = \min A, x_2 = \min(A \setminus \{x_1\}), x_3 = \min(A \setminus \{x_1, x_2\}), \text{etc.}$$

Il existe donc autant¹ de suites (x_1, \dots, x_p) strictement croissantes d'éléments de E que de parties à p éléments dans E , c'est-à-dire $\binom{n}{p}$.

(b) méthode

On transforme une suite croissante en une suite strictement croissante en incrémentant les valeurs prises.

1. On vient d'observer que l'application qui à une suite strictement croissante associe l'ensemble de ses valeurs réalise une bijection entre l'ensemble à dénombrer et l'ensemble des parties à p éléments de E .

À une suite croissante $x = (x_1, \dots, x_p)$ d'éléments de E , on associe par incrémentation de chaque valeur une suite $y = (y_1, \dots, y_p)$ déterminée par

$$y_1 = x_1, y_2 = x_2 + 1, y_3 = x_3 + 2, \dots, y_p = x_p + (p - 1).$$

La suite y est une suite strictement croissante d'éléments de $F = [1; n + p - 1]$. Inversement, une telle suite y correspond à une et une seule suite x croissante d'éléments de E . Par cette correspondance bijective, on peut affirmer qu'il y autant de suites croissantes d'éléments de E que de suites strictement croissantes d'éléments de F , c'est-à-dire $\binom{n+p-1}{p}$.

Exercice 10 ** (Anagrammes)

Un mot M long de n lettres est constitué de r lettres différentes. La j -ème lettre apparaît p_j fois dans le mot M et donc $p_1 + \dots + p_r = n$.

(a) Combien d'anagrammes différentes du mot M peut-on écrire ?

(b) Dans le développement de $(a_1 + a_2 + \dots + a_r)^n$, quel est le coefficient du terme $a_1^{p_1} a_2^{p_2} \dots a_r^{p_r}$?

Solution

(a) **méthode**

Pour construire une anagramme du mot M , on choisit p_1 emplacements où positionner la première lettre, puis p_2 emplacements parmi ceux restants, etc.

Choisir les emplacements où figurent la première lettre revient à déterminer une partie à p_1 éléments (les positions choisies) dans un ensemble à n éléments (les positions possibles). Il y a donc $\binom{n}{p_1}$ possibilités pour choisir les emplacements de la première lettre. La seconde lettre ne peut alors figurer que parmi les $n - p_1$ emplacements restants et il y a $\binom{n-p_1}{p_2}$ possibilités pour positionner celle-ci. Ainsi de suite, on obtient $\binom{n-(p_1+\dots+p_{j-1})}{p_j}$ possibilités pour positionner la j -ème lettre quand les précédentes ont pris place.

Au final, il y a

$$\binom{n}{p_1} \times \binom{n-p_1}{p_2} \times \dots \times \underbrace{\binom{n-(p_1+\dots+p_{r-1})}{p_r}}_{=1} \text{ anagrammes possibles.}$$

En exploitant l'écriture factorielle des coefficients binomiaux, on simplifie significativement cette expression¹ et l'on obtient

$$\frac{n!}{p_1! p_2! \dots p_r!} \text{ anagrammes possibles.}$$

1. On peut trouver aussi directement cette expression par le raisonnement suivant : si l'on distingue toutes les lettres du mot M , par exemple en les numérotant, former une anagramme revient à permuter les lettres, il y a $n!$ possibilités (Th. 9 p. 208). Si l'on ne distingue plus les lettres, à chacune des permutations précédentes on fait correspondre la même anagramme en permutant entre elles les premières lettres ($p_1!$ possibilités) et de même pour les autres : à chaque permutation correspond $p_1! \times \dots \times p_r!$ anagrammes identiques.

(b) méthode

On commence par étudier le cas $r = 3$ plus facilement maîtrisable que le cas général sans pour autant être aussi commun que le cas $r = 2$.

Lorsque l'on développe le produit

$$(a_1 + a_2 + a_3)^n = (a_1 + a_2 + a_3)(a_1 + a_2 + a_3) \dots (a_1 + a_2 + a_3)$$

sans combiner les facteurs sous forme de puissances, on écrit tous les mots possibles de longueur n exprimés avec trois lettres différentes. Par exemple, pour $n = 2$, ce sont les termes :

$$a_1a_1, a_1a_2, a_1a_3, a_2a_1, a_2a_2, a_2a_3, a_3a_1, a_3a_2 \text{ et } a_3a_3.$$

Plus généralement, le développement de $(a_1 + a_2 + \dots + a_r)^n$ fait apparaître tous les mots possibles de longueur n écrits avec r lettres différentes. Lorsque l'on combine les facteurs sous forme de puissances, ce sont tous les anagrammes écrits avec p_j fois a_j qui génèrent le terme $a_1^{p_1}a_2^{p_2} \dots a_r^{p_r}$. Le coefficient de ce terme est donc le *coefficient multinomial*

$$\binom{n}{p_1, p_2, \dots, p_r} = \frac{n!}{p_1!p_2!\dots p_r!}.$$

On pourra retenir la *formule du multinôme*

$$(a_1 + a_2 + \dots + a_r)^n = \sum_{p_1+p_2+\dots+p_r=n} \binom{n}{p_1, p_2, \dots, p_r} a_1^{p_1}a_2^{p_2} \dots a_r^{p_r}.$$

Exercice 11 **

Soit E un ensemble fini à n éléments. Combien existe-t-il

- (a) de relations binaires sur E ?
- (b) de relations binaires réflexives et symétriques¹ sur E ?
- (c) de relations binaires réflexives et antisymétriques sur E ?

Solution

(a) méthode

On définit le *graphe* d'une relation binaire \mathcal{R} sur un ensemble E par

$$\Gamma = \{(x, y) \in E^2 \mid x \mathcal{R} y\}.$$

Ce graphe caractérise entièrement la relation \mathcal{R} .

Le graphe d'une relation binaire est une partie de E^2 . Il y a donc autant de relations binaires que d'éléments dans $\wp(E^2)$ à savoir 2^n (Th. 6 p. 207).

1. Le nombre de relations d'équivalence sur E est l'objet du sujet 18 p. 222.

(b) Une relation binaire \mathcal{R} sur E de graphe Γ est réflexive si, et seulement si, $(x, x) \in \Gamma$ pour tout x de E . Pour construire une relation réflexive, le seul degré de liberté est de décider quels sont les couples (x, y) avec $x \neq y$ qui sont en relation, les couples (x, x) sont quant à eux obligatoirement en relation.

Une relation binaire \mathcal{R} sur E de graphe Γ est symétrique lorsque l'on a $(x, y) \in \Gamma$ si, et seulement si, $(y, x) \in \Gamma$ pour tous x et y dans E . La détermination d'une relation réflexive et symétrique revient alors au choix des paires $\{x, y\}$ avec $x \neq y$ pour lesquelles (x, y) et (y, x) appartiennent à Γ . Il existe $\binom{n}{2}$ paires d'éléments de E et déterminer celles constituées d'éléments en relation revient à choisir¹ une partie dans l'ensemble de ces paires. Il y a donc $2^{n(n-1)/2}$ relations binaires réflexives et symétriques sur E .

(c) La construction d'une relation binaire réflexive et antisymétrique passe encore par l'étude des paires $\{x, y\}$ avec $x \neq y$. La contrainte d'antisymétrie est que l'on ne peut avoir simultanément (x, y) et (y, x) éléments du graphe de \mathcal{R} . Pour chaque paire $\{x, y\}$, trois choix sont alors possibles : soit aucun des couples (x, y) et (y, x) n'appartient au graphe, soit le couple (x, y) appartient au graphe mais pas le couple (y, x) , soit l'inverse. Au final, cela détermine² $3^{n(n-1)/2}$ relations binaires réflexives et antisymétriques sur E .

6.5.3 Compositions d'un entier

Exercice 12 * (Calcul par anagramme)

Soit p, q et n des entiers naturels.

(a) Un mot est constitué de p fois le caractère « A » et q fois le caractère « B ». Combien peut-on constituer d'anagrammes de ce mot ?

(b) On suppose $p \geq 1$. En considérant les symboles « 1 » et « + », combien existe-t-il de suites $(x_1, \dots, x_p) \in \mathbb{N}^p$ vérifiant $x_1 + \dots + x_p = n$?

Solution

(a) Pour former une anagramme, il suffit de choisir les p positions du caractère « A » parmi les $p+q$ places possibles, les positions vacantes étant alors occupées par le caractère « B ». Il y a donc³

$$\binom{p+q}{p} \text{ anagrammes possibles.}$$

1. Si l'on choisit la partie vide seuls les couples (x, x) sont en relation et la relation binaire est l'égalité. Si l'on choisit la partie complète, tous les couples (x, y) sont en relation.

2. Une formalisation approfondie peut sembler nécessaire : on énumère les éléments de $E : x_1, \dots, x_n$. À chaque couple (x_k, x_ℓ) avec $k < \ell$, on associe la valeur 1, 2 ou 3 pour coder quels couples formés de x_k et x_ℓ sont en relation. L'ensemble des relations binaires réflexives et antisymétriques est alors en bijection avec l'ensemble des applications au départ d'un ensemble à $n(n-1)/2$ éléments et à valeurs dans un ensemble à 3 éléments.

3. Cette étude est un cas particulier de celle du sujet 10 p. 216.

(b) méthode

Une somme $x_1 + \dots + x_p$ peut être codée par les caractères « 1 » et « + » :

$$\underbrace{(1 + \dots + 1)}_{x_1} + \underbrace{(1 + \dots + 1)}_{x_2} + \dots + \underbrace{(1 + \dots + 1)}_{x_p}.$$

Dans le codage proposé il y a $p - 1$ signes « + » (de taille standard) et $x_1 + \dots + x_p$ signes « 1 ». Déterminer $(x_1, \dots, x_p) \in \mathbb{N}^p$ de somme égale à n revient à former une anagramme avec n fois le caractère « 1 » et $p - 1$ fois le caractère « + ». Il y a

$$\binom{n+p-1}{n}$$
 possibilités.

Exercice 13 * (Calcul par les suites croissantes)

Soit $n \in \mathbb{N}$ et $p \in \mathbb{N}^*$. Il existe¹ $\binom{n+p}{p}$ suites croissantes de longueur p constituées d'éléments de $\llbracket 0 ; n \rrbracket$.

- (a) Combien existe-t-il de suites $(x_1, \dots, x_p) \in \mathbb{N}^p$ vérifiant $x_1 + \dots + x_p \leq n$?
- (b) Même question avec la condition $x_1 + \dots + x_p = n$.

Solution

(a) méthode

À une suite $(x_1, \dots, x_p) \in \mathbb{N}^p$, on fait correspondre une suite croissante par cumuls successifs.

À chaque suite $x = (x_1, \dots, x_p) \in \mathbb{N}^p$ de somme inférieure à n , on fait correspondre une suite $y = (y_1, \dots, y_p)$ déterminée par

$$y_1 = x_1, y_2 = x_1 + x_2, \dots, y_k = x_1 + \dots + x_k, \dots, y_p = x_1 + \dots + x_p.$$

La suite y est une suite croissante d'éléments de $\llbracket 0 ; n \rrbracket$. Inversement, une telle suite y correspond à une et une seule suite x telle que voulue : cette suite x est déterminée par les différences des termes successifs de y . Par cette correspondance bijective, on peut affirmer qu'il y a autant de suites $(x_1, \dots, x_p) \in \mathbb{N}^p$ de somme inférieure à n que de suites croissantes de longueur p d'éléments de $\llbracket 0 ; n \rrbracket$, c'est-à-dire $\binom{n+p}{p}$.

(b) méthode

La condition $x_1 + \dots + x_p = n$ est remplie si $x_1 + \dots + x_p \leq n$, mais pas $x_1 + \dots + x_p \leq n - 1$.

En vertu de la formule de Pascal et la symétrie des coefficients binomiaux, le nombre de suites cherché est donc

$$\binom{n+p}{p} - \binom{n+p-1}{p} = \binom{n+p-1}{p-1} = \binom{n+p-1}{n}.$$

1. Voir sujet 9 p. 215 où l'on a simplement opéré un glissement sur l'ensemble des valeurs en considérant $\llbracket 0 ; n \rrbracket$ au lieu de $\llbracket 1 ; n+1 \rrbracket$.

Exercice 14 ** (Calcul par récurrence)

Soit $n \in \mathbb{N}$ et $p \in \mathbb{N}^*$. On note C_n^p le nombre de suites $(x_1, \dots, x_p) \in \mathbb{N}^p$ vérifiant la condition $x_1 + \dots + x_p = n$.

(a) Établir

$$C_n^{p+1} = \sum_{k=0}^n C_k^p.$$

(b) En déduire

$$C_n^p = \binom{n+p-1}{n}.$$

Solution

(a) **méthode**

|| On dénombre les suites (x_1, \dots, x_{p+1}) en discutant selon la valeur de x_{p+1} .

Les x_j étant des entiers positifs, la valeur x_{p+1} d'une suite $(x_1, \dots, x_p, x_{p+1})$ de somme égale à n est comprise entre 0 et n . Lorsque celle-ci vaut k , la suite $(x_1, \dots, x_p, x_{p+1})$ de somme n détermine de façon bijective une suite $(x_1, \dots, x_p) \in \mathbb{N}^p$ de somme $n - k$: il y a C_{n-k}^p possibilités. On en déduit à l'aide d'un renversement d'indice

$$C_n^{p+1} = \sum_{k=0}^n C_{n-k}^p = \sum_{k=0}^n C_k^p.$$

(b) On raisonne par récurrence sur $p \in \mathbb{N}$ afin d'établir $C_n^p = \binom{n+p-1}{n}$ pour tout $n \in \mathbb{N}$.

Pour $p = 1$, il existe une seule façon d'écrire $x_1 = n$ et l'on a bien $\binom{n}{1} = 1$. Supposons la propriété vraie au rang $p \geq 1$. Soit $n \in \mathbb{N}$. Par l'hypothèse de récurrence, on peut écrire

$$C_n^{p+1} = \sum_{k=0}^n \binom{p+k-1}{k}.$$

Par la formule de Pascal¹, on peut transformer cette somme en une somme télescopique

$$\binom{p+k-1}{k} = \binom{p+k}{k} - \binom{p+k-1}{k-1} \quad \text{pour tout } k \in [0; n]$$

et l'on obtient

$$C_n^{p+1} = \sum_{k=0}^n \left(\binom{p+k}{k} - \binom{p+k-1}{k-1} \right) = \underbrace{\binom{p+n}{n}}_{=0} - \binom{p-1}{-1} = \binom{n+p}{n}.$$

La récurrence est établie.

1. Rappelons que celle-ci est valable pour $k \in \mathbb{Z}$ (voir Th. 10 p. 53).

Exercice 15 ** (Nombres de combinaisons avec répétition)

On appelle *combinaison avec répétition* de longueur p formée d'éléments d'un ensemble E le choix non ordonné de p éléments dans E avec possibilité de répétitions de ceux-ci¹. Combien peut-on former de combinaisons avec répétition de longueur p sur un ensemble à n éléments ?

Solution**méthode**

On choisit le nombre d'occurrences de chaque élément de l'ensemble présent dans la combinaison.

On peut énumérer les éléments de $E : a_1, \dots, a_n$. Former une combinaison avec répétition d'éléments de E revient à choisir, pour chaque élément a_i , le nombre $x_i \in \mathbb{N}$ d'occurrences de celui-ci dans la combinaison. La longueur de la combinaison est alors simplement la somme des x_i . Le nombre de combinaisons avec répétition de longueur p sur E correspond alors au nombre de façon d'écrire $p = x_1 + \dots + x_n$ avec $(x_1, \dots, x_n) \in \mathbb{N}^n$. Dans les sujets ci-dessus, on a pu voir que ce nombre vaut $\binom{n+p-1}{p}$.

6.5.4 Dénombrements ensemblistes**Exercice 16 ***

Soit E un ensemble à $n \in \mathbb{N}$ éléments. Combien existe-t-il de couples (X, Y) constitués de parties de E vérifiant $X \subset Y$?

Solution**méthode**

On forme un couple (X, Y) en choisissant X puis en définissant Y à l'aide du complémentaire de X dans E .

Dénombrons les couples (X, Y) cherchés selon la valeur $k \in \llbracket 0; n \rrbracket$ du cardinal de la partie X .

Soit $k \in \llbracket 0; n \rrbracket$. Il y a $\binom{n}{k}$ parties X possibles à k éléments dans E . Une fois celle-ci choisie, on forme une partie Y contenant X en déterminant $Z = Y \setminus X$ qui est une partie quelconque² incluse dans $E \setminus X$. Puisque $E \setminus X$ est de cardinal $n - k$, il y a exactement 2^{n-k} parties Z possibles (Th. 6 p. 207) et donc $\binom{n}{k}2^{n-k}$ couples (X, Y) avec $X \subset Y$ et $\text{Card}(X) = k$.

Finalement, le nombre de couples cherché est

$$\sum_{k=0}^n \binom{n}{k} 2^{n-k} = (1+2)^n = 3^n.$$

1. Par exemple, les choix de 1, 1, 2 ou de 1, 2, 1 définissent la même combinaison avec répétition de longueur 3 formée d'éléments de {1, 2, 3, 4}.

2. Par exemple, on forme le couple (X, X) en choisissant $Z = \emptyset \in \wp(E \setminus X)$, on forme le couple (X, E) en choisissant $Z = E \setminus X \in \wp(E \setminus X)$, etc.

Exercice 17 **

Soit E un ensemble à n éléments avec $n \geq 2$. Combien existe-t-il de paires $\{X, Y\}$ constituées de parties de E non vides et disjointes ?

Solution**méthode**

À chaque paire $\{X, Y\}$, il correspond deux couples (X, Y) que l'on construit en commençant par choisir la partie X .

Commençons par déterminer, selon la valeur $k \in [1; n-1]$ du cardinal de X , le nombre de couples (X, Y) formés de parties de E , non vides et disjointes.

Soit $k \in [1; n-1]$. Il existe $\binom{n}{k}$ parties X à k éléments incluses dans E . L'une d'elles étant fixée, on détermine Y non vide disjointe de X en choisissant une partie non vide dans $\wp(E \setminus X)$. Il y a $2^{n-k} - 1$ choix possibles pour la partie Y et donc $\binom{n}{k}(2^{n-k} - 1)$ couples (X, Y) convenables avec X de cardinal k . Le nombre total de couples (X, Y) convenables est alors

$$\sum_{k=1}^{n-1} \binom{n}{k} (2^{n-k} - 1) = \sum_{k=1}^{n-1} \binom{n}{k} 2^{n-k} - \sum_{k=1}^{n-1} \binom{n}{k}.$$

En adjoignant aux deux sommes des termes extrêmes pour reconnaître la formule du binôme, on obtient

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

Enfin, à chaque paire $\{X, Y\}$ correspondent deux couples convenables (X, Y) et (Y, X) , et le nombre de paires cherché est donc

$$\frac{3^n + 1}{2} - 2^n.$$

Exercice 18 ** (Nombres de Bell)

On appelle partition d'un ensemble E tout ensemble constitué de parties deux à deux disjointes, non vides et de réunion égale à E . On note B_n le nombre de partitions¹ d'un ensemble fini à $n \in \mathbb{N}^*$ éléments et l'on pose $B_0 = 1$. Établir que pour tout naturel n

$$B_{n+1} = \sum_{k=0}^n \binom{n}{k} B_k.$$

1. Une relation d'équivalence est caractérisée par ses classes d'équivalences qui constituent une partition de l'ensemble : B_n détermine le nombre de relations d'équivalence sur un ensemble à n éléments.

Solution

Considérons un ensemble E à $n + 1$ éléments. Parmi ceux-ci, choisissons un élément particulier que nous nommons x .

méthode

On dénombre les partitions en discutant selon le cardinal de la partie A qui contient l'élément x .

Dans une partition de E , il existe une seule partie A contenant l'élément x et celle-ci est de cardinal $k + 1$ pour une certaine valeur de $k \in \llbracket 0 ; n \rrbracket$.

Pour $k \in \llbracket 0 ; n \rrbracket$, on construit une partition de E dont la partie contenant x est à $k + 1$ éléments en commençant par choisir k éléments dans $E \setminus \{x\}$ pour constituer A : cela offre $\binom{n}{k}$ possibilités. On complète ensuite la partie A à l'aide d'une partition de $E \setminus A$ afin de constituer une partition de E : cela offre¹ B_{n-k} possibilités. Ainsi, il y a exactement $\binom{n}{k} B_{n-k}$ partitions de E dont la partie contenant x est de cardinal $k + 1$ et, finalement,

$$B_{n+1} = \sum_{k=0}^n \binom{n}{k} B_{n-k}.$$

En renversant l'indexation puis en exploitant la symétrie des coefficients binomiaux on obtient la relation attendue

$$B_{n+1} - \sum_{j=0}^n \binom{n}{n-j} B_j = \sum_{j=0}^n \binom{n}{j} B_j.$$

6.5.5 Dénombrements d'applications

Exercice 19 ** (Nombre de surjections)

Soit E et F deux ensembles finis non vides de cardinaux respectifs p et n . On note $S_{p,n}$ le nombre de surjections de E sur F .

(a) Déterminer $S_{p,1}$, $S_{n,n}$ et $S_{p,n}$ lorsque $p < n$.

(b) On suppose $p > 1$, $n > 1$ et l'on introduit a un élément arbitraire de E . En étudiant la restriction d'une surjection au départ de $E \setminus \{a\}$, établir

$$S_{p,n} = n(S_{p-1,n} + S_{p-1,n-1}).$$

(c) En déduire que, pour tout entier $n \geq 1$ et tout entier $p \geq 1$,

$$S_{p,n} = \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} k^p.$$

¹ Si $A = E$ alors $E \setminus A$ est vide. Or on a convenu $B_0 = 1$ ce qui est cohérent avec le calcul en cours.

Solution

(a) Si F est un singleton, il n'y a qu'une application au départ de E et à valeurs dans F et celle-ci est surjective : $S_{p,1} = 1$.

Si $\text{Card}(E) = \text{Card}(F) < +\infty$, les surjections de E sur F sont aussi les bijections de E vers F ou encore les injections de E dans F : $S_{n,n} = n!$ (Th. 8 p. 207).

Si $\text{Card}(E) < \text{Card}(F)$, il n'existe pas de surjections de E sur F : $S_{p,n} = 0$.

(b) **méthode**

On discute selon que la restriction d'une surjection au départ de $E \setminus \{a\}$ réalise ou non une surjection.

Une surjection de E sur F telle que sa restriction au départ de $E \setminus \{a\}$ soit surjective peut prendre n'importe quelle valeur en a . Il y a $S_{p-1,n}$ surjections de $E \setminus \{a\}$ sur F et n choix possibles pour l'image de a , il y a donc $nS_{p-1,n}$ surjections de ce type.

Une surjection de E sur F telle que sa restriction au départ de $E \setminus \{a\}$ ne soit pas surjective définit par restriction une surjection de $E \setminus \{a\}$ sur $F \setminus \{f(a)\}$. Il y a n possibilités pour choisir la valeur $f(a)$ et $S_{p-1,n-1}$ surjections possibles de $E \setminus \{a\}$ sur $F \setminus \{f(a)\}$. Au total, il y a $nS_{p-1,n-1}$ surjections dont la restriction au départ de $E \setminus \{a\}$ n'est pas surjective.

Une surjection de E sur F entrant dans l'une ou l'autre des deux catégories dénombrées ci-dessus, on peut affirmer $S_{p,n} = n(S_{p-1,n} + S_{p-1,n-1})$.

(c) **méthode**

On exploite la formule

$$\binom{n}{k} = \frac{n(n-1)}{k(k-1)} \quad \text{pour tout } k \in \llbracket 1 ; n \rrbracket. \quad (*)$$

La relation obtenue à la question précédente invite à réaliser un raisonnement par récurrence sur le paramètre $p \in \mathbb{N}^*$.

Pour $p = 1$, si $n = 1$, on sait $S_{1,1} = 1$ et l'on vérifie :

$$\sum_{k=0}^1 (-1)^{1-k} \binom{1}{k} k = 0 + 1 = 1.$$

Si $n > 1$, on sait $S_{1,n} = 0$ et l'on vérifie à l'aide de (*) puis de la formule du binôme

$$\sum_{k=0}^n (-1)^{n-k} \binom{n}{k} k = \sum_{k=1}^n (-1)^{n-k} n \binom{n-1}{k-1} = -n(1-1)^{n-1} = 0.$$

L'identité voulue est donc vraie pour toute valeur de $n \in \mathbb{N}^*$ lorsque $p = 1$.

Supposons la propriété établie au rang $p - 1 \geq 1$. Pour $n = 1$, on sait $S_{p,1} = 1$ et l'on vérifie :

$$\sum_{k=0}^1 (-1)^{1-k} \binom{1}{k} k^p = 0 + 1^p = 1.$$

Pour $n > 1$, on exploite l'identité de la question précédente et l'hypothèse de récurrence

$$\begin{aligned} S_{p,n} &= n(S_{p-1,n} + S_{p-1,n-1}) \\ &= n \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} k^{p-1} + n \sum_{k=0}^{n-1} (-1)^{n-1-k} \binom{n-1}{k} k^{p-1}. \end{aligned}$$

On ajoute un terme nul¹ à la deuxième somme en élargissant la plage d'indexation jusqu'à n puis on combine les deux sommes avant d'employer la formule du triangle de Pascal

$$S_{p,n} = n \sum_{k=0}^n (-1)^{n-k} \left(\binom{n}{k} - \binom{n-1}{k} \right) k^{p-1} = n \sum_{k=0}^n (-1)^{n-k} \binom{n-1}{k-1} k^{p-1}.$$

Enfin, la formule (*) permet de conclure en traitant l'indice $k = 0$ séparément.

$$S_{p,n} = \underbrace{0}_{k=0} + \sum_{k=1}^n (-1)^{n-k} \binom{n}{k} k^p = \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} k^p.$$

La récurrence est établie².

Exercice 20 *** (Nombre de dérangements)

Un *dérangement* sur un ensemble E est une permutation σ de E vérifiant $\sigma(x) \neq x$ pour tout $x \in E$. On note D_n le nombre de dérangements existants sur un ensemble à $n \in \mathbb{N}^*$ éléments.

- (a) Établir $D_{n+1} = n(D_n + D_{n-1})$ pour tout $n \geq 2$.
- (b) En déduire $D_n = nD_{n-1} + (-1)^n$ pour tout $n \geq 2$.
- (c) Conclure

$$D_n = n! \sum_{k=0}^n \frac{(-1)^k}{k!} \quad \text{pour tout } n \geq 1.$$

Solution

(a) Soit E un ensemble à $n+1$ éléments et a un élément arbitrairement choisi dans E . Si σ est un dérangement de E , on sait que $b = \sigma(a)$ est un élément de E différent de a . Ceci offre n possibilités pour la valeur de b .

1. Rappelons que $\binom{n}{k}$ est nul lorsque $k > n$ ou $k < 0$. L'expression d'un coefficient binomial par quotient de factorielles n'est valable que pour $k \in \llbracket 0 ; n \rrbracket$.

2. Pourquoi faire commencer la somme à l'indice 0 dans cette formule alors que le premier terme sommé est toujours nul? Lorsque $n = 0$ (on étudie une application à valeurs dans l'ensemble vide), il n'existe pas d'applications et a fortiori pas de surjections à valeurs dans l'ensemble vide, sauf si l'ensemble de départ est vide... auquel cas l'application (que l'on appelle l'*application vide*) est bijective. Sachant $0^0 = 1$, cette affirmation est cohérente avec la formule donnant $S_{p,n}$ (et aussi avec celle donnant le cardinal de l'ensemble \mathcal{S}_E des permutations de E : $\text{Card}(\mathcal{S}_E) = p!$ avec $p = \text{Card}(E)$).

méthode

|| On dénombre les dérangements σ de E vérifiant $\sigma(a) = b$ en discutant selon que $\sigma(b) = a$ ou $\sigma(b) \neq a$.

Les dérangements de E vérifiant $\sigma(b) = a$ définissent par restriction des dérangements de $E \setminus \{a, b\}$: il existe donc D_{n-1} dérangements de cette forme.

Dénombrons maintenant les dérangements σ de E vérifiant $\sigma(a) = b$ et $\sigma(b) \neq a$. Composons cette permutation avec la permutation τ dont le seul effet est d'échanger a et b et considérons $\sigma' = \sigma \circ \tau$. On vérifie $\sigma'(b) = b$ et $\sigma'(c) \neq c$ pour tout $c \neq b$. La permutation σ' détermine donc un déangement sur $E \setminus \{b\}$. Inversement, si σ' est une permutation de E réalisant un déangement sur $E \setminus \{b\}$, il existe une unique permutation σ telle que $\sigma' = \sigma \circ \tau$ et celle-ci est un déangement de E vérifiant $\sigma(a) = \sigma'(b) = b$ et $\sigma(b) = \sigma'(a) \neq a$. Il y a donc autant de dérangements sur E vérifiant $\sigma(a) = b$ et $\sigma(b) \neq a$ que de dérangements sur $E \setminus \{b\}$, c'est-à-dire D_n .

Au final, on obtient

$$D_{n+1} = \underbrace{n}_{\text{choix de } b} \times (\underbrace{D_{n-1}}_{\text{condition } \sigma(a)=b} + \underbrace{D_n}_{\text{condition } \sigma(a)\neq b}).$$

(b) méthode

|| On vérifie l'identité par récurrence sur $n \geq 2$.

Pour $n = 2$, on a¹ $D_1 = 0$ et $D_2 = 1$: la relation voulue est valide.

Supposons la relation vraie au rang $n \geq 2$. Par la formule obtenue à la question précédente et la relation de récurrence, il vient

$$D_{n+1} = n(D_{n-1} + D_n) = (D_n - (-1)^n) + nD_n = (n+1)D_n + (-1)^{n+1}.$$

La récurrence est établie.

(c) Soit $n \geq 2$. En divisant par $n!$ la relation de la question précédente et en réorganisant les membres, on peut écrire

$$\frac{D_n}{n!} - \frac{D_{n-1}}{(n-1)!} = \frac{(-1)^n}{n!}.$$

En sommant cette relation télescopique, on obtient

$$\frac{D_n}{n!} - \frac{D_1}{1!} = \sum_{k=2}^n \left(\frac{D_k}{k!} - \frac{D_{k-1}}{(k-1)!} \right) = \sum_{k=2}^n \frac{(-1)^k}{k!}.$$

En ajoutant les termes 1 et -1 correspondant aux indices $k = 0$ et $k = 1$ dans la somme, on conclut

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

Enfin, cette identité est aussi valable² pour $n = 1$.

1. Il n'existe pas de dérangements sur $\{a\}$ et il existe un seul déangement sur $\{a, b\}$ à savoir la permutation échangeant a et b .

2. La relation vaut aussi pour $n = 0$ car l'application vide (voir note 2 p. 225) est un déangement.

6.6 Exercices d'approfondissement

Exercice 21 *

Soit \mathcal{R} une relation d'équivalence sur un ensemble E de cardinal n . On suppose que \mathcal{R} possède p classes d'équivalence et l'on note q le nombre de couples $(x, y) \in E^2$ vérifiant $x \mathcal{R} y$. Établir $n^2 \leq pq$.

Solution

méthode

On emploie l'inégalité de Cauchy-Schwarz (Th. 2 p. 394) :

$$\left(\sum_{i=1}^n a_i b_i \right)^2 \leq \left(\sum_{i=1}^n a_i^2 \right) \left(\sum_{i=1}^n b_i^2 \right).$$

Notons n_1, \dots, n_p les cardinaux respectifs des p classes d'équivalence de \mathcal{R} : on a $n = n_1 + \dots + n_p$ car les classes d'équivalences sont deux à deux disjointes de réunion E . Au surplus, deux éléments x et y sont en relation si, et seulement si, ils appartiennent à la même classe d'équivalence. Il y a n^2 couples (x, y) en relation avec x et y dans la j -ème classe d'équivalence et, au total, il y a $q = n_1^2 + \dots + n_p^2$ couples (x, y) en relation. On conclut alors par l'inégalité de Cauchy-Schwarz :

$$n^2 = \left(\sum_{j=1}^p 1 \times n_j \right)^2 \leq \left(\sum_{j=1}^p 1^2 \right) \left(\sum_{j=1}^p n_j^2 \right) = pq.$$

Exercice 22 **

Soit E un ensemble fini de cardinal $n \in \mathbb{N}$. Calculer :

$$(a) \sum_{X \subset E} \text{Card}(X) \quad (b) \sum_{X, Y \subset E} \text{Card}(X \cap Y) \quad (c) \sum_{X, Y \subset E} \text{Card}(X \cup Y).$$

Solution

méthode

On regroupe les termes sommés selon la valeur du cardinal considéré.

(a) Une partie X de E est de cardinal $k \in [0; n]$ et il y a exactement $\binom{n}{k}$ parties à k éléments dans E . Par regroupement des termes (Th. 8 p. 50)

$$\sum_{X \subset E} \text{Card}(X) = \sum_{k=0}^n \left(\underbrace{\sum_{\substack{X \subset E \\ \text{Card}(X)=k}} k}_{\binom{n}{k} \text{ termes}} \right) = \sum_{k=0}^n k \binom{n}{k} = n2^{n-1}$$

cette dernière somme se calculant de diverses manières comme cela a déjà été vu dans le sujet 8 p. 63 ou le sujet 19 p. 74.

(b) On regroupe les termes sommés selon la valeur Z de $X \cap Y$.

$$\sum_{X,Y \subset E} \text{Card}(X \cap Y) = \sum_{Z \subset E} \left(\sum_{\substack{X,Y \subset E \\ X \cap Y = Z}} \text{Card}(Z) \right) = \sum_{k=0}^n \left(\sum_{\substack{Z \subset E \\ \text{Card}(Z)=k}} \sum_{\substack{X,Y \subset E \\ X \cap Y = Z}} k \right). \quad (*)$$

Il s'agit ensuite de déterminer le nombre de couples $(X, Y) \in \wp(E)^2$ vérifiant $X \cap Y = Z$ pour une partie Z de cardinal k donnée. Dénombrons ces couples selon le nombre j d'éléments qui constituent X , j variant de k à n .

Pour former une partie X à j éléments, on adjoint $j - k$ éléments choisis dans $E \setminus Z$ à la partie Z : ceci offre $\binom{n-k}{j-k}$ possibilités. Une fois la partie X déterminée, on définit Y en adjoignant à Z des éléments qui ne sont pas dans $E \setminus X$, autrement dit, en adjoignant à Z une partie quelconque de $E \setminus X$. L'ensemble $E \setminus X$ étant de cardinal $n - j$, ceci offre 2^{n-j} parties Y possibles, puis $\binom{n-k}{j-k} 2^{n-j}$ couples (X, Y) vérifiant $X \cap Y = Z$ et $\text{Card}(X) = j$. Finalement, le nombre de couples (X, Y) d'intersection Z est

$$\sum_{j=k}^n \binom{n-k}{j-k} 2^{n-j} = \sum_{i=j-k}^{n-k} \binom{n-k}{i} 2^{n-k-i} = (1+2)^{n-k} = 3^{n-k}.$$

On peut alorsachever le calcul amorcé dans $(*)$

$$\sum_{X,Y \subset E} \text{Card}(X \cap Y) = \sum_{k=0}^n \left(\sum_{\substack{Z \subset E \\ \text{Card}(Z)=k}} k \times 3^{n-k} \right) = \sum_{k=0}^n \binom{n}{k} k 3^{n-k} = \sum_{k=1}^n \binom{n}{k} k 3^{n-k}.$$

Sachant

$$\sum_{k=1}^n k \binom{n}{k} 3^{n-k} x^{k-1} = \frac{d}{dx} ((3+x)^n) = n(3+x)^{n-1}$$

on obtient

$$\sum_{X,Y \subset E} \text{Card}(X \cap Y) = n 4^{n-1}.$$

(c) On peut reproduire un calcul analogue au précédent mais il est plus commode d'employer la formule

$$\text{Card}(X \cup Y) = \text{Card}(X) + \text{Card}(Y) - \text{Card}(X \cap Y).$$

On obtient

$$\begin{aligned} \sum_{X,Y \subset E} \text{Card}(X \cup Y) &= \underbrace{\sum_{Y \subset E} \sum_{\substack{X \subset E \\ -n2^{n-1}}} \text{Card}(X)}_{2^n \text{ termes}} + \underbrace{\sum_{X \subset E} \sum_{\substack{Y \subset E \\ =n2^{n-1}}} \text{Card}(Y)}_{2^n \text{ termes}} - \sum_{X,Y \subset E} \text{Card}(X \cap Y) \\ &= 2^n \times n 2^{n-1} + 2^n \times n 2^{n-1} - n 4^{n-1} = 3n 4^{n-1}. \end{aligned}$$

Exercice 23 ** (Formule d'inversion de Pascal)

Soit n et p des entiers naturels.

(a) Soit $k, \ell \in \mathbb{N}$ tels que $\ell \leq k \leq n$. Comparer $\binom{n}{k} \binom{k}{\ell}$ et $\binom{n}{k-\ell}$.

(b) Soit (u_n) une suite de nombres réels et (v_k) la suite déterminée par

$$v_k = \sum_{\ell=0}^k \binom{k}{\ell} u_\ell.$$

Établir la formule d'inversion

$$u_n = \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} v_k.$$

(c) Application : On note $S_{p,n}$ le nombre de surjections d'un ensemble E à p éléments sur un ensemble F à n éléments. Déterminer une expression de $S_{p,n}$ en observant

$$n^p = \sum_{k=0}^n \binom{n}{k} S_{p,k}.$$

(d) Application : On note D_n le nombre de dérangements¹ sur un ensemble à n éléments. Déterminer une expression de D_n en observant

$$n! = \sum_{k=0}^n \binom{n}{k} d_k.$$

Solution

(a) Par les expressions factorielles des coefficients binomiaux

$$\begin{aligned} \binom{n}{k} \binom{k}{\ell} &= \frac{n!}{k!(n-k)!} \cdot \frac{k!}{\ell!(k-\ell)!} = \frac{n!}{(n-k)!\ell!(k-\ell)!} \\ &= \frac{n!}{\ell!(n-\ell)!} \cdot \frac{(n-\ell)!}{(n-k)!(k-\ell)!} = \binom{n}{\ell} \binom{n-\ell}{k-\ell}. \end{aligned}$$

(b) méthode

On injecte l'expression de v_k dans la somme qui doit exprimer u_n et l'on simplifie.

Par l'expression de v_k et en échangeant les deux sommes

$$\sum_{k=0}^n (-1)^{n-k} \binom{n}{k} v_k = \sum_{k=0}^n \sum_{\ell=0}^k (-1)^{n-k} \binom{n}{k} \binom{k}{\ell} u_\ell = \sum_{\ell=0}^n \left(\sum_{k=\ell}^n (-1)^{n-k} \binom{n}{k} \binom{k}{\ell} \right) u_\ell. \quad (*)$$

1. Voir sujet 20 p. 225.

Par l'identité précédente et en écrivant $(-1)^{n-k} = (-1)^{n+k} = (-1)^{n-\ell} \times (-1)^{k-\ell}$

$$\sum_{k=\ell}^n (-1)^{n-k} \binom{n}{k} \binom{k}{\ell} = (-1)^{n-\ell} \binom{n}{\ell} \sum_{k=\ell}^n (-1)^{k-\ell} \binom{n-\ell}{k-\ell}.$$

Puis par glissement d'indice

$$\sum_{k=\ell}^n (-1)^{k-\ell} \binom{n-\ell}{k-\ell} = \sum_{j=0}^{n-\ell} (-1)^j \binom{n-\ell}{j} = (1 + (-1))^{n-\ell} = 0^{n-\ell} = \begin{cases} 1 & \text{si } \ell = n \\ 0 & \text{sinon.} \end{cases}$$

Dans le dernier membre de (*), la somme se limite alors au seul terme d'indice $\ell = n$ et l'on obtient

$$\sum_{k=0}^n (-1)^{n-k} \binom{n}{k} v_k = (-1)^0 \binom{n}{n} u_n = u_n$$

(c) On dénombre les applications de E vers F selon les cardinaux de leurs images :

$$\underbrace{w^p}_{\substack{\text{nombre} \\ \text{d'applications} \\ \text{de } E \text{ vers } F}} = \sum_{k=0}^n \underbrace{\binom{n}{k}}_{\substack{\text{nombre de choix} \\ \text{d'images à } k \text{ éléments} \\ \text{dans } F}} \underbrace{S_{p,k}}_{\substack{\text{nombre d'applications} \\ \text{de } E \text{ vers l'image} \\ \text{à } k \text{ éléments choisie}}}$$

et par la formule d'inversion, on obtient

$$S_{p,n} = \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} k^p.$$

(d) On dénombre les permutations de E selon les cardinaux des parties sur lesquelles elles opèrent un dérangement :

$$\underbrace{n!}_{\substack{\text{nombre de} \\ \text{permutations} \\ \text{de } E}} = \sum_{k=0}^n \underbrace{\binom{n}{k}}_{\substack{\text{nombre de choix} \\ \text{de parties à } k \text{ éléments} \\ \text{dans } E}} \underbrace{D_k}_{\substack{\text{nombre de permutations de } E \\ \text{dérangeant la partie} \\ \text{à } k \text{ éléments choisie}}}$$

et par la formule d'inversion, on obtient

$$D_n = \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} k! = n! \sum_{k=0}^n \frac{(-1)^{n-k}}{(n-k)!} = n! \sum_{j=0}^n \frac{(-1)^j}{j!}.$$

Exercice 24 ** (Nombres de Catalan)

Soit n et p deux entiers naturels non nuls.

On appelle chemin monotone tout déplacement sur une grille de coordonnées entières obtenu en partant de la case $(0, 0)$ et en se déplaçant successivement d'une unité vers la droite ou d'une unité vers le haut.

(a) Combien existe-t-il de chemin monotone rejoignant la case (n, p) ?

On s'intéresse désormais aux chemins monotones rejoignant la case (n, n) . On dit qu'un tel chemin est sous-diagonal si les cases (x, y) qu'il emprunte vérifient la condition $y \leq x$.

(b) On considère un chemin qui n'est pas sous-diagonal et on lui associe un nouveau chemin déterminé en changeant les mouvements vers la droite par des mouvements vers le haut, et inversement, dès son premier franchissement de la diagonale.

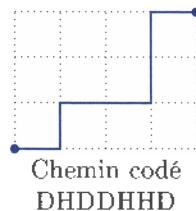
Quelle nouvelle case rejoint ce chemin ?

(c) Combien existe-t-il de chemins sous-diagonaux rejoignant la case (n, n) ?

Solution**(a) méthode**

Un chemin monotone peut être codé par une succession de lettres D et H traduisant les déplacements vers la droite ou vers le haut.

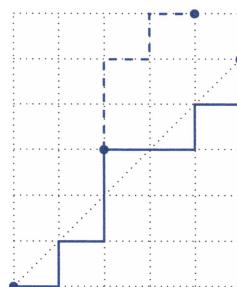
Un chemin monotone rejoint la case (n, p) si, et seulement si, son codage est constitué de n lettres D et p lettres H . Dénombrer les chemins monotones rejoignant la case (n, p) revient à dénombrer les anagrammes écrites avec n lettres D et p lettres H : on choisit les n positions occupées par la lettre D dans la liste des $n + p$ places disponibles et l'on obtient $\binom{n+p}{n}$ chemins monotones joignant la case (n, p) .

**(b) méthode**

Une petite figure peut aider à comprendre !

Illustrons ci-contre la transformation d'un chemin franchissant la diagonale.

Au premier franchissement de la diagonale, le chemin a effectué k déplacements sur la droite et $k+1$ déplacements vers le haut. Il lui reste $n-k$ déplacements sur la droite à effectuer et $n-k-1$ vers le haut pour rejoindre la case (n, n) . Cependant, on forme le nouveau chemin en inversant ces derniers mouvements. Le nouveau chemin fera donc un total de $k+(n-k-1)$ mouvements vers la droite et de $(k+1)+(n-k)$ vers le haut. Ce nouveau chemin rejoint donc la case $(n-1, n+1)$.



(c) À un chemin monotone joignant (n, n) qui n'est pas sous-diagonal on a associé ci-dessus un chemin joignant la case $(n - 1, n + 1)$. Inversement, un chemin joignant la case $(n - 1, n + 1)$ franchit nécessairement la diagonale et, en inversant les mouvements au premier franchissement, on détermine l'unique chemin joignant (n, n) non sous-diagonal qui lui est associé : il y a autant de chemins joignant (n, n) non sous-diagonaux que de chemins monotones joignant $(n - 1, n + 1)$ à savoir $\binom{2n}{n-1}$.

Il y a en tout $^{(2n)}$ chemins monotones joignant la case (n, n) et parmi ceux-ci exactement

$$\binom{2n}{n} = \binom{2n}{n-1} = \frac{1}{n+1} \binom{2n}{n}$$

chemins sous-diagonaux¹.

Exercice 25 ***

Montrer qu'un ensemble E est infini si, et seulement si, pour toute application $f: E \rightarrow E$, il existe une partie A non vide et distincte de E telle que $f(A) \subset A$.

Solution

On raisonne par double implication.

(\Leftarrow) On procède par contraposition. Soit E un ensemble fini. S'il est vide ou réduit à un élément, il n'existe pas de parties A incluses dans E vérifiant $A \neq \emptyset$ et $A \neq E$. Sinon, déterminons une application $f: E \rightarrow E$ pour laquelle il n'existe pas de parties A telles que proposées. En posant n le cardinal de E , on peut énumérer ses éléments et écrire $E = \{x_1, x_2, \dots, x_n\}$.

méthode

On définit une application f « cyclique » déterminée de sorte que l'appartenance de x_1 à A entraîne celle de x_2 , etc.

Considérons l'application $f: E \rightarrow E$ définie par

$$f(x_i) = x_{i+1} \text{ pour tout } i \in [1; n]$$

en convenant que x_{n+1} désigne x_1 .

Soit une partie A de E vérifiant $f(A) \subset A$. Si A est non vide, il existe un indice i dans $[1; n]$ tel que x_i est élément de A . L'élément $x_{i+1} = f(x_i)$ appartient aussi à A , et, en répétant ce processus, on obtient $x_i, x_{i+1}, \dots, x_n, x_1, \dots, x_{i-1}$ tous éléments de A et donc $A = E$.

Ainsi, si E est un ensemble fini, il existe une application $f: E \rightarrow E$ pour laquelle les seules parties A de E vérifiant $f(A) \subset A$ sont \emptyset et E .

(\Rightarrow) Soit E un ensemble infini, $f: E \rightarrow E$ une application quelconque et $x \in E$.

1. Dans le codage d'un chemin monotone joignant la case (n, n) il y a autant de D que de H . Un chemin est sous-diagonal, lorsqu'il y a toujours plus de D que de H dans tous les préfixes du codage : si D désigne une parenthèse ouvrante et H une parenthèse fermante, on vient de déterminer le nombre de façons de disposer correctement n paires de parenthèses.

méthode

On détermine une partie A convenable en étudiant la suite des itérés de composition : $x, f(x), f^2(x) = f(f(x)), \dots, f^n(x)$, etc.

S'il existe $n \in \mathbb{N}^*$ tel que $f^n(x) = x$ alors la partie $A = \{x, f(x), \dots, f^{n-1}(x)\}$ est non vide, distincte de l'ensemble infini E et vérifie $f(A) \subset A$.

Sinon, la partie $A = \{f(x), f^2(x), \dots\} = \{f^n(x) \mid n \in \mathbb{N}^*\}$ est non vide, distincte de E (car x n'est pas élément de A) et vérifie $f(A) \subset A$.

CHAPITRE 7

Espaces vectoriels

\mathbb{K} désigne \mathbb{R} ou \mathbb{C} .

7.1 Espaces vectoriels

7.1.1 Produit extérieur

Définition

On appelle *produit extérieur* opérant de \mathbb{K} sur E toute application de $\mathbb{K} \times E$ vers E . Ce produit extérieur est usuellement noté $(.)$ et λx désigne alors l'image du couple (λ, x) par celui-ci.

Sur \mathbb{K}^n , le produit extérieur usuel est défini par

$$\lambda(x_1, \dots, x_n) \stackrel{\text{def}}{=} (\lambda x_1, \dots, \lambda x_n) \quad \text{pour tous } \lambda \in \mathbb{K} \text{ et } (x_1, \dots, x_n) \in \mathbb{K}^n.$$

Sur l'ensemble $\mathcal{F}(X, \mathbb{K})$ des fonctions d'un ensemble X vers \mathbb{K} , le produit extérieur correspond à la multiplication par une fonction constante.

$$\lambda f: x \mapsto \lambda f(x) \quad \text{pour tous } \lambda \in \mathbb{K} \text{ et } f \in \mathcal{F}(X, \mathbb{K}).$$

Sur l'ensemble $\mathbb{K}[X]$ des polynômes en l'indéterminée X , le produit extérieur correspond à la multiplication par un polynôme constant.

7.1.2 Structure d'espace vectoriel

Définition

On appelle \mathbb{K} -espace vectoriel¹ tout triplet $(E, +, \cdot)$ formé d'un ensemble E , d'une loi de composition interne $+$ munissant E d'une structure de groupe abélien et d'un produit extérieur (\cdot) opérant de \mathbb{K} sur E vérifiant, pour tout $(x, y) \in E^2$ et tout $(\lambda, \mu) \in \mathbb{K}^2$, les identités :

$$\lambda(x + y) = \lambda x + \lambda y, \quad (\lambda + \mu)x = \lambda x + \mu x, \quad \lambda(\mu x) = (\lambda\mu)x \quad \text{et} \quad 1x = x.$$

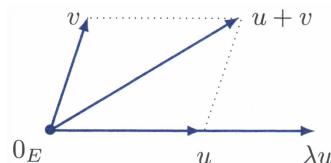
Les éléments de \mathbb{K} sont appelés *scalaires*, ceux de E sont appelés *vecteurs*. En particulier, le neutre additif de E est appelé *vecteur nul* et est noté 0 ou 0_E .

\mathbb{K}^n , $\mathcal{F}(X, \mathbb{K})$ et $\mathbb{K}[X]$ munis des opérations usuelles sont² des \mathbb{K} -espaces vectoriels.

\mathbb{K} peut aussi se comprendre comme un \mathbb{K} -espace vectoriel. Dans cette situation, scalaires et vecteurs se confondent avec les éléments de \mathbb{K} et le produit extérieur correspond à la multiplication usuelle.

\mathbb{C} peut se comprendre comme un \mathbb{R} -espace vectoriel. Les scalaires sont alors les nombres réels tandis que les vecteurs sont les nombres complexes. Plus généralement, par restriction du champ scalaire, tout \mathbb{C} -espace vectoriel peut se comprendre comme un \mathbb{R} -espace vectoriel.

Les opérations dans les espaces vectoriels peuvent être visualisées par une figure en *géométrie vectorielle*. La réalisation d'une telle figure passe par le choix arbitraire d'un point représentant le vecteur nul. Contrairement à la géométrie affine, tous les vecteurs sont alors exclusivement figurés au départ de ce point.



7.1.3 Structure produit

Théorème 1

Si E_1, \dots, E_n sont des \mathbb{K} -espaces vectoriels alors $E = E_1 \times \dots \times E_n$ est un \mathbb{K} -espace vectoriel pour les lois $+$ et (\cdot) définies par :

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) \stackrel{\text{def}}{=} (x_1 + y_1, \dots, x_n + y_n),$$

$$\lambda(x_1, \dots, x_n) \stackrel{\text{def}}{=} (\lambda x_1, \dots, \lambda x_n).$$

De plus, le vecteur nul de E est $0_E = (0_{E_1}, \dots, 0_{E_n})$.

En prenant les E_i tous égaux à \mathbb{K} , on retrouve que \mathbb{K}^n est un \mathbb{K} -espace vectoriel de vecteur nul $0_{\mathbb{K}^n} = (0, \dots, 0)$.

1. Si $\mathbb{K} = \mathbb{R}$, on parle d'*espace vectoriel réel*, si $\mathbb{K} = \mathbb{C}$, on parle d'*espace vectoriel complexe*.
2. En particulier, l'espace des suites $\mathbb{K}^\mathbb{N} = \mathcal{F}(\mathbb{N}, \mathbb{K})$ est un \mathbb{K} -espace vectoriel.

Théorème 2

Si E un \mathbb{K} -espace vectoriel et X un ensemble quelconque, l'ensemble $\mathcal{F}(X, E) = E^X$ des fonctions de X vers E est un \mathbb{K} -espace vectoriel pour les lois $+$ et $(.)$ définies par :

$$f + g : x \mapsto f(x) + g(x) \quad \text{et} \quad \lambda f : x \mapsto \lambda f(x).$$

De plus, le vecteur nul de $\mathcal{F}(X, E)$ est la fonction nulle $x \mapsto 0_E$.

En prenant $E = \mathbb{K}$, on retrouve que $\mathcal{F}(X, \mathbb{K})$ est un \mathbb{K} -espace vectoriel.

7.1.4 Combinaison linéaire

Soit E un \mathbb{K} -espace vectoriel et I un ensemble d'indexation.

Définition

Lorsque l'ensemble I est fini, on appelle *combinaison linéaire* d'une famille $(x_i)_{i \in I}$ de vecteurs de E tout vecteur x de E pouvant s'écrire

$$x = \sum_{i \in I} \lambda_i x_i \quad \text{avec} \quad (\lambda_i)_{i \in I} \in \mathbb{K}^I.$$

Les combinaisons linéaires d'une famille d'un vecteur x sont les λx pour λ parcourant \mathbb{K} . Les combinaisons linéaires d'une famille de deux vecteurs x et y sont les $\lambda x + \mu y$ pour λ et μ parcourant \mathbb{K} .

Seul le vecteur nul est combinaison linéaire de la famille vide.

Définition

Lorsque l'ensemble I est infini, on appelle *combinaison linéaire* d'une famille $(x_i)_{i \in I}$ de vecteurs de E tout vecteur combinaison linéaire d'une sous famille finie de $(x_i)_{i \in I}$.

Un tel vecteur peut encore s'écrire $\sum_{i \in I} \lambda_i x_i$ avec $(\lambda_i)_{i \in I}$ famille de scalaires à *support fini*¹.

7.2 Sous-espaces vectoriels

E désigne un \mathbb{K} -espace vectoriel.

7.2.1 Définition**Définition**

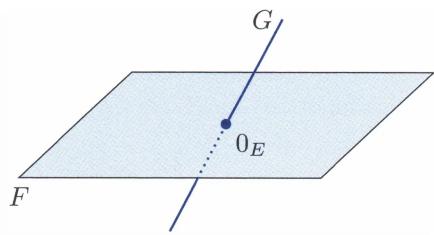
On appelle *sous-espace vectoriel* du \mathbb{K} -espace vectoriel E toute partie F non vide de E stable par combinaison linéaire², c'est-à-dire vérifiant³ $\lambda x + \mu y \in F$ pour tous $\lambda, \mu \in \mathbb{K}$ et tous $x, y \in F$.

1. Une telle famille ne comporte qu'un nombre fini de scalaires non nuls, on dit encore que c'est une famille *presque nulle*. La somme des $\lambda_i x_i$ pour i parcourant I comporte une infinité de termes mais parmi ceux-ci il ne figure qu'un nombre fini de termes non nuls.

2. Une combinaison linéaire d'une famille finie, ou infinie, de vecteurs d'un sous-espace vectoriel F appartient à F .

$\{0_E\}$ et E sont des sous-espaces vectoriels de E , ce sont ses *sous-espaces vectoriels triviaux*.

Géométriquement, les sous-espaces vectoriels non triviaux se visualisent tels des droites ou des plans passant par le point choisi pour figurer le vecteur nul.



Théorème 3

Si F est un sous-espace vectoriel d'un \mathbb{K} -espace vectoriel $(E, +, \cdot)$ alors $(F, +, \cdot)$ est un \mathbb{K} -espace vectoriel⁴ de même vecteur nul que E .

$\mathbb{K}_n[X]$ est un sous-espace vectoriel de $\mathbb{K}[X]$ donc un \mathbb{K} -espace vectoriel.

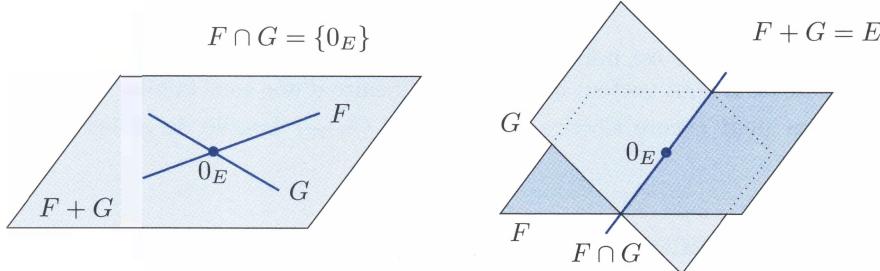
7.2.2 Opérations

Théorème 4

Si F et G sont deux sous-espaces vectoriels de E alors

$$F \cap G = \{x \in E \mid x \in F \text{ et } x \in G\} \quad \text{et} \quad F + G = \{a + b \mid a \in F \text{ et } b \in G\}$$

sont des sous-espaces vectoriels de E .



Les opérations d'intersection et de somme de sous-espaces vectoriels sont commutatives, associatives et possèdent des neutres qui sont E et $\{0_E\}$ respectivement.

En particulier, si F_1, \dots, F_n sont des sous-espaces vectoriels de E , on peut introduire les sous-espaces vectoriels

$$\begin{aligned} \bigcap_{i=1}^n F_i &= F_1 \cap \dots \cap F_n = \left\{ x \in E \mid \forall i \in [1; n], x \in F_i \right\} \\ \sum_{i=1}^n F_i &= F_1 + \dots + F_n = \left\{ \sum_{i=1}^n x_i \mid \forall i \in [1; n], x_i \in F_i \right\}. \end{aligned}$$

3. Il existe une définition équivalente « plus économique » où l'on étudie l'appartenance à F de $x + \lambda y$ pour tous $x, y \in F$ et $\lambda \in \mathbb{K}$.

4. Les lois $+$ et (\cdot) sur F sont définies par restriction des lois correspondantes sur E .

7.2.3 Sommes directes

Soit F_1, \dots, F_n des sous-espaces vectoriels de E . Tout vecteur x de la somme des espaces F_i peut s'écrire $x = x_1 + \dots + x_n$ avec $x_i \in F_i$.

Définition

On dit que la somme des espaces F_i est *directe* lorsqu'il y a unicité dans la décomposition précédente :

$$\forall x \in \sum_{i=1}^n F_i, \exists! (x_1, \dots, x_n) \in F_1 \times \dots \times F_n, \quad x = x_1 + \dots + x_n.$$

Lorsque la somme $F_1 + \dots + F_n$ est directe, celle-ci est notée

$$\bigoplus_{i=1}^n F_i \quad \text{ou} \quad F_1 \oplus \dots \oplus F_n.$$

Théorème 5

Les espaces F_1, \dots, F_n sont en somme directe si, et seulement si, il y a unicité de la décomposition du vecteur nul :

$$\forall (x_1, \dots, x_n) \in F_1 \times \dots \times F_n, \quad x_1 + \dots + x_n = 0_E \implies x_1 = \dots = x_n = 0_E.$$

Si l'on se limite à deux¹ sous-espaces vectoriels F et G , on a une caractérisation simple :

$$F \text{ et } G \text{ sont en somme directe} \iff F \cap G = \{0_E\}.$$

Si l'on considère plusieurs sous-espaces vectoriels F_1, \dots, F_n et F_{n+1} , on dispose de la propriété « d'associativité » suivante : si F_1, \dots, F_n sont en somme directe et si leur somme est en somme directe avec F_{n+1} alors les espaces F_1, \dots, F_n, F_{n+1} sont en somme directe. La réciproque étant aussi vraie, on peut écrire :

$$(F_1 \oplus \dots \oplus F_n) \oplus F_{n+1} = F_1 \oplus \dots \oplus F_n \oplus F_{n+1}.$$

7.2.4 Sous-espaces vectoriels supplémentaires

Définition

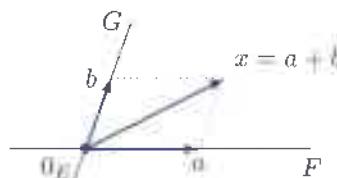
Deux sous-espaces vectoriels F et G sont dits *supplémentaires* si tout vecteur de E s'écrit de façon unique comme somme d'un vecteur de F et d'un vecteur de G :

$$\forall x \in E, \exists! (a, b) \in F \times G, \quad x = a + b.$$

Cela signifie encore que l'on peut écrire $E = F \oplus G$.

Plus généralement, on appelle *décomposition en somme directe* de l'espace E l'écriture $E = E_1 \oplus \dots \oplus E_n$ avec E_i des sous-espaces vectoriels de E .

1. Pour trois sous-espaces vectoriels ou plus, il n'existe pas de caractérisation aussi simple et il est alors préférable d'étudier l'unicité de la décomposition du vecteur nul.



7.2.5 Espace vectoriel engendré par une partie

Définition

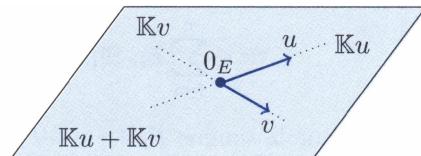
On appelle *espace vectoriel engendré* par une partie A de E l'ensemble¹ $\text{Vect}(A)$ des combinaisons linéaires des familles d'éléments de A .

Pour $u \in E$, $\text{Vect}\{u\} = \{\lambda u \mid \lambda \in \mathbb{K}\}$.

Cet espace est noté $\mathbb{K}u$.

Pour $u, v \in E$,

$$\begin{aligned}\text{Vect}\{u, v\} &= \{\lambda u + \mu v \mid (\lambda, \mu) \in \mathbb{K}^2\} \\ &= \mathbb{K}u + \mathbb{K}v.\end{aligned}$$



Plus généralement,

$$\text{Vect}\{x_1, \dots, x_n\} = \{\lambda_1 x_1 + \dots + \lambda_n x_n \mid (\lambda_i)_{1 \leq i \leq n} \in \mathbb{K}^n\} = \mathbb{K}x_1 + \dots + \mathbb{K}x_n.$$

Théorème 6

$\text{Vect}(A)$ est un sous-espace vectoriel de E contenant A .

De plus, tout sous-espace vectoriel F de E contenant A contient aussi $\text{Vect}(A)$.

Au sens de l'inclusion, $\text{Vect}(A)$ est le plus petit sous-espace vectoriel de E contenant A .

7.3 Famille de vecteurs

Soit E un \mathbb{K} -espace vectoriel et $(x_i)_{i \in I}$ une famille de vecteurs de E indexée par un ensemble I .

7.3.1 Famille génératrice

Définition

On dit que la famille $(x_i)_{i \in I}$ est *génératrice*² si tout vecteur de E est combinaison linéaire de celle-ci.

En notant $\text{Vect}(x_i)_{i \in I}$ l'espace vectoriel engendré par la partie $\{x_i \mid i \in I\}$, la famille $(x_i)_{i \in I}$ est génératrice si, et seulement si, $E = \text{Vect}(x_i)_{i \in I}$.

7.3.2 Famille libre

Définition

On dit que la famille $(x_i)_{i \in I}$ est *libre* si toute combinaison linéaire nulle des éléments de cette famille s'écrit avec des scalaires nécessairement tous nuls. Sinon, on dit que la famille est *liée* et une combinaison linéaire nulle écrite avec des scalaires non tous nuls se nomme une *relation linéaire*.

1. $\text{Vect}(A)$ est aussi l'intersection de tous les sous-espaces vectoriels de E contenant A .
2. En cas d'ambiguïté, on pourra être plus précis en affirmant que la famille est génératrice de E .

Une famille formée d'un seul vecteur est libre si, et seulement si, ce vecteur est non nul. Une famille formée de deux vecteurs est libre si, et seulement si, aucun n'est colinéaire à l'autre.

Une famille finie (x_1, \dots, x_n) de vecteurs de E est libre lorsque l'on a l'implication :

$$\forall (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n, \quad \lambda_1 x_1 + \dots + \lambda_n x_n = \mathbf{0}_E \implies \lambda_1 = \dots = \lambda_n = 0.$$

Une famille infinie est libre si, et seulement si, toutes ses sous-familles finies le sont.

Théorème 7

Si la famille $(x_i)_{i \in I}$ est libre, il y a unicité de l'écriture d'un vecteur de $\text{Vect}(x_i)_{i \in I}$ comme combinaison linéaire des vecteurs de cette famille.

7.3.3 Base

Définition

On dit qu'une famille $(e_i)_{i \in I}$ de vecteurs de E est une *base*¹ lorsque celle-ci est libre et génératrice.

Dans \mathbb{K}^n , on introduit $e_i = (0, \dots, 0, 1, 0, \dots, 0)$ où le 1 est positionné en i -ème place. La famille (e_1, \dots, e_n) est une base que l'on appelle *la base canonique*² de \mathbb{K}^n .

Dans $\mathbb{K}_n[X]$, la famille de monômes $(1, X, \dots, X^n)$ est une base que l'on appelle *la base canonique* de $\mathbb{K}_n[X]$. Dans $\mathbb{K}[X]$, la famille infinie $(X^n)_{n \in \mathbb{N}}$ est une base que l'on qualifie aussi de canonique.

La famille vide est base de l'espace nul $\{0_E\}$.

7.3.4 Coordonnées

Théorème 8

Si $(e_i)_{i \in I}$ est une base de E , tout vecteur x de E s'écrit de façon unique comme combinaison linéaire de la famille $(e_i)_{i \in I}$.

La famille de scalaires réalisant cette écriture définit la famille des *coordonnées*³ de x dans la base $(e_i)_{i \in I}$.

7.4 Espaces de dimension finie

7.4.1 Définition

Définition

On dit qu'un \mathbb{K} -espace vectoriel est de *dimension finie* s'il possède une famille génératrice formée d'un nombre fini de vecteurs.

1. En cas d'ambiguïté, on pourra être plus précis en parlant de *base de E* .
2. Le qualificatif canonique est là pour signifier que cette base est remarquablement simple car liée à la construction de l'espace.
3. On parle aussi parfois des *composantes* d'un vecteur.

On est alors assuré de l'existence de bases à cet espace par l'un ou l'autre des deux résultats suivants :

Théorème 9 (Théorème de la base extraite)

Si E est un \mathbb{K} -espace vectoriel de dimension finie, on peut extraire une base de n'importe quelle famille génératrice.

Théorème 10 (Théorème de la base incomplète)

Si E est un \mathbb{K} -espace vectoriel de dimension finie, toute famille libre peut être complétée en une base.

7.4.2 Dimension

Dans un \mathbb{K} -espace vectoriel, il y a toujours moins de vecteurs dans une famille libre que dans une famille génératrice. En conséquence, les bases sont toutes constituées du même nombre de vecteurs.

Définition

On appelle dimension d'un \mathbb{K} -espace vectoriel E de dimension finie, le nombre de vecteurs constituant les bases de celui-ci. Cette dimension est notée $\dim E$.

On a $\dim \mathbb{K}^n = n$, $\dim \mathbb{K}_n[X] = n + 1$ et l'on écrira abusivement $\dim \mathbb{K}[X] = +\infty$.

Les espaces de dimension 1 se nomment des *droites vectorielles*¹, ceux de dimension 2 s'appellent des *plans vectoriels*. L'espace nul $\{0_E\}$ est de dimension nulle.

Théorème 11

Si E_1, \dots, E_n sont des espaces vectoriels de dimensions finies, l'espace $E_1 \times \dots \times E_n$ l'est aussi et

$$\dim(E_1 \times \dots \times E_n) = \dim E_1 + \dots + \dim E_n.$$

7.4.3 Rang d'une famille de vecteurs

Définition

On appelle *rang* d'une famille $(x_i)_{i \in I}$ de vecteurs d'un espace vectoriel la dimension de l'espace engendré par cette famille. Celui-ci est noté $\text{rg}(x_i)_{i \in I}$.

Une famille de n vecteurs est libre si, et seulement si, elle est de rang n .

7.4.4 Bases en dimension finie

En dimension finie connue, on peut caractériser avec efficacité les bases :

1. En l'absence d'ambiguité, on peut parler simplement de droites.

Théorème 12

Soit E un \mathbb{K} -espace vectoriel de dimension finie n et $(e_i)_{1 \leq i \leq n}$ une famille de vecteurs de E constituée de n vecteurs exactement. On a équivalence entre :

- (i) $(e_i)_{1 \leq i \leq n}$ est une base ;
- (ii) $(e_i)_{1 \leq i \leq n}$ est une famille libre ;
- (iii) $(e_i)_{1 \leq i \leq n}$ est une famille génératrice.

7.4.5 Sous-espace vectoriel en dimension finie**Théorème 13**

Si F est un sous-espace vectoriel d'un \mathbb{K} -espace vectoriel E de dimension finie alors F est de dimension finie et $\dim F \leq \dim E$ avec égalité si, et seulement si, $F = E$.

On peut alors montrer l'égalité de deux sous-espaces vectoriels en constatant une inclusion et l'égalité de leurs dimensions (finies).

Si F est un sous-espace vectoriel d'un espace E de dimension finie, on peut compléter une base de F afin de former une base E . Une telle base est dite *adaptée à F* . Les vecteurs introduits pour compléter la base de F engendrent alors un supplémentaire de F et donc :

Théorème 14

Tout sous-espace vectoriel F d'un \mathbb{K} -espace vectoriel E de dimension finie admet au moins un¹ supplémentaire G .

De plus, en accolant une base de F et une base de G on obtient une base de $E = F \oplus G$ que l'on dit *adaptée à la supplémentarité de F et G* . On en déduit que les espaces supplémentaires de F ont tous la même dimension : $\dim E - \dim F$.

Plus généralement, si $E = E_1 \oplus \dots \oplus E_n$, on peut former une base de E en accolant des bases des espaces E_i . Une telle base est dite *adaptée à la décomposition en somme directe $E = E_1 \oplus \dots \oplus E_n$* .

7.4.6 Dimension d'une somme de sous-espaces vectoriels**Théorème 15 (Formule de Grassmann)**

Si F et G sont des sous-espaces vectoriels de dimensions finies d'un \mathbb{K} -espace vectoriel E alors $F + G$ et $F \cap G$ sont de dimensions finies et

$$\dim(F + G) = \dim F + \dim G - \dim(F \cap G).$$

On en déduit une caractérisation rapide de la supplémentarité utile lorsque les espaces sont de dimensions finies connues :

1. Il y a généralement plusieurs supplémentaires à un sous-espace donné : il ne faut pas confondre supplémentaire et complémentaire ! Le complémentaire d'un sous-espace vectoriel n'est jamais un sous-espace car il ne contient pas 0_E .

Théorème 16

Soit F et G deux sous-espaces vectoriels d'un \mathbb{K} -espace vectoriel E de dimension finie vérifiant $\dim E = \dim F + \dim G$. On a équivalence entre :

- (i) F et G sont supplémentaires ;
- (ii) $F \cap G = \{0_E\}$;
- (iii) $F + G = E$.

Enfin, on peut étudier la dimension d'une somme de plusieurs espaces :

Théorème 17

Si F_1, \dots, F_n sont des sous-espaces vectoriels de dimensions finies alors $F_1 + \dots + F_n$ est de dimension finie et

$$\dim \left(\sum_{i=1}^n F_i \right) \leq \sum_{i=1}^n \dim F_i.$$

De plus, il y a égalité si, et seulement si, les sous-espaces vectoriels F_1, \dots, F_n sont en somme directe.

On retient l'identité

$$\dim \left(\bigoplus_{i=1}^n F_i \right) = \sum_{i=1}^n \dim F_i.$$

7.5 Sous-espaces affines

Soit E un espace vectoriel.

7.5.1 Interprétation affine

Selon l'interprétation géométrique souhaitée, les éléments de E peuvent être indifféremment appelés *points* (et alors notés A , B , etc.) ou *vecteurs* (et alors notés \vec{u} , \vec{v} , etc.).

Si A et B désignent deux points de E , le vecteur \vec{u} pour lequel $B = A + \vec{u}$ est noté \overrightarrow{AB} . Il correspond au vecteur $B - A$.

Définition

Si \vec{u} est un vecteur de E , on appelle *translation* de vecteur \vec{u} l'application $t_{\vec{u}}$ de E vers E qui envoie le point M sur le point $M' = M + \vec{u}$ déterminé par $\overrightarrow{MM'} = \vec{u}$.

L'usage veut que l'on note parfois pareillement points et vecteurs même lorsque nous usons de l'interprétation affine... La translation de vecteur a détermine alors simplement l'application $x \mapsto a + x$ de E vers lui-même.

7.5.2 Sous-espaces affines

Définition

On appelle *sous-espace affine* V de E toute partie obtenue par translation d'un sous-espace vectoriel F de E .

Le sous-espace vectoriel F déterminant V est alors unique et se nomme la *direction* de V . La dimension de F définit la *dimension* du sous-espace affine V . Si $\dim V = 1$, on dit que V est une *droite affine*. Si $\dim V = 2$, on dit que V est un *plan affine*. Lorsque $\dim V = 0$, l'ensemble V est réduit à un point.

Il suffit de connaître un point et la direction d'un sous-espace affine pour le décrire :

Théorème 18

Si V est un sous-espace affine de direction F et si a est un élément de V ,

$$V = a + F = t_a(F) = \{a + x \mid x \in F\}.$$

Théorème 19

Si V et W sont des sous-espaces affines de directions F et G , leur intersection $V \cap W$ est, soit vide, soit égale à un sous-espace affine de direction $F \cap G$.

7.6 Exercices d'apprentissage

7.6.1 Sous-espaces vectoriels

Exercice 1

Soit $F = \{(x, y, z) \in \mathbb{R}^3 \mid x - y - z = 0\}$ et $G = \{(a + b, a, a + 3b) \mid a, b \in \mathbb{R}\}$.

- (a) Montrer que F et G sont des sous-espaces vectoriels de \mathbb{R}^3 .
- (b) Déterminer $F \cap G$.

Solution

(a) méthode

On vérifie que F est une partie non vide de \mathbb{R}^3 stable par combinaison linéaire.

L'ensemble F est une partie de l'espace vectoriel réel \mathbb{R}^3 et celle-ci est non vide¹ car le vecteur nul $0_{\mathbb{R}^3} = (0, 0, 0)$ en est élément.

Soit u et v deux éléments de F et λ, μ deux réels. Étudions l'appartenance à F de la combinaison linéaire $\lambda u + \mu v$.

1. Si F s'avère être un sous-espace vectoriel, le vecteur nul en est obligatoirement élément.

On peut écrire $u = (x, y, z)$ et $v = (x', y', z')$ avec $x - y - z = 0$ et $x' - y' - z' = 0$. On a alors $\lambda u + \mu v = (x'', y'', z'')$ avec $x'' = \lambda x + \mu x'$, $y'' = \lambda y + \mu y'$ et $z'' = \lambda z + \mu z'$. On vérifie aisément

$$x'' - y'' - z'' = \lambda(x - y - z) + \mu(x' - y' - z') = 0.$$

Ceci permet d'affirmer que $\lambda u + \mu v$ appartient à F et l'on peut alors conclure que F est un sous-espace vectoriel¹ de \mathbb{R}^3 .

méthode

On peut reproduire la démonstration ci-dessus pour étudier G mais aussi, avec plus d'efficacité, percevoir G égal à un « Vect », c'est-à-dire à un espace vectoriel engendré.

En écrivant

$$(a+b, a, a+3b) = (a, a, a) + (b, 0, 3b) = \underbrace{a(1, 1, 1)}_{=e_1} + \underbrace{b(1, 0, 3)}_{=e_2}$$

on exprime les éléments de G comme des combinaisons linéaires des vecteurs e_1 et e_2 introduits. On en déduit $G = \text{Vect}(e_1, e_2)$ ce qui assure que G est un sous-espace vectoriel de \mathbb{R}^3 (Th. 6 p. 240).

(b) Le vecteur $u = (x, y, z)$ appartient à $F \cap G$ si, et seulement si, il existe $a, b \in \mathbb{R}$ tels que

$$(x, y, z) = (a+b, a, a+3b) \quad \text{et} \quad x - y - z = 0.$$

Ceci conduit à étudier le système

$$\begin{cases} x = a + b \\ y = a \\ z = a + 3b \\ x - y - z = 0 \end{cases}$$

On résout celui-ci en exprimant x, y, z parmi les inconnues principales :

$$\begin{cases} x = a + b \\ y = a \\ z = a + 3b \\ x - y - z = 0 \end{cases} \iff \begin{cases} x = a + b \\ y = a \\ z = a + 3b \\ a + 2b = 0 \end{cases} \iff \begin{cases} x = -b \\ y = -2b \\ z = b \\ a = -2b \end{cases}$$

Finalement,

$$F \cap G = \{(-b, -2b, b) \mid b \in \mathbb{R}\} = \text{Vect}(-1, -2, 1) = \text{Vect}(1, 2, -1).$$

1. On peut aussi résoudre l'équation $x - y - z = 0$ et affirmer que F est l'ensemble des combinaisons linéaires des vecteurs $(1, 1, 0)$ et $(1, 0, 1)$, autrement dit, $F = \text{Vect}((1, 1, 0), (1, 0, 1))$.

Exercice 2

Soit F , G et H des sous-espaces vectoriels d'un espace vectoriel E . Montrer

$$F \cap (G + H) \supset (F \cap G) + (F \cap H) \quad \text{et} \quad F + (G \cap H) \subset (F + G) \cap (F + H).$$

Vérifier à l'aide d'une figure que ces inclusions peuvent être strictes.

Solution**(a) méthode**

Il ne faut pas confondre la somme de deux sous-espaces vectoriels avec leur union. Un vecteur appartient à $F + G$ lorsque l'on peut l'écrire $x + y$ avec x dans F et y dans G .

Soit $u \in (F \cap G) + (F \cap H)$. On peut écrire $u = x + y$ avec $x \in F \cap G$ et $y \in F \cap H$. On a alors $u = x + y \in F$ par addition dans le sous-espace vectoriel F . On a aussi $u \in G + H$ car u est la somme d'un élément de G et d'un élément de H . On en déduit l'inclusion

$$(F \cap G) + (F \cap H) \subset F \cap (G + H).$$

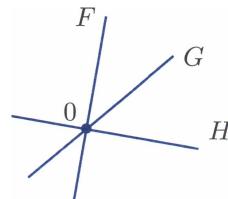
Étudions la deuxième inclusion. Soit $u \in F + (G \cap H)$. On peut écrire $u = x + y$ avec $x \in F$ et $y \in G \cap H$. D'une part, $u \in F + G$ car $x \in F$ et $y \in G$. D'autre part, $u \in F + H$ par un argument analogue. On en déduit $u \in (F + G) \cap (F + H)$ et l'on peut affirmer

$$F + (G \cap H) \subset (F + G) \cap (F + H).$$

Enfin, considérons trois droites vectorielles en position générale¹ dans le plan \mathbb{R}^2 comme illustré ci-contre.

D'une part,

$$\begin{aligned} (F \cap G) + (F \cap H) &= \{0\} + \{0\} = \{0\} \quad \text{et} \\ F \cap (G + H) &= F \cap \mathbb{R}^2 = F. \end{aligned}$$



D'autre part,

$$F + (G \cap H) = F + \{0\} = F \quad \text{et} \quad (F + G) \cap (F + H) = \mathbb{R}^2 \cap \mathbb{R}^2 = \mathbb{R}^2.$$

Il se peut donc que les inclusions soient strictes.

Exercice 3

Soit A et B deux parties d'un espace vectoriel E . Établir

$$\text{Vect}(A \cup B) = \text{Vect}(A) + \text{Vect}(B).$$

1. Bien qu'en position générale, ces trois droites vectorielles passent par le vecteur nul.

Solution**méthode**

\parallel $\text{Vect}(A \cup B)$ est un sous-espace vectoriel contenant A et contenant B . Aussi, $\text{Vect}(A \cup B)$ est inclus dans tout sous-espace vectoriel contenant à la fois A et B .

Par opérations sur les sous-espaces vectoriels, on peut affirmer que $\text{Vect}(A) + \text{Vect}(B)$ est un sous-espace vectoriel de E (Th. 4 p. 238). De plus, celui-ci contient $\text{Vect}(A)$ et donc contient A . Aussi, il contient B et donc $A \cup B \subset \text{Vect}(A) + \text{Vect}(B)$. On en déduit (Th. 6 p. 240)

$$\text{Vect}(A \cup B) \subset \text{Vect}(A) + \text{Vect}(B).$$

Inversement, A est inclus dans $A \cup B$ donc dans $\text{Vect}(A \cup B)$. On en déduit que $\text{Vect}(A)$ est inclus dans $\text{Vect}(A \cup B)$. Pareillement, on montre $\text{Vect}(B) \subset \text{Vect}(A \cup B)$. Considérons ensuite x un vecteur de $\text{Vect}(A) + \text{Vect}(B)$. On peut écrire $x = a + b$ avec $a \in \text{Vect}(A)$ et $b \in \text{Vect}(B)$. Ces deux vecteurs a et b appartiennent à $\text{Vect}(A \cup B)$ et donc, par addition dans ce sous-espace vectoriel, on peut affirmer que x appartient aussi à $\text{Vect}(A \cup B)$. Ainsi, on peut affirmer l'inclusion

$$\text{Vect}(A) + \text{Vect}(B) \subset \text{Vect}(A \cup B)$$

et conclure à l'égalité¹.

Exercice 4

Soit $F = \{f \in \mathcal{C}([0;1], \mathbb{R}) \mid f(0) = f(1) = 0\}$ et $G = \{g \in \mathcal{C}([0;1], \mathbb{R}) \mid g \text{ est affine}\}$. Montrer que F et G sont des sous-espaces vectoriels supplémentaires de l'espace réel $E = \mathcal{C}([0;1], \mathbb{R})$ des fonctions continues de $[0;1]$ vers \mathbb{R} .

Solution**méthode**

\parallel On commence² par vérifier que F et G sont des sous-espaces vectoriels de E .

L'ensemble F est une partie non vide de E car la fonction nulle en est élément. L'ensemble F est aussi stable par combinaison linéaire car si f_1 et f_2 sont des fonctions de E qui s'annulent en 0 et 1, la fonction $\lambda_1 f_1 + \lambda_2 f_2$ s'annule aussi en 0 et 1 et ce, quelles que soient les valeurs des réels λ_1 et λ_2 . On peut donc affirmer que F est un sous-espace vectoriel de E .

L'ensemble G est aussi un sous-espace vectoriel de E . En effet, une fonction affine g est de la forme $x \mapsto ax + b$ et se perçoit comme une combinaison linéaire des fonctions $g_0: x \mapsto 1$ et $g_1: x \mapsto x$ en écrivant $g = ag_1 + bg_0$. Ainsi, $G = \text{Vect}(g_0, g_1)$ avec g_0 et g_1 deux fonctions de E .

1. En particulier, si A et B sont des sous-espaces vectoriels F et G , l'égalité donne $\text{Vect}(F \cup G)$ égal à $F + G$: la somme de deux sous-espaces vectoriels détermine le plus petit sous-espace vectoriel contenant chacun.

2. Il arrive parfois que l'énoncé soit introduit par « On considère les espaces ... ». Dans ce cas il n'est pas nécessaire de vérifier que les parties sont des espaces puisque cela est « donné » par le sujet.

méthode

On montre que F et G sont supplémentaires dans E en étudiant par analyse-synthèse comment il est possible de décomposer un élément de E en la somme d'un élément de F et d'un élément de G : l'analyse produit l'unicité de l'écriture, la synthèse donne l'existence.

Soit $h \in E$.

Analyse : Supposons pouvoir écrire $h = f + g$ avec $f \in F$ et $g \in G$. On souhaite déterminer f et g en fonction de h . D'une part, on peut introduire a, b réels tels que $g: x \mapsto ax + b$. D'autre part, on sait $f(0) = f(1) = 0$. On en déduit

$$h(0) = f(0) + g(0) = b \quad \text{et} \quad h(1) = f(1) + g(1) = a + b$$

et donc $b = h(0)$ et $a = h(1) - h(0)$. Ceci détermine entièrement g puis f car $f = h - g$. L'analyse est close : s'il est possible de décomposer h en $f + g$, cette décomposition est déterminée de façon unique.

Synthèse : Considérons les fonctions f et g de E déterminées par le terme de l'analyse

$$g: x \mapsto (h(1) - h(0))x + h(0) \quad \text{et} \quad f = h - g.$$

Par ces définitions, il est clair que g est élément de G et que h est la somme de f et g . Il reste seulement à vérifier que f est élément de F ce qui s'acquiert par le petit calcul suivant :

$$f(0) = h(0) - g(0) = 0 \quad \text{et} \quad f(1) = h(1) - g(1) = 0.$$

On peut donc affirmer qu'il est possible d'écrire un élément de E comme somme d'un élément de F et d'un élément de G .

Finalement, F et G sont des sous-espaces vectoriels supplémentaires de E .

7.6.2 Liberté d'une famille de vecteurs

Exercice 5

Soit E un espace vectoriel réel.

(a) Soit x, y, z trois vecteurs de E constituant une famille libre. On pose $u = x + y$, $v = y + z$ et $w = z + x$. Montrer la liberté de la famille (u, v, w) .

(b) Soit x, y, z, t des vecteurs de E constituant une famille libre. On pose $u = x + y$, $v = y + z$, $w = z + t$ et $s = t + x$. Étudier la liberté de la famille (u, v, w, s) .

Solution
(a) méthode

On montre la liberté d'une famille de vecteurs en supposant disposer d'une combinaison linéaire nulle d'éléments de cette famille et en établissant que tous les scalaires introduits sont alors nuls.

Supposons $\alpha u + \beta v + \gamma w = 0_E$ avec α, β, γ réels. On a alors

$$(\alpha + \gamma)x + (\alpha + \beta)y + (\beta + \gamma)z = 0_E.$$

La famille (x, y, z) étant supposée libre, on peut affirmer la nullité des scalaires facteurs des vecteurs x , y et z dans la relation ci-dessus. Ceci donne les équations du système ci-dessous

$$\begin{cases} \alpha + \gamma = 0 \\ \alpha + \beta = 0 \\ \beta + \gamma = 0. \end{cases}$$

La résolution du système constitué de ces trois équations donne $\alpha = \beta = \gamma = 0$ et l'on peut conclure que la famille (u, v, w) est libre.

(b) méthode

On montre qu'une famille est liée en révélant une relation linéaire sur les éléments de cette famille, c'est-à-dire une combinaison linéaire nulle avec des scalaires non tous nuls.

On observe $u + w = v + s$. On en déduit la relation linéaire $u - v + w - s = 0_E$ qui assure que la famille (u, v, w, s) est liée.

Exercice 6

Soit u et v deux vecteurs d'un \mathbb{K} -espace vectoriel E . On dit que le vecteur v est colinéaire à u si l'on peut écrire $v = \alpha u$ avec $\alpha \in \mathbb{K}$.

(a) On suppose la famille (u, v) liée. Montrer que u est colinéaire à v ou v colinéaire à u .

(b) À quelle condition simple sur u peut-on affirmer que, lorsque la famille (u, v) est liée, le vecteur v est colinéaire à u ?

Solution

(a) méthode

|| Lorsqu'une famille est liée, on peut introduire une relation linéaire.

Il existe deux scalaires λ, μ non tous deux nuls vérifiant $\lambda u + \mu v = 0_E$. Si $\lambda \neq 0$, on peut écrire $u = \alpha v$ avec $\alpha = -\mu/\lambda$. Si $\mu \neq 0$, on peut écrire $v = \alpha u$ avec $\alpha = -\lambda/\mu$. L'un ou l'autre¹ de ces deux cas étant satisfait, on peut affirmer que l'un des vecteurs est colinéaire à l'autre. On ne peut cependant pas affirmer *a priori* lequel ce qui suscite l'intérêt de la question qui suit.

(b) méthode

|| Si l'un des vecteurs est non nul, celui-ci détermine une « direction » à laquelle doit appartenir l'autre vecteur lorsque la famille est liée.

Si le vecteur u est non nul, le scalaire μ de la relation linéaire précédente ne peut être nul. En effet, la relation $\lambda u + \mu v = 0_E$ devient sinon $\lambda u = 0_E$ ce qui entraîne $\lambda = 0$ sachant u non nul. Ceci est absurde car λ et μ ne sont pas tous deux nuls. Sachant μ non nul, on peut comme au-dessus obtenir que v est colinéaire à u .

1. Voir souvent les deux...

Exercice 7

Soit E un espace vectoriel réel de dimension 3 et $e = (e_1, e_2, e_3)$ une base de E . On pose

$$e'_1 = e_2 + 2e_3, \quad e'_2 = e_1 + e_3 \quad \text{et} \quad e'_3 = e_1 + 2e_2.$$

Montrer que la famille $e' = (e'_1, e'_2, e'_3)$ est une base de E et déterminer les coordonnées du vecteur $x = e_1 + e_2 + e_3$ dans les bases e et e' .

Solution**méthode**

En dimension connue, on peut montrer qu'une famille est une base d'un espace E en observant que c'est une famille libre constituée de $\dim E$ vecteurs de E (Th. 12 p. 243).

Etudions la liberté de la famille e' . Supposons $\lambda_1 e'_1 + \lambda_2 e'_2 + \lambda_3 e'_3 = 0_E$ avec λ_1, λ_2 et λ_3 des réels. En exprimant les vecteurs de e' en fonction des vecteurs de e il vient

$$(\lambda_2 + \lambda_3)e_1 + (\lambda_1 + 2\lambda_3)e_2 + (2\lambda_1 + \lambda_2)e_3 = 0_E.$$

La famille (e_1, e_2, e_3) étant libre, les scalaires de cette combinaison linéaire sont tous nuls ce qui produit le système

$$\begin{cases} \lambda_2 + \lambda_3 = 0 \\ \lambda_1 + 2\lambda_3 = 0 \\ 2\lambda_1 + \lambda_2 = 0. \end{cases}$$

Après résolution, on conclut $\lambda_1 = \lambda_2 = \lambda_3 = 0$. La famille e' est donc une famille libre et, puisque celle-ci est constituée de $3 = \dim E$ vecteurs de E , c'est une base de E .

Les coordonnées du vecteur x dans la base e se lisent directement sur l'écriture le définissant : ce sont 1, 1, 1.

L'obtention des coordonnées du vecteur x dans la base e' nécessite la résolution de l'équation $x = x_1 e'_1 + x_2 e'_2 + x_3 e'_3$ en les inconnues réelles x_1, x_2 et x_3 . En exprimant x et les vecteurs de e' en fonction des vecteurs de la famille e , l'équation précédente se relit

$$e_1 + e_2 + e_3 = (x_2 + x_3)e_1 + (x_1 + 2x_3)e_2 + (2x_1 + x_2)e_3.$$

Par identification¹ des scalaires en facteur des vecteurs de la famille libre (e_1, e_2, e_3) , on exprime le système

$$\begin{cases} x_2 + x_3 = 1 \\ x_1 + 2x_3 = 1 \\ 2x_1 + x_2 = 1. \end{cases}$$

Après résolution, on obtient $x_1 = 1/5$, $x_2 = 3/5$ et $x_3 = 2/5$.

1. Cette identification est possible car la différence des membres exprime une combinaison linéaire nulle dont tous les scalaires doivent être nuls.

Exercice 8

Soit E l'ensemble des fonctions $f: \mathbb{R} \rightarrow \mathbb{R}$ telles qu'il existe des réels a, b, c, d pour lesquels :

$$f(x) = (ax + b) \cos x + (cx + d) \sin x \quad \text{pour tout } x \in \mathbb{R}.$$

Montrer que E est sous-espace vectoriel de $\mathcal{F}(\mathbb{R}, \mathbb{R})$ dont on déterminera la dimension.

Solution**méthode**

La dimension d'un espace vectoriel correspond souvent « au nombre de degrés de liberté » dont on dispose pour définir un de ses éléments. Plus exactement, on calcule la dimension d'un espace vectoriel en déterminant une base de celui-ci.

Les éléments de E sont les combinaisons linéaires des fonctions

$$c_0: x \mapsto \cos x, \quad c_1: x \mapsto x \cos x, \quad s_0: x \mapsto \sin x \quad \text{et} \quad s_1: x \mapsto x \sin x.$$

On en déduit

$$E = \text{Vect}(c_0, c_1, s_0, s_1).$$

Il s'agit donc d'un sous-espace vectoriel de l'espace $\mathcal{F}(\mathbb{R}, \mathbb{R})$ et la famille (c_0, c_1, s_0, s_1) est génératrice de celui-ci. Vérifions que c'est aussi une famille libre. Supposons

$$ac_1 + bc_0 + cs_1 + ds_0 = 0$$

avec a, b, c, d réels et où le 0 en second membre désigne la fonction nulle. On a donc, pour tout réel x ,

$$(ax + b) \cos x + (cx + d) \sin x = 0. \tag{*}$$

méthode

On particularise l'équation (*) pour différentes valeurs de x afin de former un système d'inconnues (a, b, c, d) dont la seule solution est nulle.

En choisissant x respectivement égal à 0, π , $\pi/2$ et $-\pi/2$, on obtient les équations du système suivant :

$$\begin{cases} b = 0 \\ a\pi + b = 0 \\ c\pi + 2d = 0 \\ c\pi - 2d = 0 \end{cases}$$

La résolution de ce système donne facilement $a = b = c = d = 0$. On en déduit que la famille (c_0, c_1, s_0, s_1) est libre et c'est donc une base de E . On peut conclure $\dim E = 4$.

7.7 Exercices d'entraînement

7.7.1 Généralités sur les espaces vectoriels

Exercice 9 *

On dit qu'une fonction $f: \mathbb{R} \rightarrow \mathbb{R}$ est à support compact s'il existe $A \in \mathbb{R}_+$ tel que f est nulle en dehors de $[-A; A]$. Vérifier que l'ensemble \mathcal{D} des fonctions de \mathbb{R} vers \mathbb{R} de classe C^∞ à support compact est un espace vectoriel réel pour les lois usuelles.

Solution

méthode

Lorsque les lois sont connues, on peut montrer qu'un ensemble est un espace vectoriel en observant qu'il s'agit d'un sous-espace vectoriel d'un espace déjà connu (Th. 3 p. 238).

On vérifie que \mathcal{D} est un sous-espace vectoriel¹ de l'espace réel $\mathcal{F}(\mathbb{R}, \mathbb{R})$ des fonctions de \mathbb{R} vers \mathbb{R} .

\mathcal{D} est une partie de $\mathcal{F}(\mathbb{R}, \mathbb{R})$ et celle-ci est non vide car la fonction nulle en est élément².

Soit f et g deux fonctions éléments de \mathcal{D} et λ, μ deux réels. La fonction $\lambda f + \mu g$ est de classe C^∞ par opérations sur les fonctions qui le sont. Vérifions qu'elle est aussi à support compact. Les fonctions f et g l'étant, on peut introduire $A \in \mathbb{R}_+$ et $B \in \mathbb{R}_+$ vérifiant

$$\forall t \in \mathbb{R}, \quad (|t| > A \implies f(t) = 0) \text{ et } (|t| > B \implies g(t) = 0).$$

Pour $C = \max(A, B)$, on a alors

$$\forall t \in \mathbb{R}, \quad |t| > C \implies (\lambda f + \mu g)(t) = 0.$$

Ainsi, la fonction $\lambda f + \mu g$ est à support compact et c'est donc un élément de \mathcal{D} .

On peut conclure que \mathcal{D} est un sous-espace vectoriel de $\mathcal{F}(\mathbb{R}, \mathbb{R})$ et c'est donc un espace réel.

Exercice 10 *

Dans l'espace réel \mathbb{R}^3 :

- Compléter en une base la famille (u, v) avec $u = (1, 1, 1)$ et $v = (0, 1, 1)$.
- Déterminer un sous-espace supplémentaire de $F = \text{Vect}((1, 1, 0), (0, 1, -1))$.
- Déterminer une base du sous-espace $G = \{(x, y, z) \in \mathbb{R}^3 \mid x - z = 0\}$.
- Déterminer une base de \mathbb{R}^3 adaptée au sous-espace G précédent.

1. On pourrait aussi étudier \mathcal{D} sous-espace vectoriel de l'espace $C^\infty(\mathbb{R}, \mathbb{R})$ des fonctions de classe C^∞ .

2. Un exemple moins trivial de fonction C^∞ à support compact est donné dans le sujet 12 du chapitre 8 de l'ouvrage *Exercices d'analyse MPSI*.

Solution

(a) Les vecteurs u et v ne sont pas colinéaires et constituent une famille libre.

méthode

|| Toute famille libre peut être complétée en une base (Th. 10 p. 242). Les vecteurs complétant peuvent être choisis à l'intérieur d'une famille génératrice.

L'espace \mathbb{R}^3 étant de dimension 3, il suffit d'un vecteur pour compléter (u, v) en une base. On choisit ce vecteur parmi les vecteurs $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ et $e_3 = (0, 0, 1)$ de la base canonique. Le vecteur e_1 ne convient pas car $u - v = e_1$. Le vecteur e_2 convient. En effet¹, en opérant² sur les vecteurs de l'espace engendré

$$\text{Vect}(u, v, e_2) = \text{Vect}(\underbrace{e_1}_{u-v}, v, e_2) = \text{Vect}(e_1, \underbrace{e_3}_{v-e_2}, e_2) = \mathbb{R}^3.$$

Ceci assure que la famille (u, v, e_2) est génératrice et c'est donc une base de \mathbb{R}^3 . Le vecteur e_3 aurait aussi pu convenir et, plus généralement, tout vecteur n'appartenant pas à $\text{Vect}(u, v)$ convient.

(b) Les vecteurs $u_1 = (1, 1, 0)$ et $u_2 = (0, 1, -1)$ ne sont pas colinéaires et forment donc une base de l'espace F .

méthode

|| Les vecteurs complétant une base d'un sous-espace vectoriel engendrent un supplémentaire de celui-ci.

On peut compléter la famille (u_1, u_2) en une base à l'aide du vecteur e_1 de la base canonique car

$$\text{Vect}(u_1, u_2, e_1) = \text{Vect}(\underbrace{e_2}_{=u_3-e_3}, u_2, e_1) = \text{Vect}(e_2, \underbrace{-e_3}_{u_2-e_2}, e_1) = \mathbb{R}^3.$$

La droite $D = \text{Vect}(e_1)$ définit alors un supplémentaire du plan F . Plus généralement, n'importe quelle droite vectorielle qui n'est pas incluse dans F détermine un supplémentaire du plan F .

(c) méthode

|| On résout l'équation définissant G afin d'exprimer cet ensemble comme un espace engendré.

L'équation $x - z = 0$ d'inconnue $(x, y, z) \in \mathbb{R}^3$ a pour solution les triplets

$$(x, y, z) = x(1, 0, 1) + y(0, 1, 0) \quad \text{avec } x, y \in \mathbb{R}.$$

1. On peut aussi vérifier que la famille est libre en étudiant la nullité d'une combinaison linéaire.

2. On ne modifie pas l'espace engendré par une famille de vecteurs lorsque l'on permute ses vecteurs, lorsque l'on ajoute à un vecteur une combinaison linéaire des autres ou lorsque l'on multiplie un de ses vecteurs par un scalaire non nul.

L'espace G est donc l'ensemble des combinaisons linéaires des vecteurs $v_1 = (1, 0, 1)$ et $v_2 = (0, 1, 0)$. Ces derniers étant linéairement indépendants, ils forment une base de G .

(d) méthode

|| Une base adaptée à un sous-espace vectoriel est une base de l'espace dont les premiers vecteurs constituent une base du sous-espace.

Il suffit de compléter la famille (v_1, v_2) en une base de \mathbb{R}^3 . Le vecteur e_1 de la base canonique convient et (v_1, v_2, e_1) est une base de \mathbb{R}^3 adaptée à G .

Exercice 11 **

Soit F et G deux sous-espaces vectoriels d'un espace vectoriel réel E . Montrer que $F \cup G$ est un sous-espace vectoriel de E si, et seulement si¹, $F \subset G$ ou $G \subset F$.

Solution

Si $F \subset G$, la réunion de F et G est égale à G et c'est donc un sous-espace vectoriel de E . On conclut de même si $G \subset F$.

méthode

|| On établit la réciproque en raisonnant par contraposée.

Si $F \not\subset G$ et $G \not\subset F$, il existe des vecteurs $x \in F$ et $y \in G$ tels que $x \notin G$ et $y \notin F$. Étudions alors le vecteur $x + y$. Celui-ci ne peut appartenir à F car²

$$x + y \in F \implies y = (x + y) - x \in F$$

ce qui est exclu. De même, le vecteur $x + y$ ne peut être élément de G . On en déduit que la partie $F \cup G$ n'est pas stable pour l'addition des vecteurs :

$$x, y \in F \cup G \quad \text{et} \quad x + y \notin F \cup G.$$

Ainsi, $F \cup G$ n'est pas un sous-espace vectoriel.

Exercice 12 ***

Soit E un espace vectoriel de dimension finie n .

(a) Soit F et G deux sous-espaces vectoriels de E vérifiant $\dim F + \dim G > n$. Montrer que $F \cap G$ n'est pas réduit au vecteur nul.

(b) Généraliser ce résultat à plusieurs sous-espaces vectoriels F_1, \dots, F_p de E .

1. Cette étude est généralisée dans le sujet 31 p. 272.

2. Notons que l'appartenance de $x + y$ à F n'implique pas à elle seule l'appartenance de y à F : deux vecteurs peuvent avoir une somme horizontale sans qu'aucun des deux le soit. Ici, on obtient $y \in F$ par argument d'opération en employant $x + y \in F$ et $x \in F$.

Solution**(a) méthode**

La formule de Grassmann (Th. 15 p. 243) relie la dimension d'une somme et celle d'une intersection.

L'espace somme $F + G$ est inclus dans E et donc $\dim(F + G) \leq n$. La formule de Grassmann donne alors

$$\dim(F \cap G) = \underbrace{\dim F + \dim G}_{\geq n} - \underbrace{\dim(F + G)}_{\leq n} > 0.$$

Ceci assure que $F \cap G$ n'est pas réduit au vecteur nul.

(b) On généralise ce qui précède en établissant, pour F_1, \dots, F_p des sous-espaces vectoriels de E , la propriété

$$\sum_{j=1}^p \dim F_j > (p-1)n \implies \bigcap_{j=1}^p F_j \neq \{0_E\}.$$

méthode

Par récurrence sur $p \in \mathbb{N}^*$, on vérifie :

$$\dim\left(\bigcap_{j=1}^p F_j\right) \geq \sum_{j=1}^p \dim F_j - n(p-1).$$

La propriété est entendue pour $p = 1$.

Supposons la propriété vraie au rang $p \geq 1$ et considérons F_1, \dots, F_p et F_{p+1} des sous-espaces vectoriels de E . Introduisons les espaces $F = F_1 \cap \dots \cap F_p$ et $G = F_{p+1}$. Comme au-dessus, $F + G$ est de dimension inférieure à n et la formule de Grassmann donne

$$\dim(F \cap G) = \dim F + \dim G - \dim(F + G) \geq \dim F + \dim G - n.$$

On en déduit par l'hypothèse de récurrence

$$\dim\left(\bigcap_{j=1}^{p+1} F_j\right) \geq \dim\left(\bigcap_{j=1}^p F_j\right) + \dim F_{p+1} - n \geq \sum_{j=1}^p \dim F_j - np.$$

La récurrence est établie et la conclusion est dès lors immédiate

$$\begin{aligned} \sum_{j=1}^p \dim F_j > (p-1)n &\implies \dim\left(\bigcap_{j=1}^p F_j\right) > 0 \\ &\implies \bigcap_{j=1}^p F_j \neq \{0_E\}. \end{aligned}$$

Exercice 13 ***

Soit $p \in \mathbb{N}^*$ et E_p l'ensemble des suites complexes p -périodiques, c'est-à-dire l'ensemble des suites¹ $u = (u(n))$ vérifiant $u(n+p) = u(n)$ pour tout naturel n .

(a) Montrer que E_p est un \mathbb{C} -espace vectoriel de dimension finie et calculer celle-ci.

(b) Déterminer une base de E_p formée uniquement de suites géométriques.

Solution

(a) On vérifie que E_p est un sous-espace vectoriel de l'espace $\mathbb{C}^{\mathbb{N}}$ des suites complexes.

L'ensemble E_p est une partie non vide de $\mathbb{C}^{\mathbb{N}}$ car la suite nulle est évidemment périodique. Si u et v sont deux suites p -périodiques et si λ, μ désignent deux complexes, la suite $\lambda u + \mu v = (\lambda u(n) + \mu v(n))$ est p -périodique car $\lambda u(n+p) + \mu v(n+p) = \lambda u(n) + \mu v(n)$ pour tout naturel n .

méthode

On détermine la dimension de E_p en exhibant une base de cet espace : ceci invite à réfléchir sur les « degrés de liberté » possibles lors de la construction d'un élément de E_p .

Une suite u de E_p est déterminée par ses premières valeurs $u(0), \dots, u(p-1)$, les suivantes se déduisant par périodicité. On introduit une famille de suites $c = (c_0, \dots, c_{p-1})$ traduisant le choix de ces valeurs². Pour tout $0 \leq i \leq p-1$, on note e_i la suite définie par

$$e_i(n) = \begin{cases} 1 & \text{si } n = i \ [p] \\ 0 & \text{sinon.} \end{cases}$$

Les suites e_i appartiennent toutes à l'espace E_p et sont linéairement indépendantes car, si $\lambda_0 e_0 + \dots + \lambda_{p-1} e_{p-1}$ est la suite nulle, les λ_i sont tous nuls puisqu'ils correspondent aux premières valeurs de cette suite. Enfin, cette famille est génératrice car, pour tout suite u de E_p , on peut vérifier

$$u = \sum_{i=0}^{p-1} u(i) e_i.$$

Finalement, la famille $c = (c_0, \dots, c_{p-1})$ est une base de E_p ce qui permet d'affirmer que l'espace E_p est de dimension p .

(b) méthode

On commence par déterminer les suites géométriques éléments de E_p .

1. Dans ce sujet, on adopte une notation fonctionnelle des termes de la suite en écrivant $u(n)$ au lieu de u_n .

2. Dans le chapitre suivant, on introduit la notion d'isomorphisme : celle-ci permet de calculer la dimension de E_p en introduisant l'application qui à la suite u associe $(u(0), \dots, u(p-1)) \in \mathbb{C}^p$.

Soit $q \in \mathbb{C}^*$. La suite géométrique¹ (q^n) est élément de E_p si, et seulement si, $q^{n+p} = q^n$ pour tout naturel n . Ceci équivaut à affirmer que q est une racine p -ième de l'unité. On sait qu'il y a exactement p racines de l'unité distinctes ce qui invite à introduire les p suites géométriques correspondantes.

Pour $j \in \llbracket 0; p-1 \rrbracket$, considérons la suite géométrique u_j déterminée par $u_j(n) = \omega_j^n$ avec $\omega_j = e^{2i\pi j/p}$ racine p -ième de l'unité. La famille (u_0, \dots, u_{p-1}) est constituée par exactement $p = \dim E_p$ éléments de E_p , il suffit de vérifier qu'elle est libre pour affirmer que c'est une base.

Supposons

$$\lambda_0 u_0 + \dots + \lambda_{p-1} u_{p-1} = 0$$

avec $\lambda_0, \dots, \lambda_{p-1}$ des nombres complexes. Pour tout naturel n , on a l'équation

$$\sum_{j=0}^{p-1} \lambda_j \omega_j^n = 0. \quad (*)$$

méthode

$\parallel 1 + \omega + \dots + \omega^{p-1} = 0$ lorsque ω est une racine p -ième de l'unité distincte de 1.

Soit $k \in \llbracket 0; p-1 \rrbracket$ fixé. En multipliant l'équation (*) par ω_k^{-n} et en sommant les équations obtenues pour n allant de 0 à $p-1$, on obtient

$$\sum_{n=0}^{p-1} \left(\sum_{j=0}^{p-1} \lambda_j (\omega_j \omega_k^{-1})^n \right) = 0.$$

En échangeant les deux sommes

$$\sum_{j=0}^{p-1} \left(\lambda_j \sum_{n=0}^{p-1} (\omega_j \omega_k^{-1})^n \right) = 0 \quad (**)$$

où la somme contenue est géométrique de raison $\omega_j \omega_k^{-1}$. Celle-ci est une racine p -ième de l'unité et, lorsque $j \neq k$, elle est différente de 1, de sorte que

$$\sum_{n=0}^{p-1} (\omega_j \omega_k^{-1})^n = \frac{1 - (\omega_j \omega_k^{-1})^p}{1 - \omega_j \omega_k^{-1}} = 0 \text{ si } j \neq k \quad \text{et} \quad \sum_{n=0}^{p-1} (\omega_j \omega_k^{-1})^n = p \text{ si } j = k.$$

L'équation (**) se relit alors $p\lambda_k = 0$ ce qui donne $\lambda_k = 0$. On peut conclure à la liberté de la famille (u_0, \dots, u_{p-1}) constituant alors une base de E_p .

1. Une suite géométrique est plus généralement de la forme $(u_0 q^n)$ mais le facteur de colinéarité u_0 non nulle peut être omis dans cette étude : faire varier celui-ci n'influe pas sur la construction d'une base.

7.7.2 Liberté

Exercice 14 **

Soit $(u_1, \dots, u_n, u_{n+1})$ une famille de vecteurs d'un \mathbb{K} -espace vectoriel E .

(a) Etablir que si la famille (u_1, \dots, u_n) est libre et que si u_{n+1} n'appartient pas à $\text{Vect}(u_1, \dots, u_n)$ alors $(u_1, \dots, u_n, u_{n+1})$ est libre.

(b) Etablir que si la famille $(u_1, \dots, u_n, u_{n+1})$ est génératrice et que si u_{n+1} est élément de $\text{Vect}(u_1, \dots, u_n)$ alors (u_1, \dots, u_n) est génératrice.

Solution

(a) Soit $(\lambda_1, \dots, \lambda_n, \lambda_{n+1}) \in \mathbb{K}^{n+1}$. Supposons

$$\lambda_1 u_1 + \dots + \lambda_n u_n + \lambda_{n+1} u_{n+1} = 0_E. \quad (*)$$

méthode

|| Affirmer que la famille (u_1, \dots, u_n) est libre permet d'assurer que les scalaires $\lambda_1, \dots, \lambda_n$ sont nuls *une fois que l'on sait* λ_{n+1} nul.

Si $\lambda_{n+1} \neq 0$, il est possible d'écrire $u_{n+1} = \mu_1 u_1 + \dots + \mu_n u_n$ avec $\mu_i = -\lambda_i / \lambda_{n+1}$. Ceci est exclu car $u_{n+1} \notin \text{Vect}(u_1, \dots, u_n)$. On en déduit $\lambda_{n+1} = 0$. L'équation (*) se simplifie alors en $\lambda_1 u_1 + \dots + \lambda_n u_n = 0_E$ et donc $\lambda_1 = \dots = \lambda_n = 0$ car (u_1, \dots, u_n) est libre.

Finalement, tous les λ_i sont nuls et l'on peut conclure que la famille $(u_1, \dots, u_n, u_{n+1})$ est libre.

(b) méthode

|| Un vecteur de E est combinaison linéaire des u_1, \dots, u_n et de u_{n+1} et ce dernier vecteur est lui-même combinaison linéaire des u_1, \dots, u_n .

Soit x un vecteur de E . On peut écrire $x = \lambda_1 u_1 + \dots + \lambda_n u_n + \lambda_{n+1} u_{n+1}$ car la famille $(u_1, \dots, u_n, u_{n+1})$ est génératrice. Puisque $u_{n+1} \in \text{Vect}(u_1, \dots, u_n)$, on peut aussi écrire $u_{n+1} = \mu_1 u_1 + \dots + \mu_n u_n$ puis affirmer $x = \nu_1 u_1 + \dots + \nu_n u_n$ avec $\nu_i = \lambda_i + \lambda_{n+1} \mu_i$. Ainsi, x est combinaison linéaire de la famille (u_1, \dots, u_n) et, finalement, cette famille est génératrice.

Exercice 15 ***

Pour $a \in \mathbb{C}$, on note e_a l'application de \mathbb{R} vers \mathbb{C} définie par $e_a(t) = \exp(at)$.

Montrer que la famille $(e_a)_{a \in \mathbb{C}}$ est une famille libre d'éléments de l'espace $\mathcal{F}(\mathbb{R}, \mathbb{C})$.

Solution

méthode

|| On montre qu'une famille infinie est libre en vérifiant que toutes ses sous-familles finies le sont.

Par récurrence sur $n \in \mathbb{N}^*$, montrons que toute sous-famille à n éléments de $(e_a)_{a \in \mathbb{C}}$ est libre.

Pour $n = 1$, une sous-famille à un élément de $(e_a)_{a \in \mathbb{C}}$ est libre car aucune fonction de cette famille n'est nulle.

Supposons la propriété établie au rang $n - 1$ et considérons a_1, \dots, a_n, a_{n+1} des complexes deux à deux distincts. Supposons

$$\lambda_1 e_{a_1} + \cdots + \lambda_n e_{a_n} + \lambda_{n+1} e_{a_{n+1}} = 0 \quad (1)$$

avec $\lambda_1, \dots, \lambda_n, \lambda_{n+1} \in \mathbb{C}$. On dérive cette relation fonctionnelle sachant $(e_a)' = ae_a$:

$$a_1 \lambda_1 e_{a_1} + \cdots + a_n \lambda_n e_{a_n} + a_{n+1} \lambda_{n+1} e_{a_{n+1}} = 0. \quad (2)$$

La combinaison $a_{n+1}(1) - (2)$ simplifie le terme $e_{a_{n+1}}$ et donne

$$(a_{n+1} - a_1) \lambda_1 e_{a_1} + \cdots + (a_{n+1} - a_n) \lambda_n e_{a_n} = 0.$$

Par l'hypothèse de récurrence, et en exploitant que les a_i sont deux à deux distincts, on obtient $\lambda_1 = \cdots = \lambda_n = 0$ et l'on en déduit $\lambda_{n+1} = 0$.

La récurrence est établie et l'on peut affirmer que la famille $(e_a)_{a \in \mathbb{C}}$ est libre car toutes ses sous-familles finies le sont.

7.7.3 Supplémentarité

Exercice 16 *

Dans l'espace réel E des fonctions de \mathbb{R} vers \mathbb{R} , on introduit les espaces P et I constitués respectivement des fonctions paires et impaires. Montrer que ceux-ci sont supplémentaires dans E .

Solution

Inutile de vérifier que P et I sont des sous-espaces vectoriels de E car le sujet l'affirme¹.

méthode

Il n'est pas nécessaire² d'étudier si P et I sont en somme directe : on peut directement raisonner par analyse-synthèse.

Soit f une fonction élément de E .

Analyse : Supposons $f = g+h$ avec g et h des fonctions respectivement paire et impaire. Pour tout x réel, on a $f(x) = g(x) + h(x)$ mais aussi, par parité, $f(-x) = g(x) - h(x)$. La résolution du système formé par ces deux équations détermine g et h de façon unique

$$g(x) = \frac{f(x) + f(-x)}{2} \quad \text{et} \quad h(x) = \frac{f(x) - f(-x)}{2} \quad \text{pour tout } x \in \mathbb{R}.$$

1. Il n'est cependant pas difficile de vérifier que la fonction nulle est paire et qu'une combinaison linéaire de deux fonctions paires est une fonction paire.

2. Cependant, si le lecteur à l'intuition de la décomposition d'une fonction en la somme d'une fonction paire et d'une fonction impaire, il peut vérifier celle-ci et établir l'unicité de l'écriture en observant que les espaces sont en somme directe.

Synthèse : Les deux fonctions g et h proposées par les identités ci-dessus sont de somme égale à f et l'on constate aisément qu'elles ont les parités voulues.

On peut alors conclure que P et I sont des espaces supplémentaires de E .

Exercice 17 *

Soit E un espace vectoriel de dimension finie $n \geq 1$ et H un sous-espace vectoriel de E de dimension¹ $n - 1$. Montrer que si un vecteur a de E n'appartient pas à H alors $E = H \oplus \text{Vect}(a)$.

Solution

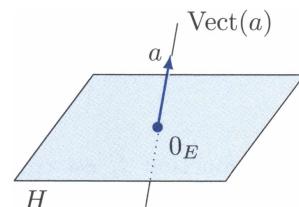
méthode

En dimension finie connue, on peut montrer que deux sous-espaces vectoriels sont supplémentaires en observant qu'ils sont en somme directe et que la somme de leurs dimensions vaut la dimension de l'espace (Th. 16 p. 244).

Le vecteur a n'appartenant pas à H , ce n'est pas le vecteur nul ce qui permet d'affirmer que l'espace $\text{Vect}(a)$ est de dimension² 1. On vérifie donc $\dim H + \dim \text{Vect}(a) = \dim E$. Montrons alors que les espaces H et $\text{Vect}(a)$ sont en somme directe en étudiant la nullité de leur intersection. Soit x un vecteur de l'intersection de H et $\text{Vect}(a)$. On a $x = \lambda a$ avec λ un scalaire. Ce dernier est nécessairement nul car sinon on peut écrire

$$a = \frac{1}{\lambda}x$$

ce qui établit que a est élément de H par opération dans le sous-espace vectoriel H . On a donc $x = 0_E$ et l'on peut affirmer³ $H \cap \text{Vect}(a) = \{0_E\}$. Les espaces H et $\text{Vect}(a)$ sont donc en somme directe et, par l'argument de dimension qui précède, ils sont supplémentaires dans E . La figure ci-contre illustre cette supplémentarité en dimension 3.



Exercice 18 **

Dans l'espace E des fonctions continues de $[-1; 1]$ vers \mathbb{R} , on considère les sous-espaces vectoriels

$$F_1 = \{f \in E \mid f \text{ est constante}\}, \quad F_2 = \{f \in E \mid \forall t \in [-1; 0], f(t) = 0\} \text{ et}$$

$$F_3 = \{f \in E \mid \forall t \in [0; 1], f(t) = 0\}.$$

Établir

$$E = F_1 \oplus F_2 \oplus F_3.$$

1. On dit que H est un *hyperplan*. Cette notion sera présentée et étudiée dans le chapitre 8. Le résultat en cours apparaîtra alors comme une conséquence immédiate du Th. 16 p. 279.

2. Vérifier la condition $a \neq 0_E$ assure que la famille (a) est libre et constitue donc une base de l'espace $\text{Vect}(a)$ ce qui permet d'affirmer que celui-ci est de dimension 1.

3. Précisément, on a plutôt établi $H \cap \text{Vect}(a) \subset \{0_E\}$ mais l'inclusion réciproque est entendue car H et $\text{Vect}(a)$ contiennent le vecteur nul puisque ce sont des sous-espaces vectoriels.

Solution**méthode**

On montre que des espaces F_1, \dots, F_n sont en somme directe en établissant que, si $x_1 + \dots + x_n = 0_E$ avec x_i dans F_i , alors chaque x_i est nul (Th. 5 p. 239).

Supposons

$$f_1 + f_2 + f_3 = 0 \quad \text{avec } f_i \in F_i \text{ pour } i \in \{1, 2, 3\}.$$

On a alors pour toute valeur de la variable x dans $[-1; 1]$

$$f_1(x) + f_2(x) + f_3(x) = 0.$$

En évaluant en 0, on obtient $f_1(0) = 0$ et la fonction constante f_1 est nulle.

En évaluant alors en $x \in]0; 1]$, on obtient $f_2(x) = 0$ et donc f_2 est nulle sur $]0; 1]$. Puisque f_2 est aussi nulle sur $[-1; 0]$, c'est la fonction nulle.

Enfin, les fonctions f_1 et f_2 étant nulles, la fonction f_3 l'est aussi. On peut alors affirmer que les espaces F_1, F_2 et F_3 sont en somme directe.

Soit $f \in E$. Décomposons f en une somme de fonctions des sous-espaces F_i . Posons¹

$$f_1: t \mapsto f(0), \quad f_2: t \mapsto \begin{cases} f(t) - f(0) & \text{si } t \in [0; 1] \\ 0 & \text{si } t \in [-1; 0] \end{cases} \quad \text{et}$$

$$f_3: t \mapsto \begin{cases} 0 & \text{si } t \in [0; 1] \\ f(t) - f(0) & \text{si } t \in [-1; 0]. \end{cases}$$

La fonction f_1 est constante. La fonction f_2 est continue, notamment en 0, et nulle sur $[-1; 0]$. La fonction f_3 est aussi continue et nulle sur $[0; 1]$. Enfin, on a évidemment f égale à la somme des f_i . On peut donc écrire

$$f = f_1 + f_2 + f_3 \quad \text{avec } f_i \in F_i \text{ pour tout } i \in \{1, 2, 3\}$$

et conclure

$$E = F_1 \oplus F_2 \oplus F_3.$$

Exercice 19 **

Soit F un sous-espace vectoriel d'un espace E de dimension finie et G un supplémentaire de F . On introduit $e = (e_1, \dots, e_p)$ une base de G et, pour $a = (a_1, \dots, a_p)$ une famille de vecteurs de F , on note

$$G_a = \text{Vect}(e_j + a_j)_{1 \leq j \leq p}.$$

Montrer que les espaces G_a déterminent tous les espaces supplémentaires de F .

1. Ces fonctions sont le fruit d'une « petite » analyse.

Solution

Commençons par établir que les espaces G_a sont des supplémentaires de F .

méthode

|| On détermine la dimension de G_a puis on vérifie que F et G_a sont en somme directe.

Etudions la liberté de la famille $(e_j + a_j)_{1 \leq j \leq p}$. Supposons

$$\lambda_1(e_1 + a_1) + \cdots + \lambda_p(e_p + a_p) = 0_E \quad \text{avec } \lambda_1, \dots, \lambda_p \text{ scalaires.}$$

En ordonnant les termes de cette identité, on écrit

$$\underbrace{\lambda_1 e_1 + \cdots + \lambda_p e_p}_{\in G} = -(\underbrace{\lambda_1 a_1 + \cdots + \lambda_p a_p}_{\in F}).$$

L'intersection des espaces F et G étant réduite au vecteur nul, on a $\lambda_1 e_1 + \cdots + \lambda_p e_p = 0_E$ puis $\lambda_1 = \cdots = \lambda_p = 0$ car la famille e est libre.

Ainsi, la famille $(e_j + a_j)_{1 \leq j \leq p}$ est libre et c'est donc une base de l'espace G_a . On en déduit $\dim G_a = p$.

Etudions ensuite l'intersection de F et G_a . Un vecteur x de G_a s'écrit

$$x = \lambda_1(e_1 + a_1) + \cdots + \lambda_p(e_p + a_p) \quad \text{avec } \lambda_1, \dots, \lambda_p \text{ scalaires.}$$

Si de plus ce vecteur appartient à F , on organise les termes de cette égalité

$$\underbrace{\lambda_1 e_1 + \cdots + \lambda_p e_p}_{\in G} = (x - (\lambda_1 a_1 + \cdots + \lambda_p a_p)) \in F$$

et l'on peut affirmer $\lambda_1 e_1 + \cdots + \lambda_p e_p = 0_E$ puis $\lambda_1 = \cdots = \lambda_p = 0$ et, enfin, $x = 0_E$.

Les espaces F et G_a sont donc en somme directe et vérifient $\dim F + \dim G_a = \dim E$, ce sont des espaces supplémentaires.

Vérifions ensuite que les espaces supplémentaires de F sont tous de la forme G_a pour une famille $a = (a_1, \dots, a_p)$ de vecteurs de F bien choisie. Soit G' un supplémentaire de F .

méthode

|| On décompose chaque vecteur e_j en somme d'un vecteur de F et d'un vecteur de G' .

Pour tout $j \in \llbracket 1 ; p \rrbracket$, on peut écrire $e_j = x_j + y_j$ avec $x_j \in F$ et $y_j \in G'$. Considérons alors l'espace G_a déterminé par la famille des $a_j = -x_j$. Pour tout $j \in \llbracket 1 ; p \rrbracket$, le vecteur $e_j + a_j$ est élément de G' et donc G_a est inclus dans G' . Au surplus, les espaces G' et G_a ont la même dimension car sont tous deux supplémentaires de F . Par inclusion et égalité des dimensions, on peut conclure $G' = G_a$ (Th. 13 p. 243).

Exercice 20 **

Soit E un \mathbb{K} -espace vectoriel de dimension finie et F_1, \dots, F_n des sous-espaces vectoriels de E vérifiant $F_1 + \dots + F_n = E$.

Montrer qu'il existe des sous-espaces vectoriels G_1, \dots, G_n tels que :

$$\forall 1 \leq i \leq n, \quad G_i \subset F_i \quad \text{et} \quad E = G_1 \oplus \dots \oplus G_n.$$

Solution**méthode**

On définit successivement chaque espace G_i comme supplémentaire dans F_i d'un espace calculé à partir des espaces G_1, \dots, G_{i-1} déjà obtenus.

Posons $G_1 = F_1$, G_2 un supplémentaire de $G_1 \cap F_2$ dans F_2 , et plus généralement, G_i un supplémentaire de $(G_1 + \dots + G_{i-1}) \cap F_i$ dans F_i . Cette construction est possible car, en dimension finie, tout sous-espace vectoriel admet un supplémentaire (Th. 14 p. 243).

Les espaces G_i sont par construction inclus dans F_i .

Montrons qu'ils sont en somme directe. Supposons $x_1 + \dots + x_n = 0_E$ avec $x_i \in G_i$. On a x_n élément de G_n , donc de F_n , et aussi $x_n = -(x_1 + \dots + x_{n-1})$ élément de la somme $G_1 + \dots + G_{n-1}$ donc de $(G_1 + \dots + G_{n-1}) \cap F_n$. Cet espace étant en somme directe avec G_n , on peut affirmer $x_n = 0_E$. On obtient alors la relation $x_1 + \dots + x_{n-1} = 0_E$ qui permet de reproduire le raisonnement pour établir successivement la nullité de chaque x_i . On obtient ainsi que les G_i sont en somme directe.

Montrons ensuite que la somme des G_i est égale à E . Soit $x \in E$. L'hypothèse d'étude permet d'écrire $x = x_1 + \dots + x_n$ avec $x_i \in F_i$. Les espaces $(G_1 + \dots + G_{i-1}) \cap F_i$ et G_i étant supplémentaires dans F_i , on peut décomposer le vecteur x_i

$$x_i = a_i + b_i \quad \text{avec} \quad a_i \in (G_1 + \dots + G_{i-1}) \cap F_i \quad \text{et} \quad b_i \in G_i.$$

Le vecteur a_i peut lui-même être décomposé en

$$a_i = a_{i,1} + \dots + a_{i,i-1} \quad \text{avec} \quad a_{i,j} \in G_j.$$

En posant $a_{i,i} = b_i \in G_i$, on obtient l'écriture suivante où l'on permute les deux sommes

$$x = \sum_{i=1}^n \left(\sum_{j=1}^i a_{i,j} \right) = \sum_{j=1}^n \left(\sum_{i=j}^n b_i \right) = \sum_{j=1}^n y_j \quad \text{avec} \quad y_j \in G_j.$$

Le vecteur x appartient donc à la somme des espaces G_j et l'on peut conclure que E est la somme directe des espaces G_1, \dots, G_n .

7.7.4 Rang d'une famille de vecteurs**Exercice 21 ***

Soit (x_1, \dots, x_n) une famille de vecteurs d'un espace vectoriel E . Établir que, pour tout $p \in \llbracket 0 ; n \rrbracket$,

$$\operatorname{rg}(x_1, \dots, x_p) \leq \operatorname{rg}(x_1, \dots, x_n) + p - n.$$

Solution**méthode**

Le rang d'une famille de vecteurs est la dimension de l'espace que celle-ci engendre.

Les combinaisons linéaires des vecteurs x_1, \dots, x_n peuvent se percevoir comme les sommes des combinaisons linéaires des vecteurs x_1, \dots, x_p et x_{p+1}, \dots, x_n . On a donc

$$\text{Vect}(x_1, \dots, x_n) = \text{Vect}(x_1, \dots, x_p) + \text{Vect}(x_{p+1}, \dots, x_n).$$

La dimension d'une somme de sous-espaces vectoriels est inférieure à la somme des dimensions de chaque espace et donc

$$\text{rg}(x_1, \dots, x_n) \leq \text{rg}(x_1, \dots, x_p) + \text{rg}(x_{p+1}, \dots, x_n).$$

Enfin, le rang d'une famille de vecteurs est plus petit que le nombre de vecteurs qui la constitue et ceci entraîne

$$\text{rg}(x_1, \dots, x_n) \leq \text{rg}(x_1, \dots, x_p) + n - p.$$

En réorganisant les membres, on obtient la comparaison voulue.

Exercice 22 **

Soit a, b, c trois réels. Déterminer dans l'espace des fonctions de \mathbb{R} vers \mathbb{R} le rang de la famille des fonctions $f_a: x \mapsto \sin(x+a)$, $f_b: x \mapsto \sin(x+b)$ et $f_c: x \mapsto \sin(x+c)$.

Solution**méthode**

Par développement du sinus, ces fonctions sont toutes combinaisons linéaires des deux fonctions sinus et cosinus.

Pour $d \in \{a, b, c\}$, on a $\sin(x+d) = \cos d \sin x + \sin d \cos x$ et les trois fonctions f_a, f_b, f_c appartiennent à l'espace $\text{Vect}(\sin, \cos)$ des combinaisons linéaires des fonctions sinus et cosinus. Au surplus, aucune de ces fonctions n'est nulle et l'on peut affirmer

$$1 \leq \text{rg}(f_a, f_b, f_c) \leq 2.$$

Si le rang vaut 1 ceci signifie que les trois fonctions sont colinéaires.

Les fonctions f_a et f_b sont colinéaires si, et seulement si, elles sont égales ou opposées car chacune varie de -1 à 1 . Ceci a lieu uniquement lorsque $a \equiv b \pmod{\pi}$. On peut alors affirmer que le rang de la famille des trois fonctions vaut 2 sauf si a, b et c sont tous égaux modulo π auquel cas le rang vaut 1.

7.7.5 L'espace des polynômes**Exercice 23 * (Polynômes de Bernstein)**

Soit $n \in \mathbb{N}$. Pour tout $k \in \llbracket 0; n \rrbracket$, on pose $P_k = X^k(1-X)^{n-k}$.

Montrer que la famille (P_0, \dots, P_n) est une base de $\mathbb{R}_n[X]$.

Solution**méthode**

|| On vérifie que la famille est libre et constituée du bon nombre d'éléments.

Commençons par souligner que les polynômes P_k sont tous de degré n : ils appartiennent bien à l'espace $\mathbb{R}_n[X]$. De plus, ceux-ci sont au nombre de $n+1$ avec $n+1$ égal à la dimension de $\mathbb{R}_n[X]$. Il suffit donc d'établir que la famille (P_0, \dots, P_n) est libre pour conclure que c'est une base de $\mathbb{R}_n[X]$ (Th. 12 p. 243).

Supposons $\lambda_0 P_0 + \dots + \lambda_n P_n = 0$ avec $\lambda_0, \dots, \lambda_n$ réels, c'est-à-dire

$$\lambda_0(1-X)^n + \lambda_1 X(1-X)^{n-1} + \dots + \lambda_n X^n = 0. \quad (*)$$

En évaluant en 0 cette identité polynomiale, on obtient immédiatement $\lambda_0 = 0$. La relation (*) peut alors être simplifiée par X ce qui donne

$$\lambda_1(1-X)^{n-1} + \lambda_2 X(1-X)^{n-2} + \dots + \lambda_n X^{n-1} = 0.$$

On évalue à nouveau¹ en 0 pour obtenir $\lambda_1 = 0$ et encore simplifier par X , etc. Ainsi, on obtient successivement $\lambda_i = 0$ pour tout indice i allant de 0 jusqu'à n : on peut conclure que la famille (P_0, \dots, P_n) est libre, c'est donc une base de $\mathbb{R}_n[X]$.

Exercice 24 ** (Polynômes de degrés étagés)

Soit $(P_n)_{n \in \mathbb{N}}$ une famille de polynômes de $\mathbb{K}[X]$ vérifiant $\deg(P_n) = n$ pour tout naturel n . Montrer que $(P_n)_{n \in \mathbb{N}}$ est une base de $\mathbb{K}[X]$.

Solution**méthode**

|| On commence par vérifier que la sous-famille $(P_k)_{0 \leq k \leq n}$ est une base de $\mathbb{K}_n[X]$.

Soit $n \in \mathbb{N}$. La famille $(P_k)_{0 \leq k \leq n}$ est constituée de $n+1 = \dim \mathbb{K}_n[X]$ polynômes tous éléments de $\mathbb{K}_n[X]$. Vérifions que celle-ci est libre.

Supposons $\lambda_0 P_0 + \dots + \lambda_n P_n = 0$ avec $\lambda_0, \dots, \lambda_n$ scalaires. En réorganisant les membres de cette identité, on peut écrire

$$\lambda_n P_n = -(\lambda_0 P_0 + \dots + \lambda_{n-1} P_{n-1}) \quad \text{avec} \quad \deg(\lambda_0 P_0 + \dots + \lambda_{n-1} P_{n-1}) \leq n-1.$$

Nécessairement le scalaire λ_n est nul car sinon $\deg(\lambda_n P_n) = n$. La relation de départ peut alors être simplifiée en $\lambda_0 P_0 + \dots + \lambda_{n-1} P_{n-1} = 0$ ce qui permet de reproduire le raisonnement et d'obtenir successivement la nullité des λ_i pour i décroissant de n à 0. La famille $(P_k)_{0 \leq k \leq n}$ est libre et c'est donc une base de $\mathbb{K}_n[X]$.

Reprendons ensuite la famille initiale $(P_n)_{n \in \mathbb{N}}$.

1. On peut être surpris d'une simplification par X suivie d'une évaluation en 0. Ceci n'a rien de singulier car on manipule ici des polynômes et non des fonctions d'une variable réelle. Si un produit XP désigne le polynôme nul avec P un polynôme, on a nécessairement $P = 0$ et donc $P(0) = 0$: c'est la démarche qui est employée ici.

méthode

|| En dimension infinie, on montre qu'une famille est une base en revenant à la définition : c'est une famille libre et génératrice.

La famille $(P_n)_{n \in \mathbb{N}}$ est libre car toutes ses sous-familles finies le sont. En effet, une sous-famille finie de $(P_n)_{n \in \mathbb{N}}$ peut être comprise comme une sous-famille de la famille libre $(P_k)_{0 \leq k \leq n}$ en considérant n assez grand.

La famille $(P_n)_{n \in \mathbb{N}}$ est génératrice car, si P désigne un polynôme de $\mathbb{K}[X]$, il existe un naturel n tel que $P \in \mathbb{K}_n[X]$ et donc P est combinaison linéaire de la sous-famille $(P_k)_{0 \leq k \leq n}$ donc aussi de la famille $(P_n)_{n \in \mathbb{N}}$.

Finalement, $(P_n)_{n \in \mathbb{N}}$ est une base¹ de $\mathbb{K}[X]$.

7.7.6 Sous-espaces affines

Exercice 25 *

Soit V et W deux sous-espaces affines de directions F et G d'un espace E . Lorsque $F \subset G$, on dit que V est *parallèle* à W . Montrer qu'alors $V \subset W$ ou bien V et W sont disjoints.

Solution**méthode**

|| Il suffit de connaître un point et sa direction pour décrire un sous-espace affine (Th. 18 p. 245).

Si V et W ne sont pas disjoints, on peut introduire un point a commun à V et W . On peut alors décrire les espaces V et W :

$$V = a + F \quad \text{et} \quad W = a + G.$$

L'inclusion de F dans G donne ensuite directement $V \subset W$.

Exercice 26 **

Soit V et W deux sous-espaces affines disjoints d'un espace vectoriel réel E .

Montrer qu'il existe deux sous-espaces affines disjoints V' et W' ayant la même direction et contenant respectivement V et W .

Solution

On introduit des points $a \in V$ et $b \in W$ afin de pouvoir décrire les sous-espaces affines considérés : $V = a + F$ et $W = b + G$.

méthode

|| La direction des sous-espaces affines V' et W' doit contenir F et G donc $F+G$. Puisqu'il serait maladroit de prendre une direction trop grande, on choisit V' et W' de direction $F+G$.

1. Les bases de ce type sont très nombreuses, on y trouve la base canonique, la base de Taylor constituée des $(X-a)^k$, la base des polynômes de Tchebychev et bien d'autres...

Posons $V' = a + (F + G)$ et $W' = b + (F + G)$. Les ensembles V' et W' sont deux sous-espaces affines de même direction contenant respectivement V et W . Vérifions qu'ils sont disjoints en raisonnant par l'absurde. S'il existe un point x commun à V' et W' , on peut introduire des vecteurs $u, u' \in F$ et $v, v' \in G$ tels que

$$x = a + (u + v) = b + (u' + v').$$

$\underbrace{a + (F + G) = V'}_{\in a + F = V} \quad \underbrace{b + (F + G) = W'}_{\in b + G = W}$

En réorganisant les membres, on peut alors écrire

$$\underbrace{a + (u - u')}_{\in a + F = V} = b + (v' - v).$$

$\underbrace{b + (F + G) = W'}_{\in b + G = W}$

Ceci détermine un élément commun à V et W ce qui est absurde car on a supposé ces sous-espaces affines disjoints.

7.8 Exercices d'approfondissement

Exercice 27 *

Montrer que les familles des fonctions $x \mapsto \cos(nx)$ et $x \mapsto \cos^n x$ pour n parcourant \mathbb{N} engendrent le même sous-espace vectoriel de $\mathcal{F}(\mathbb{R}, \mathbb{R})$.

Solution

méthode

- || Par trigonométrie, on exprime les fonctions d'une famille comme combinaison linéaire des fonctions de l'autre famille.

Soit $n \in \mathbb{N}$. Par linéarisation

$$\cos^n x = \left(\frac{e^{ix} + e^{-ix}}{2} \right)^n = \frac{1}{2^n} \sum_{k=0}^n \binom{n}{k} e^{i(n-2k)x}.$$

En regroupant¹ entre elles les exponentielles imaginaires d'angles opposés, on découpe la somme en deux cc qui amène à discuter selon la parité du nombre de termes présents. Pour tout $p \in \mathbb{N}$, on obtient

$$\begin{aligned} \cos^{2p} x &= \frac{1}{2^{2p}} \sum_{k=0}^{p-1} \binom{2p}{k} \cos((2p-2k)x) + \frac{1}{2^{2p}} \binom{2p}{p} \\ \cos^{2p+1} x &= \frac{1}{2^{2p}} \sum_{k=0}^p \binom{2p+1}{k} \cos((2p-2k+1)x). \end{aligned}$$

1. On pourrait aussi écrire que $\cos^n x$ est la partie réelle de la somme ce qui fait aussi apparaître rapidement une formule décisive en exploitant la parité de la fonction cosinus afin de résoudre les $\cos((n-2k)x)$ pour $n-2k < 0$.

Ainsi, les fonctions $x \mapsto \cos^n x$ sont combinaisons linéaires des fonctions $x \mapsto \cos(nx)$ avec $n \in \mathbb{N}$.

Inversement, on peut écrire par la formule de Moivre

$$\cos(nx) = \operatorname{Re}(e^{inx}) = \operatorname{Re}((\cos(x) + i\sin(x))^n) = \operatorname{Re}\left(\sum_{k=0}^n \binom{n}{k} \cos^{n-k}(x) i^k \sin^k(x)\right).$$

Dans la somme, les termes réels sont uniquement obtenus pour les indices k pairs et donc

$$\cos(nx) = \sum_{p=0}^{\lfloor n/2 \rfloor} \binom{n}{2p} (-1)^p \cos^{n-2p}(x) \sin^{2p}(x) \quad \text{avec} \quad \sin^{2p} x = (1 - \cos^2 x)^p.$$

En poursuivant le développement, on exprime la fonction $x \mapsto \cos(nx)$ comme combinaison linéaire des fonctions $x \mapsto \cos^n x$.

Finalement, les deux familles de fonctions engendrent le même espace.

Exercice 28 ** (Polynômes de Newton)

Pour $k \in \mathbb{N}$, on pose

$$P_k = \frac{X(X-1)\dots(X-k+1)}{k!}$$

- (a) Montrer que la famille $(P_n)_{n \in \mathbb{N}}$ est une base de $\mathbb{R}[X]$.
- (b) Vérifier que $P_k(m)$ est entier pour tout $m \in \mathbb{Z}$ et tout $k \in \mathbb{N}$.
- (c) Trouver tous les polynômes P prenant des valeurs entières sur chaque entier.

Solution

(a) On reconnaît une famille de polynômes réels de degrés étagés¹, c'est donc une base de $\mathbb{R}[X]$.

(b) Soit $m \in \mathbb{Z}$ et $k \in \mathbb{N}$.

méthode

|| Le nombre $P_k(m)$ peut être compris comme un coefficient du binôme.

Cas : $m \geq k$.

$$P_k(m) = \frac{m(m-1)\dots(m-k+1)}{k!} = \binom{m}{k} \in \mathbb{N}.$$

Cas : $0 \leq m \leq k-1$.

$$P_k(m) = 0 \in \mathbb{N}.$$

Cas : $m < 0$. On écrit $m = -p$ avec $p \in \mathbb{N}$ et

$$\begin{aligned} P_k(m) &= \frac{m(m-1)\dots(m-k+1)}{k!} = (-1)^p \frac{p(p+1)\dots(p+k-1)}{k!} \\ &= (-1)^k \binom{p+k-1}{k} \in \mathbb{Z}. \end{aligned}$$

1. Voir sujet 24 p. 266.

Dans tous les cas, les valeurs $P_k(m)$ sont entières.

(c) Soit P un polynôme solution.

méthode

On montre que les coordonnées de P dans la base $(P_n)_{n \in \mathbb{N}}$ sont toutes des nombres entiers.

Notons $(\lambda_n)_{n \in \mathbb{N}}$ la famille des coordonnées de P dans la base $(P_n)_{n \in \mathbb{N}}$. Celles-ci sont nulles à partir d'un certain rang et, en introduisant $n \in \mathbb{N}$ assez grand, on peut écrire

$$P = \lambda_0 P_0 + \lambda_1 P_1 + \cdots + \lambda_n P_n. \quad (*)$$

En évaluant (*) en 0, on obtient $\lambda_0 P_0(0) = P(0)$ car 0 est racine des polynômes P_1, \dots, P_n . Sachant $P_0(0) = 1$, il vient $\lambda_0 \in \mathbb{Z}$. En évaluant (*) en 1, on obtient ensuite $\lambda_0 P_0(1) + \lambda_1 P_1(1) = P(1)$ et l'on en déduit $\lambda_1 \in \mathbb{Z}$. Successivement, on obtient $\lambda_i \in \mathbb{Z}$ pour i allant de 0 à n en évaluant (*) en i et en exploitant $P_i(i) = 1$ et $P_j(i) = 0$ pour tout $j > i$.

Inversement, il est clair que si toutes les coordonnées d'un polynôme dans la base $(P_n)_{n \in \mathbb{N}}$ sont des nombres entiers, ce polynôme prend des valeurs entières sur les nombres entiers comme le font les P_n .

Exercice 29 **

Soit E un espace vectoriel réel de dimension finie $n \in \mathbb{N}^*$. Déterminer les applications d définies sur l'ensemble des sous-espaces vectoriels de E et à valeurs dans \mathbb{N} vérifiant, pour tous sous-espaces vectoriels F et G en somme directe,

$$d(F \oplus G) = d(F) + d(G). \quad (*)$$

Solution

L'application dimension semble être un bon candidat...

Pour $F = G = \{0_E\}$, la formule (*) donne $d(\{0_E\}) = 0$.

méthode

On vérifie que la fonction d prend la même valeur sur toutes les droites vectorielles de E .

Pour $x \neq 0_E$, posons $f(x) = d(\text{Vect}(x))$.

On a immédiatement $f(\lambda x) = f(x)$ pour tout $\lambda \in \mathbb{R}^*$ et tout $x \neq 0_E$ car les vecteurs x et λx engendrent la même droite.

Soit x et y deux vecteurs de E non colinéaires. Le plan $\text{Vect}(x, y)$ est la somme directe de $\text{Vect}(x)$ et $\text{Vect}(y)$. Le plan $\text{Vect}(x, y)$ se confond aussi avec $\text{Vect}(x, x+y)$ qui est la somme directe de $\text{Vect}(x)$ et $\text{Vect}(x+y)$. La propriété (*) donne alors

$$d(\text{Vect}(x, y)) = f(x) + f(y) = d(\text{Vect}(x, x+y)) = f(x) + f(x+y).$$

On en déduit $f(x+y) = f(y)$. Un raisonnement symétrique donne aussi $f(x+y) = f(x)$ et donc $f(x) = f(y)$.

Finalement, la fonction f est constante sur $E \setminus \{0_E\}$. Posons $\alpha \in \mathbb{N}$ la valeur de cette constante. Pour tout sous-espace vectoriel F de E , on peut alors affirmer $d(F) = \alpha \dim F$. En effet, cette égalité est vérifiée si F est l'espace nul et l'est aussi si F est de dimension $p \geq 1$ puisque, lorsque (e_1, \dots, e_p) désigne une base de F ,

$$d(F) = d\left(\bigoplus_{j=1}^p \text{Vect}(e_j)\right) = \sum_{j=1}^p f(e_j) = \alpha p.$$

Inversement, la fonction $d: F \mapsto \alpha \dim F$ avec $\alpha \in \mathbb{N}$ est évidemment solution du problème posé.

Exercice 30 **

Soit E un espace vectoriel réel de dimension finie.

Montrer que deux sous-espaces vectoriels F et G de E ont un supplémentaire commun si, et seulement si, ils ont la même dimension.

Solution

Si les sous-espaces vectoriels F et G ont un supplémentaire commun H , ils ont la même dimension car

$$\dim E = \dim F + \dim H \quad \text{et} \quad \dim E = \dim G + \dim H.$$

Inversement, montrons que si $\dim F = \dim G$ alors F et G ont un supplémentaire commun.

méthode

|| Si $F \neq G$ on peut introduire un vecteur a n'appartenant ni à F ni à G engendrant une droite en somme directe avec F et G .

Posons $n = \dim E$ et raisonnons par récurrence décroissante sur la dimension commune $p \in \llbracket 0 ; n \rrbracket$ des espaces F et G .

Si $p = n$, les espaces F et G sont égaux à E et $\{0_E\}$ détermine un supplémentaire commun.

Supposons la propriété établie au rang $p+1 \in \llbracket 1 ; n \rrbracket$. Soit F et G deux sous-espaces vectoriels de dimension p . Si $F = G$, n'importe quel supplémentaire de F est convenable et il en existe (Th. 14 p. 243). Sinon, F n'est pas inclus dans G ni G dans F . Il existe alors un vecteur x de F et un vecteur y de G tels que $x \notin G$ et $y \notin F$. Le vecteur somme $a = x + y$ n'appartient pas à F et la droite $D = \text{Vect}(a)$ est en somme directe avec F car l'intersection de ces deux espaces est réduite au vecteur nul. De même, la droite D est en somme directe avec G . On peut alors introduire les espaces de dimension $p+1$

$$F' = F \oplus D \quad \text{et} \quad G' = G \oplus D.$$

Par l'hypothèse de récurrence, ces espaces possèdent un supplémentaire commun H' et l'espace $H = D \oplus H'$ détermine alors un supplémentaire commun à F et G .

La récurrence est établie.

Exercice 31 ***

Soit F_1, \dots, F_n des sous-espaces vectoriels d'un espace vectoriel réel E .

Montrer que si l'union $F_1 \cup \dots \cup F_n$ est un sous-espace vectoriel, celle-ci est égale à l'un des espaces F_i .

Solution**méthode**

|| On vérifie que $F_1 \cup \dots \cup F_{n-1} \subset F_n$ ou $F_n \subset F_1 \cup \dots \cup F_{n-1}$.

Par l'absurde, si les inclusions précédentes sont fausses, on peut introduire un élément x de $F_1 \cup \dots \cup F_{n-1}$ qui n'appartient pas à F_n et un élément y de F_n qui n'appartient pas à $F_1 \cup \dots \cup F_{n-1}$. Pour tout $\lambda \in \mathbb{R}$, $z_\lambda = x + \lambda y$ appartient au sous-espace vectoriel $F_1 \cup \dots \cup F_n$. Cependant, ce vecteur ne peut appartenir à F_n car sinon $x = z_\lambda - \lambda y$ serait par opérations élément du sous-espace vectoriel F_n . Il existe donc un indice $i \in [1; n-1]$ tel que z_λ est élément de F_i . Les valeurs possibles de λ étant en nombre infini, on peut affirmer l'existence de réels λ et μ distincts tels que z_λ et z_μ appartiennent au même F_i . On peut alors écrire

$$y = \frac{1}{\lambda - \mu} (z_\lambda - z_\mu) \in F_i$$

ce qui contredit la définition de y : c'est absurde.

Ainsi, on a montré que, si $F_1 \cup \dots \cup F_n$ est un sous-espace vectoriel, $F_1 \cup \dots \cup F_{n-1} \subset F_n$ ou $F_n \subset F_1 \cup \dots \cup F_{n-1}$. Dans le premier cas, on peut conclure. Dans le second, on peut simplifier l'union $F_1 \cup \dots \cup F_n$ en $F_1 \cup \dots \cup F_{n-1}$ et reprendre le processus : soit cette union est incluse dans l'un des espaces, soit on peut réduire le nombre d'espaces qui la constituent sans la modifier.

CHAPITRE 8

Les applications linéaires

\mathbb{K} désigne \mathbb{R} ou \mathbb{C} et E, E' des \mathbb{K} -espaces vectoriels.

8.1 Applications linéaires

8.1.1 Généralités

Définition

On appelle *application linéaire* de E vers E' toute application $u: E \rightarrow E'$ vérifiant¹ :

$$\forall (\lambda, \mu) \in \mathbb{K}^2, \forall (x, y) \in E^2, \quad u(\lambda x + \mu y) = \lambda u(x) + \mu u(y).$$

On note $\mathcal{L}(E, E')$ l'ensemble des applications linéaires de E vers E' .

Une telle application envoie le vecteur nul de E sur le vecteur nul de E' et transforme une combinaison linéaire en la combinaison linéaire des images.

Une combinaison linéaire d'applications linéaires de E vers E' est encore une application linéaire et l'on dispose donc d'une structure d'espace vectoriel :

Théorème 1

$(\mathcal{L}(E, E'), +, .)$ est un \mathbb{K} -espace vectoriel de neutre l'application nulle $x \mapsto 0_{E'}$.

De plus, si E et E' sont de dimensions finies, l'espace $\mathcal{L}(E, E')$ est aussi de dimension finie et $\dim \mathcal{L}(E, E') = \dim E \times \dim E'$.

1. Il existe une caractérisation plus économique où l'on étudie $u(x + \lambda y) = u(x) + \lambda u(y)$.

La composée de deux applications linéaires est une application linéaire.

Définition

Un *isomorphisme* de E vers E' est une application linéaire bijective de E vers E' .

La composée de deux isomorphismes est un isomorphisme et la bijection réciproque d'un isomorphisme est aussi un isomorphisme.

8.1.2 Noyau et image

Théorème 2

L'image directe d'un sous-espace vectoriel de E par une application linéaire de E vers E' est un sous-espace vectoriel de E' .

L'image réciproque d'un sous-espace vectoriel de E' par une application linéaire de E vers E' est un sous-espace vectoriel de E .

Définition

On appelle *noyau* et *image* d'une application linéaire u de E vers E' les espaces :

$$\text{Ker}(u) = u^{-1}(\{0_{E'}\}) \quad \text{et} \quad \text{Im}(u) = u(E).$$

L'application u est surjective si, et seulement si, $\text{Im}(u) = E'$.

Théorème 3

Une application linéaire u de E vers E' est injective si, et seulement si, $\text{Ker}(u) = \{0_E\}$.

8.1.3 Équations linéaires

Définition

Une *équation linéaire* est une équation de la forme $u(x) = y$ avec $u \in \mathcal{L}(E, E')$, $y \in E'$ et d'inconnue $x \in E$.

Théorème 4

Si l'équation $u(x) = y$ admet une solution¹ x_0 , l'ensemble des solutions cette équation est le sous-espace affine passant par x_0 et dirigé par le noyau de u .

Pour déterminer le noyau de u , on résout l'équation $u(x) = 0_{E'}$ appelée *équation homogène* associée à l'équation linéaire $u(x) = y$.

1. On dit que x_0 est une *solution particulière*.

8.2 Endomorphismes

8.2.1 L'anneau des endomorphismes

Définition

On appelle *endomorphisme* de E toute application linéaire de E vers lui-même. On note $\mathcal{L}(E)$ l'ensemble des endomorphismes de E .

L'identité Id_E est un endomorphisme de E . Pour $\alpha \in \mathbb{K}$, l'application $h_\alpha = \alpha \text{Id}_E$ est aussi un endomorphisme de E appelé *homothétie* de rapport α .

$(\mathcal{L}(E), +, \circ)$ est un \mathbb{K} -espace vectoriel. Si E est dimension finie, l'espace $\mathcal{L}(E)$ l'est aussi et $\dim \mathcal{L}(E) = (\dim E)^2$. Au surplus, on dispose d'une structure d'anneau :

Théorème 5

$(\mathcal{L}(E), +, \circ)$ est un anneau de neutres l'endomorphisme nul et l'identité.

Celui-ci n'est pas commutatif dès que $\dim E > 2$.

Dans l'anneau des endomorphismes, on opère avec la composition comme avec une multiplication (non commutative). Cette structure hérite donc des propriétés¹ vues dans le chapitre 4. On note parfois uv pour $u \circ v$ et l'on comprend u^n comme un itéré de composition pour la loi \circ : $u^n = u \circ \cdots \circ u$ (n facteurs).

8.2.2 Automorphisme

Définition

On appelle *automorphisme* de E tout endomorphisme bijectif de E .

Les automorphismes sont les éléments inversibles de l'anneau $\mathcal{L}(E)$: ils constituent un groupe pour la composition des applications (Th. 8 p. 124).

Définition

L'ensemble $\text{GL}(E)$ des automorphismes de E est appelé *groupe linéaire* de E .

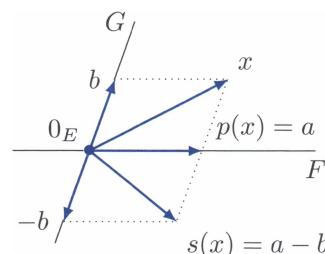
8.2.3 Projections et symétries vectorielles

Soit F et G deux sous-espaces vectoriels supplémentaires de E . Tout vecteur x de E s'écrit de façon unique $x = a + b$ avec $a \in F$ et $b \in G$. On pose alors $p(x) = a$ et $s(x) = a - b$.

Définition

L'application p est appelée *projection* sur F parallèlement à G et l'application s est appelée *symétrie* par rapport à F et parallèlement à G .

La projection p correspond à l'identité sur F et à l'application nulle sur G .



1. Par exemple, on peut appliquer la formule du binôme au calcul de $(u + v)^n$ lorsque u et v sont deux endomorphismes qui commutent.

Théorème 6

Les applications p et s sont des endomorphismes de E vérifiant

$$p^2 = p, \quad s^2 = \text{Id}_E \quad \text{et} \quad s = 2p - \text{Id}_E.$$

On peut aussi remarquer que $q = \text{Id}_E - p$ est la projection sur G parallèlement à F , on dit que c'est la *projection complémentaire* de p . Aussi, l'endomorphisme $-s$ est la symétrie par rapport à G et parallèlement à F .

On dispose d'une réciproque au théorème précédent caractérisant les projections :

Théorème 7

Si p est un endomorphisme de E vérifiant¹ $p^2 = p$ alors $F = \text{Im}(p)$ et $G = \text{Ker}(p)$ sont supplémentaires et p est la projection sur F parallèlement à G .

Aussi,

Théorème 8

Si s est un endomorphisme de E vérifiant $s^2 = \text{Id}_E$, les espaces $F = \text{Ker}(s - \text{Id}_E)$ et $G = \text{Ker}(s + \text{Id}_E)$ sont supplémentaires et s est la symétrie par rapport à F et parallèlement à G .

8.3 Détermination d'une application linéaire

8.3.1 Image linéaire d'une famille de vecteurs

Théorème 9

Soit $(x_i)_{i \in I}$ une famille de vecteurs de E et $u \in \mathcal{L}(E, E')$.

- a) Si $(x_i)_{i \in I}$ est génératrice de E , $(u(x_i))_{i \in I}$ est génératrice de $\text{Im}(u)$.
- b) Si $(x_i)_{i \in I}$ est libre et si u injective, $(u(x_i))_{i \in I}$ est libre.

En particulier, si $(x_i)_{i \in I}$ est génératrice de E et u surjective, la famille image $(u(x_i))_{i \in I}$ est génératrice de l'espace E' .

Théorème 10

Soit $u \in \mathcal{L}(E, E')$ et $(e_i)_{i \in I}$ une base de E .

- a) u est injective si, et seulement si, $(u(e_i))_{i \in I}$ est libre ;
- b) u est surjective si, et seulement si, $(u(e_i))_{i \in I}$ est génératrice de E' ;
- c) u est un isomorphisme si, et seulement si, $(u(e_i))_{i \in I}$ est une base de E' .

1. Un endomorphisme vérifiant $p^2 = p$ est appelé un *projecteur*. Ce théorème affirme qu'il n'y a pas lieu de distinguer projecteurs et projections vectorielles.

En particulier :

Théorème 11

S'il existe un isomorphisme entre deux espaces vectoriels, ceux-ci ont la même dimension.

8.3.2 Détermination par l'image d'une base

On peut construire une application linéaire en spécifiant l'image d'une base :

Théorème 12

Si $(e_i)_{i \in I}$ est une base de E et $(e'_i)_{i \in I}$ une famille de vecteurs de E' , il existe une unique application linéaire $u: E \rightarrow E'$ vérifiant $u(e_i) = e'_i$ pour tout $i \in I$.

Si deux applications linéaires u et $v \in \mathcal{L}(E, E')$ sont égales sur chacun des vecteurs d'une base de E , elles sont égales sur l'intégralité de E .

En considérant une application qui transforme une base en une autre, on peut affirmer que deux espaces de dimensions finies égales sont isomorphes¹.

8.3.3 Détermination par des restrictions linéaires

On peut aussi construire une application linéaire en désignant ses restrictions linéaires au départ d'espaces d'une décomposition en somme directe :

Théorème 13

Si E_1, \dots, E_m sont des sous-espaces vectoriels de E tels que $E = E_1 \oplus \dots \oplus E_m$ et si les u_k désignent des applications linéaires de E_k vers E' , il existe une unique application linéaire $u \in \mathcal{L}(E, E')$ prolongeant les u_k , c'est-à-dire vérifiant $u(x) = u_k(x)$ pour tout $x \in E_k$.

Si deux applications linéaires sont égales sur chacun des espaces d'une décomposition en somme directe, elles sont égales sur l'intégralité de l'espace. En particulier, on peut définir une application linéaire en précisant ses restrictions linéaires au départ de deux espaces supplémentaires.

8.4 Théorème du rang

8.4.1 Rang d'une application linéaire

Définition

Si u est une application linéaire de E vers E' , la dimension de son image est appelée *rang* de l'application linéaire u . Celui-ci est noté $\text{rg}(u)$.

1. Deux espaces de dimensions infinies peuvent ne pas être isomorphes.

Si $(e_i)_{i \in I}$ est une base de E , on a

$$\text{Im}(u) = \text{Vect}(u(e_i))_{i \in I} \quad \text{donc} \quad \text{rg}(u) = \text{rg}(u(e_i))_{i \in I}.$$

Le rang d'une application linéaire n'est pas modifié lorsque l'on compose celle-ci par un isomorphisme.

Théorème 14 (Théorème du rang)

Si u est une application linéaire de E vers E' et si S est un supplémentaire du noyau de u dans E alors u définit par restriction un isomorphisme de S vers $\text{Im}(u)$.

En particulier, si l'espace E est de dimension finie, on a la *formule du rang*

$$\dim E = \dim \text{Ker}(u) + \text{rg}(u).$$

8.4.2 Caractérisation des isomorphismes en dimension finie

Théorème 15 (Théorème d'isomorphisme)

Une application linéaire entre deux espaces de même dimension finie est bijective si, et seulement si, elle est injective, ou encore, si, et seulement si, elle est surjective.

Ce résultat s'applique en particulier aux endomorphismes d'un espace de dimension finie. Un tel endomorphisme est alors inversible si, et seulement si, il est inversible à gauche¹, ou encore, si, et seulement si, il est inversible à droite². Les inverses à droite et à gauche sont alors égaux à l'inverse de l'endomorphisme.

8.4.3 Formes linéaires et hyperplans

Définition

On appelle *forme linéaire* sur E toute application linéaire de E vers \mathbb{K} . On note E^* l'espace des formes linéaires sur E que l'on appelle le *dual* de E .

Si $(e_i)_{i \in I}$ est une base de E , l'application qui à un vecteur x associe sa coordonnée le long de l'un des vecteurs e_i est une forme linéaire.

Définition

On appelle *hyperplan* de E tout noyau d'une forme linéaire non nulle sur E .

Si H est un hyperplan noyau d'une forme linéaire non nulle φ , on dit que l'hyperplan H est défini par l'équation $\varphi(x) = 0$. Les autres formes linéaires non nulles dont H est le noyau sont de la forme $\lambda\varphi$ avec $\lambda \in \mathbb{K}^*$.

1. Un endomorphisme u de E est *inversible à gauche* s'il existe v vérifiant $v \circ u = \text{Id}_E$. Un tel endomorphisme est nécessairement injectif : voir sujet 23 p. 34.

2. Un endomorphisme inversible à droite est surjectif.

Théorème 16

Si H est un hyperplan de E , toute droite D non contenue dans H détermine un supplémentaire de H . Inversement, tout supplémentaire d'une droite de E est un hyperplan.

Si l'espace E est de dimension finie égale à n , les hyperplans de E sont exactement les espaces de dimension $n - 1$. Si (e_1, \dots, e_n) désigne une base de E et si (x_1, \dots, x_n) détermine la famille des coordonnées d'un vecteur générique x dans cette base, une équation d'un hyperplan est de la forme

$$a_1x_1 + \dots + a_nx_n = 0 \quad \text{avec } a_1, \dots, a_n \in \mathbb{K} \text{ non tous nuls.}$$

Les autres équations définissant le même hyperplan sont proportionnelles à celle-ci.

8.5 Exercices d'apprentissage

8.5.1 Généralités

Exercice 1

Etudier la linéarité des applications suivantes, préciser leur noyau et leur image, préciser aussi si celles-ci sont injectives ou surjectives :

- (a) $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ définie par $f(x, y, z) = (2x + y - z, x + y)$.
- (b) $M: \mathbb{R}[X] \rightarrow \mathbb{R}[X]$ définie par $M(P) = XP$.
- (c) $\varphi: C^1(\mathbb{R}, \mathbb{K}) \rightarrow C(\mathbb{R}, \mathbb{K})$ définie par $\varphi(f) = f' - f$.
- (d) $T: \mathbb{C}^{\mathbb{N}} \rightarrow \mathbb{C}^{\mathbb{N}}$ définie par $T((u_n)_{n \in \mathbb{N}}) = (u_{n+1})_{n \in \mathbb{N}}$.
- (e) $f: \mathbb{C} \rightarrow \mathbb{R}$ définie par $f(z) = \operatorname{Im}(z) - \operatorname{Re}(z)$.

Solution**méthode**

Pour chaque application, on vérifie l'identité $u(\lambda x + \mu y) = \lambda u(x) + \mu u(y)$. Ceci nécessite de bien percevoir les espaces¹ entre lesquels celle-ci opère ainsi que l'objet correspondant à la variable de l'application.

(a) L'application f opère entre les deux espaces vectoriels réels \mathbb{R}^3 et \mathbb{R}^2 . Pour $\lambda, \mu \in \mathbb{R}$ et $u = (x, y, z) \in \mathbb{R}^3$, $v = (x', y', z') \in \mathbb{R}^3$

$$\begin{aligned} f(\lambda u + \mu v) &= f(\lambda x + \mu x', \lambda y + \mu y', \lambda z + \mu z') \\ &= (2(\lambda x + \mu x') + \lambda y + \mu y' - (\lambda z + \mu z'), \lambda x + \mu x' + \lambda y + \mu y') \\ &= \lambda(2x + y - z, x + y) + \mu(2x' + y' - z', x' + y') = \lambda f(u) + \mu f(v). \end{aligned}$$

1. Il convient notamment d'identifier s'il s'agit d'un espace réel ou d'un espace complexe.

L'application f est donc linéaire.

méthode

On détermine le noyau de $u \in \mathcal{L}(E, E')$ en résolvant l'équation $u(x) = 0_{E'}$ d'inconnue $x \in E$.

Pour $u = (x, y, z) \in \mathbb{R}^3$

$$\begin{aligned} f(u) = 0_{\mathbb{R}^2} &\iff \begin{cases} 2x + y - z = 0 \\ x + y = 0 \end{cases} \\ &\iff \begin{cases} y = -x \\ z = x. \end{cases} \end{aligned}$$

Le noyau de f est donc

$$\text{Ker}(f) = \{(x, -x, x) \mid x \in \mathbb{R}\} = \{x(1, -1, 1) \mid x \in \mathbb{R}\} = \text{Vect}(1, -1, 1).$$

Ce noyau n'étant pas réduit au vecteur nul, l'application n'est pas injective (Th. 3 p. 274).

méthode

On obtient l'image de $u \in \mathcal{L}(E, E')$ en déterminant¹ les valeurs prises par u .

Soit $u = (x, y, z) \in \mathbb{R}^3$ et $v = (X, Y) \in \mathbb{R}^2$. Étudions l'égalité $f(u) = v$.

$$\begin{aligned} f(u) = v &\iff \begin{cases} 2x + y - z = X \\ x + y = Y \end{cases} \\ &\iff \begin{cases} y = Y - x \\ z = -X - Y + x. \end{cases} \end{aligned}$$

Il suffit alors de choisir arbitrairement x pour former un triplet u solution. Par exemple, pour $x = 0$, on obtient $u = (0, Y, -X - Y)$ tel que $f(u) = (X, Y)$. L'équation $f(u) = v$ admet donc des solutions u quelle que soit la valeur v de \mathbb{R}^2 : l'application f est surjective² et $\text{Im}(f) = \mathbb{R}^2$.

(b) L'application M opère de l'espace réel $\mathbb{R}[X]$ dans lui-même. Il s'agit d'un endomorphisme³ de $\mathbb{R}[X]$ car, pour tous $\lambda, \mu \in \mathbb{R}$ et tous $P, Q \in \mathbb{R}[X]$,

$$M(\lambda P + \mu Q) = X(\lambda P + \mu Q) = \lambda X P + \mu X Q = \lambda M(P) + \mu M(Q).$$

Le noyau de M réunit les polynômes P tels que $X P = 0$. Or un produit de deux polynômes n'est nul que lorsque l'un des facteurs est nul. Ici, le polynôme X n'est pas le

1. Affirmer que u est une application linéaire de E vers E' assure que u prend ses valeurs dans E' mais non que toute valeur de E' est prise par u .

2. Un argument de dimension plus rapide est aussi possible : le noyau de f étant de dimension 1, le théorème du rang (Th. 14 p. 278) assure que l'image de f est de dimension 2.

3. Un endomorphisme de E est non seulement une application de E vers E , c'est aussi une application linéaire.

polynôme nul et donc $XP = 0$ si, et seulement si, $P = 0$. Le noyau de M est donc réduit au polynôme nul : l'endomorphisme M est injectif.

L'image de M réunit les polynômes pouvant s'écrire XP avec $P \in \mathbb{R}[X]$, ce sont exactement les polynômes dont 0 est racine : $\text{Im}(M) = \{P \in \mathbb{R}[X] \mid P(0) = 0\}$. L'application M n'est donc pas surjective car un polynôme dont 0 n'est pas racine n'est pas une valeur prise par M .

(c) L'application φ est bien définie de l'espace $C^1(\mathbb{R}, \mathbb{K})$ vers $C(\mathbb{R}, \mathbb{K})$. Sa linéarité découle de la linéarité de la dérivation : pour tous $\lambda, \mu \in \mathbb{K}$ et tous $f, g \in C^1(\mathbb{R}, \mathbb{K})$

$$\begin{aligned}\varphi(\lambda f + \mu g) &= (\lambda f + \mu g)' - (\lambda f + \mu g) \\ &= \lambda f' + \mu g' - (\lambda f + \mu g) = \lambda\varphi(f) + \mu\varphi(g).\end{aligned}$$

Le noyau de φ est constitué des fonctions solutions de l'équation différentielle $y' = y$, à savoir les fonctions $t \mapsto \lambda e^t$ avec $\lambda \in \mathbb{K}$. L'application φ n'est donc pas injective. En revanche, cette application est surjective car le théorème de Cauchy¹ assure l'existence d'une solution à l'équation différentielle $y' - y = b(x)$ pour toute fonction b continue. De plus, cette solution est une fonction de classe C^1 . L'image de φ est donc $C(\mathbb{R}, \mathbb{K})$.

(d) L'application T est bien définie de l'espace complexe $\mathbb{C}^{\mathbb{N}}$ dans lui-même. C'est un endomorphisme car, pour tous $\lambda, \mu \in \mathbb{C}$ et tous $u, v \in \mathbb{C}^{\mathbb{N}}$,

$$\begin{aligned}T(\lambda(u_n)_{n \in \mathbb{N}} + \mu(v_n)_{n \in \mathbb{N}}) &= T((\lambda u_n + \mu v_n)_{n \in \mathbb{N}}) \\ &= (\lambda u_{n+1} + \mu v_{n+1})_{n \in \mathbb{N}} \\ &= \lambda(u_{n+1})_{n \in \mathbb{N}} + \mu(v_{n+1})_{n \in \mathbb{N}} \\ &= \lambda T((u_n)_{n \in \mathbb{N}}) + \mu T((v_n)_{n \in \mathbb{N}}).\end{aligned}$$

Le noyau de T est constitué des suites u vérifiant $u_{n+1} = 0$ pour tout $n \in \mathbb{N}$, c'est-à-dire des suites nulles à partir du rang 1. L'image de T est égale à $\mathbb{C}^{\mathbb{N}}$ car toute suite v de $\mathbb{C}^{\mathbb{N}}$ est l'image d'une suite u déterminée par $u_{n+1} = v_n$ pour tout $n \in \mathbb{N}$ et une valeur arbitraire pour u_0 . L'application T est surjective mais n'est pas injective.

(e) L'espace d'arrivée de l'application f est réel, l'espace de départ \mathbb{C} doit donc être compris comme un espace réel. Pour $\lambda, \mu \in \mathbb{R}$ et $z, z' \in \mathbb{C}$, on vérifie aisément l'identité $f(\lambda z + \mu z') = \lambda f(z) + \mu f(z')$ grâce aux linéarités réelles des fonctions partie réelle et partie imaginaire.

Le noyau de f réunit les complexes de z vérifiant $\text{Re}(z) = \text{Im}(z)$, ce sont les complexes de la droite $\text{Vect}_{\mathbb{R}}(1+i) = \{\lambda(1+i) \mid \lambda \in \mathbb{R}\}$: l'application n'est pas injective. L'image de f est égale à \mathbb{R} car, pour tout réel y , on vérifie par exemple $f(iy) = y$: l'application f est surjective.

1. Voir le chapitre 5 de l'ouvrage *Exercices d'analyse MPSI*.

Exercice 2

Soit f et g deux endomorphismes d'un espace vectoriel E . Vérifier

- | | |
|--|--|
| (a) $\text{Ker}(f) \cap \text{Ker}(g) \subset \text{Ker}(f + g)$ | (b) $\text{Im}(f) + \text{Im}(g) \supset \text{Im}(f + g)$ |
| (c) $\text{Ker}(f) \subset \text{Ker}(f^2)$ | (d) $\text{Im}(f) \supset \text{Im}(f^2)$. |

Solution**méthode**

|| On montre une inclusion $A \subset B$ en introduisant un élément arbitraire de A et en vérifiant qu'il appartient à B .

(a) Soit $x \in \text{Ker}(f) \cap \text{Ker}(g)$. On a $(f + g)(x) = f(x) + g(x) = 0_E$ et donc x est élément de $\text{Ker}(f + g)$.

(b) Soit $y \in \text{Im}(f + g)$. Il existe un antécédent x dans E pour lequel $y = (f + g)(x)$ et donc $y = f(x) + g(x)$. Le vecteur y appartient¹ alors à $\text{Im}(f) + \text{Im}(g)$.

(c) méthode

|| Pour un endomorphisme f , l'application f^2 désigne $f \circ f$.

Soit $x \in \text{Ker}(f)$. On a $f^2(x) = f(f(x)) = f(0_E) = 0_E$ et donc $x \in \text{Ker}(f^2)$.

(d) Soit $y \in \text{Im}(f^2)$. Il existe un antécédent $x \in E$ tel que $y = f^2(x) = f(f(x))$. En introduisant $a = f(x) \in E$, on obtient $y = f(a)$ et donc $y \in \text{Im}(f)$.

Exercice 3

Soit E , E' et E'' des espaces vectoriels et $f \in \mathcal{L}(E, E')$, $g \in \mathcal{L}(E', E'')$. Montrer

$$g \circ f = 0 \iff \text{Im}(f) \subset \text{Ker}(g).$$

Solution

On raisonne par double implication

(\implies) Supposons $g \circ f = 0$. Soit $y \in \text{Im}(f)$. On peut écrire $y = f(x)$ pour un certain x de E et alors

$$g(y) = g(f(x)) = (g \circ f)(x) = 0_{E''}.$$

Ainsi, y est élément du noyau de g et l'on peut affirmer $\text{Im}(f) \subset \text{Ker}(g)$.

(\impliedby) Supposons $\text{Im}(f) \subset \text{Ker}(g)$.

méthode

|| On vérifie que l'application $g \circ f$ est nulle en calculant $g(f(x))$ pour un vecteur x quelconque dans E .

Soit x un élément de E . Son image $f(x)$ est élément de $\text{Im}(f)$ donc de $\text{Ker}(g)$ et par conséquent $g(f(x)) = 0_{E''}$. Ainsi, $g \circ f$ est l'application nulle.

1. L'inclusion réciproque peut être fausse, un élément de $\text{Im}(f) + \text{Im}(g)$ s'écrit $f(x_1) + g(x_2)$ avec des antécédents x_1 et x_2 pouvant être différents.

8.5.2 Applications linéaires en dimension finie

Exercice 4

Soit f un endomorphisme d'un espace vectoriel E de dimension finie. Montrer que les assertions suivantes sont équivalentes :

- | | |
|---|---|
| (i) $E = \text{Im}(f) \oplus \text{Ker}(f)$; | (ii) $E = \text{Im}(f) + \text{Ker}(f)$; |
| (iii) $\text{Im}(f^2) = \text{Im}(f)$; | (iv) $\text{Ker}(f^2) = \text{Ker}(f)$. |

Solution

On établit un chaînage d'implications.

(i) \implies (ii) est entendue.

(ii) \implies (iii) Supposons $E = \text{Im}(f) + \text{Ker}(f)$.

méthode

|| Les inclusions $\text{Im}(f^2) \subset \text{Im}(f)$ et $\text{Ker}(f) \subset \text{Ker}(f^2)$ sont toujours vraies¹.

Etudions l'inclusion $\text{Im}(f) \subset \text{Im}(f^2)$. Soit $y \in \text{Im}(f)$. On peut introduire un antécédent x dans E tel que $y = f(x)$ et écrire $x = u + v$ avec $u \in \text{Im}(f)$ et $v \in \text{Ker}(f)$. Au surplus, on peut aussi écrire $u = f(a)$ avec $a \in E$. Il vient alors

$$y = f(f(a) + v) = f^2(a) + f(v) = f^2(a) \in \text{Im}(f^2) \quad \text{car} \quad f(v) = 0_E.$$

Ainsi, on a l'inclusion $\text{Im}(f) \subset \text{Im}(f^2)$ puis l'égalité.

(iii) \implies (iv) Supposons $\text{Im}(f^2) = \text{Im}(f)$.

méthode

|| La formule du rang (Th. 14 p. 278) permet de calculer la dimension du noyau d'une application linéaire en fonction de la dimension de son image et inversement.

Par le théorème du rang appliqué aux endomorphismes f et f^2 , on a simultanément

$$\dim E = \text{rg}(f) + \dim \text{Ker}(f) \quad \text{et} \quad \dim E = \text{rg}(f^2) + \dim \text{Ker}(f^2).$$

Les images de f et f^2 étant égales, ces endomorphismes ont le même rang et les espaces $\text{Ker}(f)$ et $\text{Ker}(f^2)$ ont donc la même dimension. Sachant de plus $\text{Ker}(f) \subset \text{Ker}(f^2)$, on conclut $\text{Ker}(f) = \text{Ker}(f^2)$ par inclusion et égalité des dimensions (Th. 13 p. 243).

(iv) \implies (i) Supposons $\text{Ker}(f) = \text{Ker}(f^2)$.

méthode

|| La formule du rang donne l'hypothèse de dimension qui permet d'établir une supplémentarité de deux espaces en montrant seulement que ceux-ci sont en somme directe (Th. 16 p. 244).

1. Voir sujet 2 p. 282.

Étudions $\text{Im}(f) \cap \text{Ker}(f)$. Soit y un élément de cette intersection. On peut écrire $y = f(x)$ avec $x \in E$ et l'on sait $f(y) = 0_E$ donc $f^2(x) = 0_E$. Ainsi, x est élément de $\text{Ker}(f^2)$ donc de $\text{Ker}(f)$. Par conséquent, $y = f(x) = 0_E$ et l'on peut affirmer¹ $\text{Im}(f) \cap \text{Ker}(f) = \{0_E\}$.

Enfin, la formule du rang donne $\dim E = \dim \text{Im}(f) + \dim \text{Ker}(f)$ et les espaces $\text{Im}(f)$ et $\text{Ker}(f)$ sont donc supplémentaires dans E .

Exercice 5

Soit E et E' deux espaces vectoriels et $u \in \mathcal{L}(E, E')$ injective.

(a) Soit F un sous-espace vectoriel de dimension finie de E . Montrer

$$\dim u(F) = \dim F.$$

(b) Soit (x_1, \dots, x_n) une famille de vecteurs de E . Montrer

$$\text{rg}(u(x_1), \dots, u(x_n)) = \text{rg}(x_1, \dots, x_n).$$

Solution

(a) Commençons par souligner que $u(F)$ est un sous-espace vectoriel de E' car image linéaire d'un sous-espace vectoriel (Th. 2 p. 274).

méthode

Une application linéaire injective transforme une famille libre en une famille libre (Th. 9 p. 276).

Soit (e_1, \dots, e_p) une base de F avec $p = \dim F$. Par l'injectivité de l'application linéaire u , on peut affirmer que la famille image $(u(e_1), \dots, u(e_p))$ est libre. Or celle-ci est aussi génératrice de l'espace $u(F)$. En effet, les vecteurs de $u(F)$ sont les images par u des combinaisons linéaires des vecteurs e_1, \dots, e_p et sont donc les combinaisons linéaires² des vecteurs $u(e_1), \dots, u(e_p)$:

$$u(F) = u(\text{Vect}(e_1, \dots, e_p)) = \text{Vect}(u(e_1), \dots, u(e_p)).$$

La famille $(u(e_1), \dots, u(e_p))$ est alors une base de $u(F)$ et par conséquent

$$\dim u(F) = p = \dim F.$$

(b) Le rang de la famille (x_1, \dots, x_n) est la dimension de $F = \text{Vect}(x_1, \dots, x_n)$. Par le même argument qu'au-dessus, l'image de F par u est l'ensemble des combinaisons linéaires des vecteurs de la famille $(u(x_1), \dots, u(x_n))$ et donc

$$\text{rg}(u(x_1), \dots, u(x_n)) = \dim u(F) = \dim F = \text{rg}(x_1, \dots, x_n).$$

1. Plus précisément, on a seulement établi $\text{Im}(f) \cap \text{Ker}(f) \subset \{0_E\}$ mais on peut affirmer que le vecteur nul appartient évidemment à $\text{Im}(f)$ et $\text{Ker}(f)$ car ce sont deux sous-espaces vectoriels.

2. Voir aussi le sujet 10 p. 290.

Exercice 6 (Interpolation de Lagrange)

Soit a_0, a_1, \dots, a_n des éléments deux à deux distincts de \mathbb{K} .

(a) Montrer que l'application $\varphi: \mathbb{K}[X] \rightarrow \mathbb{K}^{n+1}$ définie par

$$\varphi(P) = (P(a_0), P(a_1), \dots, P(a_n))$$

est linéaire et préciser son noyau.

(b) Établir que la restriction de φ au départ de $\mathbb{K}_n[X]$ réalise un isomorphisme.

Soit $y = (y_0, y_1, \dots, y_n) \in \mathbb{K}^{n+1}$.

(c) Exprimer l'unique polynôme P_0 de $\mathbb{K}_n[X]$ vérifiant $\varphi(P_0) = y$.

(d) Exprimer en fonction de P_0 tous les polynômes P vérifiant $\varphi(P) = y$.

Solution

(a) L'application φ opère entre deux \mathbb{K} -espaces vectoriels. Elle est linéaire car, si λ, μ sont des scalaires et P, Q des polynômes,

$$\begin{aligned}\varphi(\lambda P + \mu Q) &= ((\lambda P + \mu Q)(a_i))_{0 \leq i \leq n} = (\lambda P(a_i) + \mu Q(a_i))_{0 \leq i \leq n} \\ &= \lambda(P(a_i))_{0 \leq i \leq n} + \mu(Q(a_i))_{0 \leq i \leq n} = \lambda\varphi(P) + \mu\varphi(Q).\end{aligned}$$

Pour $P \in \mathbb{K}[X]$, on a $\varphi(P) = 0_{\mathbb{K}^{n+1}}$ si, et seulement si, $P(a_0) = \dots = P(a_n) = 0$. Les a_i étant deux à deux distincts, ceci revient à dire que P est un multiple du polynôme $\Pi = (X - a_0)(X - a_1) \dots (X - a_n)$ (Th. 7 p. 153). Ainsi,

$$\text{Ker}(\varphi) = \Pi \mathbb{K}[X] = \{(X - a_0)(X - a_1) \dots (X - a_n)Q \mid Q \in \mathbb{K}[X]\}.$$

(b) méthode

On peut établir qu'une application linéaire opérant entre deux espaces de même dimension finie est un isomorphisme en étudiant seulement son injectivité ou sa surjectivité (Th. 15 p. 278).

Notons φ' la restriction de φ au départ de $\mathbb{K}_n[X]$. Celle-ci est linéaire tout comme φ et son noyau se déduit¹ de celui de φ :

$$\text{Ker}(\varphi') = \text{Ker}(\varphi) \cap \mathbb{K}_n[X] = \{0\}.$$

Il en découle que l'application φ' est injective (Th. 3 p. 274). Puisque les espaces $\mathbb{K}_n[X]$ et \mathbb{K}^{n+1} ont la même dimension, on peut affirmer que φ' est un isomorphisme.

(c) Notons (e_0, \dots, e_n) la base canonique² de \mathbb{K}^{n+1} et introduisons (L_0, \dots, L_n) son image réciproque par l'isomorphisme φ' : cette famille est une base³ de $\mathbb{K}_n[X]$.

1. De façon générale, le noyau de la restriction u' d'une application linéaire $u \in \mathcal{L}(E, E')$ au départ d'un sous-espace vectoriel F de E est $\text{Ker}(u') = \text{Ker}(u) \cap F$.

2. Par commodité, on indexe celle-ci à partir de 0 pour pouvoir écrire $(y_0, \dots, y_n) = y_0 e_0 + \dots + y_n e_n$.

3. Les polynômes L_0, \dots, L_n sont appelés les *polynômes interpolateurs de Lagrange* en a_0, \dots, a_n .

méthode

|| On exprime les polynômes L_i avant d'en déduire P_0 par linéarité.
 Pour $i \in [0; n]$, le polynôme L_i est entièrement déterminé par les conditions

$$\deg(L_i) \leq n \quad \text{et} \quad L_i(a_j) = \begin{cases} 1 & \text{si } j = i \\ 0 & \text{sinon.} \end{cases}$$

Ces conditions déterminent n racines ce qui suffit à scinder le polynôme L_i . Son coefficient dominant λ est quant à lui déterminé par la valeur $L_i(a_i) = 1$:

$$L_i = \lambda \prod_{\substack{0 \leq j \leq n \\ j \neq i}} (X - a_j) \quad \text{avec} \quad \lambda = \prod_{\substack{0 \leq j \leq n \\ j \neq i}} (a_i - a_j)^{-1}.$$

Enfin, puisque l'on peut écrire $y = y_0 e_0 + \cdots + y_n e_n$, on obtient par linéarité de l'isomorphisme réciproque de φ'

$$P_0 = \varphi'^{-1}(y) = y_0 L_0 + \cdots + y_n L_n = \sum_{i=0}^n \left(y_i \prod_{\substack{0 \leq j \leq n \\ j \neq i}} \frac{X - a_j}{a_i - a_j} \right).$$

(d) méthode

|| L'équation $\varphi(P) = y$ apparaît comme une équation linéaire dont on connaît une solution particulière (Th. 4 p. 274).

L'équation $\varphi(P) = y$ se réécrit $\varphi(P) = \varphi(P_0)$ ce qui donne, compte tenu de la linéarité de φ , l'équation $\varphi(P - P_0) = 0_{\mathbb{K}^{n+1}}$. La résolution de cette dernière correspond à la détermination du noyau de φ menée ci-dessus. Les solutions de l'équation $\varphi(P) = y$ sont donc les polynômes

$$P_0 + \Pi Q \quad \text{avec} \quad Q \in \mathbb{K}[X].$$

Ces solutions décrivent le sous-espace affine passant par P_0 et dirigé par $\text{Ker}(\varphi)$.

8.6 Exercices d'entraînement

8.6.1 Généralités

Exercice 7 *

On note E l'espace vectoriel réel des applications indéfiniment dérивables de \mathbb{R} vers \mathbb{R} . Soit $D: E \rightarrow E$ et $I: E \rightarrow E$ les applications qui associent à $f \in E$ respectivement sa dérivée et sa primitive s'annulant en 0.

- (a) Vérifier que D et I sont des endomorphismes de E .
- (b) Exprimer $D \circ I$ et $I \circ D$.
- (c) Déterminer les images et noyaux de D et I .

Solution

(a) La dérivée et les primitives d'une fonction indéfiniment dérivable sont indéfiniment dérivables : les applications D et I sont bien définies de E vers lui-même. Soit $\lambda, \mu \in \mathbb{R}$ et $f, g \in E$. Par linéarité de la dérivation, on a $D(\lambda f + \mu g) = \lambda D(f) + \mu D(g)$. Aussi¹, les fonctions $I(\lambda f + \mu g)$ et $\lambda I(f) + \mu I(g)$ sont égales car désignent toutes les deux la primitive de $\lambda f + \mu g$ qui s'annule en 0.

Les applications D et I sont donc des endomorphismes de E .

(b) La dérivée d'une primitive est la fonction initiale et donc la composée $D \circ I$ est l'identité. En revanche, la primitive s'annulant en 0 d'une dérivée diffère de la fonction d'une constante : $(I \circ D)(f) = f - f(0)$ pour tout $f \in E$.

(c) **méthode**

Une application linéaire est injective si, et seulement si, son noyau est réduit au vecteur nul (Th. 3 p. 274). Une application est surjective si, et seulement si, son image est égale à son ensemble d'arrivée.

Sachant $D \circ I = \text{Id}_E$, $D \circ I$ est bijective et l'on peut affirmer² que l'application D est surjective tandis que l'application I est injective : $\text{Im}(D) = E$ et $\text{Ker}(I) = \{0\}$.

Les fonctions de dérivées nulles sont les fonctions constantes : elles constituent le noyau de l'application D .

Les valeurs prises par I sont des fonctions qui s'annulent en 0. Inversement, une telle fonction est la valeur prise par I sur sa dérivée. L'image de I est alors exactement constituée des fonctions s'annulant en 0.

Exercice 8 **

Soit f un endomorphisme d'un espace vectoriel E . Établir

- (a) $\text{Im}(f) \cap \text{Ker}(f) = \{0_E\} \iff \text{Ker}(f) = \text{Ker}(f^2)$.
- (b) $E = \text{Im}(f) + \text{Ker}(f) \iff \text{Im}(f) = \text{Im}(f^2)$.

Solution

(a) On raisonne par double implication.

(\Rightarrow) Supposons $\text{Im}(f) \cap \text{Ker}(f) = \{0_E\}$. L'inclusion $\text{Ker}(f) \subset \text{Ker}(f^2)$ est connue³. Étudions l'inclusion réciproque. Soit $x \in \text{Ker}(f^2)$, on a $f^2(x) = f(f(x)) = 0_E$ et donc $f(x)$ appartient au noyau de f . Or $f(x)$ est aussi un élément de l'image de f et donc $f(x) = 0_E$ car on a supposé $\text{Im}(f) \cap \text{Ker}(f) = \{0_E\}$. Ainsi, on obtient l'inclusion $\text{Ker}(f^2) \subset \text{Ker}(f)$ et l'on peut conclure à l'égalité des noyaux.

(\Leftarrow) Supposons $\text{Ker}(f) = \text{Ker}(f^2)$.

1. Il est possible d'exprimer $I(f)$ par une intégrale $I(f)(x) = \int_0^x f(t) dt$ pour tout $x \in \mathbb{R}$.

2. Voir sujet 23 p. 34.

3. Voir sujet 2 p. 282.

méthode

Puisque l'inclusion $\text{Ker}(f) \subset \text{Ker}(f^2)$ est toujours vraie, la force de l'hypothèse réside dans l'inclusion réciproque $\text{Ker}(f^2) \subset \text{Ker}(f)$: un élément qui annule f^2 annule f .

Soit $y \in \text{Im}(f) \cap \text{Ker}(f)$. On peut écrire $y = f(x)$ en introduisant un antécédent $x \in E$. Or $f(y) = 0_E$ et donc $f^2(x) = 0_E$. Ainsi, x est élément de $\text{Ker}(f^2)$ et donc de $\text{Ker}(f)$. On en déduit $y = f(x) = 0_E$. Finalement, $\text{Im}(f) \cap \text{Ker}(f) = \{0_E\}$.

(b) On raisonne à nouveau par double implication.

(\Rightarrow) Supposons $E = \text{Im}(f) + \text{Ker}(f)$. L'inclusion $\text{Im}(f^2) \subset \text{Im}(f)$ est connue. Étudions l'inclusion réciproque. Soit $y \in \text{Im}(f)$. On peut introduire un antécédent $x \in E$ tel que $y = f(x)$. Par l'hypothèse, on peut écrire la décomposition $x = a + b$ avec $a \in \text{Im}(f)$ et $b \in \text{Ker}(f)$. Puisque $a \in \text{Im}(f)$, on peut aussi écrire $a = f(c)$ avec $c \in E$ et alors

$$y = f(f(c) + b) = f^2(c) + f(b) = f^2(c) \in \text{Im}(f^2) \quad \text{car} \quad f(b) = 0_E.$$

Ainsi, on a $\text{Im}(f) \subset \text{Im}(f^2)$ puis l'égalité des deux espaces images.

(\Leftarrow) Supposons $\text{Im}(f) = \text{Im}(f^2)$.

méthode

L'inclusion $\text{Im}(f^2) \subset \text{Im}(f)$ étant toujours vraie, l'intérêt de l'hypothèse¹ réside dans l'affirmation qu'un élément de $\text{Im}(f)$ peut s'écrire comme une valeur prise par f^2 .

Soit $x \in E$. Le vecteur $f(x)$ est élément de l'image de f donc de celle de f^2 . On peut alors introduire $c \in E$ tel que $f(x) = f^2(c)$. Posons ensuite $a = f(c)$ et $b = x - a$. On a clairement $x = a + b$ avec $a \in \text{Im}(f)$. On a aussi $f(b) = f(x) - f(a) = f(x) - f^2(c) = 0_E$ et donc $b \in \text{Ker}(f)$. On peut alors conclure² $E = \text{Im}(f) + \text{Ker}(f)$.

Exercice 9 **

Soit a, b deux réels non nuls distincts et f un endomorphisme d'un espace réel E vérifiant

$$f^2 - (a+b)f + (ab)\text{Id}_E = 0. \quad (*)$$

(a) Montrer que f est inversible et exprimer son inverse en fonction de f .

(b) Établir que $\text{Ker}(f - a\text{Id}_E)$ et $\text{Ker}(f - b\text{Id}_E)$ sont des sous-espaces vectoriels supplémentaires de E .

(c) Vérifier $\text{Ker}(f - a\text{Id}_E) = \text{Im}(f - b\text{Id}_E)$.

1. En écrivant $\text{Im}(f) = \text{Im}(f^2)$, on ne suppose pas que les applications f et f^2 sont égales, on suppose seulement qu'elles prennent les mêmes valeurs dans leur ensemble.

2. Plus précisément, on a seulement montré $E \subset \text{Im}(f) + \text{Ker}(f)$ mais l'inclusion réciproque est entendue car $\text{Im}(f)$ et $\text{Ker}(f)$ sont des sous-espaces vectoriels de E .

Solution(a) **méthode**

|| On transforme la relation (*) en une identité du type $f \circ g = \text{Id}_E$.

En organisant les membres sachant $ab \neq 0$, on peut écrire

$$f \circ \underbrace{\left(\frac{1}{ab}f - \frac{a+b}{ab}\text{Id}_E \right)}_{=g} = \text{Id}_E.$$

On vérifie aussi¹ $g \circ f = \text{Id}_E$ et l'on peut affirmer que f est inversible d'inverse g .

(b) **méthode**

|| Noyau et image d'une application linéaire sont des sous-espaces vectoriels (Th. 2 p. 274).

En tant que noyaux d'endomorphismes, les ensembles $\text{Ker}(f - a\text{Id}_E)$ et $\text{Ker}(f - b\text{Id}_E)$ sont des sous-espaces vectoriels de E . Montrons qu'ils sont supplémentaires en raisonnant par analyse-synthèse. Soit $x \in E$.

Analyse : On suppose $x = u + v$ avec $u \in \text{Ker}(f - a\text{Id}_E)$ et $v \in \text{Ker}(f - b\text{Id}_E)$. On vérifie $f(u) = au$ et $f(v) = bv$ donc

$$\begin{cases} u + v = x \\ au + bv = f(x). \end{cases}$$

Après résolution, les vecteurs u et v sont déterminés de manière unique

$$u = \frac{1}{b-a}(bx - f(x)) \quad \text{et} \quad v = \frac{1}{b-a}(f(x) - ax).$$

Synthèse : Considérons les vecteurs u et v proposés ci-dessus. On vérifie immédiatement $u + v = x$ et l'on étudie ensuite l'appartenance de u et v aux noyaux $\text{Ker}(f - a\text{Id}_E)$ et $\text{Ker}(f - b\text{Id}_E)$. L'identité (*) permet d'écrire

$$(f - a\text{Id}_E) \circ (f - b\text{Id}_E) = 0 \quad \text{et} \quad (f - b\text{Id}_E) \circ (f - a\text{Id}_E) = 0.$$

De plus, les vecteurs u et v s'expriment

$$u = \frac{1}{a-b}(f - b\text{Id}_E)(x) \quad \text{et} \quad v = \frac{1}{b-a}(f - a\text{Id}_E)(x).$$

On a donc effectivement $u \in \text{Ker}(f - a\text{Id}_E)$ et $v \in \text{Ker}(f - b\text{Id}_E)$.

Finalement, les espaces $\text{Ker}(f - a\text{Id}_E)$ et $\text{Ker}(f - b\text{Id}_E)$ sont supplémentaires.

1. Pour affirmer que f est inversible d'inverse g on vérifie les deux égalités $f \circ g = g \circ f = \text{Id}_E$. En dimension finie, on peut se contenter d'une seule (Th. 15 p. 278).

(c) Par l'identité $(f - a\text{Id}_E) \circ (f - b\text{Id}_E) = 0$, on peut affirmer une première inclusion¹ $\text{Im}(f - b\text{Id}_E) \subset \text{Ker}(f - a\text{Id}_E)$. Inversement, pour $x \in \text{Ker}(f - a\text{Id}_E)$, on a $f(x) = ax$ et donc²

$$x = \frac{1}{a-b}(ax - bx) = \frac{1}{a-b}(f(x) - bx) = (f - b\text{Id}_E)(c) \quad \text{avec} \quad c = \frac{1}{a-b}x.$$

On obtient ainsi l'inclusion $\text{Ker}(f - a\text{Id}_E) \subset \text{Im}(f - b\text{Id}_E)$ puis l'égalité.

Exercice 10 **

Soit f une application linéaire d'un espace E vers un espace E' . Montrer que, pour toute partie A de E ,

$$f(\text{Vect}(A)) = \text{Vect}(f(A)).$$

Solution

On raisonne par double inclusion.

méthode

|| Vect(A) est un sous-espace vectoriel contenant A et inclus dans tout sous-espace vectoriel contenant A (Th. 6 p. 240).

La partie A est incluse dans Vect(A) et son image $f(A)$ est incluse dans $f(\text{Vect}(A))$. Or $f(\text{Vect}(A))$ est un sous-espace vectoriel de E' car c'est l'image d'un sous-espace vectoriel par une application linéaire (Th. 2 p. 274). On en déduit $\text{Vect}(f(A)) \subset f(\text{Vect}(A))$.

Inversement, la partie A est incluse dans $f^{-1}(\text{Vect } f(A))$ car les éléments de A sont des antécédents d'éléments de Vect($f(A)$). Aussi, $f^{-1}(\text{Vect } f(A))$ est un sous-espace vectoriel de E en tant qu'image réciproque d'un sous-espace vectoriel par une application linéaire. On en déduit Vect(A) $\subset f^{-1}(\text{Vect } f(A))$. Ceci assure que les valeurs prises par f sur Vect(A) appartiennent à Vect($f(A)$) et donc $f(\text{Vect}(A)) \subset \text{Vect}(f(A))$.

Exercice 11 ***

Soit E, E', E'' trois espaces vectoriels et $u \in \mathcal{L}(E, E')$, $v \in \mathcal{L}(E', E'')$ et $w = v \circ u$.

À quelles conditions sur u et v peut-on affirmer que w est un isomorphisme ?

Solution

Supposons que w soit un isomorphisme. L'application composée $v \circ u$ étant injective, l'application u est nécessairement injective³. De même, l'application v doit être surjective car $v \circ u$ l'est.

méthode

|| L'application u réalise un isomorphisme sur son image tandis que l'application v induit un isomorphisme de tout supplémentaire de son noyau vers E'' (Th. 14 p. 278) : on étudie la supplémentarité de $\text{Im}(u)$ et $\text{Ker}(v)$.

1. Voir sujet 3 p. 282.

2. $\frac{1}{a-b}(f - b\text{Id}_E)$ est le projecteur sur $\text{Ker}(f - a\text{Id}_E)$ parallèlement à $\text{Ker}(f - b\text{Id}_E)$.

3. Voir sujet 23 p. 34.

Montrons par analyse-synthèse que l'image de u et le noyau de v sont supplémentaires dans l'espace E' . Soit $y \in E'$

Analyse : Supposons $y = a + b$ avec $a \in \text{Im}(u)$ et $b \in \text{Ker}(v)$. On peut écrire $a = u(x)$ avec $x \in E$ et l'on a $v(b) = 0_{E''}$. On en déduit $v(y) = v(a) = v(u(x)) = w(x)$. Ceci détermine le vecteur x égal à $w^{-1}(v(y))$ puis les vecteurs a et b : $a = u(x)$ et $b = y - a$. Ainsi, si l'écriture $y = a + b$ est possible, elle est unique.

Synthèse : Posons $x = w^{-1}(v(y))$, $a = u(x)$ et $b = y - a$. On a évidemment $y = a + b$ avec $a \in \text{Im}(u)$. Reste à vérifier l'appartenance de b à $\text{Ker}(v)$. On a $v(b) = v(y) - v(a)$. Or $v(a) = v(u(x)) = w(x)$ donc $v(a) = v(y)$ puis $v(b) = 0_{E''}$.

On peut donc affirmer $\text{Im}(u) \oplus \text{Ker}(v) = E'$.

Inversement, supposons u injective, v surjective et les espaces $\text{Im}(u)$ et $\text{Ker}(v)$ supplémentaires dans E' . Par injectivité, l'application u induit un isomorphisme de E vers son image. Parallèlement, l'application v induit un isomorphisme de $\text{Im}(u)$ (supplémentaire de $\text{Ker}(v)$) vers son image $\text{Im}(v) = E''$. Par composition de ces deux isomorphismes, w est un isomorphisme de E vers E'' .

Finalement, w est un isomorphisme si, et seulement si,

$$\begin{cases} u \text{ injective} \\ v \text{ surjective} \\ \text{Im}(u) \oplus \text{Ker}(v) = E' \end{cases}$$

8.6.2 Projections vectorielles

Exercice 12

Dans l'espace \mathbb{R}^3 , on considère le plan P d'équation $x - y + z = 0$ et la droite vectorielle D engendrée par $u = (1, 3, 1)$.

- (a) On note p la projection sur P parallèlement à D . Exprimer $p(x, y, z)$.
- (b) On note s la symétrie par rapport à P parallèlement à D . Exprimer $s(x, y, z)$.

Solution

Commençons par souligner que le plan P et la droite D sont supplémentaires car le vecteur u n'appartient pas à l'hyperplan P (Th. 16 p. 279).

- (a) Soit $(x, y, z) \in \mathbb{R}^3$. Exprimons $(x', y', z') = p(x, y, z)$.

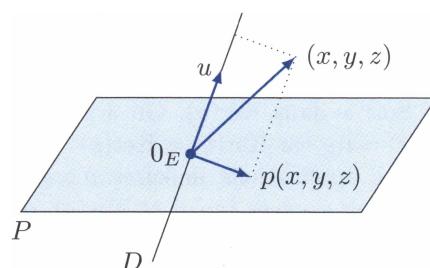
méthode

|| $p(x, y, z)$ est l'unique vecteur de P tel
que $p(x, y, z) - (x, y, z)$ appartient à D .

L'appartenance de $p(x, y, z)$ à P fournit l'équation

$$x' - y' + z' = 0. \quad (*)$$

L'appartenance à D de $p(x, y, z) - (x, y, z)$ signifie que l'on peut écrire ce vecteur λu avec $\lambda \in \mathbb{R}$. Ceci se traduit par le système



$$\begin{cases} x' - x = \lambda \\ y' - y = 3\lambda \\ z' - z = \lambda \end{cases} \quad (**)$$

Par (*), on détermine $\lambda = x - y + z$ et (**) permet alors d'exprimer $p(x, y, z)$:

$$\begin{cases} x' = 2x - y + z \\ y' = 3x - 2y + 3z \\ z' = x - y + 2z \end{cases}$$

(b) méthode

|| La symétrie s se déduit de p par l'identité $s = 2p - \text{Id}_{\mathbb{R}^3}$.

Après quelques calculs, $(x', y', z') = s(x, y, z)$ est exprimé par

$$\begin{cases} x' = 3x - 2y + 2z \\ y' = 6x - 5y + 6z \\ z' = 2x - 2y + 3z \end{cases}$$

Exercice 13 *

Soit p et q deux projecteurs d'un espace vectoriel E .

- (a) Montrer que p et q ont le même noyau si, et seulement si, $p \circ q = p$ et $q \circ p = q$.
- (b) Énoncer une condition semblable pour que p et q possèdent la même image.

Solution

(a) Raisonnons par double implication.

(\Leftarrow) Supposons $p \circ q = p$ et $q \circ p = q$. Pour tout x appartenant au noyau de p , on a $q(x) = q(p(x)) = q(0_E) = 0_E$ et donc x appartient au noyau de q . Ainsi, $\text{Ker}(p) \subset \text{Ker}(q)$. Un raisonnement symétrique fournit l'inclusion réciproque ce qui donne l'égalité.

(\Rightarrow) Supposons $\text{Ker}(p) = \text{Ker}(q)$.

méthode

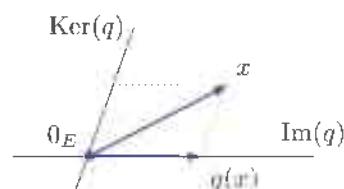
|| On peut montrer l'égalité de deux applications linéaires en constatant que celles-ci sont égales sur des espaces supplémentaires (Th. 13 p. 277).

Puisque q est un projecteur, on sait $\text{Im}(q) \oplus \text{Ker}(q) = E$.

Soit $x \in \text{Im}(q)$. Le vecteur x est invariant pour la projection q et donc $p(q(x)) = p(x)$.

Soit x dans $\text{Ker}(q)$. On a $p(q(x)) = p(0_E) = 0_E$ et $p(x) = 0_E$ car $\text{Ker}(q) = \text{Ker}(p)$.

Les applications linéaires $p \circ q$ et p sont égales sur les deux espaces $\text{Im}(q)$ et $\text{Ker}(q)$, elles sont donc égales sur leur somme E . Un raisonnement symétrique fournit $q \circ p = q$.



(b) **méthode**

|| L'image d'une projection réunit les vecteurs invariants par celle-ci.

Supposons $\text{Im}(p) = \text{Im}(q)$. Sachant $\text{Im}(p) = \text{Ker}(p - \text{Id}_E)$, on peut affirmer que les valeurs prises par q annulent $p - \text{Id}_E$ et donc $(p - \text{Id}_E) \circ q = 0$. On en tire l'égalité $p \circ q = q$. De façon semblable, on obtient aussi $q \circ p = p$.

Inversement, l'égalité $p \circ q = q$ entraîne $\text{Im}(q) \subset \text{Im}(p)$ et l'égalité $q \circ p = p$ entraîne $\text{Im}(p) \subset \text{Im}(q)$. Ainsi, les projecteurs p et q ont la même image si, et seulement si¹, $p \circ q = q$ et $q \circ p = p$.

Exercice 14 *

Soit f_1, \dots, f_n des endomorphismes d'un espace vectoriel E vérifiant :

$$f_1 + \cdots + f_n = \text{Id}_E \quad \text{et} \quad \forall 1 \leq i \neq j \leq n, \quad f_i \circ f_j = 0.$$

(a) Montrer que chaque f_i est une projection vectorielle.

(b) Établir $\bigoplus_{i=1}^n \text{Im}(f_i) = E$.

Solution

(a) Soit $i \in [1; n]$.

méthode

|| On montre que f_i est une projection vectorielle en vérifiant $f_i^2 = f_i$ (Th. 7 p. 276).

En écrivant l'identité de E égale à la somme des f_j , on obtient après simplifications

$$f_i = f_i \circ \text{Id}_E = f_i \circ \sum_{j=1}^n f_j = \sum_{j=1}^n (f_i \circ f_j) = \underbrace{\sum_{\substack{j=1 \\ j \neq i}}^n (f_i \circ f_j)}_{=0} + f_i \circ f_i = f_i \circ f_i.$$

Ainsi, l'endomorphisme f_i est une projection vectorielle.

(b) Commençons par établir que la somme est directe. Supposons $x_1 + \cdots + x_n = 0_E$ avec chaque x_i dans $\text{Im}(f_i)$. En appliquant f_i à cette égalité, on obtient $f_i(x_i) - x_i = 0_E$ car $f_i(x_j) = 0_E$ pour tout $j \neq i$ puisque l'égalité $f_i \circ f_j = 0$ donne $\text{Im}(f_j) \subset \text{Ker}(f_i)$. Ainsi, les espaces $\text{Im}(f_i)$ sont en somme directe. Au surplus, pour tout $x \in E$, on peut écrire

$$x = \text{Id}_E(x) = \sum_{i=1}^n f_i(x) \in \sum_{i=1}^n \text{Im}(f_i).$$

On peut alors conclure

$$\bigoplus_{i=1}^n \text{Im}(f_i) = E.$$

1. En introduisant, les projecteurs complémentaires $p' = \text{Id}_E - p$ et $q' = \text{Id}_E - q$, on échange les espaces images et noyaux : on peut alors mettre en correspondance les deux études qui viennent d'être menées.

Exercice 15 **

Soit p et q deux projecteurs d'un espace vectoriel E .

- (a) Montrer que $p + q$ est un projecteur si, et seulement si¹, $pq = qp = 0$.
- (b) Préciser alors $\text{Im}(p + q)$ et $\text{Ker}(p + q)$.

Solution

(a) Établissons l'équivalence par double implication.

(\implies) Supposons $pq = qp = 0$. On a alors par développement

$$(p + q)^2 = p^2 + \underbrace{pq + qp}_{=0} + q^2 = p + q \quad \text{car } p^2 = p \text{ et } q^2 = q.$$

Ainsi, $p + q$ est un projecteur.

(\impliedby) Supposons $p + q$ projecteur. Par les mêmes calculs que ceux menés au-dessus, l'égalité $(p + q)^2 = p + q$ donne $pq + qp = 0$.

méthode

|| On compose l'identité $pq + qp = 0$ par p à droite et à gauche.

Sachant $p^2 = p$, il vient $p(pq + qp) = pq + pqp$ et $(pq + qp)p = pqp + qp$. Or $pq + qp$ est nul et donc $pq + pqp = 0 = pqp + qp$. On en déduit $pq = qp$ et l'on peut conclure $pq = qp = 0$ car la somme de pq et qp est nulle.

(b) L'inclusion $\text{Im}(p + q) \subset \text{Im}(p) + \text{Im}(q)$ est toujours² vraie. Inversement, pour tout x de $\text{Im}(p) + \text{Im}(q)$, on a $x = a + b$ avec $a \in \text{Im}(p)$ et $b \in \text{Im}(q)$. Les vecteurs de l'image d'une projection sont invariants et donc $p(a) = a$ et $q(b) = b$. Aussi, l'égalité $pq = 0$ donne $\text{Im}(q) \subset \text{Ker}(p)$ et donc $p(b) = 0_E$. On a de même $q(a) = 0_E$ et donc $p(x) + q(x) = a + b = x$ ce qui assure³ $x \in \text{Im}(p + q)$.

On sait aussi $\text{Ker}(p) \cap \text{Ker}(q) \subset \text{Ker}(p + q)$. Inversement, pour tout $x \in \text{Ker}(p + q)$, on a $p(x) + q(x) = 0_E$ donc $p^2(x) + p(q(x)) = 0_E$ puis $p(x) = 0_E$ car $p^2 = p$ et $pq = 0$. Ainsi, $x \in \text{Ker}(p)$ et l'on montre de même $x \in \text{Ker}(q)$.

Finalement, on conclut

$$\text{Im}(p + q) = \text{Im}(p) + \text{Im}(q) \quad \text{et} \quad \text{Ker}(p + q) = \text{Ker}(p) \cap \text{Ker}(q).$$

Exercice 16 **

Soit B un polynôme non constant de $\mathbb{K}[X]$ et r l'application de $\mathbb{K}[X]$ vers lui-même qui à $A \in \mathbb{K}[X]$ associe le reste R de la division euclidienne de A par B .

- (a) Vérifier que r est un endomorphisme de $\mathbb{K}[X]$.
- (b) Calculer $r \circ r$ et préciser la transformation géométrique réalisée par r .

1. Dans l'anneau $(\mathcal{L}(E), +, \circ)$, l'écriture pq est une façon concise de signifier $p \circ q$.

2. Voir sujet 2 p. 282.

3. L'image d'un projecteur p est constituée de vecteurs invariants : pour montrer qu'un élément x appartient à l'image, il est naturel d'étudier si l'égalité $p(x) = x$ est satisfaite.

Solution

(a) L'application r est bien définie de l'espace $\mathbb{K}[X]$ vers lui-même. Soit $\lambda_1, \lambda_2 \in \mathbb{K}$ et $A_1, A_2 \in \mathbb{K}[X]$. Les divisions euclidiennes de A_1 et A_2 par B permettent d'écrire

$$A_1 = BQ_1 + R_1 \quad \text{et} \quad A_2 = BQ_2 + R_2$$

avec Q_1, Q_2 des polynômes et $R_1 = r(A_1)$, $R_2 = r(A_2)$ des polynômes de degrés strictement inférieurs à celui de B . On a alors

$$\lambda_1 A_1 + \lambda_2 A_2 = B(\underbrace{\lambda_1 Q_1 + \lambda_2 Q_2}_{=Q \in \mathbb{K}[X]}) + \underbrace{\lambda_1 R_1 + \lambda_2 R_2}_{=R \in \mathbb{K}[X]}.$$

méthode

|| Pour identifier le reste et le quotient d'une division euclidienne, il faut disposer d'une identité $A = BQ + R$ mais aussi de la condition $\deg(R) < \deg(B)$.

Le degré d'une combinaison linéaire de polynômes étant inférieur aux degrés des polynômes intervenant dans celle-ci

$$\deg(R) \leq \max(\deg(R_1), \deg(R_2)) < \deg(B).$$

Le reste de la division euclidienne de $\lambda_1 A_1 + \lambda_2 A_2$ par B est donc exactement le polynôme R ce qui permet d'écrire

$$r(\lambda_1 A_1 + \lambda_2 A_2) = R = \lambda_1 R_1 + \lambda_2 R_2 = \lambda_1 r(A_1) + \lambda_2 r(A_2).$$

L'application r est donc bien un endomorphisme de $\mathbb{K}[X]$.

(b) Les valeurs prises par l'endomorphisme r sont des polynômes de degrés strictement inférieurs à celui de B . Aussi, lorsqu'un polynôme A est de degré strictement inférieur à celui de B , la division euclidienne de A par B s'écrit simplement $A = B \times 0 + A$ et donc $r(A) = A$. On en déduit $r \circ r = r$: l'endomorphisme r est un projecteur.

méthode

|| Un projecteur projette sur son image parallèlement à son noyau et ces deux espaces sont supplémentaires (Th. 7 p. 276).

Par l'étude ci-dessus, on a vu que l'image de r est incluse dans l'espace des polynômes de degrés strictement inférieurs à celui de B et que, inversement, un tel polynôme est sa propre image. L'image de r est donc $\mathbb{K}_{n-1}[X]$ avec $n = \deg(B) \in \mathbb{N}^*$. Le noyau de r est quant à lui constitué des polynômes divisibles par B . On peut alors conclure que r est la projection vectorielle sur $\mathbb{K}_{n-1}[X]$ parallèlement à $B\mathbb{K}[X]$.

8.6.3 Applications linéaires en dimension finie

Exercice 17 *

Soit f un endomorphisme d'un \mathbb{K} -espace vectoriel E de dimension finie $n \geq 1$.

On suppose qu'il existe¹ un entier $p \geq 1$ tel que $f^p = 0$ et l'on considère le plus petit entier p vérifiant cette propriété.

- (a) Soit $x \notin \text{Ker}(f^{p-1})$. Montrer la liberté de $(x, f(x), f^2(x), \dots, f^{p-1}(x))$.
- (b) En déduire que $f^n = 0$.

Solution

(a) Par définition de l'entier p , on peut affirmer que l'endomorphisme f^{p-1} est non nul et l'on peut donc introduire un vecteur x n'appartenant pas à $\text{Ker}(f^{p-1})$, c'est-à-dire un vecteur tel que $f^{p-1}(x) \neq 0_E$. Supposons

$$\lambda_0 x + \lambda_1 f(x) + \cdots + \lambda_{p-1} f^{p-1}(x) = 0_E \quad \text{avec } \lambda_0, \dots, \lambda_{p-1} \in \mathbb{K}. \quad (*)$$

méthode

|| On simplifie l'identité (*) en composant celle-ci plusieurs fois par f .

En composant (*) une première fois par f , l'équation se réduit en

$$\lambda_0 f(x) + \lambda_1 f^2(x) + \cdots + \lambda_{p-2} f^{p-1}(x) = 0_E \quad \text{car } f^p(x) = 0_E.$$

En composant plusieurs fois par f , on obtient les équations du système suivant :

$$\left\{ \begin{array}{l} \lambda_0 x + \lambda_1 f(x) + \cdots + \lambda_{p-2} f^{p-2}(x) + \lambda_{p-1} f^{p-1}(x) = 0_E \\ \lambda_0 f(x) + \cdots + \lambda_{p-3} f^{p-2}(x) + \lambda_{p-2} f^{p-1}(x) = 0_E \\ \vdots \\ \lambda_0 f^{p-2}(x) + \lambda_1 f^{p-1}(x) = 0_E \\ \lambda_0 f^{p-1}(x) = 0_E. \end{array} \right.$$

Sachant $f^{p-1}(x) \neq 0_E$, la dernière équation donne $\lambda_0 = 0$. Ceci permet de simplifier l'équation précédente qui devient $\lambda_1 f^{p-1}(x) = 0_E$ et donne $\lambda_1 = 0$. Ainsi, on remonte le système pour obtenir la nullité de tous les λ_i : la famille $(x, f(x), \dots, f^{p-1}(x))$ est libre.

(b) Comme une famille libre comporte moins d'éléments que la dimension de l'espace, on peut affirmer $p \leq n$. Or $f^p = 0$ et donc² $f^n = f^{n-p} \circ f^p = 0$.

Exercice 18 *

Soit f et g deux endomorphismes d'un espace E de dimension finie vérifiant

$$f^2 + f \circ g = \text{Id}_E.$$

Montrer que f et g commutent.

1. On dit que l'endomorphisme est *nilpotent*.
2. L'écriture est possible car l'exposant $n - p$ est positif.

Solution**méthode**

On montre que f est inversible et l'on détermine son inverse par le théorème d'isomorphisme (Th. 15 p. 278).

On peut écrire $f \circ h = \text{Id}_E$ avec $h = f + g \in \mathcal{L}(E)$. L'endomorphisme f est donc inversible à droite et le théorème d'isomorphisme assure alors que f est inversible d'inverse h . En particulier, on a aussi $h \circ f = \text{Id}_E$ ce qui fournit l'égalité $f^2 + g \circ f = \text{Id}_E$. Celle-ci, combinée à l'hypothèse de départ, permet de conclure $f \circ g = g \circ f$.

Exercice 19 **

Soit f un endomorphisme d'un \mathbb{K} -espace vectoriel E de dimension $n \in \mathbb{N}^*$. On suppose qu'il existe un vecteur $x_0 \in E$ pour lequel la famille $(x_0, f(x_0), \dots, f^{n-1}(x_0))$ est une base de E et l'on introduit

$$\mathcal{C}_f = \{g \in \mathcal{L}(E) \mid g \circ f = f \circ g\}.$$

(a) Observer que

$$\mathcal{C}_f = \{a_0 \text{Id}_E + a_1 f + \dots + a_{n-1} f^{n-1} \mid (a_0, \dots, a_{n-1}) \in \mathbb{K}^n\}.$$

(b) Justifier que \mathcal{C}_f est un sous-espace vectoriel de $\mathcal{L}(E)$ dont on précisera la dimension.

Solution

(a) On raisonne par double inclusion.

Soit ¹ $h = a_0 \text{Id}_E + a_1 f + \dots + a_{n-1} f^{n-1} \in \mathcal{L}(E)$ avec $a_0, \dots, a_{n-1} \in \mathbb{K}$. On vérifie par le calcul que h commute avec f

$$h \circ f = a_0 f + a_1 f^2 + \dots + a_{n-1} f^n = f \circ h.$$

Ainsi, on dispose d'une première inclusion

$$\{a_0 \text{Id}_E + a_1 f + \dots + a_{n-1} f^{n-1} \mid (a_0, \dots, a_{n-1}) \in \mathbb{K}^n\} \subset \mathcal{C}_f.$$

Inversément, soit $g \in \mathcal{C}_f$.

méthode

Par analyse, si l'endomorphisme g s'écrit $a_0 \text{Id}_E + a_1 f + \dots + a_{n-1} f^{n-1}$, les scalaires a_0, a_1, \dots, a_{n-1} se comprennent comme les coordonnées du vecteur $g(x_0)$ dans la base $(x_0, f(x_0), \dots, f^{n-1}(x_0))$.

En notant $(a_0, a_1, \dots, a_{n-1})$ la famille des coordonnées de $g(x_0)$ dans la base introduite, on peut écrire

$$\begin{aligned} g(x_0) &= a_0 x_0 + a_1 f(x_0) + \dots + a_{n-1} f^{n-1}(x_0) \\ &= (a_0 \text{Id}_E + a_1 f + \dots + a_{n-1} f^{n-1})(x_0). \end{aligned}$$

¹ On dit que l'endomorphisme h est un *polynôme* en f .

Notons h l'endomorphisme $a_0\text{Id}_E + a_1f + \cdots + a_{n-1}f^{n-1}$ introduit de sorte que l'identité précédente s'exprime $g(x_0) = h(x_0)$. Il s'agit d'étendre cette égalité à tout vecteur de E .

méthode

On peut montrer que deux applications linéaires sont égales en constatant que celles-ci coïncident sur une base (Th. 12 p. 277).

On a $g(x_0) = h(x_0)$ et donc, en composant par f , $f(g(x_0)) = f(h(x_0))$. Or g et h commutent avec f et il vient $g(f(x_0)) = h(f(x_0))$. De façon plus générale, on peut écrire pour k naturel

$$g(f^k(x_0)) = f^k(g(x_0)) = f^k(h(x_0)) = h(f^k(x_0))$$

car g et h commutent avec f^k . Ainsi, les applications linéaires g et h prennent les mêmes valeurs sur chaque vecteur de la base $(x_0, f(x_0), \dots, f^{n-1}(x_0))$, elles sont donc égales. On peut alors conclure à l'inclusion réciproque

$$\mathcal{C}_f \subset \{a_0\text{Id}_E + a_1f + \cdots + a_{n-1}f^{n-1} \mid (a_0, \dots, a_{n-1}) \in \mathbb{K}^n\}$$

puis l'égalité.

(b) Ce qui précède fournit $\mathcal{C}_f = \text{Vect}(\text{Id}_E, f, \dots, f^{n-1})$. L'ensemble \mathcal{C}_f est donc un sous-espace vectoriel de $\mathcal{L}(E)$.

méthode

On détermine la dimension de \mathcal{C}_f en proposant une base de cet espace.

La famille $(\text{Id}_E, f, \dots, f^{n-1})$ est génératrice de \mathcal{C}_f . Vérifions que c'est aussi une famille libre.

Soit $\lambda_0, \dots, \lambda_{n-1}$ des scalaires tels que $\lambda_0\text{Id}_E + \lambda_1f + \cdots + \lambda_{n-1}f^{n-1} = 0$. Pour tout vecteur x de E , on a l'égalité $\lambda_0x + \lambda_1f(x) + \cdots + \lambda_{n-1}f^{n-1}(x) = 0_E$. Ceci vaut en particulier en x_0 et donc $\lambda_0x_0 + \lambda_1f(x_0) + \cdots + \lambda_{n-1}f^{n-1}(x_0) = 0_E$. Or la famille $(x_0, f(x_0), \dots, f^{n-1}(x_0))$ est libre et donc $\lambda_0 = \lambda_1 = \cdots = \lambda_{n-1} = 0$.

Finalement, la famille $(\text{Id}_E, f, \dots, f^{n-1})$ est une base de \mathcal{C}_f qui est donc un espace de dimension n .

8.6.4 Rang d'une application linéaire

Exercice 20 *

Soit f et g deux endomorphismes d'un espace E de dimension finie vérifiant

$$\text{Im}(f) + \text{Im}(g) = \text{Ker}(f) + \text{Ker}(g) = E.$$

Montrer que ces sommes sont directes.

Solution**méthode**

|| On exploite la formule de Grassmann et la formule du rang.

La formule de Grassmann donne

$$\begin{aligned}\dim E + \dim(\text{Im}(f) \cap \text{Im}(g)) &= \dim \text{Im}(f) + \dim \text{Im}(g) \text{ et} \\ \dim E + \dim(\text{Ker}(f) \cap \text{Ker}(g)) &= \dim \text{Ker}(f) + \dim \text{Ker}(g).\end{aligned}$$

La formule du rang donne

$$\begin{aligned}\dim \text{Im}(f) + \dim \text{Ker}(f) &= \dim E \text{ et} \\ \dim \text{Im}(g) + \dim \text{Ker}(g) &= \dim E.\end{aligned}$$

En sommant ces quatre égalités et en simplifiant, on obtient

$$\dim(\text{Im}(f) \cap \text{Im}(g)) + \dim(\text{Ker}(f) \cap \text{Ker}(g)) = 0.$$

Ces deux dimensions sont donc nulles et alors

$$\text{Im}(f) \cap \text{Im}(g) = \text{Ker}(f) \cap \text{Ker}(g) = \{0_E\}.$$

Les sommes des images et des noyaux sont donc toutes les deux directes.

Exercice 21 *

Soit $n \in \mathbb{N}$ et $\Delta: \mathbb{K}_{n+1}[X] \rightarrow \mathbb{K}_n[X]$ l'application définie¹ par

$$\Delta(P) = P(X+1) - P(X).$$

(a) Montrer que Δ définit une application linéaire.

(b) Déterminer le noyau de Δ et établir que l'application Δ est surjective.

Solution**(a) méthode**

|| On vérifie que l'application Δ est bien définie à valeurs dans $\mathbb{K}_n[X]$.

Soit P un polynôme de $\mathbb{K}_{n+1}[X]$. Les deux polynômes $P(X+1)$ et $P(X)$ ont les mêmes degrés et les mêmes coefficients dominants : il y a simplification du terme de plus haut degré lors du calcul de $P(X+1) - P(X)$. On en déduit que $\Delta(P)$ est de degré au plus n et l'application Δ prend effectivement ses valeurs dans $\mathbb{K}_n[X]$. De plus, pour $\lambda, \mu \in \mathbb{K}$ et $P, Q \in \mathbb{K}_{n+1}[X]$, on vérifie :

$$\begin{aligned}\Delta(\lambda P + \mu Q) &= (\lambda P + \mu Q)(X+1) - (\lambda P + \mu Q)(X) \\ &= \lambda(P(X+1) - P(X)) + \mu(Q(X+1) - Q(X)) \\ &= \lambda\Delta(P) + \mu\Delta(Q).\end{aligned}$$

L'application Δ définit une application linéaire de $\mathbb{K}_{n+1}[X]$ vers $\mathbb{K}_n[X]$.

1. L'écriture $P(X+1)$ fait référence à la composition de deux polynômes et non à un produit : on remplace X par $X+1$ dans l'expression du polynôme P . L'écriture $P(X)$ fait directement référence à P .

(b) Soit $P \in \mathbb{K}_{n+1}[X]$. On a $\Delta(P) = 0$ si, et seulement si, $P(X+1) = P(X)$. Un tel polynôme est périodique et donc constant¹. Le noyau de Δ est donc constitué des polynômes constants.

méthode

Par la formule du rang (Th. 14 p. 278), on peut lier la dimension de l'image d'une application linéaire à celle de son noyau.

Le noyau de Δ étant de dimension 1

$$\text{rg}(\Delta) = \dim \underbrace{\mathbb{K}_{n+1}[X]}_{\text{espace de départ}} - \dim \text{Ker}(\Delta) = n + 2 - 1 = n + 1 = \dim \underbrace{\mathbb{K}_n[X]}_{\text{espace d'arrivée}}.$$

On en déduit que l'application linéaire Δ est surjective.

Exercice 22 **

Soit F et G deux sous-espaces vectoriels d'un espace E de dimension finie n . Enoncer une condition nécessaire et suffisante portant sur F et G pour qu'il existe un endomorphisme u de E tel que $\text{Ker}(u) = F$ et $\text{Im}(u) = G$.

Solution

Par la formule du rang, la condition $\dim F + \dim G = n$ est nécessaire. Vérifions qu'elle est aussi suffisante.

méthode

On construit un endomorphisme en fixant l'image d'une base (Th. 12 p. 277).

Posons $p = \dim F$ et $q = \dim G$ avec $p + q = n$. Considérons (e_1, \dots, e_p) une base de F que l'on complète en (e_1, \dots, e_n) base de E . Introduisons aussi (e'_1, \dots, e'_q) une base de G . Enfin, considérons l'endomorphisme u déterminé par

$$\begin{aligned} \forall j \in \llbracket 1 ; p \rrbracket, \quad & u(e_j) = 0_E \text{ et} \\ \forall k \in \llbracket 1 ; q \rrbracket, \quad & u(e_{p+k}) = e'_k. \end{aligned}$$

$$\begin{array}{ccccccc} e_1 & \cdots & e_p & e_{p+1} & \cdots & e_n \\ \downarrow & & \downarrow & \downarrow & & \downarrow \\ 0_E & & 0_E & e'_1 & & e'_q \end{array}$$

Par construction, l'endomorphisme u s'annule assurément sur F et toutes les vecteurs de G sont des valeurs prises par u . On dispose donc des inclusions

$$F \subset \text{Ker}(u) \quad \text{et} \quad G \subset \text{Im}(u).$$

Cependant, $\dim F + \dim G = n = \dim \text{Ker}(u) + \dim \text{Im}(u)$. Les espaces F et $\text{Ker}(u)$ d'une part, les espaces G et $\text{Im}(u)$ d'autre part, ont donc les mêmes dimensions et sont par conséquent égaux.

Exercice 23 **

Soit f et g deux endomorphismes d'un espace E de dimension finie.

(a) Établir $|\text{rg}(f) - \text{rg}(g)| \leq \text{rg}(f + g) \leq \text{rg}(f) + \text{rg}(g)$.

(b) Établir $\text{rg}(f) + \text{rg}(g) - n \leq \text{rg}(f \circ g) \leq \min(\text{rg}(f), \text{rg}(g))$.

1. Voir sujet 13 p. 176.

Solution

(a) On sait $\text{Im}(f+g) \subset \text{Im}(f) + \text{Im}(g)$ donc

$$\text{rg}(f+g) \leq \dim(\text{Im}(f) + \text{Im}(g)) \leq \dim \text{Im}(f) + \dim \text{Im}(g) = \text{rg}(f) + \text{rg}(g).$$

En écrivant $f = (f+g) + (-g)$ et sachant $\text{rg}(-g) = \text{rg}(g)$, il vient

$$\text{rg}(f) \leq \text{rg}(f+g) + \text{rg}(g) \quad \text{donc} \quad \text{rg}(f) - \text{rg}(g) \leq \text{rg}(f+g).$$

En inversant les rôles de f et g , on a aussi $\text{rg}(g) - \text{rg}(f) \leq \text{rg}(f+g)$ et donc

$$|\text{rg}(f) - \text{rg}(g)| \leq \text{rg}(f+g).$$

(b) L'image de $f \circ g$ est incluse dans l'image de f et donc $\text{rg}(f \circ g) \leq \text{rg}(f)$.

méthode

|| L'image de $f \circ g$ est l'image de la restriction de f au départ de $\text{Im}(g)$.

Notons $f' \in \mathcal{L}(\text{Im}(g), E)$ cette restriction : $\text{rg}(f \circ g) = \text{rg}(f')$. En appliquant la formule du rang (Th. 14 p. 278) à l'application linéaire f' , on obtient

$$\text{rg}(f \circ g) = \text{rg}(f') = \dim \text{Im}(g) - \dim \text{Ker}(f') = \text{rg}(g) - \dim \text{Ker}(f'). \quad (*)$$

On en déduit la comparaison $\text{rg}(f \circ g) \leq \text{rg}(g)$ puis $\text{rg}(f \circ g) \leq \min(\text{rg}(f), \text{rg}(g))$.

Aussi, puisque le noyau de f' est inclus dans celui de f , l'égalité (*) donne

$$\text{rg}(f \circ g) \geq \text{rg}(g) - \dim \text{Ker}(f)$$

avec $\dim \text{Ker}(f) = n - \text{rg}(f)$. On en déduit la comparaison

$$\text{rg}(f \circ g) \geq \text{rg}(f) + \text{rg}(g) - n.$$

Notons qu'en appliquant la formule du rang à f , g et $f \circ g$, il vient alors

$$\dim \text{Ker}(f \circ g) \leq \dim \text{Ker}(f) + \dim \text{Ker}(g).$$

Exercice 24 ** (Images et noyaux itérés d'un endomorphisme)

Soit f un endomorphisme d'un espace vectoriel E . Pour tout $p \in \mathbb{N}$, on introduit les images et noyaux de $f^p = f \circ \cdots \circ f$ (p facteurs) :

$$I_p = \text{Im}(f^p) \quad \text{et} \quad N_p = \text{Ker}(f^p).$$

(a) Montrer que les suites (I_p) et (N_p) sont respectivement décroissante et croissante au sens de l'inclusion.

On suppose dans ce qui suit que l'espace E est de dimension finie.

(b) Justifier l'existence d'un rang $r \in \mathbb{N}$ tel que $I_{r+1} = I_r$.

(c) Vérifier que les deux suites (I_p) et (N_p) sont alors constantes à partir du rang r .

(d) Établir $I_r \oplus N_r = E$.

Solution

(a) Soit $y \in \text{Im}(f^{p+1})$. Il existe $x \in E$ tel que $y = f^{p+1}(x) = f^p(f(x))$ et donc y est une valeur prise par f^p . Ainsi, on a l'inclusion $I_{p+1} \subset I_p$.

Soit $x \in \text{Ker}(f^p)$. On a $f^p(x) = 0_E$ donc $f^{p+1}(x) = f(f^p(x)) = f(0_E) = 0_E$ et x annule l'endomorphisme f^{p+1} . Ainsi, on a l'inclusion $N_p \subset N_{p+1}$.

Plus généralement, pour f et g endomorphismes, on a les inclusions

$$\text{Im}(g \circ f) \subset \text{Im}(g) \quad \text{et} \quad \text{Ker}(g) \subset \text{Ker}(f \circ g).$$

Ces inclusions sont ici particularisées au cas $g = f^n$.

(b) méthode

|| La suite des dimensions des I_p est décroissante.

La suite $(\dim I_p)$ est une suite décroissante d'entiers naturels : elle ne peut être strictement décroissante et il existe donc un rang $r \in \mathbb{N}$ tel que $\dim I_r = \dim I_{r+1}$. Par inclusion et égalité de leurs dimensions, les espaces I_r et I_{r+1} sont égaux.

(c) méthode

|| On propage l'égalité $I_{p+1} = I_p$ aux rangs $p \geq r$ en écrivant $f^p = f^{p-r} \circ f^r$.

Soit $p \geq r$. On sait l'inclusion $I_{p+1} \subset I_p$. Inversement, considérons $y \in I_p$. Il existe x dans E tel que

$$y = f^p(x) = f^{p-r}(f^r(x)).$$

Puisque $f^r(x)$ est élément de I_r avec $I_r = I_{r+1}$, on peut introduire un élément a de E tel que $f^r(x) = f^{r+1}(a)$ et alors $y = f^{p+1}(a)$. Le vecteur y est donc élément de I_{p+1} . Ainsi, on obtient $I_p \subset I_{p+1}$ puis l'égalité. La suite (I_p) est alors constante à partir du rang r .

Les dimensions des espaces images I_p étant les mêmes au delà du rang r , la formule du rang assure que, parallèlement, les dimensions des espaces noyaux N_p sont elles aussi identiques au delà du rang r . Par inclusion et égalité des dimensions, la suite (N_p) est constante à partir du rang r .

(d) méthode

|| Par la formule du rang, on peut établir la supplémentarité en montrant seulement que la somme est directe (Th. 16 p. 244).

Soit $x \in I_r \cap N_r$. Il existe un antécédent $a \in E$ tel que $x = f^r(a)$ et l'on a $f^r(x) = 0$. On en déduit $f^{2r}(a) = 0_E$ et a est donc élément de $N_{2r} = N_r$. Le vecteur x est par conséquent nul. Ainsi, $I_r \cap N_r = \{0_E\}$. De plus, la formule du rang donne

$$\dim I_r + \dim N_r = \dim E$$

et l'on peut affirmer que les espaces I_r et N_r sont supplémentaires.

Exercice 25 ***

Soit u un endomorphisme d'un espace de dimension finie E .

Montrer qu'il existe un endomorphisme v de E tel que $uv = 0$ et $u + v \in \text{GL}(E)$ si, et seulement si, les espaces $\text{Im}(u)$ et $\text{Ker}(u)$ sont supplémentaires.

Solution

On raisonne par double implication.

(\Rightarrow) On suppose $uv = 0$ et $u + v$ inversible.

méthode

Par la bijectivité de $u + v$, on peut écrire un vecteur de E comme somme d'un vecteur de $\text{Im}(u)$ et d'un vecteur de $\text{Im}(v)$.

Soit x un vecteur de E . Il existe un antécédent $a \in E$ tel que $x = u(a) + v(a)$. Le vecteur $u(a)$ appartient à l'image de u tandis que le vecteur $v(a)$ appartient à l'image de v et donc au noyau de u car $u \circ v$ est l'application nulle. On en déduit $E = \text{Im}(u) + \text{Ker}(u)$ et cette somme est directe en vertu de la formule du rang.

(\Leftarrow) On suppose $\text{Im}(u) \oplus \text{Ker}(u) = E$.

méthode

On introduit la projection v sur $\text{Ker}(u)$ parallèlement à $\text{Im}(u)$.

Puisque v prend ses valeurs dans $\text{Ker}(u)$, on peut affirmer $uv = 0$. Aussi, si x est élément de $\text{Ker}(u + v)$, on a $u(x) + v(x) = 0_E$ donc

$$u(x) \in \text{Im}(u) \cap \text{Im}(v) = \text{Im}(u) \cap \text{Ker}(u) = \{0_E\}.$$

Les vecteurs $u(x)$ et $v(x)$ sont donc nuls et $x \in \text{Ker}(u) \cap \text{Ker}(v) = \text{Ker}(u) \cap \text{Im}(u)$ ce qui permet de conclure que x est nul. L'endomorphisme $u + v$ est alors injectif et c'est donc un automorphisme car l'espace E est de dimension finie.

8.6.5 Formes linéaires et hyperplans**Exercice 26 ****

Soit E un espace de dimension finie $n \geq 1$ et F un sous-espace vectoriel distinct de E .

(a) Montrer que F peut s'écrire comme une intersection d'un nombre fini d'hyperplans.

(b) Quel est le nombre minimum d'hyperplans nécessaire ?

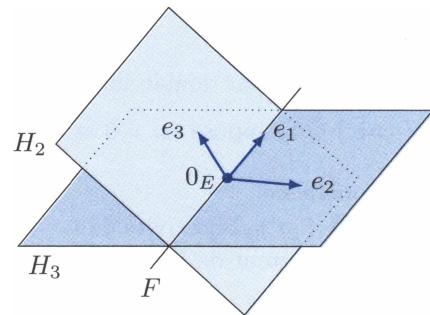
Solution

(a) Notons p la dimension de F . Soit (e_1, \dots, e_p) une base de F que l'on complète en $e = (e_1, \dots, e_n)$ base de E .

méthode

|| On introduit les formes linéaires coordonnées dans la base e .

Notons φ_i la forme linéaire qui à $x \in E$ associe la coordonnée d'indice i du vecteur x dans la base e . L'application φ_i est une forme linéaire non nulle et son noyau H_i est un hyperplan. Puisque les vecteurs de F sont ceux dont les coordonnées d'indices $p+1$ à n dans la base e sont nulles, on a $F = H_{p+1} \cap \dots \cap H_n$. Ainsi, on exprime F comme l'intersection de $n-p$ hyperplans.



(b) Il ne peut pas y avoir moins d'hyperplans pour décrire F . En effet, vérifions par récurrence que l'intersection de q hyperplans de E détermine un espace de dimension au moins égale à $n-q$.

Pour $q=0$ ou $q=1$, la propriété est entendue. Supposons la propriété vraie au rang $q > 0$. Soit H_1, \dots, H_q, H_{q+1} des hyperplans de E . Posons $G = H_1 \cap \dots \cap H_q$. L'hypothèse de récurrence assure $\dim G \geq n-q$. Aussi, $G + H_{q+1}$ est un sous-espace vectoriel de E et donc de dimension inférieure à n . La formule de Grassmann donne alors

$$\dim(G \cap H_{q+1}) = \underbrace{\dim G}_{\geq n-q} + \underbrace{\dim H_{q+1}}_{=n-1} - \underbrace{\dim(G + H_{q+1})}_{\leq n} \geq n-q-1.$$

La récurrence est établie.

Exercice 27 ***

Soit $\varphi_1, \dots, \varphi_n$ des formes linéaires sur un espace E de dimension $n \in \mathbb{N}^*$. Montrer que la famille $(\varphi_1, \dots, \varphi_n)$ constitue une base du dual de E si, et seulement si,

$$\bigcap_{i=1}^n \text{Ker}(\varphi_i) = \{0_E\}.$$

Solution

Notons que le dual $E^* = \mathcal{L}(E, \mathbb{K})$ est un espace de dimension n ce qui correspond à la longueur de la famille $(\varphi_1, \dots, \varphi_n)$. Raisonnons par double implication.

(\Rightarrow) Supposons la famille $(\varphi_1, \dots, \varphi_n)$ base de E^* .

méthode

|| Pour $x \neq 0_E$, il existe une forme linéaire φ sur E vérifiant $\varphi(x) \neq 0$.

Soit x un vecteur de l'intersection des noyaux des φ_i . Si, par l'absurde, x est non nul, le vecteur $e_1 = x$ constitue une famille libre que l'on peut compléter en une base (e_1, \dots, e_n) .

La forme linéaire φ lisant la première coordonnée dans cette base vérifie $\varphi(x) = 1$ et donc $\varphi(x) \neq 0$. Or φ est combinaison linéaire des formes linéaires φ_i et ces dernières s'annulent toutes en x : c'est absurde. On en déduit que seul le vecteur nul peut appartenir à l'intersection des noyaux des φ_i .

(\Leftarrow) Raisonnons par contraposition. Si la famille $(\varphi_1, \dots, \varphi_n)$ n'est pas une base, c'est une famille liée. L'un des éléments de cette famille est alors combinaison linéaire des autres. Quitte à reprendre l'indexation des formes linéaires, on peut supposer que φ_n est combinaison linéaire des $\varphi_1, \dots, \varphi_{n-1}$. On a alors

$$\bigcap_{i=1}^{n-1} \text{Ker}(\varphi_i) \subset \text{Ker}(\varphi_n) \quad \text{donc} \quad \bigcap_{i=1}^n \text{Ker}(\varphi_i) = \bigcap_{i=1}^{n-1} \text{Ker}(\varphi_i).$$

Cependant, on peut établir par une récurrence semblable à celle vue dans le sujet précédent l'inégalité

$$\dim \left(\bigcap_{i=1}^q \text{Ker}(\varphi_i) \right) \geq n - q \quad \text{pour tout } q \in [0; n].$$

En particulier, l'espace $\bigcap_{i=1}^{n-1} \text{Ker}(\varphi_i)$ ne peut être nul car de dimension au moins 1.

8.7 Exercices d'approfondissement

Exercice 28 *

Soit F et G deux sous-espaces de dimensions finies d'un espace vectoriel E . Retrouver la formule de Grassmann en appliquant le théorème du rang à la fonction

$$\sigma: \begin{cases} F \times G \rightarrow E \\ (x, y) \mapsto x + y. \end{cases}$$

Solution

L'application σ est linéaire au départ de l'espace $F \times G$ qui est de dimension finie. Par définition de la somme de deux sous-espaces vectoriels, son espace image est $F + G$. La formule du rang donne alors

$$\dim(F \times G) = \dim(F + G) + \dim \text{Ker}(\sigma) \tag{*}$$

avec $\dim(F \times G) = \dim F + \dim G$ (Th. 11 p. 242).

méthode

On vérifie que $\text{Ker}(\sigma)$ est isomorphe à $F \cap G$.

Le noyau de σ est constitué des couples $(x, y) \in F \times G$ vérifiant $x + y = 0_E$. Ces couples sont de la forme $(x, -x)$ avec $x \in F \cap G$. Considérons alors l'application ϕ de $F \cap G$

vers $\text{Ker}(\sigma)$ définie par $\phi(x) = (x, -x)$. Celle-ci est bien définie, linéaire, injective et surjective : c'est un isomorphisme. On en déduit que les espaces $F \cap G$ et $\text{Ker}(\sigma)$ ont la même dimension (Th. 11 p. 277). La relation (*) se relit alors

$$\dim F + \dim G = \dim(F + G) + \dim(F \cap G).$$

Exercice 29 **

Soit u un endomorphisme d'un espace vectoriel E . A quelle(s) condition(s) un sous-espace vectoriel F de E vérifie-t-il $u^{-1}(u(F)) = u(u^{-1}(F))$?

Solution

méthode

Les inclusions qui suivent sont toujours vraies¹

$$u(u^{-1}(F)) \subset F \quad \text{et} \quad F \subset u^{-1}(u(F)).$$

Si $u(u^{-1}(F)) = u^{-1}(u(F))$ alors $u(u^{-1}(F)) = F$ et $u^{-1}(u(F)) = F$.

L'inclusion $u^{-1}(u(F)) \subset F$ entraîne $\text{Ker}(u) \subset F$ car le vecteur nul est élément de $u(F)$. L'inclusion $F \subset u(u^{-1}(F))$ entraîne quant à elle $F \subset \text{Im}(u)$.

Inversement, supposons $\text{Ker}(u) \subset F \subset \text{Im}(u)$. Pour $x \in u^{-1}(u(F))$, on a $u(x) \in u(F)$ et il existe donc $a \in F$ tel que $u(x) = u(a)$. On a alors $x - a \in \text{Ker}(u)$ puis on peut écrire $x = a + (x - a) \in F$ par addition dans F . Ainsi, $u^{-1}(u(F)) \subset F$ puis $u^{-1}(u(F)) = F$.

Aussi, pour $y \in F$, il existe $a \in E$ tel que $y = u(a)$ car $F \subset \text{Im}(u)$. Puisque $y \in F$, on a $a \in u^{-1}(F)$ puis $y = u(a) \in u(u^{-1}(F))$. Ainsi, $F \subset u(u^{-1}(F))$ puis $F = u(u^{-1}(F))$.

Finalement,

$$u^{-1}(u(F)) = u(u^{-1}(F)) \iff \text{Ker}(u) \subset F \subset \text{Im}(u).$$

Exercice 30 ***

Soit u un endomorphisme non bijectif d'un espace E de dimension finie. Montrer qu'il existe un isomorphisme φ de E tel que $v = \varphi \circ u$ soit nilpotent².

Solution

Le noyau de u n'est pas réduit au vecteur nul, on peut introduire une base (e_1, \dots, e_p) (avec $p \geq 1$) de $\text{Ker}(u)$ que l'on complète en (e_1, \dots, e_n) base de E (avec $n = \dim E$). Si φ est un isomorphisme, l'endomorphisme $v = \varphi \circ u$ annule nécessairement les vecteurs e_1, \dots, e_p .

méthode

On détermine φ de sorte que v opère un glissement envoyant e_{p+1}, \dots, e_n sur respectivement e_p, \dots, e_{n-1} : en itérant v on obtient l'endomorphisme nul.

1. Voir sujet 27 p. 38.

2. Autrement dit, il existe $p \in \mathbb{N}^*$ tel que l'itéré $v^p = v \circ \dots \circ v$ est nul.

Pour définir un isomorphisme φ , il suffit de fixer l'image d'une base égale à une base. Au surplus, on veut ici $\varphi(u(e_i)) = e_{i-1}$ pour tout i compris entre $p+1$ et n . Il s'agit alors d'introduire des bases de E adéquates pour construire φ .

L'espace $\text{Vect}(e_{p+1}, \dots, e_n)$ est un supplémentaire de $\text{Ker}(u)$, l'endomorphisme u est injectif sur celui-ci et transforme donc la famille (e_{p+1}, \dots, e_n) en $(u(e_{p+1}), \dots, u(e_n))$ base de $\text{Im}(u)$. On peut compléter celle-ci en une base de E adaptée à $\text{Im}(u)$ en introduisant des vecteurs f_1, \dots, f_p bien choisis. Considérons ensuite l'application linéaire φ déterminée sur cette base par

$$\begin{aligned}\varphi(u(e_i)) &= e_{i-1} \quad \text{pour } i \in [p+1; n] \text{ et} \\ \varphi(f_1) &= e_n, \varphi(f_2) = e_{n-1}, \dots, \varphi(f_p) = e_{p+1}.\end{aligned}$$

$$\begin{array}{ccccccccc} u(e_{p+1}) & \cdots & u(e_n) & f_1 & f_2 & \cdots & f_p \\ \downarrow & & \downarrow & \downarrow & \downarrow & & \downarrow \\ e_p & & e_{n-1} & e_n & e_1 & & e_{p-1} \end{array}$$

À l'ordre près des vecteurs, l'application linéaire φ transforme la base adaptée à $\text{Im}(u)$ en la base (e_1, \dots, e_n) , c'est donc un isomorphisme. Au surplus, $v = \varphi \circ u$ vérifie :

$$\begin{aligned}\forall j \in [1; p], v(e_j) &= 0_E \text{ et} \\ \forall j \in [p+1; n], v(e_j) &= e_{j-1}.\end{aligned}$$

$$\begin{array}{ccccccccc} e_1 & \cdots & e_p & e_{p+1} & \cdots & e_{n-1} & e_n \\ \downarrow & & \downarrow & \curvearrowright & \curvearrowright & \curvearrowright & \downarrow \\ 0_E & & 0_E & & & & 0_E \end{array}$$

Il est alors facile d'itérer l'endomorphisme v pour constater qu'à chaque itération le noyau s'agrandit d'un vecteur :

$$\begin{aligned}\text{Ker}(v) &= \text{Vect}(e_1, \dots, e_p), \\ \text{Ker}(v^2) &= \text{Vect}(e_1, \dots, e_{p+1}), \\ \text{Ker}(v^3) &= \text{Vect}(e_1, \dots, e_{p+2}), \text{ etc.}\end{aligned}$$

En particulier, $\text{Ker}(v^{n-p+1}) = E$ et donc $v^{n-p+1} = 0$: l'endomorphisme v est nilpotent.

Exercice 31 *** (Factorisation par un endomorphisme)

Soit f et g deux endomorphismes d'un espace E de dimension finie.

(a) Montrer

$$\text{Im}(g) \subset \text{Im}(f) \iff \exists h \in \mathcal{L}(E), g = f \circ h.$$

(b) Montrer

$$\text{Ker}(f) \subset \text{Ker}(g) \iff \exists h \in \mathcal{L}(E), g = h \circ f.$$

Solution

(a) On raisonne par double implication.

(\Leftarrow) Si l'on peut écrire $g = f \circ h$, il est entendu que les valeurs prises par g sont aussi des valeurs prises par f et donc $\text{Im}(g) \subset \text{Im}(f)$.

(\Rightarrow) Supposons $\text{Im}(g) \subset \text{Im}(f)$. Pour définir h solution, il serait bon de pouvoir « inverser » l'endomorphisme f .

méthode

|| L'application linéaire f induit un isomorphisme entre tout supplémentaire de son noyau et son image.

Soit S un supplémentaire de $\text{Ker}(f)$ dans E (il en existe car l'espace E est de dimension finie) et φ l'isomorphisme induit par f au départ de S et à valeurs dans $\text{Im}(f)$. L'application composée $h = \varphi^{-1} \circ g$ est bien définie car g prend ses valeurs dans $\text{Im}(g) \subset \text{Im}(f)$ et φ^{-1} est définie sur $\text{Im}(f)$. Par composition, l'application h est linéaire et peut se comprendre comme un endomorphisme de E . Enfin, φ^{-1} prend ses valeurs dans S et f se confond avec φ au départ de S . Ceci permet d'écrire

$$f \circ h = f \circ \varphi^{-1} \circ g = (\underbrace{\varphi \circ \varphi^{-1}}_{=\text{Id}_{\text{Im}(f)}}) \circ g = g.$$

L'endomorphisme h est solution.

(b) Raisonnons encore par double implication.

(\Leftarrow) Si l'on peut écrire $g = h \circ f$, il est immédiat que les vecteurs annulant f annulent aussi g et donc $\text{Ker}(f) \subset \text{Ker}(g)$.

(\Rightarrow) Supposons $\text{Ker}(f) \subset \text{Ker}(g)$.

méthode

|| La condition $g = h \circ f$ détermine h sur l'image de f , on complète la définition de h en choisissant arbitrairement h sur un supplémentaire de $\text{Im}(f)$.

On introduit à nouveau l'isomorphisme φ induit par f au départ d'un supplémentaire S de $\text{Ker}(f)$ et à valeurs dans $\text{Im}(f)$. Soit aussi F un supplémentaire de $\text{Im}(f)$ dans E . Considérons l'application linéaire h déterminée par ses restrictions linéaires : h égale à $g \circ \varphi^{-1}$ sur $\text{Im}(f)$ et h égale à l'application nulle¹ au départ de F . Vérifions que l'endomorphisme h convient.

Pour $x \in \text{Ker}(f)$, $(h \circ f)(x) = h(0_E) = 0_E = g(x)$ car $\text{Ker}(f) \subset \text{Ker}(g)$.

Pour $x \in S$, $(h \circ f)(x) = h(\varphi(x)) = g(x)$ car h se confond avec $g \circ \varphi^{-1}$ sur $\text{Im}(f)$.

Les applications linéaires $h \circ f$ et g sont égales sur deux espaces supplémentaires, elles sont donc égales sur E .

1. Ou toute autre application linéaire de F vers E , cela est sans conséquence pour la suite.

CHAPITRE 9

Matrices

\mathbb{K} désigne \mathbb{R} ou \mathbb{C} .

Les entiers n, p, q introduits dans ce chapitre sont supposés strictement positifs.

9.1 Calcul matriciel

9.1.1 Espaces des matrices

Définition

On note $\mathcal{M}_{n,p}(\mathbb{K})$ l'ensemble des matrices de type (n, p) à coefficients dans \mathbb{K} , c'est-à-dire l'ensemble des familles $A = (a_{i,j})_{1 \leq i \leq n, 1 \leq j \leq p}$ d'éléments¹ de \mathbb{K} .

Une telle matrice est usuellement figurée par un tableau à n lignes et p colonnes

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix} \in \mathcal{M}_{n,p}(\mathbb{K}).$$

Lorsque $n = 1$, ce tableau se résume à une ligne et l'on parle de *matrice ligne*.

Lorsque $p = 1$, on parle de *matrice colonne*.

Lorsque tous les coefficients sont nuls, on dit que la matrice est nulle et l'on note $O_{n,p}$, ou simplement 0 , la *matrice nulle* de type (n, p) .

1. On dit que $a_{i,j}$ est le *coefficent général* de la matrice A , le premier indice est l'indice de ligne et le second l'indice de colonne.

Pour $i \in [1; n]$ et $j \in [1; p]$, la *matrice élémentaire* d'indice (i, j) et de type (n, p) est la matrice $E_{i,j}$ déterminée par le tableau ci-dessous¹ :

$$E_{i,j} = \begin{array}{c|ccccc|c} & & & \xrightarrow{\text{p colonnes}} & & & \\ & (0) & & & (0) & & & \\ & & 1 & \longleftarrow & & & \\ \xleftarrow{j} & (0) & & \uparrow & (0) & & \\ & & & & & & \\ & & & & & & \end{array} i \quad \left[\begin{array}{c} n \text{ lignes} \\ \downarrow \\ j \end{array} \right]$$

Lorsque λ est un élément de \mathbb{K} et $A = (a_{i,j})$, $B = (b_{i,j})$ des matrices de type (n, p) à coefficients dans \mathbb{K} , on définit la *matrice produit* de A par le scalaire λ et la *matrice somme* de A et B par

$$\lambda A \stackrel{\text{def}}{=} (\lambda a_{i,j})_{1 \leq i \leq n, 1 \leq j \leq p} \quad \text{et} \quad A + B \stackrel{\text{def}}{=} (a_{i,j} + b_{i,j})_{1 \leq i \leq n, 1 \leq j \leq p}.$$

On vérifie alors

Théorème 1

$(\mathcal{M}_{n,p}(\mathbb{K}), +, \cdot)$ est un \mathbb{K} -espace vectoriel dont l'élément nul est $O_{n,p}$.

De plus, cet espace est de dimension finie np et la famille des matrices élémentaires $(E_{i,j})_{1 \leq i \leq n, 1 \leq j \leq p}$ en constitue une base appelée *base canonique* de $\mathcal{M}_{n,p}(\mathbb{K})$.

9.1.2 Produit matriciel

Pour $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{j,k}) \in \mathcal{M}_{p,q}(\mathbb{K})$, on définit la *matrice produit* de A par B en posant

$$AB = (c_{i,k}) \in \mathcal{M}_{n,q}(\mathbb{K}) \quad \text{avec} \quad c_{i,k} \stackrel{\text{def}}{=} \sum_{j=1}^p a_{i,j} b_{j,k} \quad \text{pour tout } (i, k) \in [1; n] \times [1; q].$$

Le produit matriciel AB n'est possible que lorsque le nombre de colonnes de A correspond au nombre de lignes de B . En outre, on peut retenir que le produit d'une matrice de type (n, p) par une matrice de type (p, q) définit une matrice de type (n, q) .

Le produit matriciel est associatif. Il est aussi bilinéaire dans le sens où l'on peut écrire

$$A(\lambda B + \mu C) = \lambda AB + \mu AC \quad \text{et} \quad (\lambda B + \mu C)A = \lambda BA + \mu CA$$

pour $\lambda, \mu \in \mathbb{K}$ et A, B, C matrices de « types convenables ».

On appelle *matrice unité* (ou *matrice identité*) de taille n la matrice de type (n, n) dont tous les coefficients sont nuls sauf ceux d'indices (i, i) avec $i \in [1; n]$ qui sont égaux à 1 :

$$I_n = \begin{pmatrix} 1 & & (0) \\ (0) & \ddots & \\ & & 1 \end{pmatrix}$$

1. Dans ce tableau, les (0) figurent des regroupements de coefficients tous nuls.

La matrice unité de taille convenable est élément neutre à droite et à gauche pour le produit matriciel : $A \times I_p = A = I_n \times A$ pour toute matrice A de type (n, p) .

Notons que le produit matriciel n'est pas commutatif.

9.1.3 L'anneau des matrices carrées

Définition

Une *matrice carrée* de taille n est une matrice de type (n, n) . On note $\mathcal{M}_n(\mathbb{K})$ l'ensemble des matrices carrées de taille n à coefficients dans \mathbb{K} .

Les coefficients d'indices (i, i) d'une matrice carrée $A = [a_{i,j}]$ s'appellent les *coefficients diagonaux* de A et la famille $(a_{1,1}, \dots, a_{n,n})$ détermine la *diagonale* de A .

Théorème 2

$(\mathcal{M}_n(\mathbb{K}), +, \times)$ est un anneau de neutres O_n et I_n .

$\mathcal{M}_n(\mathbb{K})$ étant un anneau, on peut y utiliser les formules du binôme (Th. 6 p. 124) et de factorisation géométrique (Th. 7 p. 124). L'anneau $\mathcal{M}_n(\mathbb{K})$ n'étant pas commutatif lorsque $n > 2$, on sera attentif à vérifier l'hypothèse de commutativité nécessaire à l'usage de ces formules.

Dans un anneau, les éléments inversibles sont remarquables et constituent un groupe multiplicatif (Th. 8 p. 124) :

Définition

On dit que $A \in \mathcal{M}_n(\mathbb{K})$ est *inversible* s'il existe $B \in \mathcal{M}_n(\mathbb{K})$ vérifiant $AB = BA = I_n$. Cette matrice B est alors unique et s'appelle l'*inverse* de A . On la note A^{-1} .

La matrice unité I_n est inversible et $I_n^{-1} = I_n$.

Théorème 3

Une matrice carrée est inversible si, et seulement si, elle est inversible à droite, ou encore si, et seulement si, elle est inversible à gauche.

Les inverses à droite et à gauche sont alors égaux à l'inverse de la matrice.

Théorème 4

L'ensemble $GL_n(\mathbb{K})$ des matrices inversibles de $\mathcal{M}_n(\mathbb{K})$ est un groupe multiplicatif de neutre I_n .

$(GL_n(\mathbb{K}), \times)$ est le *groupe linéaire* d'ordre n .

9.1.4 Matrices carrées remarquables

Définition

Une *matrice diagonale* de taille n est une matrice carrée $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ dont tous les coefficients $a_{i,j}$ sont nuls lorsque $i \neq j$:

$$A = \begin{pmatrix} a_{1,1} & & (0) \\ (0) & \ddots & \\ & & a_{n,n} \end{pmatrix}.$$

Le produit de deux matrices diagonales définit une matrice diagonale :

$$\begin{pmatrix} a_{1,1} & & (0) \\ (0) & \ddots & \\ & & a_{n,n} \end{pmatrix} \begin{pmatrix} b_{1,1} & & (0) \\ (0) & \ddots & \\ & & b_{n,n} \end{pmatrix} = \begin{pmatrix} a_{1,1}b_{1,1} & & (0) \\ (0) & \ddots & \\ & & a_{n,n}b_{n,n} \end{pmatrix}.$$

Soulignons aussi que ce produit commute.

Définition

Une *matrice triangulaire supérieure* (resp. *triangulaire inférieure*) de taille n est une matrice carrée $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ dont tous les coefficients $a_{i,j}$ sont nuls lorsque $i > j$ (resp. $i < j$)¹ :

$$A = \begin{pmatrix} a_{1,1} & & (*) \\ (0) & \ddots & \\ & & a_{n,n} \end{pmatrix} \quad (\text{resp. } \begin{pmatrix} a_{1,1} & & (0) \\ (*) & \ddots & \\ & & a_{n,n} \end{pmatrix}).$$

Le produit de deux matrices triangulaires supérieures (resp. inférieures) est une matrice triangulaire supérieure (resp. inférieure) et ses coefficients diagonaux sont remarquables :

$$\begin{pmatrix} a_{1,1} & & (*) \\ (0) & \ddots & \\ & & a_{n,n} \end{pmatrix} \begin{pmatrix} b_{1,1} & & (*)' \\ (0) & \ddots & \\ & & b_{n,n} \end{pmatrix} = \begin{pmatrix} a_{1,1}b_{1,1} & & (*)'' \\ (0) & \ddots & \\ & & a_{n,n}b_{n,n} \end{pmatrix}.$$

Une matrice diagonale (resp. triangulaire) est inversible si, et seulement si, ses coefficients diagonaux sont non nuls. Son inverse est alors une matrice diagonale (resp. triangulaire).

9.1.5 Transposition

Définition

On appelle *matrice transposée* de $A = (a_{i,j})$ de type (n, p) , la matrice $A^T = (a_{j,i}^T)$ de type (p, n) déterminée par

$$\forall (i, j) \in \llbracket 1 ; n \rrbracket \times \llbracket 1 ; p \rrbracket, \quad a_{j,i}^T \stackrel{\text{def}}{=} a_{i,j}.$$

Cette matrice est constituée des mêmes coefficients que A mais ceux-ci sont disposés en ligne plutôt qu'en colonne. La matrice transposée de A est quelquefois notée A^T .

1. Dans cette description le symbole * signifie : des coefficients quelconques.

L'opération de transposition est linéaire et vérifie ${}^t({}^tA) = A$ et ${}^t(AB) = {}^tB {}^tA$ pour A et B matrices de types compatibles.

Théorème 5

Si $A \in M_n(\mathbb{K})$ est inversible, sa transposée tA l'est aussi et $({}^tA)^{-1} = {}^t(A^{-1})$.

9.1.6 Matrices par blocs

Il est possible de décrire une matrice en regroupant ses coefficients par blocs. On peut par exemple définir une matrice M en réalisant un découpage par blocs 2×2

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

à partir de A, B, C, D quatre matrices telles que A et B d'une part, C et D d'autre part, ont le même nombre de lignes et que A et C d'une part, B et D d'autre part, ont le même nombre de colonnes.

On peut alors opérer sur les décompositions par blocs sous réserve de compatibilité des types matriciels. Par exemple, on peut écrire

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} + \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} = \begin{pmatrix} A + A' & B + B' \\ C + C' & D + D' \end{pmatrix} \text{ et}$$

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \times \begin{pmatrix} A' & B' \\ C' & D' \end{pmatrix} = \begin{pmatrix} AA' + BC' & AB' + BD' \\ CA' + DC' & CB' + DD' \end{pmatrix}.$$

Ce type de formule se généralise à des découpages par blocs de différents types. Le calcul d'un produit par blocs est alors analogue au produit matriciel sous la réserve d'être attentif à la compatibilité des types matriciels et à l'ordre des facteurs puisque la multiplication matricielle n'est pas commutative.

9.2 Représentations matricielles

E et F désignent des \mathbb{K} -espaces vectoriels de dimensions finies.

9.2.1 Matrice des coordonnées d'un vecteur

Si E est muni d'une base $e = (e_1, \dots, e_n)$, tout vecteur x de E s'écrit de façon unique

$$x = \lambda_1 e_1 + \dots + \lambda_n e_n \quad \text{avec} \quad (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n.$$

Les $\lambda_1, \dots, \lambda_n$ déterminent les coordonnées du vecteur x dans la base e .

Définition

On appelle *matrice du vecteur* x dans la base e la matrice colonne

$$\text{Mat}_e(x) = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathcal{M}_{n,1}(\mathbb{K}).$$

Plus généralement :

Définition

On appelle *matrice d'une famille* (x_1, \dots, x_p) de vecteurs de E dans la base e , la matrice de type (n, p) dont les colonnes sont formées par les coordonnées des vecteurs x_1, \dots, x_p dans la base e :

$$\text{Mat}_e(x_1, \dots, x_p) = \left(\begin{array}{ccc|c} & x_1 & & x_p \\ & \downarrow & & \downarrow \\ & \vdots & \cdots & \vdots \\ & \downarrow & & \downarrow \end{array} \right) \begin{matrix} \text{ coordonnées} \\ \text{lues dans } e \end{matrix}$$

En particulier, la matrice de la famille e dans la base e est la matrice I_n .

9.2.2 Matrice d'une application linéaire

Soit $e = (e_1, \dots, e_p)$ et $f = (f_1, \dots, f_n)$ des bases des espaces E et F .

Définition

On appelle *matrice d'une application linéaire* $u \in \mathcal{L}(E, F)$ relative aux bases e et f la matrice de la famille $(u(e_1), \dots, u(e_p))$ dans la base f :

$$\text{Mat}_{e,f}(u) \stackrel{\text{def}}{=} \text{Mat}_f(u(e_1), \dots, u(e_p)) \in \mathcal{M}_{n,p}(\mathbb{K})$$

$$\text{Mat}_{e,f}(u) = \left(\begin{array}{ccc|c} u(e_1) & & u(e_p) & \\ \downarrow & \cdots & \downarrow & \\ \vdots & \cdots & \vdots & \\ \downarrow & & \downarrow & \end{array} \right) \begin{matrix} \text{ coordonnées} \\ \text{lues dans } f \end{matrix}$$

Les bases e et f correspondent aux *bases de départ* et *d'arrivée* de la représentation matricielle. Cette matrice permet de calculer facilement les coordonnées des images des vecteurs :

Théorème 6

Si u est une application linéaire de E vers F , la matrice de u relative aux bases e et f est l'unique matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ telle que, pour tous vecteurs $x \in E$ et $y \in F$,

$$y = u(x) \iff Y = AX \quad \text{avec} \quad X = \text{Mat}_e(x) \text{ et } Y = \text{Mat}_f(y).$$

La représentation matricielle d'une application linéaire caractérise entièrement celle-ci :

Théorème 7

L'application qui à $u \in \mathcal{L}(E, F)$ associe sa matrice relative aux bases e et f est un isomorphisme de l'espace vectoriel $\mathcal{L}(E, F)$ vers $\mathcal{M}_{n,p}(\mathbb{K})$.

Il est donc possible de déterminer une application linéaire par le choix de sa matrice représentative dans des bases données.

Au surplus, le produit matriciel traduit la composition des applications linéaires :

Théorème 8

En introduisant un espace vectoriel G muni d'une base g , on a, pour tous $u \in \mathcal{L}(E, F)$ et $v \in \mathcal{L}(F, G)$,

$$\text{Mat}_{e,g}(v \circ u) = \text{Mat}_{f,g}(v) \text{Mat}_{e,f}(u).$$

En particulier, u est un isomorphisme si, et seulement si, sa matrice représentative est une matrice carrée inversible.

9.2.3 Matrice d'un endomorphisme

Soit $e = (e_1, \dots, e_n)$ une base de E . Pour figurer un endomorphisme, il est usuel de choisir une base d'arrivée identique à celle de départ.

Définition

On appelle *matrice d'un endomorphisme* $u \in \mathcal{L}(E)$ relative à la base e , la matrice de l'application linéaire u relative aux bases e au départ et e à l'arrivée. Cette matrice est notée $\text{Mat}_e(u)$, il s'agit d'une matrice carrée de taille n .

En particulier, la matrice de l'endomorphisme Id_E est égale à la matrice unité I_n dans n'importe quelle base de E .

Comme pour les applications linéaires, le calcul matriciel est compatible avec le calcul relatif aux endomorphismes.

9.2.4 Matrice d'une forme linéaire

Soit $e = (e_1, \dots, e_n)$ une base de E . Une forme linéaire sur E est une application linéaire au départ de E et à valeurs dans \mathbb{K} . Il est usuel de munir l'espace d'arrivée \mathbb{K} de sa base canonique uniquement constituée du nombre 1.

Définition

On appelle *matrice d'une forme linéaire* $\varphi \in \mathcal{L}(E, \mathbb{K})$ relative à la base e , la matrice représentative de l'application linéaire φ relative à la base e au départ et à la base (1) à l'arrivée. Cette matrice est une matrice ligne de longueur n .

Si L figure une forme linéaire φ sur E et si X désigne les coordonnées d'un vecteur x de E dans la même base, le calcul LX détermine une matrice à un seul coefficient correspondant à la valeur $\varphi(x)$.

9.2.5 Application linéaire canoniquement associée à une matrice

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$. Par l'isomorphisme introduit dans le théorème 7, on peut affirmer qu'il existe une unique application linéaire $a \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ figurée par la matrice A dans les bases canoniques de \mathbb{K}^p et de \mathbb{K}^n .

Définition

L'application a est nommée *application linéaire canoniquement associée* à la matrice A .

Si $x = (x_1, \dots, x_p)$ et $y = (y_1, \dots, y_n)$ désignent des éléments de \mathbb{K}^p et \mathbb{K}^n , on a

$$y = a(x) \iff Y = AX$$

avec X et Y les colonnes constituées respectivement des éléments x_1, \dots, x_p et y_1, \dots, y_n . En réalisant l'identification du vecteur x avec la colonne X et du vecteur y avec la colonne Y , on peut écrire que l'application canoniquement associée à A est l'application qui à $x \in \mathbb{K}^p$ associe $y \in \mathbb{K}^n$ déterminé par $y = Ax$.

Définition

On appelle *noyau d'une matrice* $A \in \mathcal{M}_{n,p}(\mathbb{K})$ le noyau de l'application linéaire qui lui est canoniquement associée. On le note $\text{Ker}(A)$.

Ce noyau est le sous-espace vectoriel de \mathbb{K}^p formé des solutions de l'équation¹ $Ax = 0$: les lignes de la matrice A donnent un système d'équations déterminant son noyau.

Définition

On appelle *image d'une matrice* $A \in \mathcal{M}_{n,p}(\mathbb{K})$ l'image de l'application linéaire qui lui est canoniquement associée. On la note $\text{Im}(A)$.

Cette image est le sous-espace vectoriel formé des vecteurs Ax pour x parcourant \mathbb{K}^p , c'est aussi le sous-espace vectoriel de \mathbb{K}^n engendré par les colonnes de A .

9.2.6 Rang d'une matrice

Définition

On appelle *rang d'une matrice* A la dimension de son image. On le note $\text{rg}(A)$.

Ce rang est aussi le rang de la famille des colonnes de la matrice A , il est assurément inférieur aux nombres de lignes et de colonnes de A .

Théorème 9

Si A est la matrice d'une famille (x_1, \dots, x_p) de vecteurs de E dans une base, le rang de A est aussi le rang de la famille (x_1, \dots, x_p) .

Si A est la matrice d'une application linéaire $u \in \mathcal{L}(E, F)$ relative à des bases de E et F , le rang de A est aussi le rang de l'application linéaire u .

1. Le zéro écrit dans cette équation est volontairement ambigu : il peut à la fois se comprendre comme le vecteur nul de \mathbb{K}^n ou comme la colonne nulle de $\mathcal{M}_{n,1}(\mathbb{K})$.

Par l'application linéaire canoniquement associée, on peut énoncer une formule du rang

$$\underbrace{p}_{\text{nombre de colonnes}} = \operatorname{rg}(A) + \dim \operatorname{Ker}(A) \quad \text{pour } A \in \mathcal{M}_{n,p}(\mathbb{K}).$$

Aussi, le rang d'un produit de deux matrices est inférieur aux rangs de chacun des facteurs

$$\operatorname{rg}(AB) \leq \min(\operatorname{rg}(A), \operatorname{rg}(B)) \quad \text{pour } A \in \mathcal{M}_{n,p}(\mathbb{K}), B \in \mathcal{M}_{p,q}(\mathbb{K}).$$

On en déduit :

Théorème 10

On ne modifie pas le rang d'une matrice lorsque l'on multiplie celle-ci par une matrice inversible.

Théorème 11

Une matrice carrée est inversible si, et seulement si, son noyau est réduit à l'espace nul, ou encore, si, et seulement si, son rang est égal à sa taille.

9.3 Changements de bases

9.3.1 Matrice de passage

Soit e et e' deux bases d'un même espace vectoriel E de dimension n .

Définition

On appelle *matrice de passage* de e à e' la matrice figurant la famille e' dans la base e

$$P^{e'} = \operatorname{Mat}_e e' \in \mathcal{M}_n(\mathbb{K}).$$

Cette matrice figure aussi l'application linéaire identité relativement aux bases e' au départ et e à l'arrivée. À ce titre une matrice de passage est inversible et son inverse est la matrice de passage de e' à e .

9.3.2 Formules de changement de base

Théorème 12

Si P est la matrice de passage d'une base e à une base e' d'un espace vectoriel E alors, pour tout vecteur x de E ,

$$X = P X' \quad \text{avec} \quad X = \operatorname{Mat}_e(x) \quad \text{et} \quad X' = \operatorname{Mat}_{e'}(x).$$

Théorème 13

Si P est la matrice de passage d'une base e à une base e' d'un espace vectoriel E et si Q est la matrice de passage d'une base f à une base f' d'un espace vectoriel F alors, pour toute application linéaire $u \in \mathcal{L}(E, F)$,

$$A' = Q^{-1}AP \quad \text{avec} \quad A = \text{Mat}_{e,f}(u) \quad \text{et} \quad A' = \text{Mat}_{e',f'}(u).$$

En particulier, lorsque u est un endomorphisme de E ,

$$A' = P^{-1}AP \quad \text{avec} \quad A = \text{Mat}_e(u) \quad \text{et} \quad A' = \text{Mat}_{e'}(u).$$

9.3.3 Matrices équivalentes et rang**Définition**

On dit qu'une matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ est *équivalente* à une matrice B de même type lorsqu'il existe $P \in \text{GL}_p(\mathbb{K})$ et $Q \in \text{GL}_n(\mathbb{K})$ telles que $A = QBP^{-1}$.

Les différentes matrices d'une même application linéaire sont toutes équivalentes.

Théorème 14

Si $u \in \mathcal{L}(E, F)$ est de rang r , il existe une base e de E et une base f de F telles que la matrice de u dans les bases e et f est la matrice décrite par blocs

$$J_r = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{M}_{n,p}(\mathbb{K})$$

(où les 0 désignent des blocs nuls de tailles appropriées).

Une matrice $A \in \mathcal{M}_{n,p}(\mathbb{K})$ est alors de rang r si, et seulement si, elle est équivalente à la matrice¹ J_r de même type :

$$A = QJ_rP^{-1} \quad \text{avec} \quad P \in \text{GL}_p(\mathbb{K}) \text{ et } Q \in \text{GL}_n(\mathbb{K}).$$

Deux matrices sont alors équivalentes si, et seulement si, elles ont le même rang.

Par transposition, la matrice J_r est transformée en une matrice de rang r . On en déduit que le rang d'une matrice est invariant par transposition.

Le rang d'une matrice est celui de la famille de ses colonnes. Par transposition, c'est aussi le rang de la famille de ses lignes. Si l'on retire un certain nombre de lignes et/ou de colonnes à une matrice, on forme ce que l'on appelle une *matrice extraite*. Une telle matrice est de rang inférieur à la matrice dont elle est issue. Plus précisément

Théorème 15

Le rang d'une matrice est la taille maximale d'une matrice carrée inversible extraite de celle-ci.

1. La matrice J_r est appelée *matrice canonique* de rang r de type (n, p) .

9.3.4 Matrices semblables et trace

Définition

On appelle *trace d'une matrice* $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ la somme de ses coefficients diagonaux :

$$\text{tr}(A) = a_{1,1} + \cdots + a_{n,n}.$$

La trace définit une forme linéaire sur $\mathcal{M}_n(\mathbb{K})$. Elle vérifie :

Théorème 16

Pour toutes matrices $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,n}(\mathbb{K})$,

$$\text{tr}(AB) = \text{tr}(BA).$$

Définition

On dit qu'une matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$ est *semblable à* $B \in \mathcal{M}_n(\mathbb{K})$, s'il existe $P \in \text{GL}_n(\mathbb{K})$ telle que $B = P^{-1}AP$.

Les différentes matrices d'un même endomorphisme sont toutes semblables. De plus, deux matrices semblables ont la même trace¹ ce qui permet d'introduire la définition suivante :

Définition

On appelle *trace d'un endomorphisme* u de E la trace commune aux matrices figurant cet endomorphisme. On la note $\text{tr}(u)$.

La trace définit une forme linéaire sur $\mathcal{L}(E)$. Elle vérifie $\text{tr}(u \circ v) = \text{tr}(v \circ u)$ pour toutes applications linéaires $u \in \mathcal{L}(E, F)$ et $v \in \mathcal{L}(F, E)$.

Théorème 17

Si p est une projection vectorielle de l'espace E , on a $\text{tr}(p) = \text{rg}(p)$.

9.4 Opérations élémentaires et systèmes linéaires

9.4.1 Opérations élémentaires

Définition

Les *opérations élémentaires* transforment une matrice en opérant sur ses rangées. Les opérations sur les lignes sont de trois types :

- $L_i \leftarrow \alpha L_i$: on multiplie la ligne i par un scalaire $\alpha \neq 0$;
- $L_i \leftarrow L_i + \lambda L_j$: on ajoute à la ligne i un multiple de la ligne j avec $j \neq i$;
- $L_i \leftrightarrow L_j$: on échange les lignes d'indices i et j .

Les opérations élémentaires sur les colonnes sont analogues et sont codées $C_i \leftarrow \alpha C_i$, $C_i \leftarrow C_i + \lambda C_j$ et $C_i \leftrightarrow C_j$.

¹. Elles ont aussi le même rang mais ces conditions ne sont pas suffisantes pour affirmer que deux matrices sont semblables.

Les opérations élémentaires sur les colonnes d'une matrice A (resp. les lignes) peuvent s'interpréter comme la multiplication à droite (resp. à gauche) de A par une matrice inversible « adaptée ». Ces opérations conservent l'image (resp. le noyau) de la matrice. En particulier :

Théorème 18

Les opérations élémentaires conservent le rang d'une matrice.

Les opérations élémentaires permettent aussi de calculer l'inverse¹ d'une matrice.

9.4.2 Systèmes linéaires

Considérons un système d'équations linéaires à n équations et p inconnues de la forme

$$(\Sigma): \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,p}x_p = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,p}x_p = b_2 \\ \vdots \quad \vdots \quad \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,p}x_p = b_n. \end{cases}$$

On introduit la matrice $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et les vecteurs $x = (x_1, \dots, x_p)$ de \mathbb{K}^p et $b = (b_1, \dots, b_n)$ de \mathbb{K}^n que l'on identifie avec les colonnes formées des mêmes coefficients. Résoudre le système (Σ) revient à résoudre l'équation $Ax = b$.

Définition

|| L'équation $Ax = b$ est appelée *équation matricielle* associée au système (Σ) .

L'équation $Ax = b$ est une équation linéaire d'équation homogène associée $Ax = 0$. Résoudre celle-ci revient à déterminer le noyau de A . Celui-ci est un sous-espace vectoriel de dimension $p - r$ avec $r = \text{rg}(A)$.

Lorsque le système (Σ) est compatible et que x_0 en est une solution particulière, on peut décrire le sous-espace affine des solutions : $x_0 + \text{Ker}(A)$. (Th. 4 p. 274).

9.4.3 Systèmes de Cramer

Un *système carré* est un système d'équations linéaires ayant autant d'équations que d'inconnues. Un tel système peut ne pas être compatible ou posséder une infinité de solution. Cependant,

Théorème 19

Un système carré d'équation matricielle $Ax = b$ possède une et une seule solution si, et seulement si, la matrice A est inversible.

Un tel système se nomme un *système de Cramer*.

1. Voir sujet 5 p. 325.

9.5 Exercices d'apprentissage

9.5.1 Calculs matriciels

Exercice 1

On note $E_{i,j}$ et $E_{k,\ell}$ les matrices élémentaires de $\mathcal{M}_{n,p}(\mathbb{K})$ et $\mathcal{M}_{p,q}(\mathbb{K})$ d'indices (i,j) et (k,ℓ) convenables. Calculer $E_{i,j} \times E_{k,\ell}$.

Solution

Commençons par souligner que le produit des matrices élémentaires proposées est possible et détermine une matrice de type (n, q) .

méthode

|| On pose le produit des tableaux figurant $E_{i,j}$ et $E_{k,\ell}$.

Chaque ligne nulle de la matrice $E_{i,j}$ induit une ligne nulle sur la matrice produit. De même, chaque colonne nulle de la matrice $E_{k,\ell}$ définit une colonne nulle sur la matrice produit. Par conséquent, dans la matrice $E_{i,j} \times E_{k,\ell}$, seul le coefficient d'indice (i, ℓ) est susceptible d'être non nul. Plus précisément, il est non nul si, et seulement si, les deux 1 des tableaux figurant $E_{i,j}$ et $E_{k,\ell}$ se croisent, c'est-à-dire si, et seulement si, $j = k$ auquel cas il vaut 1.

Finalement, $E_{i,j}E_{k,\ell} = \mathbf{0}_{n,q}$ si $j \neq k$ et $E_{i,j}E_{k,\ell} = E_{i,\ell} \in \mathcal{M}_{n,q}(\mathbb{K})$ si $j = k$. On pourra retenir l'écriture synthétique

$$E_{i,j}E_{k,\ell} = \delta_{j,k}E_{i,\ell} \quad \text{avec} \quad \delta_{j,k} \text{ symbole de Kronecker : } \delta_{j,k} = \begin{cases} 1 & \text{si } j = k \\ 0 & \text{sinon.} \end{cases}$$

$$\left(\begin{array}{c|ccccc} (0) & & & & & \ell \\ \downarrow & & & & & \\ (0) & & & & 1 \leftarrow k & (0) \end{array} \right) \left(\begin{array}{c|ccccc} (0) & & (0) & & & \\ i \rightarrow 1 & & (0) & & & \\ \uparrow & & (0) & & & \end{array} \right) \left(\begin{array}{c|ccccc} \dots & & \dots & & & \\ \vdots & & \vdots & & & \end{array} \right)$$

Exercice 2

Calculer A^n pour $n \in \mathbb{N}$ et les matrices A suivantes :

$$(a) A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \quad (b) A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad (c) A = \begin{pmatrix} 1 & 1 \\ 0 & 3 \end{pmatrix}.$$

Solution

méthode

|| On calcule les premières puissances de A afin de proposer une forme générale.

(a) On observe $A^2 = I_2$ ce qui invite à discuter selon la parité de n .

Cas : n est pair. On peut écrire $n = 2p$ avec $p \in \mathbb{N}$ et alors $A^n = (A^2)^p = I_2$.

Cas : n est impair. On écrit $n = 2p + 1$ avec $p \in \mathbb{N}$ puis $A^n = (A^2)^p A = A$.

(b) On observe

$$A^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, A^1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, A^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, A^3 = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}, \text{ etc.}$$

On vérifie alors par récurrence l'identité

$$A^n = \begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix}.$$

(c) Le calcul des premières puissances de A laisse observer une écriture

$$A^n = \begin{pmatrix} 1 & u_n \\ 0 & 3^n \end{pmatrix} \quad \text{avec } u_n \text{ à déterminer.}$$

L'égalité $A^{n+1} = A^n A$ permet d'exprimer u_{n+1} en fonction de u_n : $u_{n+1} = 3u_n + 1$. La suite (u_n) est une suite arithmético-géométrique et, après résolution¹, il vient

$$u_n = \frac{3^n - 1}{2}.$$

Exercice 3

Une matrice $M \in \mathcal{M}_n(\mathbb{R})$ est dite *symétrique* (resp. *antisymétrique*) lorsque ${}^t M = M$ (resp. ${}^t M = -M$). On note $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ les ensembles constitués des matrices symétriques et des matrices antisymétriques de $\mathcal{M}_n(\mathbb{R})$.

- (a) Montrer que $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ sont des espaces supplémentaires de $\mathcal{M}_n(\mathbb{R})$.
- (b) Préciser leurs dimensions respectives.

Solution

(a) Commençons par vérifier que les parties $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ sont des sous-espaces vectoriels de $\mathcal{M}_n(\mathbb{R})$. Elles contiennent chacune la matrice nulle et sont stables par combinaison linéaire. En effet, pour $\lambda, \mu \in \mathbb{R}$, $A, B \in \mathcal{M}_n(\mathbb{R})$ et $\varepsilon = \pm 1$, si ${}^t A = \varepsilon A$ et ${}^t B = \varepsilon B$ alors ${}^t (\lambda A + \mu B) = \varepsilon (\lambda A + \mu B)$ par linéarité de l'opération de transposition.

méthode

|| La supplémentarité de deux espaces vectoriels peut être établie par analyse-synthèse.

Soit $M \in \mathcal{M}_n(\mathbb{R})$.

Analyse : On suppose $M = A + B$ avec $A \in \mathcal{S}_n(\mathbb{R})$ et $B \in \mathcal{A}_n(\mathbb{R})$. En transposant cette égalité, il vient ${}^t M = {}^t A + {}^t B = A - B$. On en déduit

$$A = \frac{1}{2}(M + {}^t M) \quad \text{et} \quad B = \frac{1}{2}(M - {}^t M). \tag{*}$$

1. Voir sujet 1 du chapitre 6 de l'ouvrage *Exercices d'analyse MPSI*.

Ceci assure l'unicité de l'écriture de M comme somme d'une matrice symétrique et d'une matrice antisymétrique.

Synthèse : Considérons les deux matrices A et B déterminées par (*). La matrice M est bien la somme de A et B et, en exploitant ${}^t({}^tM) = M$, on observe ${}^tA = A$ et ${}^tB = -B$: les matrices A et B sont respectivement symétrique et antisymétrique.

Finalement, les espaces $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ sont supplémentaires¹ dans $\mathcal{M}_n(\mathbb{R})$.

(b) méthode

|| La dimension d'un espace est le nombre de vecteurs constituant ses bases.

En taille 3, les matrices symétriques et antisymétriques sont respectivement de la forme

$$\begin{pmatrix} a & d & e \\ d & b & f \\ e & f & c \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}.$$

Les matrices symétriques sont combinaisons linéaires de $E_{1,1}$, $E_{2,2}$, $E_{3,3}$, $E_{1,2} + E_{2,1}$, $E_{1,3} + E_{3,1}$ et $E_{2,3} + E_{3,2}$ tandis que les matrices antisymétriques sont combinaisons linéaires de $E_{1,2} - E_{2,1}$, $E_{1,3} - E_{3,1}$ et $E_{2,3} - E_{3,2}$.

En taille n , une matrice symétrique $A = (a_{i,j})$ vérifie $a_{i,j} = a_{j,i}$ pour tous i et j compris entre 1 et n . Elle peut s'écrire

$$A = \sum_{i=1}^n a_{i,i} E_{i,i} + \sum_{1 \leq i < j \leq n} a_{i,j} (E_{i,j} + E_{j,i}).$$

Une telle matrice est combinaison linéaire des matrices $E_{i,i}$ pour $1 \leq i \leq n$ et des matrices $E_{i,j} + E_{j,i}$ pour $1 \leq i < j \leq n$. Ces dernières constituent une famille libre donc une base. En effet, supposons

$$\sum_{i=1}^n \lambda_{i,i} E_{i,i} + \sum_{1 \leq i < j \leq n} \lambda_{i,j} (E_{i,j} + E_{j,i}) = \mathbf{O}_n \quad \text{avec } \lambda_{i,j} \in \mathbb{R}.$$

Pour $i \leq j$, les réels $\lambda_{i,j}$ déterminent les coefficients d'indices (i,j) de la matrice en premier membre, ils sont donc tous nuls. Il suffit ensuite de dénombrer les éléments de cette base pour obtenir la dimension de $\mathcal{S}_n(\mathbb{R})$

$$\dim \mathcal{S}_n(\mathbb{R}) = n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2}.$$

Par supplémentarité, on en déduit la dimension de $\mathcal{A}_n(\mathbb{R})$

$$\dim \mathcal{A}_n(\mathbb{R}) = \dim \mathcal{M}_n(\mathbb{R}) - \dim \mathcal{S}_n(\mathbb{R}) = n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}.$$

1. L'endomorphisme $T: \mathcal{M}_n(\mathbb{R}) \rightarrow \mathcal{M}_n(\mathbb{R})$ défini par $T(A) = {}^tA$ vérifie $T \circ T = \text{Id}$. Le Th. 8 p. 276 assure alors que $\text{Ker}(T - \text{Id}) = \mathcal{S}_n(\mathbb{R})$ et $\text{Ker}(T + \text{Id}) = \mathcal{A}_n(\mathbb{R})$ sont des espaces supplémentaires et l'on peut interpréter la transposition comme la symétrie par rapport au premier et parallèlement au second.

On peut confirmer ce résultat en observant que les matrices antisymétriques $A = (a_{i,j})$ sont les matrices vérifiant $a_{i,i} = 0$ et $a_{j,i} = -a_{i,j}$. Celles-ci peuvent s'écrire

$$A = \sum_{1 \leq i < j \leq n} a_{i,j} (E_{i,j} - E_{j,i})$$

ce qui permet de proposer une base de $\mathcal{A}_n(\mathbb{R})$.

Exercice 4

Calculer le rang des matrices suivantes :

$$(a) \begin{pmatrix} 1 & -1 & 2 & 1 & 0 \\ 0 & -1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad (b) \begin{pmatrix} 1 & 1 & 1 & -1 \\ -1 & -1 & 0 & 1 \\ 1 & 0 & 2 & -1 \end{pmatrix} \quad (c) \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ -1 & 1 & 0 & -1 \\ 1 & -1 & 0 & 1 \end{pmatrix}.$$

Solution

(a) La matrice considérée a une forme *échelonnée*¹ qui permet d'en déterminer le rang. Par opérations élémentaires sur les colonnes (ce qui conserve le rang), on peut annuler les coefficients à droite du 1 sur la première ligne :

$$\text{rg} \begin{pmatrix} 1 & -1 & 2 & 1 & 0 \\ 0 & -1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \text{rg} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{via } \begin{cases} C_2 \leftarrow C_2 + C_1 \\ C_3 \leftarrow C_3 - 2C_1 \\ C_4 \leftarrow C_4 - C_1. \end{cases}$$

De même, on annule les coefficients à droite du -1 sur la seconde ligne puis ceux à droite du 1 sur la troisième ligne. Par passage à l'opposé de la deuxième colonne, on peut aussi transformer le -1 en un 1

$$\text{rg} \begin{pmatrix} 1 & -1 & 2 & 1 & 0 \\ 0 & -1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \text{rg} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \text{rg} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

La dernière matrice est de rang 3 car les colonnes d'indices 1, 2 et 4 sont des colonnes élémentaires indépendantes et les deux autres sont nulles.

En pratique, il n'est pas usuel de détailler les étapes qui viennent d'être présentées : reconnaître une forme échelonnée suffit à affirmer la valeur du rang car on peut transformer par opérations élémentaires une telle matrice en une matrice constituée de colonnes élémentaires et de colonnes nulles.

1. On dit qu'un matrice est *échelonnée* lorsque le nombre de coefficients nuls en début de ligne augmente d'au moins une unité ligne par ligne. Les premiers coefficients à droite de ces zéros sont les *pivots* sur lesquels on prend appui pour annuler le restant de la ligne par opérations sur les colonnes.

(b) méthode

Par opérations élémentaires sur les rangées, on peut transformer une matrice en une matrice échelonnée pour laquelle le rang est immédiat à déterminer.

On prend appui sur le premier coefficient de la matrice pour annuler par opérations sur les lignes les coefficients figurant en dessous

$$\text{rg} \begin{pmatrix} 1 & 1 & 1 & -1 \\ -1 & -1 & 0 & 1 \\ 1 & 0 & 2 & -1 \end{pmatrix} \stackrel{L_1 \leftarrow L_1 + L_1}{=} \text{rg} \begin{pmatrix} 1 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \end{pmatrix} \stackrel{L_2 \leftrightarrow L_3}{=} \text{rg} \begin{pmatrix} 1 & 1 & 1 & -1 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = 3.$$

(c) On commence par échanger les deux premières lignes afin de faire figurer en tête de la matrice un coefficient permettant d'annuler ceux dessous

$$\text{rg} \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ -1 & 1 & 0 & -1 \\ 1 & -1 & 0 & 1 \end{pmatrix} \stackrel{L_1 \leftrightarrow L_2}{=} \text{rg} \begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ -1 & 1 & 0 & -1 \\ 1 & -1 & 0 & 1 \end{pmatrix} \stackrel{L_3 \leftarrow L_3 + L_1}{=} \text{rg} \begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & -1 & 0 & 1 \end{pmatrix} \\ \stackrel{L_3 \leftarrow L_3 - L_2}{=} \text{rg} \begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 1 \end{pmatrix} = 2.$$

Exercice 5

Montrer que la matrice suivante est inversible et calculer son inverse

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & -1 & 1 \\ -1 & 1 & -1 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R}).$$

Solution**méthode**

Les opérations élémentaires sur les lignes qui transforment une matrice inversible en la matrice unité transforment parallèlement la matrice unité en l'inverse recherché.

On figure parallèlement la matrice A et la matrice \mathbf{I}_n .

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 2 & -1 & 1 & 0 & 1 & 0 \\ -1 & 1 & -1 & 0 & 0 & 1 \end{array} \right).$$

Les opérations $L_2 \leftarrow L_2 - 2L_1$ et $L_3 \leftarrow L_3 + L_1$ annulent les coefficients en dessous du 1 de la première colonne

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{array} \right).$$

L'opération $L_3 \leftarrow L_3 + L_2$ finit de transformer la matrice A en une matrice triangulaire

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & -2 & 1 & 0 \\ 0 & 0 & -1 & -1 & 1 & 1 \end{array} \right).$$

À ce stade on est assuré que la matrice A est de rang 3, c'est donc une matrice inversible. Les opérations $L_2 \leftarrow -L_2$ et $L_3 \leftarrow -L_3$ transforment les coefficients diagonaux en des 1

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 2 & -1 & 0 \\ 0 & 0 & 1 & 1 & -1 & -1 \end{array} \right).$$

Enfin, les opérations $L_2 \leftarrow L_2 - L_3$ et $L_1 \leftarrow L_1 - L_3$ finissent de transformer le bloc de gauche en la matrice unité et le bloc de droite décrit alors la matrice inverse de A .

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & -1 & -1 \end{array} \right).$$

On peut aussi inverser la matrice A en introduisant le système d'équation matricielle $AX = Y$ avec X et Y colonnes de coefficients x_1, x_2, x_3 et y_1, y_2, y_3 . Résoudre ce système afin d'exprimer X en fonction de Y revient à écrire $Y = A^{-1}X$ ce qui révèle la matrice inverse de A :

$$\left\{ \begin{array}{l} x_1 + x_3 = y_1 \\ 2x_1 - x_2 + x_3 = y_2 \\ -x_1 + x_2 - x_3 = y_3 \end{array} \right. \iff \left\{ \begin{array}{l} x_1 = y_2 + y_3 \\ x_2 = y_1 + y_3 \\ x_3 = y_1 - y_2 - y_3 \end{array} \right.$$

9.5.2 Matrices et applications linéaires

Exercice 6

Soit f l'endomorphisme de \mathbb{R}^3 figuré dans la base canonique par la matrice

$$A = \begin{pmatrix} -1 & 3 & -3 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}.$$

On introduit les vecteurs $e_1 = (1, 1, 0)$, $e_2 = (-1, 1, 1)$ et $e_3 = (0, 1, 1)$.

- (a) Montrer que $e = (e_1, e_2, e_3)$ constitue une base de \mathbb{R}^3 .
- (b) Ecrire la matrice de f dans cette base.
- (c) Sans calculs, déterminer une base de $\text{Ker}(f)$ et de $\text{Im}(f)$.

Solution

(a) La famille e est constituée de trois vecteurs d'un espace de dimension 3 : il suffit de vérifier qu'elle est libre pour pouvoir affirmer qu'elle forme une base (Th. 12 p. 243). Supposons $\lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 e_3 = 0_{\mathbb{R}^3}$ avec $\lambda_1, \lambda_2, \lambda_3$ des réels. Cette égalité produit le système

$$\begin{cases} \lambda_1 - \lambda_2 = 0 \\ \lambda_1 + \lambda_2 + \lambda_3 = 0 \\ \lambda_2 + \lambda_3 = 0. \end{cases}$$

Après résolution, on constate que seul le triplet $(\lambda_1, \lambda_2, \lambda_3) = (0, 0, 0)$ est solution de ce système. La famille e est donc libre et constitue une base de l'espace \mathbb{R}^3 .

(b) méthode

On remplit la matrice figurant f dans la base e en calculant les images des vecteurs $f(e_1), f(e_2)$ et $f(e_3)$.

Si X désigne la colonne des coordonnées d'un vecteur $x = (x_1, x_2, x_3)$ dans la base canonique, le produit AX détermine la colonne figurant $f(x)$ (Th. 6 p. 314). Les produits matriciels suivants déterminent alors les images des vecteurs e_1, e_2, e_3 par f :

$$\begin{pmatrix} -1 & 3 & -3 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} -1 & 3 & -3 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 3 & -3 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

méthode

Ce ne sont pas directement ces colonnes qui remplissent la matrice de l'endomorphisme f dans la base e mais les coordonnées dans e des vecteurs associés.

Les calculs matriciels précédents donnent $f(e_1) = 2e_1$, $f(e_2) = -e_2$ et $f(e_3) = 0_{\mathbb{R}^3}$. On en déduit

$$\text{Mat}_e(f) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

(c) La matrice figurant f est de rang 2 et donc l'endomorphisme f est aussi de rang 2. Il suffit de déterminer deux vecteurs indépendants dans l'image de f pour en constituer une base : les vecteurs e_1 et e_2 peuvent¹ convenir car les calculs qui précèdent assurent qu'ils appartiennent à l'image de f .

Par la formule du rang, le noyau de f est de dimension 1 et puisque e_3 en est un vecteur non nul, il constitue une base de $\text{Ker}(f)$.

1. Aussi, les colonnes de la matrice A correspondent aux images des vecteurs de la base canonique. Elles définissent donc des éléments de l'image de f pouvant servir à en constituer une base.

Exercice 7

Soit E un espace vectoriel réel muni d'une base $e = (e_1, e_2, e_3)$ et f l'endomorphisme de E figuré dans la base e par la matrice

$$A = \begin{pmatrix} -3 & -1 & 1 \\ 4 & 2 & -1 \\ 2 & 2 & -1 \end{pmatrix}.$$

On pose $e'_1 = e_2 + e_3$, $e'_2 = e_1 - e_2 + e_3$ et $e'_3 = e_1 - e_2$.

- (a) Vérifier que $e' = (e'_1, e'_2, e'_3)$ est une base de E .
- (b) Former la matrice D de f dans la base e' .
- (c) Exprimer la matrice de passage P de e à e' et son inverse P^{-1} .
- (d) Quelle relation relie les matrices A , D et P ?
- (e) En déduire une expression de A^n pour $n \in \mathbb{N}$.

Solution

(a) La famille e' est une famille de trois vecteurs de l'espace E de dimension 3 : on peut vérifier que c'est une base de E en étudiant sa liberté ou, plus directement, en calculant son rang (Th. 9 p. 316)

$$\begin{aligned} \text{rg}(e'_1, e'_2, e'_3) &= \text{rg} \text{Mat}_e(e'_1, e'_2, e'_3) = \text{rg} \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 1 & 0 \end{pmatrix} \xrightarrow[L_1 \leftrightarrow L_2]{} \text{rg} \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \\ &\xrightarrow[L_3 \leftarrow L_3 - L_1]{} \text{rg} \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix} \xrightarrow[L_3 \leftarrow L_3 - 2L_2]{} \text{rg} \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix} = 3. \end{aligned}$$

La famille $e = (e_1, e_2, e_3)$ est donc une base de E .

(b) méthode

On remplit la matrice D par colonne à l'aide des coordonnées dans e' des images des vecteurs de e' .

Si X désigne la colonne des coordonnées dans e d'un vecteur x de E , le produit AX fournit la colonne des coordonnées du vecteur $f(x)$ dans la base e . Les produits suivants déterminent alors les images des vecteurs e'_i :

$$\begin{pmatrix} -3 & -1 & 1 \\ 4 & 2 & -1 \\ 2 & 2 & -1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -3 & -1 & 1 \\ 4 & 2 & -1 \\ 2 & 2 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} -3 & -1 & 1 \\ 4 & 2 & -1 \\ 2 & 2 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 0 \end{pmatrix}.$$

Sur ces calculs, on lit $f(e'_1) = e'_1$, $f(e'_2) = -e'_2$ et $f(e'_3) = -2e'_3$. On peut alors former la matrice D

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

(c) **méthode**

|| La matrice de passage de e à e' figure les vecteurs de e' dans la base e .

$$P = \text{Mat}_e(e'_1, e'_2, e'_3) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 1 & 0 \end{pmatrix}.$$

On retrouve la matrice dont on a calculé le rang ci-dessus.

méthode

|| La matrice P^{-1} est la matrice de passage de e' à e : on renverse le système exprimant les vecteurs de e' en fonction de ceux de e .

Après résolution

$$\begin{cases} e_2 + e_3 = e'_1 \\ e_1 - e_2 + e_3 = e'_2 \\ e_1 - e_2 = e'_3 \end{cases} \iff \begin{cases} e_1 = e'_1 - e'_2 + 2e'_3 \\ e_2 = e'_1 - e'_2 + e'_3 \\ e_3 = e'_2 - e'_3 \end{cases} \text{ donc } P^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ -1 & -1 & 1 \\ 2 & 1 & -1 \end{pmatrix}.$$

(d) Par la formule de changement de base (Th. 13 p. 318), on sait $A = PDP^{-1}$.

(e) **méthode**

|| Il est facile de calculer D^n et l'on peut en déduire A^n .

Soit $n \in \mathbb{N}$. On vérifie par récurrence

$$D^n = \begin{pmatrix} 1 & 0 & 0 \\ 0 & (-1)^n & 0 \\ 0 & 0 & (-2)^n \end{pmatrix}$$

En simplifiant les facteurs $P^{-1}P$ en I_n , on a aussi

$$A^n = \underbrace{(PDP^{-1})(PDP^{-1}) \dots (PDP^{-1})}_{n \text{ facteurs } PDP^{-1}} = PD\underbrace{(P^{-1}P)}_{=I_n} \dots \underbrace{(P^{-1}P)}_{=I_n} DP^{-1} = PD^n P^{-1}$$

On en déduit après quelques calculs¹

$$A^n = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} + (-1)^n \begin{pmatrix} -1 & -1 & 1 \\ 1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix} + (-2)^n \begin{pmatrix} 2 & 1 & -1 \\ -2 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Exercice 8

Déterminer le noyau et l'image de la matrice

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & -1 & 0 \end{pmatrix} \in \mathcal{M}_3(\mathbb{R}).$$

1. En décomposant D^n en une combinaison linéaire des matrices élémentaires $E_{1,1}$, $E_{2,2}$ et $E_{3,3}$, on forme par des calculs simples l'écriture proposée où A^n s'exprime comme combinaison linéaire de trois matrices de projection de rang 1.

Solution

Rappelons que l'image et le noyau d'une matrice A de $\mathcal{M}_{n,p}(\mathbb{K})$ sont l'image et le noyau de l'application linéaire qui lui est canoniquement associée. Cette application linéaire est celle qui à $x = (x_1, \dots, x_p) \in \mathbb{K}^p$ associe Ax où l'on identifie x et la colonne constituée des éléments x_1, \dots, x_p .

méthode

|| La résolution du système associé à l'équation $Ax = 0$ détermine le noyau de A .

Soit $x = (x_1, x_2, x_3) \in \mathbb{R}^3$. Le système associé à l'équation $Ax = 0$ s'écrit

$$\begin{cases} x_1 + x_3 = 0 \\ x_2 + x_3 = 0 \\ x_1 - x_2 = 0 \end{cases} \quad \text{soit encore} \quad \begin{cases} x_3 = -x_1 \\ x_2 = x_1 \end{cases}$$

Le noyau de A est donc constitué des triplets $(x_1, x_1, -x_1)$ avec x_1 quelconque :

$$\text{Ker}(A) = \{x_1(1, 1, -1) \mid x_1 \in \mathbb{R}\} = \text{Vect}(1, 1, -1).$$

Le noyau de A est la droite vectorielle engendrée par $(1, 1, -1)$.

Par la formule du rang, il vient $\text{rg}(A) = 3 - 1 = 2$: il suffit donc de connaître deux éléments indépendants dans l'image de A pour déterminer une base de celle-ci.

méthode

|| Les colonnes d'une matrice constituent une famille génératrice de son image.

Les deux premières colonnes de la matrice A n'étant pas colinéaires, elles forment une base de l'image A

$$\text{Im}(A) = \text{Vect}((1, 0, 1), (0, 1, -1)).$$

Exercice 9

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,q}(\mathbb{K})$. Montrer

$$AB = \mathbf{O}_{n,q} \iff \text{Im}(B) \subset \text{Ker}(A).$$

Solution

Cette affirmation est la transposition matricielle d'une étude vectorielle déjà menée¹.

méthode

|| Les démonstrations vectorielles peuvent souvent être transposées aux matrices.

Raisonnons par double implication.

(\Rightarrow) Supposons $AB = \mathbf{O}_{n,q}$. Pour tout y de l'image de B , il existe x dans \mathbb{K}^q tel que $y = Bx$ et alors $Ay = ABx = 0$. Ainsi, y appartient au noyau de A et donc $\text{Im}(B) \subset \text{Ker}(A)$.

(\Leftarrow) Supposons $\text{Im}(B) \subset \text{Ker}(A)$. Pour tout x de \mathbb{K}^q , on a $Bx \in \text{Im}(B)$ donc $Bx \in \text{Ker}(A)$ et par conséquent $ABx = 0$. Ainsi, la matrice AB est nulle puisque l'application linéaire qui lui est canoniquement associée est nulle.

1. Voir sujet 3 p. 282. Dans un esprit semblable, on peut établir pour $A, B \in \mathcal{M}_n(\mathbb{K})$ les inclusions $\text{Im}(A^2) \subset \text{Im}(A)$, $\text{Ker}(A) \subset \text{Ker}(A^2)$, $\text{Ker}(A) \cap \text{Ker}(B) \subset \text{Ker}(A+B)$, $\text{Im}(A+B) \subset \text{Im}(A) + \text{Im}(B)$, ...

9.6 Exercices d'entraînement

9.6.1 Les matrices carrées

Exercice 10 *

Déterminer toutes les matrices M de $\mathcal{M}_2(\mathbb{R})$ vérifiant $M^2 = \mathbf{I}_2$.

Solution

méthode

|| On recherche M par coefficients inconnus.

En écrivant $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ l'équation $M^2 = \mathbf{I}_2$ équivaut¹ au système

$$(\Sigma): \begin{cases} a^2 + bc = 1 \\ (a+d)b = 0 \\ (a+d)c = 0 \\ d^2 + bc = 1. \end{cases}$$

On mène la résolution de (Σ) en discutant selon la nullité de $a+d$.

Cas : $a+d \neq 0$. Le système donne $b=c=0$ et $a^2=d^2=1$. On a donc $a=d=1$ ou $a=-d=-1$. Ces solutions déterminent les matrices \mathbf{I}_2 et $-\mathbf{I}_2$ qui vérifient effectivement l'équation $M^2 = \mathbf{I}_2$.

Cas : $a+d=0$. Le système (Σ) équivaut au suivant :

$$\begin{cases} d = -a \\ bc = 1 - a^2. \end{cases}$$

Les solutions associées sont les matrices

$$\begin{pmatrix} a & b \\ c & -a \end{pmatrix} \quad \text{avec } (a, b, c) \in \mathbb{R}^3 \text{ vérifiant } a^2 + bc = 1.$$

Ce sujet illustre que les résolutions d'équations dans le cadre matriciel peuvent produire « plus de solutions » que dans le cadre numérique.

Exercice 11 *

On considère la matrice réelle

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

(a) Calculer A^n pour $n \in \mathbb{N}$.

(b) Calculer A^n pour $n \in \mathbb{Z}$.

1. L'équation $M^2 = \mathbf{I}_2$ équivaut aussi à $(M - \mathbf{I}_2)(M + \mathbf{I}_2) = \mathbf{O}_2$. Cependant, cette écriture ne permet pas d'affirmer $M = \pm \mathbf{I}_2$ car un produit de deux matrices non nulles peut être égal à la matrice nulle!

Solution(a) **méthode**

|| On écrit¹ $A = I_3 + N$ puis on applique la formule du binôme.

En écrivant $A = I_3 + N$, les puissances de la matrice N introduite sont remarquables

$$N = \begin{pmatrix} 0 & 2 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}, N^2 = \begin{pmatrix} 0 & 0 & 4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, N^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ et } N^k = O_3 \text{ pour } k \geq 3.$$

Les matrices I_3 et N commutent et l'on peut donc appliquer la formule du binôme

$$\begin{aligned} A^n &= (I_3 + N)^n = \sum_{k=0}^n \binom{n}{k} N^k = \binom{n}{0} I_3 + \binom{n}{1} N + \binom{n}{2} N^2 + \sum_{k=3}^n \binom{n}{k} N^k \underset{=O_3}{=} O_3 \\ &= I_3 + nN + \frac{n(n-1)}{2} N^2. \end{aligned}$$

On peut alors conclure

$$A^n = \begin{pmatrix} 1 & 2n & n(2n+1) \\ 0 & 1 & 2n \\ 0 & 0 & 1 \end{pmatrix}. \quad (*)$$

(b) La matrice A est inversible car triangulaire à coefficients diagonaux non nuls, on peut donc introduire A^n pour $n \in \mathbb{Z}$.

méthode

|| L'expression de A^n obtenue pour $n \in \mathbb{N}$ peut être étendue à $n \in \mathbb{Z}$.

Soit n un entier négatif et $p = -n \in \mathbb{N}$. La matrice A^n est l'inverse de A^p et en vérifiant

$$\begin{pmatrix} 1 & 2n & n(2n+1) \\ 0 & 1 & 2n \\ 0 & 0 & 1 \end{pmatrix} \underbrace{\begin{pmatrix} 1 & 2p & p(2p+1) \\ 0 & 1 & 2p \\ 0 & 0 & 1 \end{pmatrix}}_{=A^p} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

on peut affirmer que l'identité (*) est encore valable pour $n \in \mathbb{Z}$.

Exercice 12 **

Soit $n \in \mathbb{N}$ avec $n \geq 2$.

- (a) Déterminer les matrices commutant avec toutes celles de $\mathcal{M}_n(\mathbb{K})$.
- (b) Déterminer les matrices commutant avec toutes celles de $\mathrm{GL}_n(\mathbb{K})$.
- (c) Soit $A \in \mathcal{M}_n(\mathbb{K})$. On suppose que, pour toutes matrices M et N de $\mathcal{M}_n(\mathbb{K})$,

$$A = MN \implies A = NM.$$

Montrer qu'il existe $\lambda \in \mathbb{K}$ tel que $A = \lambda I_n$.

1. On peut aussi calculer les premières puissances de A et proposer une formule que l'on valide par récurrence.

Solution**(a) méthode**

On vérifie que seules les *matrices scalaires*, c'est-à-dire les matrices de la forme λI_n avec $\lambda \in \mathbb{K}$, commutent avec toutes les matrices de $\mathcal{M}_n(\mathbb{K})$.

Soit $\lambda \in \mathbb{K}$. Pour toute matrice M de $\mathcal{M}_n(\mathbb{K})$, on vérifie immédiatement $(\lambda I_n)M = \lambda M$ et $\lambda M = M(\lambda I_n)$. Ainsi, les matrices scalaires commutent avec toutes les matrices de $\mathcal{M}_n(\mathbb{K})$.

Inversement, soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ vérifiant $AM = MA$ pour tout $M \in \mathcal{M}_n(\mathbb{K})$.

méthode

L'hypothèse $AM = MA$ est délicate à analyser dans sa généralité. Particulariser celle-ci à des matrices M simples permet de réunir des conditions sur les coefficients de A .

Soit i et j deux indices de $[1 ; n]$. Considérons la matrice élémentaire $M = E_{i,j}$. Dans la matrice $AE_{i,j}$ les colonnes d'indices différents de j sont nulles tandis que la colonne d'indice j correspond à celle d'indice i de la matrice A . Parallèlement, dans la matrice $E_{i,j}A$ toutes les lignes d'indices différents de i sont nulles tandis que la ligne d'indice i correspond à celle d'indice j de la matrice A :

$$AE_{i,j} = \begin{pmatrix} & & \downarrow^j \\ a_{1,i} & & \\ (0) & \vdots & (0) \\ & a_{n,i} & \end{pmatrix} \quad \text{et} \quad E_{i,j}A = \begin{pmatrix} (0) \\ a_{j,1} & \cdots & a_{j,n} \\ (0) \end{pmatrix}^i$$

Pour $i \neq j$, l'identification des coefficients d'indice (i, j) d'une part, et d'indice (j, j) d'autre part, donne les égalités $a_{i,j} = a_{j,j}$ et $a_{j,i} = 0$. On en déduit que la matrice A est diagonale à coefficients diagonaux identiques : c'est une matrice scalaire.

(b) méthode

On se ramène à la situation précédente en introduisant $I_n + E_{i,j}$.

Les matrices scalaires commutent avec toutes les matrices inversibles. Inversement, considérons $A \in \mathcal{M}_n(\mathbb{K})$ commutant avec toutes les matrices inversibles. Pour $i \neq j$ choisis dans $[1 ; n]$, la matrice $I_n + E_{i,j}$ est inversible car triangulaire à coefficients diagonaux non nuls. La matrice A commute donc avec celle-ci ce qui entraîne qu'elle commute avec la matrice élémentaire $E_{i,j}$. Comme au-dessus, on peut en déduire que la matrice A est scalaire.

Finalement, les matrices commutant avec toutes les matrices inversibles sont les matrices scalaires.

(c) Soit $M \in \mathrm{GL}_n(\mathbb{K})$. On peut écrire $A = MN$ avec $N = M^{-1}A$. On a donc $A = NM = M^{-1}AM$ puis $MA = AM$. Ainsi, la matrice A commute avec toutes les matrices inversibles et l'on peut conclure que A est une matrice scalaire.

Exercice 13 **

Soit E l'ensemble des matrices de la forme

$$M(a, b, c) = \begin{pmatrix} a & b & c \\ b & a+c & b \\ c & b & a \end{pmatrix} \quad \text{avec } a, b, c \in \mathbb{R}.$$

(a) Montrer que $(E, +, .)$ est un espace vectoriel réel dont on précisera la dimension.

(b) Montrer que E est stable pour le produit matriciel.

Soit A une matrice inversible de E .

(c) En considérant l'application $f: M \mapsto AM$ définie sur E , montrer que l'inverse de A est élément de E .

Solution(a) **méthode**

|| On décrit E comme l'espace vectoriel engendré par trois matrices.

Pour $(a, b, c) \in \mathbb{R}^3$, une matrice $M(a, b, c)$ peut s'écrire $aI_3 + bJ + cK$ avec

$$J = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \text{et} \quad K = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

L'ensemble E est donc l'ensemble des combinaisons linéaires réelles des trois matrices I_3 , J et K : $E = \text{Vect}(I_3, J, K)$ est un sous-espace vectoriel de $\mathcal{M}_3(\mathbb{R})$, donc un espace vectoriel réel, et la famille (I_3, J, K) est génératrice de celui-ci. De plus, cette famille est libre car, pour tous réels λ , μ et ν , l'étude des coefficients donne

$$\lambda I_3 + \mu J + \nu K = O_3 \implies \lambda = \mu = \nu = 0.$$

La famille (I_3, J, K) est donc une base de E et par conséquent $\dim E = 3$.

(b) **méthode**

|| Il suffit de vérifier l'appartenance à E des matrices JK , KJ , J^2 et K^2 .

Par calcul, on obtient $JK = KJ = J$, $J^2 = I_3 + K$ et $K^2 = I_3$. Par conséquent, le produit de deux matrices éléments de E est une combinaison linéaire des trois matrices I_3 , J et K : c'est un élément de E .

(c) **méthode**

|| On vérifie que f est un automorphisme de l'espace E .

L'application f est linéaire définie sur E et à valeurs dans E : c'est un endomorphisme de E . On étudie son noyau. Soit M une matrice de E telle $f(M) = O_3$, c'est-à-dire telle que $AM = O_3$. En introduisant l'inverse de A , on peut écrire $M = A^{-1}(AM)$ et

donc $M = \mathbf{O}_3$. Ainsi, le noyau de f est réduit à la matrice nulle. L'endomorphisme f est alors injectif dans un espace vectoriel de dimension finie, c'est un automorphisme (Th. 15 p. 278).

Par surjectivité de f , il existe une matrice $B \in E$ telle que $f(B) = AB = \mathbf{I}_3$. En multipliant à gauche par A^{-1} , on conclut $A^{-1} = B \in E$.

9.6.2 Matrices carrées inversibles

Exercice 14 *

Soit $A, B \in \mathcal{M}_n(\mathbb{K})$ vérifiant $AB = A + B$. Montrer que A et B commutent.

Solution

méthode

|| Si M et N sont des matrices carrées vérifiant $MN = \mathbf{I}_3$, celles-ci commutent car inverses l'une de l'autre (Th. 3 p. 311).

En ajoutant \mathbf{I}_n aux membres de l'équation $AB = A + B$, il vient $(\mathbf{I}_n - A)(\mathbf{I}_n - B) = \mathbf{I}_n$. On en déduit que la matrice $\mathbf{I}_n - A$ est inversible et que $\mathbf{I}_n - B$ est son inverse. L'égalité $(\mathbf{I}_n - B)(\mathbf{I}_n - A) = \mathbf{I}_n$ entraîne alors $B\mathbf{A} = \mathbf{A} + B$ et l'on peut conclure que $AB = BA$: les matrices A et B commutent.

Exercice 15 **

On dit que $N \in \mathcal{M}_n(\mathbb{K})$ est une matrice *nilpotente* s'il existe $p \in \mathbb{N}^*$ tel que $N^p = \mathbf{O}_n$.

(a) On suppose que N est une matrice nilpotente de $\mathcal{M}_n(\mathbb{R})$. Vérifier que la matrice $\mathbf{I}_3 - N$ est inversible.

(b) Soit de plus $A \in \mathcal{M}_n(\mathbb{K})$ telle que $AN = NA$. Montrer que A et $A + N$ sont simultanément inversibles.

Solution

(a) méthode

|| On détermine¹ un inverse à la matrice $\mathbf{I}_3 - N$ en exploitant la formule de factorisation géométrique (Th. 7 p. 124).

On introduit $p \in \mathbb{N}^*$ tel que $N^p = \mathbf{O}_n$ et l'on peut écrire

$$\mathbf{I}_n = \mathbf{I}_n - N^p = (\mathbf{I}_n - N) \sum_{k=0}^{p-1} N^k \quad \text{car } \mathbf{I}_n \text{ et } N \text{ commutent.}$$

On en déduit que la matrice $\mathbf{I}_n - N$ est inversible et $\mathbf{I}_n + N + \cdots + N^{p-1}$ est son inverse.

1. Cette question est un cas particulier de l'étude vue dans le sujet 24 p. 139.

(b) Supposons A inversible.

méthode

|| On vérifie que $A^{-1}N$ est nilpotente.

En multipliant l'égalité $AN = NA$ par A^{-1} à droite et à gauche, on obtient l'identité $NA^{-1} = A^{-1}N$. Ainsi, les matrices A^{-1} et N commutent. En introduisant à nouveau p tel que $N^p = \mathbf{O}_n$, on peut écrire par commutation $(A^{-1}N)^p = (A^{-1})^p N^p = \mathbf{O}_n$: la matrice $A^{-1}N$ est donc nilpotente. Il en est de même de son opposée $-A^{-1}N$ et l'étude de la question précédente assure que $\mathbf{I}_n + A^{-1}N$ est inversible. En multipliant à gauche par A , on peut conclure que $A + N$ est inversible en tant que produit de deux matrices inversibles.

Inversement, supposons $A + N$ inversible. Puisque $-N$ est nilpotente et commute avec $A + N$, on peut exploiter l'étude ci-dessus pour affirmer que $A = (A + N) + (-N)$ est inversible.

Exercice 16 *** (Matrice à diagonale strictement dominante)

Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ vérifiant

$$|a_{i,i}| > \sum_{j \neq i} |a_{i,j}| \quad \text{pour tout } i \in \llbracket 1 ; n \rrbracket.$$

Montrer que la matrice A est inversible.

Solution

méthode

|| Une matrice est inversible lorsque son noyau est réduit à la colonne nulle (Th. 11 p. 317).

Soit $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ un élément du noyau de la matrice A . On a $Ax = 0$ (dans cette écriture, on identifie x avec la colonne formée des mêmes coefficients) et donc

$$\sum_{j=1}^n a_{i,j} x_j = 0 \quad \text{pour tout } i \in \llbracket 1 ; n \rrbracket.$$

Afin d'exploiter l'hypothèse de travail, on isole le terme d'indice i de la somme et l'on passe en valeurs absolues afin d'écrire

$$|a_{i,i}| |x_i| = \left| - \sum_{j \neq i} a_{i,j} x_j \right| \leq \sum_{j \neq i} |a_{i,j}| |x_j|. \quad (*)$$

méthode

|| On introduit un indice i_0 tel que $|x_{i_0}|$ est le maximum des $|x_1|, \dots, |x_n|$.

Les $|x_j|$ étant tous inférieurs à $|x_{i_0}|$, l'inégalité (*) permet d'écrire

$$|a_{i_0, i_0}| |x_{i_0}| \leq \sum_{j \neq i_0} |a_{i_0, j}| \underbrace{|x_j|}_{\leq |x_{i_0}|} \leq \sum_{j \neq i_0} |a_{i_0, j}| |x_{i_0}| = \left(\sum_{j \neq i_0} |a_{i_0, j}| \right) |x_{i_0}|. \quad (**)$$

Si par l'absurde $|x_{i_0}| > 0$, on simplifie (**) par $|x_{i_0}|$ et cela produit une inégalité qui contredit l'hypothèse du sujet. On en déduit $|x_{i_0}| = 0$ puis $x_1 = \dots = x_n = 0$ car $|x_{i_0}|$ a été introduit comme le maximum des $|x_1|, \dots, |x_n|$.

Finalement, le noyau de A est réduit à l'élément nul et la matrice A est inversible.

9.6.3 Matrices et applications linéaires

Exercice 17 *

Soit f l'endomorphisme de \mathbb{R}^3 figuré dans la base canonique par la matrice

$$A = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

- (a) Déterminer le noyau et l'image de f .
- (b) Vérifier que ces espaces sont supplémentaires et exprimer la matrice de f dans une base adaptée à cette complémentarité.
- (c) Décrire f comme la composition de transformations vectorielles simples.

Solution

(a) méthode

On détermine le noyau de f en résolvant l'équation matricielle $AX = 0$ d'inconnue une colonne X .

Soit $(x, y, z) \in \mathbb{R}^3$ et X la colonne de coefficients x, y et z . On a $f(x, y, z) = 0_{\mathbb{R}^3}$ si, et seulement si, $AX = 0$, c'est-à-dire

$$\begin{cases} x + z = 0 \\ -x + 2y + z = 0 \\ x + z = 0 \end{cases}$$

Après résolution, on obtient

$$\text{Ker}(f) = \{(x, x, -x) \mid x \in \mathbb{R}\} = \text{Vect}(1, 1, -1).$$

Le noyau de f est donc la droite vectorielle engendrée par $(1, 1, -1)$.

Par la formule du rang, on peut affirmer que l'image de f est un plan.

méthode

|| Les colonnes de A déterminent des éléments de l'image de f .

Les colonnes de A sont données par les images des vecteurs de la base canonique. On peut donc affirmer que les vecteurs $(1, -1, 1)$ et $(0, 2, 0)$ appartiennent à l'image de f . Ils sont linéairement indépendants et définissent une base de $\text{Im}(f)$. Cependant, l'image de f est un espace vectoriel, on peut « simplifier » par opérations ces vecteurs : $(0, 1, 0)$ et $(1, 0, 1)$ sont aussi des vecteurs formant une base de l'image¹ de f .

(b) méthode

|| Dans un espace de dimension finie, deux sous-espaces sont supplémentaires si, et seulement si, la réunion d'une base de l'un et d'une base de l'autre constitue une base de l'espace.

Posons $e_1 = (1, 1, -1)$, $e_2 = (0, 1, 0)$ et $e_3 = (1, 0, 1)$. Les familles (e_1) et (e_2, e_3) sont respectivement bases du noyau et de l'image de f . On vérifie que la famille (e_1, e_2, e_3) est une base de \mathbb{R}^3 en calculant son rang (Th. 9 p. 316)

$$\text{rg}(e_1, e_2, e_3) = \text{rg} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \underset{L_2 \leftarrow L_2 - L_1}{=} \underset{L_3 \leftarrow L_3 + L_1}{=} \text{rg} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{pmatrix} = 3.$$

Les espaces $\text{Ker}(f)$ et $\text{Im}(f)$ sont donc supplémentaires et la base e est adaptée à cette supplémentarité.

On forme la matrice de f dans la base e en calculant les coordonnées dans e des images $f(e_1)$, $f(e_2)$ et $f(e_3)$. On sait déjà $f(e_1) = 0_{\mathbb{R}^3}$ et les produits matriciels

$$\begin{pmatrix} 1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix}$$

donnent $f(e_2) = 2e_2$ et $f(e_3) = 2e_3$. La matrice de f dans e est donc²

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

(c) La matrice D peut s'écrire

$$D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

L'endomorphisme f est donc la composition de l'homothétie vectorielle de rapport 2 avec la projection vectorielle sur $\text{Im}(f)$ parallèlement à $\text{Ker}(f)$.

1. L'image de f apparaît comme le plan d'équation $x - z = 0$ dans \mathbb{R}^3 .

2. Dans toute base adaptée à la supplémentarité de $\text{Ker}(f)$ et $\text{Im}(f)$, la matrice de f est égale à la matrice D . Cette affirmation n'est pas vraie en général mais est une conséquence de la transformation vectorielle réalisée par f .

Exercice 18 *

Déterminer les transformations vectorielles de \mathbb{R}^3 réalisées par les endomorphismes figurés dans la base canonique par les matrices :

$$(a) A = \begin{pmatrix} 3 & -4 & 2 \\ 1 & -1 & 1 \\ -1 & 2 & 0 \end{pmatrix}$$

$$(b) B = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 1 & 1 \\ -1 & 0 & 0 \end{pmatrix}.$$

Solution**méthode**

Les projections vectorielles d'un espace E sont les endomorphismes p de E vérifiant $p^2 = p$. Les symétries vectorielles sont les endomorphismes s tels que $s^2 = \text{Id}_E$.

(a) Notons f l'endomorphisme canoniquement associé à la matrice A . On a $A^2 = A$ et donc $f^2 = f$: f est une projection vectorielle. Plus précisément, l'image et le noyau de f sont supplémentaires et f est la projection sur l'image parallèlement au noyau. On détermine le noyau en résolvant le système associé à l'équation matricielle $AX = 0$:

$$\begin{cases} 3x - 4y + 2z = 0 \\ x - y + z = 0 \\ -x + 2y = 0. \end{cases}$$

On obtient

$$\text{Ker}(f) = \{(2y, y, -y) \mid y \in \mathbb{R}\} = \text{Vect}(2, 1, -1).$$

Par la formule du rang¹, on peut affirmer que l'image de f est un plan. On obtient une base de celui-ci en considérant deux vecteurs indépendants appartenant à l'image, par exemple, deux vecteurs déterminés par les colonnes de A . On peut aussi déterminer simplement ce plan par une équation : les vecteurs de l'image d'une projection étant invariants par celle-ci, l'image de f s'obtient en résolvant le système associé à l'équation matricielle $AX = X$. Ce système se résume à la seule équation $x - 2y + z = 0$.

Finalement, f est la projection sur le plan d'équation $x - 2y + z = 0$ parallèlement à la droite $\text{Vect}(2, 1, -1)$.

(b) Notons g l'endomorphisme canoniquement associé à la matrice B . On observe $B^2 = I_3$ et donc $g^2 = \text{Id}_{\mathbb{R}^3}$: g est une symétrie vectorielle. Plus précisément, les noyaux de $g - \text{Id}_{\mathbb{R}^3}$ et $g + \text{Id}_{\mathbb{R}^3}$ sont supplémentaires et g est la symétrie par rapport au premier et parallèlement au second. On détermine ces noyaux en résolvant les systèmes associés aux équations matricielles $(B - I_3)X = 0$ et $(B + I_3)X = 0$:

$$\begin{cases} -x - z = 0 \\ x + z = 0 \\ -x - z = 0 \end{cases} \quad \text{et} \quad \begin{cases} x - z = 0 \\ x + 2y + z = 0 \\ -x + z = 0 \end{cases}$$

¹ On peut aussi calculer la trace de A et exploiter le Th. 17 p. 319 pour affirmer que f projette sur un plan.

Le noyau de $g - \text{Id}_{\mathbb{R}^3}$ est le plan d'équation $x + z = 0$ tandis que le noyau de $g + \text{Id}_{\mathbb{R}^3}$ est la droite $\{(x, -x, x) \mid x \in \mathbb{R}\} = \text{Vect}(1, -1, 1)$.

Finalement, g est la symétrie vectorielle par rapport au plan d'équation $x + z = 0$ et parallèlement à la droite $\text{Vect}(1, -1, 1)$.

Exercice 19 **

Soit $A = (a_{i,j})_{0 \leq i, j \leq n} \in \mathcal{M}_{n+1}(\mathbb{R})$ la matrice dont le coefficient général¹ est donné par le coefficient binomial :

$$a_{i,j} = \binom{j}{i} \quad \text{pour } (i, j) \in \llbracket 0 ; n \rrbracket^2.$$

Soit φ l'endomorphisme de $\mathbb{R}_n[X]$ représenté par la matrice A dans la base canonique $(1, X, \dots, X^n)$.

- (a) Exprimer simplement $\varphi(P)$ pour tout $P \in \mathbb{R}_n[X]$.
- (b) Montrer que A est inversible et calculer A^{-1} .

Solution

(a) méthode

¶ Pour comprendre ce sujet, il peut être pertinent d'étudier le cas $n = 3$.

Rappelons que $\binom{i}{j} = 0$ lorsque $i > j$. Quand $n = 3$, la matrice A s'écrit

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Les colonnes de cette matrice permettent de lire les images des vecteurs de la base canonique

$$\begin{aligned} \varphi(1) &= 1 \\ \varphi(X) &= 1 + X \\ \varphi(X^2) &= 1 + 2X + X^2 = (X + 1)^2 \\ \varphi(X^3) &= 1 + 3X + 3X^2 + X^3 = (X + 1)^3. \end{aligned}$$

Revenons au cas général. Pour tout $0 \leq j \leq n$,

$$\varphi(X^j) = \sum_{i=0}^n \binom{j}{i} X^i = \sum_{i=0}^j \binom{j}{i} X^i + \underbrace{\sum_{i=j+1}^n \binom{j}{i} X^i}_{=0} = (X + 1)^j.$$

1. On notera que, dans ce sujet, lignes et colonnes sont indexées à partir du rang 0.

Pour $P = a_0 + a_1X + \cdots + a_nX^n \in \mathbb{R}_n[X]$, il vient par combinaison linéaire

$$\varphi(P) = a_0 + a_1(X+1) + \cdots + a_n(X+1)^n = P(X+1).$$

Finalement, φ est l'endomorphisme qui envoie P sur le polynôme composé $P(X+1)$.

(b) La matrice A est inversible car triangulaire à coefficients diagonaux non nuls. L'endomorphisme φ est donc un isomorphisme.

méthode

|| L'inverse d'une matrice carrée figure un isomorphisme réciproque.

Considérons l'endomorphisme ψ de $\mathbb{R}_n[X]$ qui envoie P sur $P(X-1)$. On vérifie que les composées $\psi \circ \varphi$ et $\varphi \circ \psi$ sont égales à l'identité ce qui assure¹ que ψ est l'isomorphisme réciproque de φ . La matrice A^{-1} est donc la matrice de ψ dans la base $(1, X, \dots, X^n)$:

$$A^{-1} = \left((-1)^{j-i} a_{i,j} \right)_{0 \leq i, j \leq n}.$$

Exercice 20 **

Soit E un espace vectoriel réel de dimension 3 muni d'une base $e = (e_1, e_2, e_3)$.

On considère les matrices

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix} \quad \text{et} \quad D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Soit f l'endomorphisme de E figuré par la matrice A dans la base e .

(a) Montrer que l'endomorphisme f peut être représenté par la matrice D .

(b) En déduire une matrice $P \in \mathrm{GL}_3(\mathbb{R})$ telle que $A = PDP^{-1}$.

(c) On considère les suites réelles (x_n) , (y_n) et (z_n) déterminées par :

$$\begin{cases} x_0 = 1 \\ y_0 = 0 \\ z_0 = 0 \end{cases} \quad \text{et} \quad \forall n \in \mathbb{N}, \quad \begin{cases} x_{n+1} = y_n + z_n \\ y_{n+1} = x_n + z_n \\ z_{n+1} = x_n - y_n \end{cases}$$

Exprimer x_n , y_n et z_n pour tout $n \geq 1$.

Solution

(a) méthode

Une base dans laquelle l'endomorphisme f est figuré par D est formée de trois vecteurs non nuls solutions des équations $f(x) = 0_E$, $f(x) = -x$ et $f(x) = x$.

1. Puisque φ est un endomorphisme en dimension finie ou parce que l'on sait déjà que φ est un isomorphisme, il suffit en fait d'une seule composée égale à l'identité pour pouvoir affirmer $\varphi^{-1} = \psi$.

Soit $x = x_1e_1 + x_2e_2 + x_3e_3 \in E$. Le vecteur x est solution de l'équation $f(x) = 0_E$, $f(x) = -x$ ou $f(x) = x$ si, et seulement si, la colonne X de coefficients x_1, x_2, x_3 est solution des équations matricielles $AX = 0$, $(A + I_3)X = 0$ ou $(A - I_3)X = 0$. Ceci conduit à l'étude des systèmes :

$$\begin{cases} y + z = 0 \\ x + z = 0 \\ x - y = 0 \end{cases}, \quad \begin{cases} x + y + z = 0 \\ x + y + z = 0 \\ x - y + z = 0 \end{cases} \text{ et } \begin{cases} -x + y + z = 0 \\ x - y + z = 0 \\ x - y - z = 0 \end{cases}$$

Après résolution de ces trois systèmes, on peut introduire les vecteurs

$$\begin{cases} e'_1 = e_1 + e_2 - e_3 \\ e'_2 = e_1 - e_3 \\ e'_3 = e_1 + e_2 \end{cases}$$

vérifiant respectivement $f(e'_1) = 0_E$, $f(e'_2) = -e'_2$ et $f(e'_3) = e'_3$. Ces trois vecteurs déterminent une base de E comme le confirme le calcul de rang qui suit

$$\operatorname{rg}(e'_1, e'_2, e'_3) = \operatorname{rg} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ -1 & -1 & 0 \end{pmatrix} \stackrel{L_2 \leftarrow L_2 - L_1}{=} \stackrel{L_3 \leftarrow L_3 + L_1}{=} \begin{pmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 3.$$

Par construction, D est la matrice de f dans la base $e' = (e'_1, e'_2, e'_3)$.

(b) Par la formule de changement de base, on sait¹ $A = PDP^{-1}$ avec P la matrice de passage de e à e'

$$P = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ -1 & -1 & 0 \end{pmatrix}.$$

(c) méthode

On introduit X_n la colonne dont les coefficients sont x_n, y_n, z_n de sorte que

$$X_0 = {}^t(1 \ 0 \ 0) \quad \text{et} \quad \forall n \in \mathbb{N}, X_{n+1} = AX_n.$$

Par une récurrence immédiate, on observe $X_n = A^n X_0$ ce qui rapporte le problème à celui du calcul de A^n . Par la relation $A = PDP^{-1}$, on obtient $A^n = PD^nP^{-1}$ avec

$$D^n = \begin{pmatrix} 0 & 0 & 0 \\ 0 & (-1)^n & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{et, après calcul,} \quad P^{-1} = \begin{pmatrix} -1 & 1 & -1 \\ 1 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

On en déduit

$$A^n = \begin{pmatrix} 1 + (-1)^n & -(-1)^n & 1 \\ 1 & 0 & 1 \\ -(-1)^n & (-1)^n & 0 \end{pmatrix} \quad \text{puis} \quad \begin{cases} x_n = 1 + (-1)^n \\ y_n = 1 \\ z_n = (-1)^{n+1} \end{cases} \quad \text{pour } n \geq 1.$$

1. Les matrices A et D sont donc semblables.

Exercice 21 ***

Soit f un endomorphisme d'un espace réel E de dimension finie vérifiant $f^2 = 0$. Montrer qu'il existe une base de E dans laquelle la matrice de f s'écrit par blocs

$$\begin{pmatrix} 0 & I_r \\ 0 & 0 \end{pmatrix} \quad \text{avec } r \in \mathbb{N}$$

où les 0 désignent des blocs nuls de tailles appropriées.

Solution**méthode**

On raisonne par analyse-synthèse. Dans l'analyse, on détermine la valeur de r et les conditions que doivent remplir les vecteurs d'une base convenable. Dans la synthèse, on construit une base réunissant ces conditions.

Analyse : Supposons que $e = (e_1, \dots, e_n)$ soit une base de E dans laquelle l'endomorphisme f est figuré comme voulu. La valeur de r détermine le rang de la matrice et correspond donc au rang de f . Les vecteurs e_1, \dots, e_{n-r} appartiennent au noyau de f car les premières colonnes de la matrice sont nulles. Ces vecteurs forment alors une base de $\text{Ker}(f)$ puisque la formule du rang donne $\dim \text{Ker}(f) = n - r$. Au surplus, les vecteurs e_{n-r+1}, \dots, e_n sont envoyés respectivement sur e_1, \dots, e_r et ces derniers vecteurs doivent donc être éléments de l'image de f et en constituer une base.

Synthèse : Posons $r = \text{rg}(f)$ et introduisons (e_1, \dots, e_r) une base de l'image de f . Puisque l'endomorphisme f^2 est nul, on peut affirmer que l'image de f est incluse dans le noyau de f . La famille (e_1, \dots, e_r) peut alors se comprendre comme une famille libre d'éléments de $\text{Ker}(f)$, on complète celle-ci en une base de $\text{Ker}(f)$: (e_1, \dots, e_{n-r}) . Enfin, les vecteurs initiaux e_1, \dots, e_r appartenant à l'image de f , on peut introduire des antécédents e_{n-r+1}, \dots, e_n de sorte que $f(e_{n-r+i}) = e_i$ pour tout $i \in [1 ; r]$.

Considérons alors la famille $e = (e_1, \dots, e_n)$ constituée de l'ensemble de ces vecteurs. Vérifions que celle-ci est une base de E . Il s'agit d'une famille de $n = \dim E$ vecteurs, il suffit d'en étudier la liberté. Soit $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$ tel que

$$\lambda_1 e_1 + \dots + \lambda_n e_n = 0_E. \quad (*)$$

En composant (*) par f , on obtient

$$\lambda_{n-r+1} e_1 + \dots + \lambda_n e_r = 0_E \quad (**)$$

car les vecteurs e_1, \dots, e_{n-r} appartiennent au noyau de f alors que e_{n-r+1}, \dots, e_n sont envoyés respectivement sur e_1, \dots, e_r . La famille (e_1, \dots, e_r) étant libre, (**) donne $\lambda_{n-r+1} = \dots = \lambda_n = 0$. L'égalité (*) se simplifie alors en

$$\lambda_1 e_1 + \dots + \lambda_{n-r} e_{n-r} = 0_E$$

et l'on peut affirmer $\lambda_1 = \dots = \lambda_{n-r} = 0$ car (e_1, \dots, e_{n-r}) est une base de $\text{Ker}(f)$.

Finalement, la famille e est une base de E et, par construction, la matrice de f dans cette base est telle que voulue.

9.6.4 Trace d'une matrice carrée

Exercice 22 *

Soit A et B deux matrices de $\mathcal{M}_n(\mathbb{K})$ vérifiant $AB - BA = A$.

Calculer $\text{tr}(A^p)$ pour tout $p \in \mathbb{N}$.

Solution
méthode

|| On exploite l'identité $\text{tr}(AB) = \text{tr}(BA)$ (Th. 16 p. 319).

Par linéarité

$$\text{tr}(A) = \text{tr}(AB - BA) = \text{tr}(AB) - \text{tr}(BA) = 0.$$

On généralise ce calcul :

$$\text{tr}(A^p) = \text{tr}(A^{p-1}(AB - BA)) = \text{tr}(A^p B) - \text{tr}(A^{p-1} B A) = 0$$

car¹

$$\text{tr}(A^{p-1} B A) = \text{tr}((A^{p-1} B) A) = \text{tr}(A(A^{p-1} B)) = \text{tr}(A^p B).$$

Exercice 23 **

Soit T une forme linéaire sur $\mathcal{M}_n(\mathbb{K})$ vérifiant $T(AB) = T(BA)$ pour toutes matrices A et B de $\mathcal{M}_n(\mathbb{K})$. Etablir que T est colinéaire à la forme linéaire trace.

Solution
méthode

|| Deux applications linéaires sont égales lorsqu'elles coïncident sur une base (Th. 12 p. 277).

Considérons la base canonique de l'espace $\mathcal{M}_n(\mathbb{K})$ constituée des matrices élémentaires $E_{i,j}$ avec $1 \leq i, j \leq n$. En exploitant le résultat du sujet 1 p. 321, on peut écrire pour $i \neq j$

$$E_{i,i} E_{i,j} = E_{i,j} \quad \text{et} \quad E_{i,j} E_{i,i} = O_n.$$

On a donc $T(E_{i,j}) = T(O_n) = 0$. Aussi,

$$E_{i,j} E_{j,i} = E_{i,i} \quad \text{et} \quad E_{j,i} E_{i,j} = E_{j,j}$$

et donc $T(E_{i,i}) = T(E_{j,j})$.

En notant λ la valeur commune prise par T sur les matrices élémentaires $E_{i,i}$, on peut affirmer l'égalité $T = \lambda \cdot \text{tr}()$ car ces deux applications linéaires prennent les mêmes valeurs sur la base des matrices élémentaires.

1. On est attentif à l'organisation du calcul : la propriété $\text{tr}(AB) = \text{tr}(BA)$ ne permet pas d'établir l'égalité $\text{tr}(ABC) = \text{tr}(ACB)$.

9.6.5 Rang

Exercice 24 *

Calculer le rang des matrices suivantes en fonction des paramètres réels a , b et c :

$$(a) A = \begin{pmatrix} 1 & 1 & 1 \\ b+c & c+a & a+b \\ bc & ca & ab \end{pmatrix}$$

$$(b) B = \begin{pmatrix} 1 & 1 & 1 \\ a & b & c \\ a^3 & b^3 & c^3 \end{pmatrix}$$

Solution

(a) méthode

En réduisant autant que possible la discussion selon les valeurs de paramètres, on fait apparaître par opérations élémentaires une matrice échelonnée.

$$\text{rg}(A) \underset{\substack{L_2 \leftarrow L_2 - (b+c)L_1 \\ L_3 \leftarrow L_3 - bcL_1}}{=} \text{rg} \begin{pmatrix} 1 & 1 & 1 \\ 0 & a-b & a-c \\ 0 & c(a-b) & b(a-c) \end{pmatrix} \underset{L_3 \leftarrow L_3 - cL_2}{=} \text{rg} \begin{pmatrix} 1 & 1 & 1 \\ 0 & a-b & a-c \\ 0 & 0 & (b-c)(a-c) \end{pmatrix}.$$

Si a , b et c sont deux à deux distincts, la matrice A est de rang 3. Si $a \neq b$ et $c \in \{a, b\}$, la matrice est de rang 2. Il en est de même, si $a = b$ et $c \neq a$. Enfin, si $a = b = c$, la matrice est de rang 1. En résumé

$$\text{rg}(A) = \text{Card } \{a, b, c\}.$$

(b) On exploite la deuxième ligne pour faire apparaître un zéro sur la troisième ligne avant d'employer la première ligne pour opérer sur la seconde et de poursuivre la transformation

$$\text{rg}(B) \underset{\substack{L_3 \leftarrow L_3 - a^2L_2 \\ L_2 \leftarrow L_2 - aL_1}}{=} \text{rg} \begin{pmatrix} 1 & 1 & 1 \\ 0 & b-a & c-a \\ 0 & b(b^2-a^2) & c(c^2-a^2) \end{pmatrix} \underset{L_3 \leftarrow L_3 - b(b+a)L_2}{=} \text{rg} \begin{pmatrix} 1 & 1 & 1 \\ 0 & b-a & c-a \\ 0 & 0 & x \end{pmatrix}$$

avec $x = (c-a)(c-b)(a+b+c)$.

Si a , b et c sont deux à deux distincts et si $a+b+c \neq 0$, la matrice B est de rang 3. Si a , b et c sont deux à deux distincts et si $a+b+c = 0$, la matrice B est de rang 2. Elle est aussi de rang 2, si exactement deux éléments parmi a, b, c sont égaux. Enfin, la matrice B est de rang 1 lorsque $a = b = c$.

Exercice 25 **

Soit $A \in \mathcal{M}_n(\mathbb{K})$ une matrice de rang 1.

- (a) Etablir l'existence de deux colonnes $X, Y \in \mathcal{M}_{n,1}(\mathbb{K})$ vérifiant $A = Y^t X$.
- (b) En déduire l'existence de $\lambda \in \mathbb{K}$ tel que $A^2 = \lambda A$ et vérifier $\lambda = \text{tr}(A)$.

Solution

(a) Avant de résoudre la question posée, étudions à quoi ressemble une matrice $Y^t X$ avec X et Y colonnes de hauteur n . En introduisant x_1, \dots, x_n les coefficients de X et y_1, \dots, y_n ceux de Y

$$Y^t X = \begin{pmatrix} x_1 y_1 & \cdots & x_j y_1 & \cdots & x_n y_1 \\ \vdots & & \vdots & & \vdots \\ x_1 y_n & \cdots & x_j y_n & \cdots & x_n y_n \end{pmatrix}.$$

On voit que les colonnes d'une telle matrice sont colinéaires à la colonne Y , la colonne X servant à définir les coefficients de colinéarité.

méthode

|| Le rang d'une matrice est le rang de la famille de ses colonnes.

Les colonnes C_1, \dots, C_n de la matrice A définissent une famille de rang 1, il existe donc parmi celles-ci une colonne C_k telle que toutes les colonnes de A soit colinéaires à C_k . Pour tout $j \in [1; n]$, on peut écrire $C_j = \lambda_j C_k$ avec $\lambda_j \in \mathbb{K}$ (et en particulier $\lambda_k = 1$). En posant $Y = C_k$ et X la colonne des coefficients $\lambda_1, \dots, \lambda_n$, on constate¹ $A = Y^t X$.

(b) Par l'écriture qui précède, on a $A^2 = Y^t X Y^t X$. Or le produit $Y^t X$ est le produit d'une ligne par une colonne. Ceci détermine une matrice ne comportant qu'un seul coefficient. En notant λ celui-ci, on peut écrire $A^2 = Y \times (\lambda) \times X = \lambda Y^t X = \lambda A$.

Enfin, λ peut se comprendre comme la trace de la matrice $Y^t X$, c'est donc aussi la trace de $Y^t X$ (Th. 16 p. 319), c'est-à-dire la trace de A .

Exercice 26 **

Soit $A \in \mathcal{M}_{3,2}(\mathbb{R})$ et $B \in \mathcal{M}_{2,3}(\mathbb{R})$ deux matrices vérifiant

$$AB = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Déterminer les rangs de A et B et calculer BA .

Solution

Le rang d'une matrice est inférieur au nombre de lignes et au nombre de colonnes qui la constituent. On en déduit $\text{rg}(A) \leq 2$ et $\text{rg}(B) \leq 2$. De plus, $\text{rg}(AB) = 2$ et le rang d'un produit est inférieur aux rangs de ses facteurs. On en déduit $\text{rg}(A) \geq 2$ et $\text{rg}(B) \geq 2$.

Finalement, les matrices A et B sont de rang 2.

méthode

|| On remarque $ABA = AB$ et l'on simplifie cette relation.

1. On vient ici de produire un couple (X, Y) solution. Il n'y a pas unicité de cette solution : on peut montrer que les couples $(\alpha X, \frac{1}{\alpha} Y)$ avec $\alpha \neq 0$ déterminent les autres solutions.

La matrice AB est la matrice d'un projecteur et donc $(AB)^2 = AB$ ce qui donne $ABAB = AB$. En réorganisant les membres, on obtient $A(BA - I_2)B = O_3$ et donc¹ $\text{Im}((BA - I_2)B) \subset \text{Ker}(A)$. Cependant, la matrice A comporte deux colonnes et est de rang 2, son noyau est donc réduit au vecteur nul et par conséquent $(BA - I_2)B = O_{2,3}$. Ceci entraîne $\text{Im}(B) \subset \text{Ker}(BA - I_2)$. Cependant, la matrice B comporte deux lignes et est de rang 2, son image réunit donc tous les vecteurs de \mathbb{R}^2 et par conséquent $BA - I_2 = O_2$.

Finalement, $BA = I_2$.

Exercice 27 ***

Soit $M \in \mathcal{M}_n(\mathbb{K})$ une matrice de rang r que l'on suppose pouvoir écrire par blocs

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \quad \text{avec} \quad A \in \text{GL}_r(\mathbb{K}).$$

(a) Montrer que, pour tout $X \in \mathcal{M}_{n-r,1}(\mathbb{K})$, il existe $Y \in \mathcal{M}_{r,1}(\mathbb{K})$ telle que

$$M \begin{pmatrix} 0 \\ X \end{pmatrix} = M \begin{pmatrix} Y \\ 0 \end{pmatrix}$$

où les 0 désignent des colonnes nulles de tailles appropriées.

(b) En déduire que $D = CA^{-1}B$.

Solution

(a) méthode

On étudie les valeurs prises par l'application qui à Y associe $M \begin{pmatrix} Y \\ 0 \end{pmatrix}$.

Introduisons l'image de la matrice M comprise comme un espace de colonnes :

$$\text{Im}(M) = \{MZ \mid Z \in \mathcal{M}_{n,1}(\mathbb{K})\}.$$

C'est un espace de dimension $\text{rg}(M) = r$. Considérons ensuite l'application linéaire φ qui à $Y \in \mathcal{M}_{r,1}(\mathbb{K})$ associe

$$M \begin{pmatrix} Y \\ 0 \end{pmatrix} = \begin{pmatrix} AY \\ CY \end{pmatrix}.$$

Les valeurs prises par φ appartiennent à l'image de M : $\text{Im}(\varphi) \subset \text{Im}(M)$. Or l'application linéaire φ est injective car A est inversible :

$$\begin{pmatrix} AY \\ CY \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies \begin{aligned} AY &= 0 \\ CY &= 0 \end{aligned} \implies Y = 0.$$

Le rang de φ est donc égal à la dimension r de son espace de départ $\mathcal{M}_{r,1}(\mathbb{K})$. Par inclusion et égalité des dimensions, on peut affirmer $\text{Im}(\varphi) = \text{Im}(M)$.

1. Voir sujet 9 p. 330.

On peut alors conclure : pour tout $X \in \mathcal{M}_{n-r,1}(\mathbb{K})$, le produit $M(X)$ détermine un élément de l'image de M et il existe donc une colonne $Y \in \mathcal{M}_{r,1}(\mathbb{K})$ telle que

$$M \begin{pmatrix} 0 \\ X \end{pmatrix} = \varphi(Y) = M \begin{pmatrix} Y \\ 0 \end{pmatrix}.$$

(b) L'égalité matricielle précédente se relit

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} 0 \\ X \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} Y \\ 0 \end{pmatrix} \quad \text{donc} \quad \begin{pmatrix} BX \\ DX \end{pmatrix} = \begin{pmatrix} AY \\ CY \end{pmatrix}.$$

On en déduit $Y = A^{-1}BX$ puis $DX = CA^{-1}BX$. Cette dernière égalité étant vraie pour toute colonne $X \in \mathcal{M}_{n-r,1}(\mathbb{K})$, on peut affirmer que les matrices D et $CA^{-1}B$ sont égales¹ : elles définissent chacune la même application linéaire canoniquement associée.

9.7 Exercices d'approfondissement

Exercice 28 *

Soit a_0, \dots, a_{2n} des points du plan complexe.

Montrer qu'il existe d'uniques points z_0, \dots, z_{2n} tels que,

z_i est le milieu de $[z_i ; z_{i+1}]$ pour tout $i \in [0 ; 2n]$

(en convenant de poser $z_{2n+1} = z_0$).

Solution

Le milieu du segment $[z_i ; z_{i+1}]$ est $\frac{1}{2}(z_i + z_{i+1})$. Le problème étudié revient alors à résoudre le système

$$\left\{ \begin{array}{l} z_0 + z_1 = 2a_0 \\ \vdots \\ z_{2n-1} + z_{2n} = 2a_{2n-1} \\ z_0 + z_{2n} = 2a_{2n}. \end{array} \right.$$

méthode

On vérifie que le système est de Cramer en montrant l'inversibilité de sa matrice.

1. Même pour $X \neq 0$, on ne peut pas simplifier la relation $DX = CA^{-1}BX$ par X car X est une colonne ! Cependant, en faisant varier X , on remarque que les applications $x \mapsto Dx$ et $x \mapsto CA^{-1}Bx$ sont égales sur \mathbb{K}^n . On peut aussi considérer X égale à une colonne élémentaire arbitraire et vérifier que les matrices D et $CA^{-1}B$ sont égales colonne par colonne.

La matrice exprimant le système étudié est

$$A = \begin{pmatrix} 1 & 1 & (0) \\ (0) & \ddots & \vdots \\ 1 & (0) & 1 \end{pmatrix} \in \mathcal{M}_{2n+1}(\mathbb{R}).$$

Par l'opération $L_{2n+1} \leftarrow L_{2n+1} - L_1 + L_2 - \cdots - L_{2n-1} + L_{2n}$, on obtient

$$\text{rg} \begin{pmatrix} 1 & 1 & (0) \\ (0) & \ddots & \vdots \\ 1 & (0) & 1 \end{pmatrix} = \text{rg} \begin{pmatrix} 1 & 1 & (0) \\ (0) & \ddots & \vdots \\ 0 & 1 & 2 \end{pmatrix} = 2n+1.$$

La matrice A est donc inversible et le système étudié est de Cramer : il possède une solution unique¹

Exercice 29 *

Établir que $\text{Vect}\{AB - BA \mid A, B \in \mathcal{M}_n(\mathbb{R})\}$ est un hyperplan de $\mathcal{M}_n(\mathbb{R})$.

Solution

méthode

|| On compare l'espace considéré et le noyau de la trace.

Pour tous A et $B \in \mathcal{M}_n(\mathbb{R})$, on a $\text{tr}(AB - BA) = \text{tr}(AB) - \text{tr}(BA) = 0$ et donc $AB - BA$ appartient au sous-espace noyau de la trace. On en déduit

$$\text{Vect}\{AB - BA \mid A, B \in \mathcal{M}_n(\mathbb{R})\} \subset \text{Ker}(\text{tr}).$$

Or la trace est une forme linéaire non nulle sur $\mathcal{M}_n(\mathbb{R})$ et son noyau est un hyperplan. Par conséquent,

$$\dim \text{Vect}\{AB - BA \mid A, B \in \mathcal{M}_n(\mathbb{R})\} \leq n^2 - 1.$$

On détermine ensuite $n^2 - 1$ éléments linéairement indépendants appartenant à l'espace $H = \text{Vect}\{AB - BA \mid A, B \in \mathcal{M}_n(\mathbb{R})\}$.

Pour $i, j \in [1; n]$ distincts et $A = E_{i,j}$, $B = E_{j,i}$, on a $AB - BA = E_{i,j}$. Ainsi, les matrices élémentaires $E_{i,j}$ avec $i \neq j$ appartiennent à H .

Pour $i \in [1; n-1]$ et $A = E_{i,n}$, $B = E_{n,i}$, on a $AB - BA = E_{i,i} - E_{n,n} = F_i$. Les matrices F_i appartiennent aussi à H .

1. Avec un nombre pair de points a_0, \dots, a_{2n+1} le problème est plus complexe et nécessite la résolution du système : il ne possède des solutions que si $a_0 - a_1 + \cdots + a_{2n} - a_{2n+1} = 0$ auquel cas il en possède une infinité.

Enfin, la famille formée des matrices $E_{i,j}$ (pour $i, j \in [1; n]$ distincts) et des matrices F_i (pour $i \in [1; n-1]$) est libre¹ et est constituée de $n^2 - 1$ éléments appartenant à $H = \text{Vect}\{AB - BA \mid A, B \in \mathcal{M}_n(\mathbb{R})\}$. On en déduit

$$\dim \text{Vect}\{AB - BA \mid A, B \in \mathcal{M}_n(\mathbb{R})\} \geq n^2 - 1.$$

Finalement, $\text{Vect}\{AB - BA \mid A, B \in \mathcal{M}_n(\mathbb{R})\}$ est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$ de dimension $n^2 - 1$, c'est un hyperplan.

Exercice 30 **

Soit $A \in \mathcal{M}_n(\mathbb{R})$. Résoudre l'équation $X + {}^t X = \text{tr}(X)A$ d'inconnue $X \in \mathcal{M}_n(\mathbb{R})$.

Solution

Soit $X \in \mathcal{M}_n(\mathbb{R})$. La matrice $X + {}^t X$ exprimant le premier membre est symétrique².

méthode

|| On discute la résolution selon que la matrice A est symétrique ou non.

Si la matrice A n'est pas symétrique, les solutions X de l'équation sont à rechercher parmi les matrices de trace nulle. L'équation se simplifie alors en $X + {}^t X = \mathbf{0}_n$ dont les solutions sont exactement les matrices antisymétriques³.

Supposons désormais la matrice A symétrique. Si l'équation étudiée possède une solution X , on obtient par la trace des deux membres l'équation $2 \text{tr}(X) = \text{tr}(X) \text{tr}(A)$.

méthode

|| On discute la résolution selon la valeur de $\text{tr}(A)$.

Cas : $\text{tr}(A) \neq 2$. Les solutions de l'équation sont à nouveau à rechercher parmi les matrices de trace nulle et l'on revient à la résolution précédente.

Cas : $\text{tr}(A) = 2$. Il est remarquable que la matrice A figure parmi les solutions de l'équation. Au surplus, on vérifie que l'ensemble des solutions de l'équation possède une structure d'espace vectoriel. On exploite ces propriétés pour se ramener à une matrice de trace nulle. Si $X \in \mathcal{M}_n(\mathbb{R})$ est solution, la matrice Y déterminée par

$$Y = X - \frac{1}{2} \text{tr}(X)A$$

vérifie $Y + {}^t Y = \mathbf{0}_n$: il s'agit d'une matrice antisymétrique. La matrice X peut alors s'écrire $X = \lambda A + Y$ avec $\lambda \in \mathbb{R}$ et Y matrice antisymétrique. Inversement, une telle matrice est solution.

En résumé :

Si A n'est pas symétrique ou si $\text{tr}(A) \neq 2$, les solutions sont les matrices antisymétriques.

Si A est symétrique et si $\text{tr}(A) = 2$, les solutions sont les matrices $X = \lambda A + Y$ avec λ un réel et Y une matrice antisymétrique.

1. Ces éléments constituent une base de l'hyperplan noyau de la trace sur $\mathcal{M}_n(\mathbb{R})$.

2. Voir sujet 3 p. 322.

3. Celles-ci ont aussi été présentées dans le sujet 3 p. 322.

Exercice 31 ***

Soit p_1, \dots, p_n des projecteurs d'un espace E de dimension finie.

Montrer que $p_1 + \dots + p_n$ est un projecteur si, et seulement si, $p_i \circ p_j = 0$ pour tous i et $j \in \llbracket 1 ; n \rrbracket$ tels que $i \neq j$.

Solution

Raisonnons par double implication.

(\Leftarrow) Supposons $p_i \circ p_j = 0$ pour tous $i, j \in \llbracket 1 ; n \rrbracket$ tels que $i \neq j$. On peut simplifier les termes nuls dans le calcul qui suit

$$\left(\sum_{i=1}^n p_i \right)^2 = \left(\sum_{i=1}^n p_i \right) \circ \left(\sum_{j=1}^n p_j \right) = \sum_{i=1}^n \left(\sum_{j=1}^n p_i \circ p_j \right) = \sum_{i=1}^n \underbrace{\sum_{\substack{j=1 \\ =0 \text{ si } j \neq i}}^n p_i \circ p_j}_{=p_i} = \sum_{i=1}^n p_i.$$

On en déduit que l'endomorphisme $p_1 + \dots + p_n$ est un projecteur de E .

(\Rightarrow) Supposons $p = p_1 + \dots + p_n$ projecteur de E .

méthode

En exploitant la trace, on montre que l'image de p est la somme directe des images des p_i .

On a immédiatement l'inclusion $\text{Im}(p_1 + \dots + p_n) \subset \text{Im}(p_1) + \dots + \text{Im}(p_n)$. Or la dimension de l'image d'un projecteur est égale à sa trace (Th. 17 p. 319) et donc, par linéarité de la trace,

$$\begin{aligned} \dim \text{Im}(p_1 + \dots + p_n) &= \text{tr}(p_1 + \dots + p_n) = \text{tr}(p_1) + \dots + \text{tr}(p_n) \\ &= \dim \text{Im}(p_1) + \dots + \dim \text{Im}(p_n) \end{aligned}$$

Cependant, l'inclusion précédente donne aussi

$$\begin{aligned} \dim \text{Im}(p_1 + \dots + p_n) &\leq \dim(\text{Im}(p_1) + \dots + \text{Im}(p_n)) \\ &\leq \dim \text{Im}(p_1) + \dots + \dim \text{Im}(p_n) \end{aligned}$$

On en déduit

$$\begin{aligned} \text{Im}(p_1 + \dots + p_n) &= \text{Im}(p_1) + \dots + \text{Im}(p_n) \text{ et} \\ \dim(\text{Im}(p_1) + \dots + \text{Im}(p_n)) &= \dim \text{Im}(p_1) + \dots + \dim \text{Im}(p_n). \end{aligned}$$

L'image de la projection p est donc la somme directe des espaces $\text{Im}(p_1), \dots, \text{Im}(p_n)$ (Th. 17 p. 244).

Soit $j \in \llbracket 1 ; n \rrbracket$. Pour tout $x_j \in \text{Im}(p_j)$, on a $x_j \in \text{Im}(p)$ donc $p(x_j) = x_j$. Parallèlement, on a aussi $p(x_j) = p_1(x_j) + \dots + p_n(x_j)$. On peut donc écrire

$$\underbrace{p_1(x_j)}_{\in \text{Im}(p_1)} + \dots + \underbrace{p_j(x_j)}_{\in \text{Im}(p_j)} + \dots + \underbrace{p_n(x_j)}_{\in \text{Im}(p_n)} = \underbrace{0_E}_{\in \text{Im}(p_1)} + \dots + \underbrace{x_j}_{\in \text{Im}(p_j)} + \dots + \underbrace{0_E}_{\in \text{Im}(p_n)}.$$

Par unicité de l'écriture d'un vecteur d'une somme directe, on obtient $p_i(x_j) = 0_E$ pour chaque i différent de j . Ainsi, $\text{Im}(p_j) \subset \text{Ker}(p_i)$ et donc $p_i \circ p_j = 0$ pour tous $i \neq j$.

Exercice 32 ***

Soit n un entier supérieur à 2.

(a) Soit φ une forme linéaire sur $\mathcal{M}_n(\mathbb{R})$. Montrer qu'il existe une unique matrice $A \in \mathcal{M}_n(\mathbb{R})$ telle que $\varphi(M) = \text{tr}(AM)$ pour toute matrice $M \in \mathcal{M}_n(\mathbb{R})$.

(b) En déduire que tout hyperplan de $\mathcal{M}_n(\mathbb{R})$ contient une matrice inversible.

Solution

(a) On montre l'unicité et l'existence de la matrice A en raisonnant par analyse-synthèse.

Analyse : Supposons qu'il existe une matrice $A = (a_{i,j})$ pour laquelle $\varphi(M) = \text{tr}(AM)$ pour tout $M \in \mathcal{M}_n(\mathbb{R})$. L'identité doit être en particulier valable pour les matrices élémentaires. Soit $(i,j) \in \llbracket 1; n \rrbracket^2$. On a

$$\varphi(E_{i,j}) = \text{tr}(AE_{i,j}) = a_{j,i} \quad (*)$$

car $AE_{i,j}$ est la matrice¹ dont les colonnes d'indices différents de j sont nulles et dont la colonne d'indice j correspond à celle d'indice i de la matrice A . Les identités (*) déterminent alors entièrement les coefficients de la matrice A .

Synthèse : Considérons la matrice $A = (a_{i,j})$ de coefficient général $a_{i,j} = \varphi(E_{j,i})$. Par les calculs qui précèdent, on peut affirmer $\varphi(E_{i,j}) = \text{tr}(AE_{i,j})$ pour tout $(i,j) \in \llbracket 1; n \rrbracket^2$. Les applications linéaires φ et $M \mapsto \text{tr}(AM)$ sont alors égales sur une base de $\mathcal{M}_n(\mathbb{R})$, elles sont donc égales sur tout l'espace.

(b) Soit H un hyperplan de $\mathcal{M}_n(\mathbb{R})$, c'est-à-dire le noyau d'une forme linéaire non nulle φ . Par la question qui précède, on peut introduire une matrice $A \in \mathcal{M}_n(\mathbb{R})$ permettant d'exprimer φ et alors, pour tout $M \in \mathcal{M}_n(\mathbb{R})$,

$$M \in H \iff \text{tr}(AM) = 0.$$

Il s'agit ensuite de construire une matrice M inversible telle que $\text{tr}(AM) = 0$.

méthode

|| On traite séparément le cas où la matrice A est diagonale.

Cas : La matrice A n'est pas diagonale. Il existe $i \neq j$ tels que $\text{tr}(AE_{i,j}) = a_{j,i} \neq 0$. La matrice $M = I_n + \lambda E_{i,j}$ avec $\lambda = -\text{tr}(A)/a_{j,i}$ est alors convenable. Il s'agit en effet d'une matrice inversible car triangulaire à coefficients diagonaux non nuls et, par linéarité, $\text{tr}(AM) = \text{tr}(A) - \lambda \text{tr}(AE_{i,j}) = 0$.

Cas : La matrice A est diagonale. La matrice N suivante est inversible et convient car la diagonale de AN est alors nulle :

$$N = \begin{pmatrix} 0 & \cdots & 0 & 1 \\ 1 & & & 0 \\ & \ddots & & \\ (0) & & 1 & 0 \end{pmatrix} \quad \text{et} \quad AN = \begin{pmatrix} 0 & \cdots & 0 & a_{1,1} \\ a_{2,2} & & & 0 \\ & \ddots & & \\ (0) & & a_{n,n} & 0 \end{pmatrix}$$

1. Voir sujet 12 p. 332.

CHAPITRE 10

Déterminants

\mathbb{K} désigne \mathbb{R} ou \mathbb{C} et E un \mathbb{K} -espace vectoriel de dimension finie $n \geq 1$.

10.1 Groupe symétrique

10.1.1 Permutation de $\{1, \dots, n\}$

Définition

|| Une *permutation* de $\{1, \dots, n\}$ est une bijection de $\{1, \dots, n\}$ vers lui-même. On peut représenter une telle permutation σ à l'aide d'un tableau à deux lignes figurant la correspondance réalisée

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}.$$

Par la bijectivité de σ , les éléments $1, 2, \dots, n$ figurent une fois et une seule sur la seconde ligne.

Théorème 1

L'ensemble S_n des permutations de $\{1, \dots, n\}$ est un groupe pour la loi \circ de composition des applications. Le neutre de ce groupe est la permutation identité Id .

Ce groupe possède $n!$ éléments et n'est pas commutatif dès que $n \geq 3$.

Définition

|| Le groupe S_n est appelé *groupe symétrique* d'ordre n .

10.1.2 Cycles

Soit p un entier au moins égal à 2 et a_1, \dots, a_p des éléments deux à deux distincts de $\{1, \dots, n\}$.

Sur $\{1, \dots, n\}$, on définit une permutation c en posant $c(a_i) = a_{i+1}$ pour $i \in [1; p-1]$, $c(a_p) = a_1$ et $c(x) = x$ pour tout x de $\{1, \dots, n\}$ différent des a_i :

$$\begin{array}{cccccc} a_1 & a_2 & \dots & a_{p-1} & a_p \\ \curvearrowright & \curvearrowright & & \curvearrowright & \end{array}$$

Définition

On dit que c est un *cycle de longueur p* (ou encore un *p -cycle*). On note

$$c = (a_1 \ a_2 \ \dots \ a_p).$$

L'ensemble $\{a_1, a_2, \dots, a_n\}$ constitue le *support* du cycle c .

Un cycle c de longueur p vérifie¹ $c^p = \text{Id}$. L'inverse d'un cycle est encore un cycle :

$$(a_1 \ a_2 \ \dots \ a_p)^{-1} = (a_p \ a_{p-1} \ \dots \ a_1).$$

Deux cycles de supports disjoints commutent :

$$\{a_1, \dots, a_p\} \cap \{b_1, \dots, b_q\} = \emptyset \implies (a_1 \ \dots \ a_p) \circ (b_1 \ \dots \ b_q) = (b_1 \ \dots \ b_q) \circ (a_1 \ \dots \ a_p).$$

Théorème 2

Toute permutation σ de $\{1, \dots, n\}$ peut s'écrire comme un produit de cycles à supports disjoints. De plus, cette décomposition est unique à l'ordre près des facteurs.

10.1.3 Transpositions

Définition

On appelle *transposition* tout cycle de longueur 2.

Pour $i, j \in \{1, \dots, n\}$ distincts, une transposition $\tau = (i \ j)$ a pour seul effet d'échanger les deux éléments i et j .

Une transposition τ vérifie $\tau^2 = \text{Id}$ et se confond avec son inverse.

Théorème 3

Tout cycle de longueur p peut s'écrire comme un produit de $p-1$ transpositions :

$$(a_1 \ a_2 \ \dots \ a_p) = (a_1 \ a_2) \circ (a_2 \ a_3) \circ \dots \circ (a_{p-1} \ a_p).$$

En conséquence, toute permutation de $\{1, \dots, n\}$ peut se décomposer en un produit de transpositions.

1. La puissance est à comprendre comme un itéré de composition $c^p = c \circ c \circ \dots \circ c$ (p facteurs).

10.1.4 Signature

Il existe une application notée ε de S_n vers $\{1, -1\}$ telle que $\varepsilon(\tau) = -1$ pour toute transposition τ et telle que $\varepsilon(\sigma \circ \sigma') = \varepsilon(\sigma)\varepsilon(\sigma')$ pour toutes permutations σ et σ' .

Définition

|| L'application ε est appelée *signature*.

La signature d'un cycle de longueur p vaut $(-1)^{p-1}$.

10.2 Déterminants

10.2.1 Applications multilinéaires

Soit E_1, \dots, E_n et E' des \mathbb{K} -espaces vectoriels.

Définition

On dit qu'une application

$$\left\{ \begin{array}{l} E_1 \times \cdots \times E_n \rightarrow E' \\ (x_1, \dots, x_n) \mapsto \varphi(x_1, \dots, x_n) \end{array} \right.$$

est *multilinéaire* lorsque celle-ci est linéaire en chacune de ses variables. Ceci signifie que, pour tout $i \in \{1, \dots, n\}$, lorsque l'on fixe x_j dans E_j pour chaque $j \neq i$, l'application

$$x_i \mapsto \varphi(x_1, \dots, x_i, \dots, x_n)$$

est linéaire.

Lorsque $n = 2$, on parle communément d'*application bilinéaire*.

Une application multilinéaire φ s'annule en (x_1, \dots, x_n) dès qu'il figure un vecteur nul dans cette famille.

L'ensemble des applications multilinéaires de $E_1 \times \cdots \times E_n$ vers E' est un espace vectoriel car sous-espace vectoriel de l'espace des fonctions de $E_1 \times \cdots \times E_n$ vers E' .

10.2.2 Formes n -linéaires alternées

Définition

|| On appelle *forme n -linéaire* sur E toute application multilinéaire de E^n vers \mathbb{K} .

Définition

|| Une forme n -linéaire sur E est dite *alternée* lorsque celle-ci s'annule sur toute famille comportant deux fois le même vecteur.

Théorème 4

Une forme n -linéaire alternée s'annule sur les familles liées.

Une forme n -linéaire alternée φ est *antisymétrique*¹ dans le sens où, pour toute famille (x_1, \dots, x_n) de vecteurs de E et pour toute transposition $\tau = (i \ j)$ de $\{1, \dots, n\}$,

$$\varphi(x_{\tau(1)}, \dots, x_{\tau(n)}) = -\varphi(x_1, \dots, x_n).$$

On en déduit :

Théorème 5

Si φ est une forme n -linéaire alternée alors, pour toute famille (x_1, \dots, x_n) de vecteurs de E et toute permutation $\sigma \in S_n$,

$$\varphi(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \varepsilon(\sigma) \varphi(x_1, \dots, x_n).$$

10.2.3 Déterminant d'une famille de vecteurs dans une base

Soit $e = (e_1, \dots, e_n)$ une base de l'espace E .

Théorème 6

Il existe une unique forme n -linéaire alternée φ vérifiant $\varphi(e_1, \dots, e_n) = 1$.

Celle-ci est notée \det_e et si (x_1, \dots, x_n) est une famille de vecteurs de E figurée par la matrice $A = (a_{i,j})$ dans la base e , la valeur $\det_e(x_1, \dots, x_n)$ est donnée par

$$\det_e(x_1, \dots, x_n) \stackrel{\text{def}}{=} \sum_{\sigma \in S_n} \left(\varepsilon(\sigma) \prod_{i=1}^n a_{\sigma(i), i} \right).$$

Définition

|| L'application \det_e est appelée *déterminant dans la base e* .

Théorème 7

L'ensemble des formes n -linéaires alternées sur un espace de dimension n est une droite vectorielle.

Toute forme n -linéaire alternée sur E est donc colinéaire à \det_e . En particulier :

Théorème 8

Si e' est une base de E , on dispose de la formule de changement de base

$$\det_e(x_1, \dots, x_n) = \det_e e' \times \det_{e'}(x_1, \dots, x_n)$$

pour toute famille (x_1, \dots, x_n) de vecteurs de E .

1. Inversement, une forme n -linéaire vérifiant la propriété d'antisymétrie est nécessairement alternée.

Théorème 9

Une famille (x_1, \dots, x_n) de vecteurs de E est une base si, et seulement si, $\det_e(x_1, \dots, x_n) \neq 0$.

10.2.4 Déterminant d'un endomorphisme**Théorème 10**

Si u est un endomorphisme de E , il existe un unique scalaire λ tel que

$$\det_e(u(x_1), \dots, u(x_n)) = \lambda \det_e(x_1, \dots, x_n)$$

pour toute base e de E et toute famille (x_1, \dots, x_n) de vecteurs de E .

Ce scalaire est noté $\det(u)$ et est déterminé par l'égalité

$$\det(u) = \det_e(u(e_1), \dots, u(e_n)) \quad \text{pour toute base } e = (e_1, \dots, e_n) \text{ de } E.$$

Définition

Le scalaire $\det(u)$ est appelé *déterminant de l'endomorphisme u* .

Le déterminant de l'endomorphisme Id_E est égal à 1.

Théorème 11

Pour tous u et v endomorphismes de E ,

$$\det(u \circ v) = \det(u) \times \det(v).$$

De plus, un endomorphisme u est un automorphisme si, et seulement si, $\det(u) \neq 0$.

10.2.5 Déterminant d'une matrice carrée**Définition**

On appelle *déterminant d'une matrice carrée* $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{K})$ le déterminant de l'endomorphisme qui lui est canoniquement associé. Celui-ci est noté $\det(A)$ et est donné par

$$\det(A) \stackrel{\text{def}}{=} \sum_{\sigma \in S_n} \left(\varepsilon(\sigma) \prod_{i=1}^n a_{\sigma(i)i} \right).$$

Théorème 12

Pour toutes matrices A et B de $\mathcal{M}_n(\mathbb{K})$,

$$\det(AB) = \det(A) \times \det(B).$$

De plus, une matrice carrée A est inversible si, et seulement si, $\det(A) \neq 0$.

Le déterminant d'une matrice est aussi le déterminant de la famille de ses colonnes dans la base canonique : on dit que le déterminant d'une matrice est une forme n -linéaire alternée en la famille de ses colonnes.

En particulier, si l'on multiplie chaque colonne de $A \in \mathcal{M}_n(\mathbb{K})$ par un même scalaire λ , on obtient la formule¹

$$\det(\lambda A) = \lambda^n \det(A).$$

Théorème 13

Pour toute matrice A de $\mathcal{M}_n(\mathbb{K})$,

$$\det({}^t A) = \det(A).$$

Par transposition, le déterminant d'une matrice est aussi une forme n -linéaire alternée en la famille de ses lignes.

Enfin, savoir calculer le déterminant d'une matrice suffit pour calculer le déterminant d'un endomorphisme ou d'une famille de vecteurs dans une base :

Théorème 14

Si $A \in \mathcal{M}_n(\mathbb{K})$ représente une famille (x_1, \dots, x_n) de vecteurs de E dans une base e ,

$$\det_e(x_1, \dots, x_n) = \det(A).$$

Si $A \in \mathcal{M}_n(\mathbb{K})$ figure un endomorphisme u de E dans une base e ,

$$\det(u) = \det(A).$$

10.3 Calculs de déterminants

10.3.1 Premiers calculs

On peut exprimer le déterminant d'une matrice $A = (a_{i,j})$ de $\mathcal{M}_n(\mathbb{K})$ par un tableau écrit entre deux barres verticales² :

$$\begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix}_{[n]}$$

Si la matrice est de taille 1, son déterminant est simplement égal à son coefficient.

Si la matrice est de taille 2, on peut calculer son déterminant par un produit en croix :

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

1. Lorsque $n = 2$, le déterminant n'est pas linéaire.

2. L'indice $[n]$ est utile lorsque la taille du tableau est ambiguë. On peut l'omettre sinon.

Si la matrice est triangulaire, son déterminant est le produit de ses coefficients diagonaux :

$$\begin{vmatrix} a_{1,1} & & \text{(*)} \\ (0) & & a_{n,n} \end{vmatrix} = a_{1,1} \times \cdots \times a_{n,n}.$$

10.3.2 Opérations élémentaires

Théorème 15

Soit $i, j \in [1 ; n]$ distincts et $\lambda \in \mathbb{K}$. L'opération :

- $C_i \leftarrow C_i + \lambda C_j$ ne modifie pas le déterminant ;
- $C_i \leftarrow \lambda C_i$ multiplie le déterminant par λ ;
- $C_i \leftrightarrow C_j$ multiplie le déterminant par -1 ;

Les opérations sur les lignes ont des effets similaires et sont codées $L_i \leftarrow L_i + \lambda L_j$, etc.

Si l'on réordonne les colonnes (resp. les lignes) par une permutation σ , le déterminant est multiplié par $\varepsilon(\sigma)$.

10.3.3 Déterminants triangulaires par blocs

Théorème 16

Si $M \in \mathcal{M}_n(\mathbb{K})$ est une matrice triangulaire supérieure par blocs de la forme

$$M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \quad \text{avec} \quad A \in \mathcal{M}_p(\mathbb{K}) \text{ et } C \in \mathcal{M}_{n-p}(\mathbb{K})$$

alors

$$\det(M) = \det(A) \times \det(C).$$

Ce résultat se généralise aux matrices triangulaires inférieures par blocs et aux décompositions comportant plusieurs blocs diagonaux carrés.

10.3.4 Développement selon une rangée

Définition

Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ et $i, j \in [1 ; n]$. On appelle *mineur d'indice* (i, j) de A le déterminant $\Delta_{i,j}$ de la matrice obtenue en supprimant la i -ème ligne et la j -ème colonne :

$$\Delta_{i,j} \stackrel{\text{def}}{=} \begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix}_{(i-1)(n-1)}$$

On appelle *cofacteur d'indice* (i, j) de A le scalaire $A_{i,j} = (-1)^{i+j} \Delta_{i,j}$.

Théorème 17

Pour tout $i \in \llbracket 1 ; n \rrbracket$, on peut développer le déterminant de A selon la i -ème ligne par la formule :

$$\det(A) = \sum_{j=1}^n a_{i,j} A_{i,j} = \sum_{j=1}^n (-1)^{i+j} a_{i,j} \Delta_{i,j}.$$

Pour tout $j \in \llbracket 1 ; n \rrbracket$, on peut développer le déterminant de A selon la j -ème colonne par la formule :

$$\det(A) = \sum_{i=1}^n a_{i,j} A_{i,j} = \sum_{i=1}^n (-1)^{i+j} a_{i,j} \Delta_{i,j}.$$

10.3.5 Déterminant de Vandermonde

Pour $(a_1, \dots, a_n) \in \mathbb{K}^n$, on définit le *déterminant de Vandermonde*

$$V_n(a_1, \dots, a_n) \stackrel{\text{déf}}{=} \begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix}$$

Théorème 18

$$V_n(a_1, \dots, a_n) = \prod_{1 \leq i < j \leq n} (a_j - a_i).$$

10.4 Applications**10.4.1 Comatrice**

On suppose $n \geq 2$.

Définition

On appelle *comatrice* de $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ la matrice $\text{Com}(A) = (A_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ dont les coefficients sont les cofacteurs de A .

Théorème 19

Pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$, on a

$${}^t(\text{Com}(A)) A = A {}^t(\text{Com}(A)) = \det(A) I_n.$$

Si A est inversible

$$A^{-1} = \frac{1}{\det(A)} {}^t(\text{Com}(A)).$$

10.4.2 Orientation d'un espace réel

Soit E un espace vectoriel réel de dimension $n > 1$.

Définition

On dit que deux bases e et e' ont *même orientation* si $\det_e e' > 0$. Sinon, ces bases sont dites d'orientations opposées.

On définit ainsi une relation d'équivalence sur l'ensemble des bases de E .

Définition

Orienter un espace réel, c'est choisir arbitrairement une base de celui-ci et qualifier celle-ci de *base orientée de référence*. Toute base de l'espace de même orientation que la base orientée de référence est dite *directe*, les autres sont dites *indirectes*.

Selon le choix de la base orientée de référence, il y a deux orientations possibles sur E . Lorsque l'on oriente \mathbb{R}^n par le choix de la base canonique, on dit que l'espace \mathbb{R}^n est muni de son *orientation canonique*.

Lorsque le déterminant d'un endomorphisme u est strictement positif (resp. strictement négatif), il transforme une base e en une base $u(e)$ de même orientation (resp. d'orientation opposée). En effet,

$$\det_e(u(e_1), \dots, u(e_n)) = \det(u).$$

10.5 Exercices d'apprentissage

Sauf précision contraire, n désigne un entier naturel non nul dans l'ensemble des sujets qui suivent.

10.5.1 Permutation de $\{1, \dots, n\}$

Exercice 1

Décomposer les permutations suivantes en produit de cycles à supports disjoints et en déduire leur signature

$$(a) \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 5 & 3 & 7 & 8 & 2 & 4 & 1 & 6 \end{pmatrix} \quad (b) \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 8 & 6 & 3 & 7 & 4 & 5 & 2 & 1 \end{pmatrix}$$

$$(c) \sigma = \begin{pmatrix} 1 & 2 & & n-1 & n \\ n & n-1 & \ddots & 2 & 1 \end{pmatrix}.$$

Solution

Si l'on applique plusieurs fois une permutation σ à partir d'un élément i , les éléments successifs $i, \sigma(i), \sigma^2(i), \dots$ reviennent à l'élément initial : ces éléments constituent l'*orbite* de i sous l'action de σ .

(a) **méthode**

|| La détermination des différentes orbites de σ fournit sa décomposition en cycles à supports disjoints.

L'orbite engendrée par 1 définit la succession : 1, 5, 2, 3, 7.

L'orbite engendrée par 4 définit la succession : 4, 8, 6.

Tous les éléments de $\{1, 2, \dots, 8\}$ ayant été parcourus, la¹ décomposition de σ en cycles à supports disjoints s'écrit

$$\sigma = (1 \ 5 \ 2 \ 3 \ 7) \circ (4 \ 8 \ 6).$$

méthode

|| La signature d'un produit est le produit des signatures et la signature d'un cycle de longueur p vaut $(-1)^{p-1}$.

La signature de σ est donc

$$\varepsilon(\sigma) = \varepsilon((1 \ 5 \ 2 \ 3 \ 7)) \times \varepsilon((4 \ 8 \ 6)) = (-1)^4 \times (-1)^2 = 1.$$

(b) L'orbite engendrée par 1 définit la succession : 1, 8.

L'orbite engendrée par 2 définit la succession : 2, 6, 5, 4, 7.

L'orbite engendrée par 3 se réduit à l'élément 3 (c'est un point fixe).

La décomposition de σ en cycles à supports disjoints s'écrit

$$\sigma = (1 \ 8) \circ (2 \ 6 \ 5 \ 4 \ 7)$$

et la signature de σ vaut $\varepsilon(\sigma) = (-1) \times (-1)^4 = -1$.

(c) Les orbites de σ sont composées par les couples $(k, n+1-k)$ avec $0 \leq k \leq n/2$.

On poursuit en discutant selon la parité de n .

Cas : n pair. On écrit $n = 2p$ et la permutation σ est la composée de p transpositions :

$$\sigma = (1 \ n) \circ (2 \ n-1) \circ \cdots \circ (p \ p+1).$$

Cas : n impair. On écrit $n = 2p+1$ et la permutation σ est encore la composée de p transpositions :

$$\sigma = (1 \ n) \circ (2 \ n-1) \circ \cdots \circ (p \ p+2).$$

Dans les deux cas, la signature de σ vaut² donc $(-1)^p$ avec $p = \lfloor n/2 \rfloor$.

1. Cette décomposition est unique à l'ordre près des facteurs (Th. 2 p. 354) mais aussi à la description des cycles près : les cycles $(4 \ 8 \ 6)$, $(8 \ 6 \ 4)$ et $(6 \ 4 \ 8)$ sont identiques.

2. On peut aussi écrire $\varepsilon(\sigma) = (-1)^{n(n-1)/2}$.

10.5.2 Calculs de déterminants

Exercice 2

Calculer les déterminants d'ordre n suivants :

$$(a) A_n = \begin{vmatrix} 1 & n & n & \cdots & n \\ n & 2 & n & & n \\ n & n & 3 & \cdots & n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ n & n & n & \cdots & n \end{vmatrix}_{[n]}$$

$$(b) B_n = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2 & \cdots & 2 \\ 1 & 2 & 3 & \cdots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & n \end{vmatrix}_{[n]}$$

$$(c) C_n = \begin{vmatrix} 1 & 0 & 1 & \cdots & 1 \\ 1 & 1 & 0 & & \vdots \\ \vdots & \vdots & \vdots & \ddots & 1 \\ 1 & & 1 & 0 & \vdots \\ 0 & 1 & \cdots & 1 & 1 \end{vmatrix}_{[n]}$$

$$(d) D_n = \begin{vmatrix} 1 & 1 & 0 & \cdots & 0 \\ 0 & 1 & 1 & & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & & & 1 & 1 \\ 1 & 0 & \cdots & 0 & 1 \end{vmatrix}_{[n]}$$

Solution

(a) méthode

Par opérations élémentaires, on transforme le déterminant étudié en le déterminant d'une matrice triangulaire.

On retranche la dernière colonne à chacune des précédentes

$$\left| \begin{array}{ccccc} 1 & n & n & \cdots & n \\ n & 2 & n & & n \\ n & n & 3 & \cdots & n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ n & n & n & \cdots & n \end{array} \right| \xrightarrow[C_1 \leftrightarrow C_2 \leftrightarrow C_n]{} \left| \begin{array}{ccccc} 1-n & & (0) & n \\ 0 & 2-n & & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & n \\ (0) & & -1 & n \\ 0 & n & & \vdots & \vdots \end{array} \right|$$

Le déterminant d'une matrice triangulaire est le produit des coefficients diagonaux et donc

$$A_n = \prod_{k=1}^{n-1} (k-n) \times n = (-1)^{n-1} n!$$

(b) On fait apparaître un déterminant triangulaire en retranchant à chaque ligne la précédente. Cependant, chaque opération élémentaire transforme la matrice sur laquelle agit l'opération suivante.

méthode

Il faut être attentif à l'ordre dans lequel les opérations élémentaires sont effectuées.

On retranche à chaque ligne la précédente en commençant par la dernière :

$$\left| \begin{array}{cccc|c} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2 & \cdots & 2 \\ 1 & 2 & 3 & \cdots & 3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 2 & 3 & \cdots & n \end{array} \right| \xrightarrow{L_n \leftarrow L_n - L_{n-1}} \left| \begin{array}{cccc|c} 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & \cdots & 2 & 2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 2 & \cdots & n & n \\ 0 & 0 & \cdots & 0 & 1 \end{array} \right| = \cdots = \left| \begin{array}{cccc|c} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & \ddots & \vdots \\ 0 & \cdots & 1 & 1 \\ 0 & \cdots & 0 & 1 & 1 \end{array} \right|$$

Par le produit des coefficients diagonaux, on obtient $B_n = 1$.

(c) méthode

|| On somme toutes les colonnes sur la première afin de faire apparaître une colonne constante.

On réalise l'opération $C_1 \leftarrow C_1 + C_2 + \cdots + C_n$ puis on retranche à chaque ligne la précédente en commençant par la dernière

$$\left| \begin{array}{cccc|c} 1 & 0 & 1 & \cdots & 1 \\ 1 & 1 & 0 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & & 1 & 0 & 0 \\ 0 & 1 & \cdots & 1 & 1 \end{array} \right|_{[n]} = \left| \begin{array}{cccc|c} n-1 & 0 & 1 & \cdots & 1 \\ n-1 & 1 & \cdots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ n-1 & 1 & \cdots & 1 & 0 \\ n-1 & 1 & \cdots & 1 & 1 \end{array} \right|_{[n]} = \left| \begin{array}{cccc|c} n-1 & 0 & 1 & \cdots & 1 \\ 0 & 1 & -1 & \cdots & (0) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \\ 0 & 0 & 0 & \cdots & 1 \end{array} \right|_{[n]}$$

On en déduit $C_n = (n-1)$.

(d) Par les opérations $L_n \leftarrow L_n - L_1$, $L_n \leftarrow L_n + L_2$, etc. il est possible de transformer le déterminant définissant D_n en celui d'une matrice triangulaire, cependant :

méthode

|| Il est quelquefois plus commode de développer selon une rangée (Th. 17 p. 360).

On développe le déterminant définissant D_n selon la dernière ligne

$$D_n = (-1)^{n+1} \times 1 \times \left| \begin{array}{cc|c} 1 & & (0) \\ 1 & \cdots & (0) \\ (0) & 1 & 1 \end{array} \right|_{[n-1]} + (-1)^{2n} \times 1 \times \left| \begin{array}{cc|c} 1 & & (0) \\ (0) & \cdots & (0) \\ (0) & 1 & 1 \end{array} \right|_{[n-1]}$$

Les deux mineurs apparus sont triangulaires et l'on peut conclure $D_n = 1 + (-1)^{n+1}$

Exercice 3

(a) On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{R})$ est *antisymétrique* lorsque ${}^t A = -A$. Montrer qu'une matrice antisymétrique de taille impaire n'est pas inversible.

(b) On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{C})$ est *transconjuguée* lorsque¹ ${}^t A = \bar{A}$. Que peut-on dire du déterminant de A ?

Solution

(a) Soit $A \in \mathcal{M}_n(\mathbb{R})$.

méthode

|| Une matrice est inversible si, et seulement si, son déterminant est non nul. D'une part, le déterminant d'une matrice est égal au déterminant de sa transposée

$$\det({}^t A) = \det(A).$$

D'autre part, le déterminant est multilinéaire en la famille des colonnes²

$$\det(-A) = (-1)^n \det(A).$$

Si la matrice A est antisymétrique et de taille impaire, on obtient $\det(A) = -\det(A)$ et donc $\det(A) = 0$. La matrice A n'est pas inversible.

(b) méthode

|| **Le déterminant de la matrice conjuguée est le conjugué du déterminant.**

La formule définissant le déterminant d'une matrice $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{C})$ est

$$\det(A) = \sum_{\sigma \in S_n} \left(\varepsilon(\sigma) \prod_{i=1}^n a_{\sigma(i), i} \right).$$

Par conjugaison d'une somme et de produits

$$\overline{\det(A)} = \sum_{\sigma \in S_n} \left(\varepsilon(\sigma) \prod_{i=1}^n \overline{a_{\sigma(i), i}} \right) = \det(\bar{A}).$$

Si la matrice A est transconjuguée, on obtient

$$\det(A) = \det({}^t A) = \det(\bar{A}) = \overline{\det(A)}$$

et le déterminant de A est réel.

1. \bar{A} désigne la *matrice conjuguée* de A : ses coefficients sont les conjugués respectifs de ceux de A .
 2. On peut aussi écrire $\det(-A) = \det(-I_n) \det(A)$ avec $\det(-I_n) = (-1)^n$.

10.6 Exercices d'entraînement

10.6.1 Calculs de déterminants

Exercice 4 *

Exprimer sous forme factorisée

$$\begin{vmatrix} a & b & c \\ a^2 & b^2 & c^2 \\ a^4 & b^4 & c^4 \end{vmatrix} \quad \text{puis} \quad \begin{vmatrix} a+b & b+c & c+a \\ a^2+b^2 & b^2+c^2 & c^2+a^2 \\ a^4+b^4 & b^4+c^4 & c^4+a^4 \end{vmatrix}$$

Solution

On factorise a , b et c en colonne

$$\begin{vmatrix} a & b & c \\ a^2 & b^2 & c^2 \\ a^4 & b^4 & c^4 \end{vmatrix} = abc \begin{vmatrix} 1 & 1 & 1 \\ a & b & c \\ a^3 & b^3 & c^3 \end{vmatrix}$$

On fait apparaître des zéros sur la première colonne par $L_3 \leftarrow L_3 - a^2L_2$ et $L_2 \leftarrow L_2 - aL_1$

$$\begin{vmatrix} a & b & c \\ a^2 & b^2 & c^2 \\ a^4 & b^4 & c^4 \end{vmatrix} = abc \begin{vmatrix} 1 & 1 & 1 \\ 0 & b-a & c-a \\ 0 & b(b^2-a^2) & c(c^2-a^2) \end{vmatrix}$$

On développe ensuite selon la première colonne puis on factorise $b-a$ et $c-a$ sur chacune des deux colonnes

$$\begin{aligned} \begin{vmatrix} a & b & c \\ a^2 & b^2 & c^2 \\ a^4 & b^4 & c^4 \end{vmatrix} &= abc \times (-1)^2 \times 1 \times \begin{vmatrix} b-a & c-a \\ b(b^2-a^2) & c(c^2-a^2) \end{vmatrix} \\ &= abc(b-a)(c-a) \begin{vmatrix} 1 & 1 \\ b(b+a) & c(c+a) \end{vmatrix} \\ &= abc(b-a)(c-a)(c(c+a) - b(b+a)) \\ &= abc(b-a)(c-a)(c-b)(a+b+c). \end{aligned}$$

méthode

Le déterminant est multilinéaire en la famille des colonnes : on peut séparer une colonne en somme de deux colonnes et exprimer le déterminant comme la somme des déterminants des deux matrices ainsi formées.

En séparant la première colonne en deux

$$\begin{vmatrix} a+b & b+c & c+a \\ a^2+b^2 & b^2+c^2 & c^2+a^2 \\ a^4+b^4 & b^4+c^4 & c^4+a^4 \end{vmatrix} = \begin{vmatrix} a & b+c & c+a \\ a^2 & b^2+c^2 & c^2+a^2 \\ a^4 & b^4+c^4 & c^4+a^4 \end{vmatrix} + \begin{vmatrix} b & b+c & c+a \\ b^2 & b^2+c^2 & c^2+a^2 \\ b^4 & b^4+c^4 & c^4+a^4 \end{vmatrix}$$

On répète cette opération avec les deuxième et troisième colonnes et l'on simplifie tous les déterminants comportant deux fois la même colonne car ils sont nuls. Il reste

$$\begin{vmatrix} a+b & b+c & c+a \\ a^2+b^2 & b^2+c^2 & c^2+a^2 \\ a^4+b^4 & b^4+c^4 & c^4+a^4 \end{vmatrix} = \begin{vmatrix} a & b & c \\ a^2 & b^2 & c^2 \\ a^4 & b^4 & c^4 \end{vmatrix} + \begin{vmatrix} b & c & a \\ b^2 & c^2 & a^2 \\ b^4 & c^4 & a^4 \end{vmatrix}$$

Enfin, par permutation des colonnes du second déterminant selon le cycle $(1 \ 2 \ 3)$ de signature $(-1)^{3-1} = 1$

$$\begin{vmatrix} a+b & b+c & c+a \\ a^2+b^2 & b^2+c^2 & c^2+a^2 \\ a^4+b^4 & b^4+c^4 & c^4+a^4 \end{vmatrix} = 2 \begin{vmatrix} a & b & c \\ a^2 & b^2 & c^2 \\ a^4 & b^4 & c^4 \end{vmatrix} = 2abc(b-a)(c-a)(c-b)(a+b+c).$$

Exercice 5 *

Soit $a, b \in \mathbb{K}$. Calculer

$$\begin{vmatrix} a & & (b) \\ & \ddots & \\ (b) & & a \end{vmatrix}_{[n]}$$

Solution

méthode

Les lignes de la matrice sont de somme constante : on ajoute à la première colonne la somme des autres.

L'opération $C_1 \leftarrow C_1 + C_2 + \dots + C_n$ conserve le déterminant

$$\begin{vmatrix} a & & (b) \\ & \ddots & \\ (b) & & a \end{vmatrix}_{[n]} = \begin{vmatrix} a+(n-1)b & b & \cdots & b \\ a+(n-1)b & a-b & & (b) \\ a+(n-1)b & (b) & \ddots & \\ & & & a \end{vmatrix}_{[n]}$$

On retranche la première ligne à chacune des suivantes afin de faire apparaître un déterminant d'une matrice triangulaire

$$\begin{vmatrix} a & & (b) \\ & \ddots & \\ (b) & & a \end{vmatrix}_{[n]} = \begin{vmatrix} a+(n-1)b & b & \cdots & b \\ 0 & a-b & & (0) \\ 0 & (0) & a-b & \\ & & & a-b \end{vmatrix}_{[n-1]} = (a+(n-1)b)(a-b)^{n-1}$$

Exercice 6 *

Calculer pour tout $n \geq 1$

$$D_n = \begin{vmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & (0) \\ 1 & (0) & 1 \end{vmatrix}_{[n]}$$

Solution**méthode**

|| On forme une relation de récurrence¹ en développant le déterminant selon une rangée.

Soit $n \geq 2$. On développe le déterminant selon la deuxième ligne

$$D_n = (-1)^3 \times 1 \times \underbrace{\begin{vmatrix} 1 & \cdots & 1 \\ & \ddots & 0 \\ 0 & \cdots & 1 \end{vmatrix}_{[n-1]}}_{D_1} + (-1)^4 \times 1 \times \underbrace{\begin{vmatrix} 1 & \cdots & 1 \\ & \ddots & 0 \\ 1 & \cdots & 1 \end{vmatrix}_{[n-1]}}_{D_{n-1}}$$

Ainsi, $D_n = D_{n-1} - 1$. Sachant $D_1 = 1$, on conclut $D_n = 2 - n$ pour tout $n \in \mathbb{N}^*$.

Exercice 7 *

Soit a_1, \dots, a_n des réels. Calculer $\det(\sin(a_i + a_j))_{1 \leq i, j \leq n}$

Solution**méthode**

|| Les colonnes sont combinaisons linéaires de deux colonnes remarquables.

On développe les sinus

$$\sin(a_i + a_j) = \sin a_i \cos a_j + \sin a_j \cos a_i.$$

La j -ème colonne C_j de la matrice $(\sin(a_i + a_j))$ peut alors se décomposer en

$$C_j = \cos(a_j)S + \sin(a_j)C \quad \text{avec} \quad S = \begin{pmatrix} \sin a_1 \\ \vdots \\ \sin a_n \end{pmatrix} \quad \text{et} \quad C = \begin{pmatrix} \cos a_1 \\ \vdots \\ \cos a_n \end{pmatrix}.$$

Les colonnes de la matrice sont donc toutes combinaisons linéaires de C et S : elles appartiennent à l'espace $\text{Vect}(C, S)$.

Cas : $n \geq 3$. Les colonnes forment une famille liée et le déterminant étudié est nul.

Cas : $n = 1$. Le déterminant vaut $\sin(2a_1)$.

Cas : $n = 2$. On obtient après calcul

$$\begin{vmatrix} \sin(2a_1) & \sin(a_1 + a_2) \\ \sin(a_1 + a_2) & \sin(2a_2) \end{vmatrix} = -\sin^2(a_1 - a_2).$$

1. On peut aussi faire apparaître un déterminant triangulaire en retranchant à la première colonne la somme des autres.

Exercice 8 **

Soit a et b deux nombres réels non tous deux nuls. Calculer pour tout $n \geq 1$

$$D_n = \begin{vmatrix} a+b & ab & (0) \\ 1 & \ddots & \ddots \\ (0) & 1 & ab \end{vmatrix}_{[n]}$$

Solution**méthode**

|| On forme une relation de récurrence linéaire double par développement.

Soit $n \in \mathbb{N}$ avec $n \geq 3$. On développe le déterminant par rapport à la première ligne

$$D_n = (a+b) \begin{vmatrix} a+b & ab & (0) \\ 1 & \ddots & \ddots \\ (0) & 1 & ab \end{vmatrix}_{[n-1]} - ab \begin{vmatrix} 1 & ab & 0 & \cdots & 0 \\ 0 & a+b & ab & \cdots & (0) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & (0) & \ddots & \ddots & ab \\ \vdots & & \ddots & 1 & a+b \end{vmatrix}_{[n-1]}$$

On développe le second déterminant par rapport à sa première colonne

$$D_n = (a+b) \underbrace{\begin{vmatrix} a+b & ab & (0) \\ 1 & \ddots & \ddots \\ (0) & 1 & ab \end{vmatrix}_{[n-1]}}_{=D_{n-1}} - ab \underbrace{\begin{vmatrix} a+b & ab & (0) \\ 1 & \ddots & \ddots \\ (0) & 1 & ab \end{vmatrix}_{[n-2]}}_{=D_{n-2}}$$

La suite $(D_n)_{n \geq 1}$ satisfait la relation de récurrence $D_n - (a+b)D_{n-1} + abD_{n-2} = 0$. C'est une relation de récurrence linéaire double¹ d'équation caractéristique $r^2 - (a+b)r + ab = 0$ de racines a et b . On poursuit la résolution en discutant selon que les racines a et b sont égales ou non.

Cas : $a = b$. Le terme général de la suite $(D_n)_{n \geq 1}$ s'exprime $D_n = (\lambda n + \mu)a^n$ avec λ et μ réels. On détermine λ et μ par les valeurs initiales $D_1 = 2a$ et $D_2 = 3a^2$. On obtient $\lambda = \mu = 1$ et l'on conclut $D_n = (n+1)a^n$ pour tout $n \geq 1$.

Cas : $a \neq b$. Le terme général de la suite s'exprime $D_n = \lambda a^n + \mu b^n$ avec λ et μ réels. Sachant $D_1 = a+b$ et $D_2 = a^2 + ab + b^2$, on obtient après résolution

$$D_n = \frac{a^{n+1} - b^{n+1}}{a - b} \quad \text{pour tout } n \geq 1.$$

1. Voir Th 19 du chapitre 11 de l'ouvrage *Exercices d'analyse MPSI*.

Exercice 9 **

Soit a, b et $\lambda_1, \lambda_2, \dots, \lambda_n$ des réels. On souhaite calculer le déterminant de la matrice

$$M_{a,b} = \begin{pmatrix} \lambda_1 & a & \cdots & a \\ b & \lambda_2 & & \vdots \\ \vdots & & \ddots & a \\ b & \cdots & b & \lambda_n \end{pmatrix} \in \mathcal{M}_n(\mathbb{R}).$$

On introduit le polynôme $P = (\lambda_1 - X) \dots (\lambda_n - X)$ et la matrice J de $\mathcal{M}_n(\mathbb{R})$ dont tous les coefficients sont égaux à 1.

- (a) Montrer que la fonction qui à un réel x associe $\det(M_{a,b} + xJ)$ est affine.
- (b) En déduire une expression de $\det(M_{a,b})$ lorsque $a \neq b$ en fonction de P .
- (c) Calculer $\det(M_{a,a})$.

Solution

(a) Soit x un réel. Étudions, sans le calculer exactement, le déterminant de $M_{a,b} + xJ$. On retranche la première colonne à chacune des suivantes afin de réduire le nombre de x figurant dans l'expression de la matrice

$$\det(M_{a,b} + xJ) = \begin{vmatrix} \lambda_1 + x & a+x & \cdots & a+\lambda_1 \\ & \lambda_2 + x & & \vdots \\ (b+x) & & \ddots & a-\lambda_1 \\ & & & \lambda_n + x \end{vmatrix}_{[n]} = \begin{vmatrix} \lambda_1 + x & a-\lambda_1 & \cdots & a-\lambda_1 \\ b+x & \lambda_2 - b & & (a-b) \\ \vdots & & \ddots & \vdots \\ b+x & (0) & & \lambda_n - b \end{vmatrix}_{[n]}$$

Lorsque l'on développe le déterminant du dernier membre selon sa première colonne, on obtient une somme d'expressions affines (les coefficients) multipliées par des constantes (les cofacteurs associés). On en déduit que $\det(M_{a,b} + xJ)$ est une expression affine de la variable x :

$$\det(M_{a,b} + xJ) = \alpha x + \beta \quad \text{avec } \alpha, \beta \in \mathbb{R}.$$

(b) méthode

¶ Une fonction affine est déterminée par la connaissance de deux valeurs.

Pour $x = -a$, la matrice $M_{a,b} - aJ$ est triangulaire inférieure tandis que pour $x = -b$, la matrice $M_{a,b} - bJ$ est triangulaire supérieure. On peut donc calculer les déterminants associés et former le système

$$\begin{cases} \beta - \alpha a = P(a) \\ \beta - \alpha b = P(b) \end{cases}$$

Après résolution

$$\det(M_{a,b}) = \beta = \frac{bP(a) - aP(b)}{b - a}$$

(c) L'étude qui précède ne peut pas être reproduite quand $a = b$.

méthode

|| On passe à la limite le résultat obtenu pour $b \neq a$.

Le déterminant de $M_{a,b}$ peut se comprendre comme une fonction polynomiale en le paramètre b , il dépend donc continûment de b ce qui permet d'écrire

$$\det(M_{a,a}) = \lim_{\substack{b \rightarrow a \\ b \neq a}} \det(M_{a,b}) = \lim_{\substack{b \rightarrow a \\ b \neq a}} \frac{bP(a) - aP(b)}{b - a}.$$

Pour résoudre la forme indéterminée de cette limite, on écrit

$$\frac{bP(a) - aP(b)}{b - a} = P(a) - a \frac{P(b) - P(a)}{b - a}.$$

On peut alors conclure

$$\det(M_{a,a}) = P(a) - aP'(a).$$

Exercice 10 **

Soit a_1, \dots, a_n et $\lambda_1, \dots, \lambda_n$ des nombres complexes. Calculer le déterminant de la matrice

$$M = \begin{pmatrix} a_1 + \lambda_1 & a_2 & \cdots & a_n \\ a_1 & a_2 + \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_n \\ a_1 & \cdots & a_{n-1} & a_n + \lambda_n \end{pmatrix}.$$

Solution

méthode

|| On décompose chaque colonne en la somme d'une colonne constante et d'une colonne élémentaire.

On écrit la première colonne

$$\begin{pmatrix} a_1 + \lambda_1 \\ a_1 \\ \vdots \\ a_1 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_1 \\ \vdots \\ a_1 \end{pmatrix} + \begin{pmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = a_1 C + \lambda_1 E_1$$

avec C la colonne uniquement composée de 1 et E_1 une colonne élémentaire. On décompose de façon semblable chaque colonne de la matrice et l'on peut exprimer le déterminant de la matrice en écrivant¹

$$\det(M) = \det(a_1 C + \lambda_1 E_1 \mid a_2 C + \lambda_2 E_2 \mid \dots \mid a_n C + \lambda_n E_n).$$

1. En second membre, on comprend le déterminant de la matrice constituée des colonnes $a_j C + \lambda_j E_j$.

On développe ce calcul par la multilinéarité du déterminant en la famille des colonnes et l'on simplifie les déterminants où figurent deux fois la même colonne car ils sont nuls.

$$\det(M) = \det(\lambda_1 E_1 | \dots | \lambda_n E_n) + \sum_{i=1}^n a_i \det(\lambda_1 E_1 | \dots | \overset{i}{\overbrace{C}} | \dots | \lambda_n E_n).$$

Le premier déterminant est simplement diagonal et le i -ème déterminant de la somme se calcule en développant selon la i -ème ligne. Finalement,

$$\det(M) = \prod_{i=1}^n \lambda_i + \sum_{i=1}^n \left(a_i \prod_{\substack{1 \leq k \leq n \\ k \neq i}} \lambda_k \right).$$

Exercice 11 **

Calculer le déterminant de la matrice $A = (a_{i,j})_{0 \leq i,j \leq n}$ de coefficient général le coefficient binomial

$$a_{i,j} = \binom{i+j}{j} \quad \text{pour tous } i, j \in \llbracket 0 ; n \rrbracket.$$

Solution

Il s'agit ici de calculer

$$\det(A) = \begin{vmatrix} \binom{0}{0} & \binom{1}{0} & \binom{2}{0} & \cdots & \binom{n}{0} \\ \binom{1}{0} & \binom{2}{1} & \binom{3}{1} & \cdots & \binom{n+1}{1} \\ \binom{2}{0} & \binom{3}{1} & \binom{4}{2} & \cdots & \binom{n+2}{2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \binom{n}{0} & \binom{n+1}{1} & \binom{n+2}{2} & \cdots & \binom{2n}{n} \end{vmatrix}_{[n+1]}$$

méthode

|| On exploite la formule de Pascal sous la forme : $\binom{i+j+1}{j} - \binom{i+j}{j} = \binom{i+j}{j-1}$.

En retranchant à chaque ligne la précédente en commençant par la dernière, on obtient

$$\det(A) = \begin{vmatrix} \binom{0}{0} & \binom{1}{0} & \binom{2}{0} & \cdots & \binom{n}{0} \\ 0 & \binom{1}{0} & \binom{2}{1} & \cdots & \binom{n}{1} \\ 0 & \binom{2}{0} & \binom{3}{1} & \cdots & \binom{n+1}{1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \binom{n}{0} & \binom{n+1}{1} & \cdots & \binom{2n-1}{n-1} \end{vmatrix}_{[n+1]}$$

On développe par rapport à la première colonne et l'on retranche à nouveau à chaque

ligne la précédente en commençant par la dernière

$$\det(A) = \underbrace{\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}}_{=1} \times \begin{vmatrix} \binom{1}{0} & \binom{2}{1} & \cdots & \binom{n}{n-1} \\ \binom{2}{0} & \binom{3}{1} & \cdots & \binom{n+1}{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{n}{0} & \binom{n+1}{1} & \cdots & \binom{2n-1}{n-1} \end{vmatrix}_{[n]} = \begin{vmatrix} \binom{1}{0} & \binom{2}{1} & \cdots & \binom{n}{n-1} \\ 0 & \binom{2}{0} & \cdots & \binom{n}{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \binom{n}{0} & \cdots & \binom{2n-2}{n-2} \end{vmatrix}_{[n]}$$

On répète ce calcul jusqu'à conclure

$$\det(A) = \begin{vmatrix} \binom{n-1}{0} & \binom{n}{1} \\ 0 & \binom{n}{0} \end{vmatrix} = 1.$$

Exercice 12 ***

Soit a_1, \dots, a_n des nombres complexes. Calculer

$$D_n = \begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-2} & a_1^n \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-2} & a_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-2} & a_n^n \end{vmatrix}_{[n]}$$

Solution

méthode

|| On introduit le déterminant de Vandermonde $V_{n+1}(a_1, \dots, a_n, x)$.

Soit $x \in \mathbb{C}$. Le déterminant de Vandermonde de la famille (a_1, \dots, a_n, x) est

$$V_{n+1}(a_1, \dots, a_n, x) = \begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} & a_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} & a_n^n \\ 1 & x & x^2 & \cdots & x^{n-1} & x^n \end{vmatrix}_{[n+1]}$$

L'application $x \mapsto V_{n+1}(a_1, \dots, a_n, x)$ est une fonction polynomiale et l'on obtient la valeur de chacun de ses coefficients en développant le déterminant selon la dernière ligne. En particulier, le coefficient de x^{n-1} est le cofacteur $(-1)^{2n+1} D_n$.

Or on sait (Th. 18 p. 360)

$$V_{n+1}(a_1, \dots, a_n, x) = \left(\prod_{1 \leq i < j \leq n} (a_j - a_i) \right) \left(\prod_{i=1}^n (x - a_i) \right).$$

Le coefficient de x^{n-1} dans le développement du second membre est

$$-\left(\sum_{i=1}^n a_i \right) \left(\prod_{1 \leq i < j \leq n} (a_j - a_i) \right).$$

Par unicité des coefficients décrivant une fonction polynomiale, on conclut

$$D_n = \left(\sum_{i=1}^n a_i \right) \left(\prod_{1 \leq i < j \leq n} (a_j - a_i) \right).$$

Exercice 13 ***

Pour $n \geq 2$, soit a_1, \dots, a_n et b_1, \dots, b_n des réels.

Calculer le déterminant de $M = ((a_i + b_j)^{n-1})_{1 \leq i, j \leq n}$.

Solution
méthode

|| On écrit M comme le produit de deux matrices.

Par la formule du binôme

$$(a_i + b_j)^{n-1} = \sum_{k=0}^{n-1} \binom{n-1}{k} a_i^k b_j^{n-1-k}$$

Afin d'interpréter ce calcul comme celui du coefficient général d'un produit de deux matrices, on réalise un glissement d'indice dans la somme

$$(a_i + b_j)^{n-1} = \sum_{k=1}^n \underbrace{\binom{n-1}{k-1}}_{=a_{i,k}} \underbrace{a_i^{k-1} b_j^{n-1-(k-1)}}_{=b_{k,j}}.$$

La matrice M est donc le produit des matrices $A = (a_{i,k})$ et $B = (b_{k,j})$ de $\mathcal{M}_n(\mathbb{R})$.
Ecrivons les tableaux associés :

$$A = \begin{pmatrix} \binom{n-1}{0} a_1^0 & \binom{n-1}{1} a_1^1 & \cdots & \binom{n-1}{n-1} a_1^{n-1} \\ \binom{n-1}{0} a_2^0 & \binom{n-1}{1} a_2^1 & \cdots & \binom{n-1}{n-1} a_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{n-1}{0} a_n^0 & \binom{n-1}{1} a_n^1 & \cdots & \binom{n-1}{n-1} a_n^{n-1} \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} b_1^{n-1} & b_2^{n-1} & \cdots & b_n^{n-1} \\ b_1^{n-2} & b_2^{n-2} & \cdots & b_n^{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ b_1^0 & b_2^0 & \cdots & b_n^0 \end{pmatrix}$$

Après factorisation des coefficients binomiaux, le déterminant de A se déduit d'un déterminant de Vandermonde :

$$\det(A) = \left(\prod_{j=0}^{n-1} \binom{n-1}{j} \right) \left(\prod_{1 \leq i < j \leq n} (a_j - a_i) \right).$$

Le déterminant de B est aussi associé à un déterminant de Vandermonde que l'on reconnaît après transposition et renversement de l'ordre des colonnes par la permutation

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ n & n-1 & \cdots & 2 & 1 \end{pmatrix} \quad \text{avec}^1 \quad \varepsilon(\sigma) = (-1)^{n(n-1)/2}.$$

¹. Voir sujet 1 p. 361.

On obtient

$$\det(B) = (-1)^{n(n-1)/2} \prod_{1 \leq i < j \leq n} (b_j - b_i) = \prod_{1 \leq i < j \leq n} (b_i - b_j).$$

Finalement,

$$\det(M) = \prod_{j=0}^{n-1} \binom{n-1}{j} \prod_{1 \leq i < j \leq n} ((a_j - a_i)(b_i - b_j)).$$

10.6.2 Déterminant d'une matrice

Exercice 14 *

Soit $A, B \in \mathcal{M}_n(\mathbb{R})$ telles que $AB = BA$. Établir $\det(A^2 + B^2) \geq 0$.

Solution

méthode

|| On introduit la matrice complexe $A + iB$ et sa matrice conjuguée.

Puisque les matrices réelles A et B commutent, on peut écrire

$$(A + iB)\overline{(A + iB)} = (A + iB)(A - iB) = A^2 - iAB + iBA + B^2 = A^2 + B^2.$$

Le déterminant d'une matrice conjuguée est le conjugué du déterminant¹ et donc

$$\det(A^2 + B^2) = \det(A + iB) \det(\overline{A + iB}) = \det(A + iB) \overline{\det(A + iB)}.$$

On peut alors conclure

$$\det(A^2 + B^2) = |\det(A + iB)|^2 \geq 0.$$

Exercice 15 **

Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{R})$ vérifiant

$$\begin{cases} |a_{1,1}| + \dots + |a_{1,n}| \leq 1 \\ |a_{n,1}| + \dots + |a_{n,n}| \leq 1. \end{cases}$$

Montrer $|\det(A)| \leq 1$.

1. Voir sujet 3 p. 365.

Solution**méthode**

|| On raisonne par récurrence¹ sur la taille $n \in \mathbb{N}^*$ de la matrice.

Le résultat est immédiat pour $n = 1$.

Supposons la propriété voulue vraie pour un rang $n \geq 1$. Soit $A = (a_{i,j}) \in \mathcal{M}_{n+1}(\mathbb{R})$ dont la somme des valeurs absolues des coefficients sur chaque ligne est inférieure à 1. En développant le déterminant de A selon la première ligne, on écrit

$$\det(A) = \sum_{j=1}^{n+1} (-1)^{1+j} a_{1,j} \Delta_{1,j}$$

avec $\Delta_{1,j}$ le mineur d'indice $(1, j)$ de la matrice A . Pour chaque $j \in [1; n]$, la somme des valeurs absolues des coefficients des lignes de la matrice définissant le mineur $\Delta_{1,j}$ est inférieure à 1. Par l'hypothèse de récurrence au rang n , on a $|\Delta_{1,j}| \leq 1$. On en déduit

$$|\det(A)| \leq \sum_{j=1}^{n+1} |a_{1,j}| |\underbrace{\Delta_{1,j}}_{\leq 1}| \leq \sum_{j=1}^{n+1} |a_{1,j}| \leq 1.$$

La récurrence est établie.

Exercice 16 ***

Soit A une matrice carrée de taille n paire dont les coefficients diagonaux sont des entiers pairs et les coefficients non diagonaux des entiers impairs. Montrer que la matrice A est inversible.

Solution

Commençons par remarquer que, si une matrice carrée A est à coefficients entiers, son déterminant est un entier en vertu de la formule

$$\det(A) = \sum_{\sigma \in \mathcal{S}_n} \left(\varepsilon(\sigma) \prod_{i=1}^n a_{\sigma(i), i} \right).$$

méthode

|| On montre que l'entier $\det(A)$ a la parité d'un déterminant simple à calculer.

Pour tous $i, j \in [1; n]$, on a $a_{i,j} \equiv 0 \pmod{2}$ si $i = j$ et $a_{i,j} \equiv 1 \pmod{2}$ sinon. Considérons alors la matrice $B = (b_{i,j})$ déterminée par $b_{i,j} = 0$ si $i = j$ et $b_{i,j} = 1$ sinon. Pour toute permutation $\sigma \in \mathcal{S}_n$, on a par produit de congruences

$$\prod_{i=1}^n a_{\sigma(i), i} \equiv \prod_{i=1}^n b_{\sigma(i), i} \pmod{2}$$

1. On peut aussi développer $\prod_{i=1}^n \sum_{j=1}^n |a_{i,j}|$ et observer que l'on y retrouve tous les termes produits de la formule définissant le déterminant A accompagnés de quelques autres, tous positifs.

puis par somme de congruences

$$\det(A) = \det(B) [2].$$

Enfin, le déterminant de B vaut¹ $(-1)^{n-1}(n-1)$ ce qui est un entier impair. On en déduit que $\det(A)$ est un entier impair et, en particulier, il n'est pas nul : la matrice A est inversible.

10.6.3 Déterminant par blocs

Exercice 17 **

Soit $A, B \in \mathcal{M}_n(\mathbb{R})$.

(a) Montrer

$$\begin{vmatrix} A & B \\ B & A \end{vmatrix} = \det(A+B)\det(A-B).$$

(b) Justifier

$$\begin{vmatrix} A & -B \\ B & A \end{vmatrix} \geq 0.$$

Solution

(a) **méthode**

Les opérations élémentaires sur les colonnes d'une matrice (resp. les lignes) permettent d'opérer sur les blocs colonnes (resp. les blocs lignes).

Par les opérations $C_j \leftarrow C_j + C_{n+j}$ pour $j \in \llbracket 1 ; n \rrbracket$, on ajoute la seconde colonne de blocs à la première sans modifier la valeur du déterminant

$$\begin{vmatrix} A & B \\ B & A \end{vmatrix} = \begin{vmatrix} B+A & B \\ A+B & A \end{vmatrix}.$$

Par les opérations $L_{n+i} \leftarrow L_{n+i} - L_i$ pour $i \in \llbracket 1 ; n \rrbracket$, on retranche la première ligne de blocs à la seconde

$$\begin{vmatrix} A & B \\ B & A \end{vmatrix} = \begin{vmatrix} A+B & B \\ O_n & A-B \end{vmatrix}.$$

Enfin, le déterminant d'une matrice triangulaire par blocs est le produit des déterminants des blocs diagonaux (Th. 16 p. 359)

$$\begin{vmatrix} A & B \\ B & A \end{vmatrix} = \det(A+B)\det(A-B).$$

1. C'est un cas particulier du sujet 5 p. 367 avec $a = 0$ et $b = 1$.

(b) méthode

|| On opère de façon analogue en faisant intervenir le nombre i complexe.

On ajoute à la première colonne de blocs i fois la seconde puis on retranche à la deuxième ligne de blocs i fois la première

$$\begin{vmatrix} A & -B \\ B & A \end{vmatrix} = \begin{vmatrix} A - iB & -B \\ B + iA & A \end{vmatrix} = \begin{vmatrix} A - iB & -B \\ 0_n & A + iB \end{vmatrix}$$

$$= \det(A + iB) \det(A - iB) = \det(A + iB) \overline{\det(A + iB)}$$

$$= |\det(A + iB)|^2 \geq 0.$$

Exercice 18 **

Soit $A, B, C, D \in \mathcal{M}_n(\mathbb{R})$ avec D inversible.

(a) Établir

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(AD - BD^{-1}CD).$$

(b) Comment simplifier cette formule si C et D commutent ou si B et D commutent ?

Solution

(a) méthode

|| On multiplie la matrice étudiée par une matrice triangulaire par blocs afin que le produit obtenu soit lui aussi triangulaire par blocs.

On a

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} I_n & 0_n \\ -D^{-1}C & I_n \end{pmatrix} = \begin{pmatrix} A - BD^{-1}C & B \\ 0_n & D \end{pmatrix}.$$

En calculant le déterminant des deux membres on obtient

$$\begin{aligned} \det \begin{pmatrix} A & B \\ C & D \end{pmatrix} &= \det(A - BD^{-1}C) \det(D) \\ &= \det(AD - BD^{-1}CD). \end{aligned} \tag{*}$$

(b) Si C et D commutent

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(AD - \underbrace{BD^{-1}CD}_{=DC}) = \det(AD - BC).$$

Si B et D commutent, on intervertit dans (*) les facteurs $\det(A - BD^{-1}C)$ et $\det(D)$

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(D) \det(A - BD^{-1}C) = \det(DA - \underbrace{DBD^{-1}C}_{=BD}) = \det(DA - BC).$$

10.6.4 Déterminant d'un endomorphisme

Exercice 19 **

Pour P polynôme de $\mathbb{R}_{2n+1}[X]$, on pose

$$f(P) = (2n+1)XP - (X^2 - 1)P'.$$

- (a) Vérifier que f définit un endomorphisme de $\mathbb{R}_{2n+1}[X]$.
- (b) Calculer le déterminant de f .

Solution

- (a) L'application f est à valeurs dans $\mathbb{R}_{2n+1}[X]$ car, par opérations sur les degrés, $\deg((2n+1)XP - (X^2 - 1)P') \leq \max((2n+1)XP, (X^2 - 1)P') \leq \deg(P) + 1$.

De plus, si $\deg(P) = 2n+1$, les termes en X^{2n+2} se simplifient lors du calcul.

Enfin, l'application f est linéaire car on vérifie sans peine

$$f(\lambda P + \mu Q) = \lambda f(P) + \mu f(Q) \quad \text{pour tous } \lambda, \mu \in \mathbb{R} \text{ et } P, Q \in \mathbb{R}_{2n+1}[X].$$

méthode

|| Le déterminant d'un endomorphisme est le déterminant d'une matrice figurant celui-ci.

Introduisons¹ $\mathcal{B} = (1, X, \dots, X^{2n+1})$ la base canonique de $\mathbb{R}_{2n+1}[X]$. On a

$$f(1) = (2n+1)X \quad \text{et} \quad f(X^k) = (2n+1-k)X^{k+1} + kX^{k-1} \quad \text{pour } k \in \llbracket 1 ; 2n+1 \rrbracket.$$

La matrice de f dans \mathcal{B} s'écrit alors

$$A = \begin{pmatrix} 0 & 1 & & (0) \\ 2n+1 & 0 & 2 & \\ & 2n & & \\ (0) & & 0 & 2n+1 \\ & 1 & 0 & \end{pmatrix} \in \mathcal{M}_{2n+2}(\mathbb{R}).$$

On développe le déterminant selon la première ligne

$$\det(A) = (-1)^{1+2} \times 1 \times \begin{vmatrix} 2n+1 & 2 & 0 & \cdots & 0 \\ 0 & 0 & 3 & \cdots & (0) \\ 2n & 1 & 0 & \cdots & 4 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (0) & & 0 & 2n+1 & \\ 0 & & 1 & 0 & \end{vmatrix}_{[2n+1]}$$

1. Si l'on choisit de figurer f dans la base de Taylor formée des $(X-1)^k$ pour $k \in \llbracket 0 ; 2n+1 \rrbracket$, on obtient une matrice triangulaire inférieure permettant un calcul plus facile du déterminant!

puis on développe selon la première colonne

$$\det(A) = (-1) \times 1 \times (2n+1) \times (-1)^{1+1} \begin{vmatrix} 0 & 3 & & (0) \\ 2n-1 & 0 & 4 & \\ & 2n-2 & \ddots & \\ (0) & & 0 & 2n+1 \\ & & 1 & 0 \end{vmatrix}_{[2n]}$$

On renouvelle l'opération de proche en proche

$$\det(A) = (-1)^n \times (1 \times 3 \times \cdots \times (2n-1)) \times ((2n+1) \times (2n-1) \times \cdots \times 3) \begin{vmatrix} 0 & (2n+1) \\ 1 & 0 \end{vmatrix}$$

Finalement, on exprime le produit d'entiers impairs à l'aide de nombres factoriels¹ et l'on conclut

$$\det(f) = (-1)^{n+1} (1 \times 3 \times \cdots \times (2n+1))^2 = (-1)^{n+1} \left(\frac{(2n+1)!}{2^n n!} \right)^2.$$

Exercice 20 **

Soit f un endomorphisme d'un espace vectoriel réel E de dimension finie vérifiant $f^3 + f = 0$.

- (a) Vérifier que l'image et le noyau de f sont supplémentaires.
- (b) Montrer que l'endomorphisme f est de rang pair.

Solution

(a) Par la formule du rang, on sait $\dim \text{Im}(f) + \dim \text{Ker}(f) = \dim E$: il suffit alors d'établir que l'intersection des deux espaces est réduite au vecteur nul (Th. 16 p. 244). Soit $x \in \text{Im}(f) \cap \text{Ker}(f)$. On peut écrire $x = f(a)$ avec $a \in E$ et l'on a $f(x) = 0_E$. On en déduit $f^2(a) = 0_E$ puis $f^3(a) = f(0_E) = 0_E$. Or $f^3(a) = -f(a)$ car $f^3 + f = 0$ et donc $x = f(a) = 0_E$. Ainsi, $\text{Im}(f) \cap \text{Ker}(f) = \{0_E\}$ et l'on peut conclure que les espaces $\text{Im}(f)$ et $\text{Ker}(f)$ sont supplémentaires.

(b) méthode

|| On figure f dans une base adaptée à la supplémentarité de $\text{Im}(f)$ et $\text{Ker}(f)$.

Soit e une base formée en accolant une base de $\text{Im}(f)$ et une base de $\text{Ker}(f)$. La matrice de f dans e est de la forme

$$M = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \quad \text{avec} \quad A \in \mathcal{M}_r(\mathbb{K}) \text{ et } r = \text{rg}(f)$$

où les 0 désignent des blocs nuls de tailles appropriées.

1. Voir sujet 5 p. 61.

En effet :

- les r premiers vecteurs de e sont envoyés sur des vecteurs de l'image de f et ces derniers ont des coordonnées nulles le long des vecteurs de la base de $\text{Ker}(f)$;
- les $n - r$ derniers vecteurs de e sont envoyés sur le vecteur nul car appartiennent au noyau de f .

De plus, la matrice A est inversible car $\text{rg}(f) = \text{rg}(M) = \text{rg}(A) = r$.

Enfin, l'égalité $f^3 + f = 0$ donne $M^3 + M = \mathbf{O}_n$ donc, par opérations par blocs, $A^3 + A = \mathbf{O}_r$. On peut simplifier cette dernière égalité en multipliant par A^{-1} pour obtenir $A^2 = -\mathbf{I}_r$. En appliquant la fonction déterminant à chacun des deux membres, il vient

$$\underbrace{(\det(A))^2}_{\geq 0} = (-1)^r.$$

On peut alors conclure que le rang r de l'endomorphisme f est un entier pair.

Exercice 21 ***

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Calculer déterminant et trace de l'endomorphisme f de $\mathcal{M}_n(\mathbb{K})$ défini par $f(M) = AM$.

Solution

méthode

|| On « figure » la matrice de f dans la base canonique de $\mathcal{M}_n(\mathbb{K})$.

La base canonique de $\mathcal{M}_n(\mathbb{K})$ est constituée des matrices élémentaires $E_{i,j}$ pour i et j allant de 1 à n . On peut décomposer la matrice A dans cette base

$$A = \sum_{i=1}^n \left(\sum_{j=1}^n a_{i,j} E_{i,j} \right).$$

On sait¹ $E_{i,j} E_{k,\ell} = \delta_{j,k} E_{i,\ell}$ et donc

$$f(E_{k,\ell}) = AE_{k,\ell} = \sum_{i=1}^n \left(\sum_{j=1}^n a_{i,j} E_{i,j} E_{k,\ell} \right) = \sum_{i=1}^n a_{i,k} E_{i,\ell}.$$

Ordonnons les matrices élémentaires de la base canonique en regroupant celles-ci selon l'indice de colonne :

$$E_{1,1}, E_{2,1}, \dots, E_{n,1},$$

$$E_{1,2}, E_{2,2}, \dots, E_{n,2},$$

$$\vdots$$

$$E_{1,n}, E_{2,n}, \dots, E_{n,n}.$$

1. Voir sujet 1 p. 321.

Pour $\ell \in \llbracket 1 ; n \rrbracket$, les images par f des éléments $E_{1,\ell}, \dots, E_{n,\ell}$ s'expriment comme combinaisons linéaires de ces mêmes éléments par des coefficients qui font apparaître les colonnes successives de la matrice A . On en déduit que la matrice de f dans la base canonique ainsi ordonnée est une matrice diagonale par blocs où figurent sur la diagonale n répétitions de la matrice A . On en déduit $\det(f) = (\det(A))^n$ et $\text{tr}(f) = n \text{tr}(A)$.

Exercice 22 ***

Soit f un endomorphisme d'un espace vectoriel E de dimension n et $e = (e_1, \dots, e_n)$ une base de E . Montrer que, pour tout $(x_1, \dots, x_n) \in E^n$,

$$\sum_{k=1}^n \det_e(x_1, \dots, f(x_k), \dots, x_n) = \text{tr}(f) \det_e(x_1, \dots, x_n).$$

Solution
méthode

|| On montre que le premier membre définit une forme n -linéaire alternée sur E .

Posons $\varphi: E^n \rightarrow \mathbb{K}$ l'application définie par

$$\varphi(x_1, \dots, x_n) = \sum_{k=1}^n \det_e(x_1, \dots, f(x_k), \dots, x_n).$$

Par multilinearité du déterminant d'une famille de vecteurs dans une base et par linéarité de f , il est entendu que l'application φ est linéaire en chacune de ses n variables. Vérifions son caractère alterné.

Soit (x_1, \dots, x_n) dans E^n tel que $x_i = x_j$ pour $i < j$ deux indices éléments de $\llbracket 1 ; n \rrbracket$. Lorsque $k \in \llbracket 1 ; n \rrbracket$ est distinct de i et j , le terme $\det_e(x_1, \dots, f(x_k), \dots, x_n)$ est nul car le déterminant est calculé sur une famille comportant deux fois le même vecteur. En simplifiant ces termes, $\varphi(x_1, \dots, x_n)$ se résume à la somme de deux termes

$$\varphi(x_1, \dots, x_n) = \det_e(x_1, \dots, f(x_i), \dots, x_j, \dots, x_n) + \det_e(x_1, \dots, x_i, \dots, f(x_j), \dots, x_n).$$

En échangeant les termes d'indices i et j dans le second déterminant, la propriété d'antisymétrie induit un passage à l'opposé

$$\varphi(x_1, \dots, x_n) = \det_e(x_1, \dots, f(x_i), \dots, x_j, \dots, x_n) - \det_e(x_1, \dots, f(x_j), \dots, x_i, \dots, x_n).$$

Or $x_i = x_j$ et l'on peut simplifier : $\varphi(x_1, \dots, x_n) = 0$.

L'application φ étant une forme n -linéaire alternée, il existe $\lambda \in \mathbb{K}$ tel que $\varphi = \lambda \cdot \det_e$ (Th. 7 p. 356). Reste à vérifier $\lambda = \text{tr}(f)$.

Sachant $\det_e e = 1$, on peut calculer λ par l'égalité

$$\lambda = \varphi(e_1, \dots, e_n) = \sum_{k=1}^n \det_e(e_1, \dots, f(e_k), \dots, e_n).$$

Introduisons alors $A = (a_{i,j})$ la matrice figurant f dans la base e . Les coordonnées du vecteur $f(e_k)$ dans la base sont fournies par la k -ième colonne de la matrice A . On calcule alors chaque déterminant de la somme en développant selon la k -ième ligne

$$\lambda = \sum_{k=1}^n \left| \begin{array}{cccc|c} 1 & & a_{1,k} & & \\ & \ddots & \vdots & & k \\ & & 1 & a_{k,k} & \\ & & & \ddots & 1 \\ & & & & a_{n,k} \end{array} \right| = \sum_{k=1}^n (-1)^{2k} \times a_{k,k} \times 1 = \text{tr}(f).$$

10.6.5 Applications

Exercice 23 *

Soit a_0, a_1, \dots, a_n des réels deux à deux distincts.

Montrer que la famille des polynômes $(X + a_j)^n$ pour $j = 0, 1, \dots, n$ constitue une base de $\mathbb{R}_n[X]$.

Solution

Les polynômes $(X + a_j)^n$ sont de degré n et à coefficients réels, ils appartiennent chacun à l'espace $\mathbb{R}_n[X]$.

méthode

La non-nullité du déterminant d'une famille de vecteurs caractérise une base (Th. 9 p. 357).

Calculons le déterminant de la famille des $(X + a_j)^n$ dans la base canonique \mathcal{B} formée des monômes $1, X, \dots, X^n$. Par la formule du binôme

$$(X + a_j)^n = \sum_{i=0}^n \binom{n}{i} a_j^{n-i} X^i$$

et donc

$$\det_{\mathcal{B}}((X + a_0)^n, (X + a_1)^n, \dots, (X + a_n)^n) = \begin{vmatrix} \binom{n}{0} a_0^n & \binom{n}{0} a_1^n & \cdots & \binom{n}{0} a_n^n \\ \binom{n}{1} a_0^{n-1} & \binom{n}{1} a_1^{n-1} & \cdots & \binom{n}{1} a_n^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{n}{n} a_0^0 & \binom{n}{n} a_1^0 & \cdots & \binom{n}{n} a_n^0 \end{vmatrix}_{[n+1]}$$

Pour chaque $i \in \llbracket 0 ; n \rrbracket$, on peut factoriser $\binom{n}{i}$ de la ligne d'indice i . On renverse ensuite l'ordre des lignes et l'on transpose la matrice afin de faire apparaître un déterminant de Vandermonde en la famille (a_0, a_1, \dots, a_n) . Ce dernier est non nul ce qui assure que la famille des $(X + a_j)^n$ pour $j = 0, 1, \dots, n$ est une base de $\mathbb{R}_n[X]$.

Exercice 24 *

Soit $a, b \in \mathbb{K}$. Calculer le rang de la matrice

$$M(a, b) = \begin{pmatrix} a & (b) \\ (b) & a \end{pmatrix} \in \mathcal{M}_n(\mathbb{K}).$$

Solution**méthode**

Le rang d'une matrice est la taille maximale d'une matrice carrée inversible extraite de celle-ci (Th. 15 p. 318).

On a déjà calculé le déterminant de $M(a, b)$ dans le sujet 5 p. 367 :

$$\det(M(a, b)) = (a + (n - 1)b)(a - b)^{n-1}.$$

Cas : $a = b$. La matrice $M(a, a)$ est de rang 1 sauf si $a = 0$ où c'est la matrice nulle de rang 0.

Cas : $a \neq b$ et $a + (n - 1)b = 0$. La matrice $M(a, b)$ n'est pas inversible elle est donc de rang strictement inférieur à n . Cependant, la matrice extraite de $M(a, b)$ par suppression de la dernière ligne et de la dernière colonne est de déterminant

$$(b - a)^{n-2} \underbrace{(a + (n - 2)b)}_{= -b \neq 0} \neq 0.$$

La matrice $M(a, b)$ est donc de rang $n - 1$.

Cas : $a \neq b$ et $a + (n - 1)b \neq 0$. La matrice est inversible et donc de rang n .

Exercice 25 **

Soit $A = (a_{i,j})$ une matrice carrée de taille n à coefficients dans \mathbb{Z} .

Montrer que la matrice A est inversible d'inverse une matrice à coefficients entiers si, et seulement si, $\det(A) = \pm 1$.

Solution

Rappelons¹ que le déterminant d'une matrice à coefficients entiers est un nombre entier.

Raisonnons par double implication.

(\Rightarrow) Si A est inversible et si A^{-1} est à coefficients entiers, on a

$$\underbrace{\det(A)}_{\in \mathbb{Z}} \underbrace{\det(A^{-1})}_{\in \mathbb{Z}} = \det(I_n) = 1.$$

Puisque 1 et -1 sont les seuls diviseurs de 1, on obtient $\det(A) = 1$ ou -1 .

1. Voir sujet 16 p. 376.

(\Leftarrow) Supposons $\det(A) = \pm 1$.

méthode

|| L'inverse d'une matrice peut être exprimé par la comatrice (Th. 19 p. 360).

Le déterminant de A étant non nul, la matrice A est inversible et

$$A^{-1} = \frac{1}{\det(A)} {}^t(\text{Com}(A)) = \pm {}^t(\text{Com}(A)).$$

Or les coefficients de la comatrice sont les cofacteurs de A et ceux-ci sont au signe près égaux aux mineurs de A . Chacun de ces mineurs est un déterminant d'une matrice à coefficients entiers, c'est donc un entier. On en déduit que la comatrice de A , et donc l'inverse de A , est à coefficients entiers.

Exercice 26 ** (Identités de Diophante et des quatre carrés d'Euler)

(a) Soit a, b, a', b' des nombres entiers. Calculer de deux façons :

$$\begin{vmatrix} a & -b \\ b & a \end{vmatrix} \begin{vmatrix} a' & -b' \\ b' & a' \end{vmatrix}$$

Exprimer des entiers a'' et b'' tels que

$$(a^2 + b^2)(a'^2 + b'^2) = a''^2 + b''^2.$$

(b) Soit a, b, c, d des réels. Calculer le déterminant de

$$M(a, b, c, d) = \begin{pmatrix} a & -b & c & -d \\ b & a & d & c \\ -c & -d & a & b \\ d & -c & -b & a \end{pmatrix}.$$

(c) Soit a, b, c, d et a', b', c', d' des nombres entiers. Déterminer a'', b'', c'', d'' entiers tels que

$$(a^2 + b^2 + c^2 + d^2)(a'^2 + b'^2 + c'^2 + d'^2) = a''^2 + b''^2 + c''^2 + d''^2.$$

Solution

(a) Un calcul direct de chacun des deux déterminants donne

$$\begin{vmatrix} a & -b \\ b & a \end{vmatrix} \begin{vmatrix} a' & -b' \\ b' & a' \end{vmatrix} = (a^2 + b^2)(a'^2 + b'^2).$$

Le calcul du produit des deux matrices puis du déterminant en résultant donne

$$\begin{vmatrix} a & -b \\ b & a \end{vmatrix} \begin{vmatrix} a' & -b' \\ b' & a' \end{vmatrix} = \begin{vmatrix} aa' - bb' & -(ab' + ba') \\ ba' + ab' & aa' - bb' \end{vmatrix} = (aa' - bb')^2 + (ab' + ba')^2.$$

On en déduit que $a'' = aa' - bb'$ et $b'' = ab' + ba'$ définissent des entiers convenables.

(b) méthode

Par opérations élémentaires faisant intervenir le nombre complexe i , on transforme le déterminant étudié en celui-ci d'une matrice triangulaire par blocs.

Par les opérations $C_1 \leftarrow C_1 + iC_3$ et $C_2 \leftarrow C_2 - iC_4$, on écrit

$$\begin{vmatrix} a & -b & c & -d \\ b & a & d & c \\ -c & -d & a & b \\ d & -c & -b & a \end{vmatrix} = \begin{vmatrix} a+ic & -b+id & c & -d \\ b+id & a-ic & d & c \\ -c+ia & -d-ib & a & b \\ d-ib & -c-ia & -b & a \end{vmatrix}$$

Par $L_3 \leftarrow L_3 - iL_1$ et $L_4 \leftarrow L_4 + iL_2$, on fait apparaître des zéros en lignes 3 et 4

$$\begin{vmatrix} a & -b & c & -d \\ b & a & d & c \\ -c & -d & a & b \\ d & -c & -b & a \end{vmatrix} = \begin{vmatrix} a+ic & -b+id & c & -d \\ b+id & a-ic & d & c \\ 0 & 0 & a-ic & b+id \\ 0 & 0 & -b+id & a+ic \end{vmatrix}$$

On peut alors calculer le déterminant de cette matrice triangulaire par blocs

$$\det(M(a, b, c, d)) = (|a+ic|^2 + |b+id|^2)^2 = (a^2 + b^2 + c^2 + d^2)^2.$$

(c) On pose le produit matriciel des deux matrices, soit directement, soit en raisonnant par blocs 2×2 à l'aide des calculs réalisés dans première question. On obtient alors l'égalité

$$M(a, b, c, d)M(a', b', c', d') = M(a'', b'', c'', d'')$$

avec

$$\begin{cases} a'' = aa' - bb' - cc' - dd' \\ b'' = ab' + ba' + cd' - dc' \\ c'' = ac' - bd' + ca' + db' \\ d'' = ad' + bc' - cb' + da'. \end{cases}$$

En considérant les déterminants de ces matrices, il vient

$$(a^2 + b^2 + c^2 + d^2)^2 (a'^2 + b'^2 + c'^2 + d'^2)^2 = (a''^2 + b''^2 + c''^2 + d''^2)^2.$$

Enfin, les quantités étant toutes positives, on conclut

$$(a^2 + b^2 + c^2 + d^2)(a'^2 + b'^2 + c'^2 + d'^2) = a''^2 + b''^2 + c''^2 + d''^2$$

avec $a'', b'', c'', d'' \in \mathbb{Z}$.

10.7 Exercices d'approfondissement

Exercice 27 * (Formules de Cramer)

On considère le système d'équations linéaires

$$(\Sigma) \quad \begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = b_2 \\ \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\ a_{n,1}x_1 + a_{n,2}x_2 + \cdots + a_{n,n}x_n = b_n \end{cases}$$

d'équation matricielle $AX = B$.

- (a) Montrer que ce système est de Cramer si, et seulement si, $\det(A) \neq 0$.
- (b) Montrer que sa solution (x_1, \dots, x_n) est alors déterminée par

$$x_j = \frac{\begin{vmatrix} a_{1,1} & \cdots & b_1 & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & b_n & \cdots & a_{n,n} \end{vmatrix}}{\det(A)} \quad \text{pour tout } j = 1, \dots, n.$$

Solution

(a) Le système (Σ) est de Cramer si, et seulement si, la matrice A qui l'exprime est inversible : ceci revient à dire $\det(A) \neq 0$.

(b) Lorsque le système est de Cramer, il possède une solution unique (x_1, \dots, x_n) .

méthode

|| On exprime B comme combinaison linéaire des colonnes de la matrice A .

Notons C_1, \dots, C_n les colonnes de la matrice A . L'égalité $AX = B$ avec X colonne de coefficients x_1, \dots, x_n donne l'identité $B = x_1C_1 + \cdots + x_nC_n$.

Soit $j \in \llbracket 1 ; n \rrbracket$. Par multilinéarité du déterminant en la famille des colonnes, il vient

$$\begin{aligned} \begin{vmatrix} a_{1,1} & \cdots & b_1 & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & b_n & \cdots & a_{n,n} \end{vmatrix} &= \det(C_1 \quad \dots \mid \sum_{i=1}^n x_i C_i \mid \dots \quad C_n) \\ &= \sum_{i=1}^n x_i \det(C_1 \mid \dots \mid C_i \mid \dots \mid C_n). \end{aligned}$$

Dans la somme, le déterminant est nul lorsque $i \neq j$ car la matrice comporte deux fois la colonne C_i : une fois à la position i et une fois à la position j . Après simplification, il ne

reste dans la somme que le terme d'indice $j = i$ où le déterminant se reconnaît comme étant celui de la matrice A . On peut alors conclure¹

$$\begin{vmatrix} a_{1,1} & \cdots & b_1 & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & b_n & \cdots & a_{n,n} \end{vmatrix}^j = x_j \det(A).$$

Exercice 28 **

Soit n un entier supérieur à 2 et $A \in \mathcal{M}_n(\mathbb{K})$.

- (a) Calculer le rang de la comatrice de A en fonction de celui de A .
- (b) Déterminer $\text{Com}(\text{Com}(A))$.

Solution

(a) On sait (Th. 19 p. 360)

$$A^t(\text{Com}(A)) = \det(A)I_n. \quad (*)$$

Si la matrice A est inversible, la transposée de la comatrice de A , et donc la comatrice de A , est aussi inversible. C'est alors une matrice de rang n .

Si la matrice A n'est pas inversible, l'égalité (*) devient

$$A^t(\text{Com}(A)) = O_n. \quad (**)$$

méthode

Le rang d'une matrice est la taille maximale d'une matrice inversible extraite de celle-ci.

Si $\text{rg}(A) \leq n - 2$, il n'existe aucune matrice inversible de taille $n - 1$ issue de A . Tous les mineurs de A sont donc nuls et, par conséquent, la comatrice de A est nulle.

Si $\text{rg}(A) = n - 1$, il existe au moins un mineur non nul dans la matrice A et la comatrice de A n'est pas nulle. De plus, l'égalité (**) donne

$$\text{Im}({}^t\text{Com}(A)) \subset \text{Ker}(A) \quad \text{avec} \quad \dim \text{Ker}(A) = n - \text{rg}(A) = 1.$$

La transposée de la comatrice de A , et donc la comatrice de A , est alors de rang 1.

En résumé

$$\text{rg}(\text{Com}(A)) = \begin{cases} n & \text{si } \text{rg}(A) = n \\ 1 & \text{si } \text{rg}(A) = n - 1 \\ 0 & \text{si } \text{rg}(A) \leq n - 2. \end{cases}$$

1. Les formules obtenues permettent de résoudre les systèmes de Cramer mais sont assez peu utilisées pour les résolutions pratiques lorsque $n \geq 3$. En revanche, celles-ci sont utiles pour les études théoriques où l'on souhaite savoir comment les solutions d'un système dépendent des coefficients de celui-ci.

(b) Lorsqu'une matrice carrée M est inversible, on sait

$$M^{-1} = \frac{1}{\det(M)} {}^t(\text{Com}(M)) \quad \text{donc} \quad \text{Com}(M) = \det(M)({}^t M)^{-1}.$$

Si la matrice A est inversible, la comatrice de A l'est aussi et

$$\text{Com}(\text{Com}(A)) = \det(\text{Com}(A))({}^t \text{Com}(A))^{-1} = \frac{\det(\text{Com}(A))}{\det(A)} A.$$

Aussi, en calculant le déterminant des deux membres de (*), on obtient

$$\det(A) \det(\text{Com}(A)) = (\det(A))^n$$

et donc

$$\text{Com}(\text{Com}(A)) = (\det(A))^{n-2} A. \quad (\Delta)$$

Lorsque la matrice A n'est pas inversible, la comatrice de A est de rang inférieur à 1. Si $n \geq 3$, la comatrice de la comatrice de A est alors la matrice nulle. Si $n = 2$, on peut simplement mener le calcul à partir des coefficients de A :

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \text{Com}(A) = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} \quad \text{et} \quad \text{Com}(\text{Com}(A)) = A.$$

Dans tous les cas, la formule (Δ) est valable.

Exercice 29 ** (Inégalités de réordonnement)

Soit a_1, \dots, a_n et b_1, \dots, b_n des réels triés par ordre croissant :

$$a_1 \leq a_2 \leq \dots \leq a_n \quad \text{et} \quad b_1 \leq b_2 \leq \dots \leq b_n.$$

Déterminer

$$\min_{\sigma \in S_n} \left(\sum_{i=1}^n a_i b_{\sigma(i)} \right) \quad \text{et} \quad \max_{\sigma \in S_n} \left(\sum_{i=1}^n a_i b_{\sigma(i)} \right).$$

Solution

méthode

|| Pour x, y, x' et y' réels tels que $x \leq y$ et $x' \leq y'$, on a $xy' + x'y \leq xx' + yy'$.

En effet, la différence des deux membres est positive

$$xx' + yy' - (xy' + x'y) = \underbrace{(y-x)}_{\geq 0} \underbrace{(y'-x')}_{\geq 0} \geq 0.$$

Pour $\sigma \in S_n$, la quantité $a_1 b_{\sigma(1)} + \dots + a_n b_{\sigma(n)}$ semble alors maximale lorsque les grandes quantités sont multipliées entre elles et, à l'inverse, elle semble minimale lorsque les grandes quantités multiplient les petites. Montrons

$$\sum_{i=1}^n a_i b_{n-i} \leq \sum_{i=1}^n a_i b_{\sigma(i)} \leq \sum_{i=1}^n a_i b_i.$$

méthode

|| On établit la majoration en raisonnant par récurrence.

La propriété est immédiate pour $n = 1$.

Supposons la propriété vraie à un rang $n - 1 \geq 1$. Au rang suivant, on introduit des réels a_1, \dots, a_n et b_1, \dots, b_n triés par ordre croissant et σ une permutation des entiers allant de 1 à n .

Cas : $\sigma(n) = n$. L'application σ définit par restriction une permutation des entiers allant de 1 à $n - 1$ et il suffit d'appliquer directement l'hypothèse de récurrence à partir des sous-familles des réels (a_1, \dots, a_{n-1}) et (b_1, \dots, b_{n-1})

$$\sum_{i=1}^n a_i b_{\sigma(i)} = \sum_{i=1}^{n-1} a_i b_{\sigma(i)} + a_n b_n \leq \sum_{i=1}^{n-1} a_i b_i + a_n b_n = \sum_{i=1}^n a_i b_i.$$

Cas : $\sigma(n) \neq n$. Introduisons l'indice k tel que $\sigma(k) = n$ et isolons dans la somme les termes d'indices n et k

$$\begin{aligned} \sum_{i=1}^n a_i b_{\sigma(i)} &= \left(\sum_{i \neq k, n} a_i b_{\sigma(i)} \right) + a_k b_n + a_n b_{\sigma(n)} \\ &\leq \left(\sum_{i \neq k, n} a_i b_{\sigma(i)} \right) + a_k b_{\sigma(n)} + a_n b_n = \sum_{i=1}^n a_i b_{\varphi(i)} \end{aligned}$$

avec φ la permutation de $[1 ; n]$ définie par la composition $\varphi = (n \ \sigma(n)) \circ \sigma$. Puisque $\varphi(n) = n$, on revient au cas précédent et l'on peut conclure

$$\sum_{i=1}^n a_i b_{\sigma(i)} \leq \sum_{i=1}^n a_i b_{\varphi(i)} \leq \sum_{i=1}^n a_i b_i.$$

La récurrence est établie.

méthode

|| La minoration se déduit de la majoration par passage à l'opposé.

La famille $(-b_n, \dots, -b_1)$ étant croissante, on peut affirmer que, pour toute permutation σ ,

$$\sum_{i=1}^n a_i (-b_{\sigma(i)}) \leq \sum_{i=1}^n -a_i b_{n-i}$$

et donc

$$\sum_{i=1}^n a_i b_{\sigma(i)} \geq \sum_{i=1}^n a_i b_{n-i}.$$

Finalement,

$$\min_{\sigma \in S_n} \left(\sum_{i=1}^n a_i b_{\sigma(i)} \right) = \sum_{i=1}^n a_i b_{n-i} \quad \text{et} \quad \max_{\sigma \in S_n} \left(\sum_{i=1}^n a_i b_{\sigma(i)} \right) = \sum_{i=1}^n a_i b_i.$$

Exercice 30 * (Déterminant de Cauchy)**

Soit a_1, \dots, a_n et b_1, \dots, b_n des réels tels que $a_i + b_j \neq 0$ pour tous i et j de $\llbracket 1 ; n \rrbracket$.

Calculer

$$\det\left(\frac{1}{a_i + b_j}\right)_{1 \leq i, j \leq n}$$

Solution

Notons D_n le déterminant étudié

$$D_n = \det\left(\frac{1}{a_i + b_j}\right)_{1 \leq i, j \leq n} = \begin{vmatrix} \frac{1}{a_1 + b_1} & \frac{1}{a_1 + b_2} & \cdots & \frac{1}{a_1 + b_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{a_n + b_1} & \frac{1}{a_n + b_2} & \cdots & \frac{1}{a_n + b_n} \end{vmatrix}_{(n)}$$

méthode

|| On forme une relation de récurrence exprimant D_n en fonction de D_{n-1} .

Soit $j \in \llbracket 1 ; n-1 \rrbracket$. Par l'opération élémentaire $C_j \leftarrow C_j - C_n$, la colonne d'indice j devient

$$\begin{pmatrix} \frac{1}{a_1 + b_j} - \frac{1}{a_1 + b_n} \\ \vdots \\ \frac{1}{a_n + b_j} - \frac{1}{a_n + b_n} \end{pmatrix} = \begin{pmatrix} \frac{b_n - b_j}{(a_1 + b_j)(a_1 + b_n)} \\ \vdots \\ \frac{b_n - b_j}{(a_n + b_j)(a_n + b_n)} \end{pmatrix}.$$

Cette colonne peut être factorisée par $b_n - b_j$. Au surplus, après l'ensemble des opérations élémentaires $C_1 \leftarrow C_1 - C_n, \dots, C_{n-1} \leftarrow C_{n-1} - C_n$, la ligne d'indice $i \in \llbracket 1 ; n \rrbracket$ peut aussi être factorisée par $1/(a_i + b_n)$. On obtient alors

$$D_n = \frac{(b_n - b_1) \dots (b_n - b_{n-1})}{(a_1 + b_n) \dots (a_n + b_n)} \begin{vmatrix} \frac{1}{a_1 + b_1} & \cdots & \frac{1}{a_1 + b_{n-1}} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \frac{1}{a_n + b_1} & \cdots & \frac{1}{a_n + b_{n-1}} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \frac{1}{a_n + b_1} & \cdots & \frac{1}{a_n + b_{n-1}} & 1 \end{vmatrix}_{(n)}$$

Étudions le déterminant ainsi introduit. Par l'opération élémentaire $L_i \leftarrow L_i - L_n$, la ligne d'indice $i \in \llbracket 1 ; n-1 \rrbracket$ devient

$$\left(\frac{a_n - a_i}{(a_i + b_1)(a_n + b_1)} \cdots \frac{a_n - a_i}{(a_i + b_{n-1})(a_n + b_{n-1})} \quad 0 \right).$$

Cette ligne est multiple de $a_n - a_i$. Aussi, après la réalisation de l'ensemble des opérations élémentaires $L_1 \leftarrow L_1 - L_n, \dots, L_{n-1} \leftarrow L_{n-1} - L_n$, la colonne d'indice $j \in \llbracket 1 ; n-1 \rrbracket$

peut être factorisée par $1/(a_n + b_j)$ et l'on obtient

$$\begin{vmatrix} \frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_{n-1}} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \frac{1}{a_{n-1}+b_1} & \cdots & \frac{1}{a_{n-1}+b_{n-1}} & 1 \\ \frac{1}{a_n+b_1} & \cdots & \frac{1}{a_n+b_{n-1}} & 1 \end{vmatrix} = \frac{(a_n - a_1) \dots (a_n - a_{n-1})}{(a_n + b_1) \dots (a_n + b_{n-1})} \begin{vmatrix} \frac{1}{a_1+b_1} & \cdots & \frac{1}{a_1+b_{n-1}} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ \frac{1}{a_{n-1}+b_1} & \cdots & \frac{1}{a_{n-1}+b_{n-1}} & 0 \\ 1 & \cdots & 1 & 1 \end{vmatrix}$$

En développant ce dernier déterminant selon la dernière colonne et en combinant l'ensemble des résultats précédents, on obtient

$$D_n = \frac{\prod_{j=1}^{n-1} ((a_n - a_j)(b_n - b_j))}{\prod_{j=1}^{n-1} ((a_j + b_n)(a_n + b_j)) \times (a_n + b_n)} D_{n-1}.$$

On exprime ensuite D_{n-1} en fonction de D_{n-2} , etc. À terme, on obtient la formule suivante que l'on pourra valider en raisonnant par récurrence

$$D_n = \frac{\prod_{1 \leq i < j \leq n} ((a_j - a_i)(b_j - b_i))}{\prod_{1 \leq i, j \leq n} (a_i + b_j)}$$

CHAPITRE 11

Espaces préhilbertiens réels

E désigne un espace vectoriel réel.

11.1 Produit scalaire

11.1.1 Définition

Définition

- Un *produit scalaire* sur l'espace réel E est une application $\varphi: E \times E \rightarrow \mathbb{R}$ vérifiant :
- 1) φ est bilinéaire, c'est-à-dire φ est linéaire en chacune de ses deux variables ;
 - 2) φ est *symétrique*, c'est-à-dire $\varphi(x, y) = \varphi(y, x)$ pour tous x et y de E ;
 - 3) φ est *définie positive*, c'est-à-dire $\varphi(x, x) > 0$ pour tout $x \in E$ non nul¹.

On dit qu'un produit scalaire est une forme bilinéaire symétrique définie positive.

Définition

- On appelle *espace préhilbertien* tout couple (E, φ) formé d'un espace vectoriel réel E et d'un produit scalaire φ sur E .

Dans un espace préhilbertien E , il est usuel de noter $(x | y)$, $\langle x, y \rangle$ ou $x \cdot y$ le produit scalaire de deux vecteurs x et y . Nous privilégions ici la notation $(x | y)$.

Sur $E = \mathbb{R}^n$, on définit un produit scalaire (appelé *produit scalaire canonique*) en posant

$$(x | y) = \sum_{k=1}^n x_k y_k = x_1 y_1 + \cdots + x_n y_n \quad \text{pour tous } x, y \in \mathbb{R}^n.$$

1. En pratique, on vérifie souvent $\varphi(x, x) \geq 0$ puis l'implication $\varphi(x, x) = 0 \implies x = 0_E$

Sur $E = C([a ; b], \mathbb{R})$ (avec $a < b$), on définit un produit scalaire en posant

$$(f | g) = \int_a^b f(t)g(t) dt \quad \text{pour tous } f, g \in C([a ; b], \mathbb{R}).$$

11.1.2 Norme euclidienne

E désigne un espace préhilbertien de produit scalaire $(\cdot | \cdot)$.

Définition

On appelle *norme euclidienne* sur E l'application $\|\cdot\| : E \rightarrow \mathbb{R}_+$ définie par

$$\|x\| = \sqrt{(x | x)}.$$

Sur $E = \mathbb{R}^n$ muni du produit scalaire canonique

$$\|x\| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Théorème 1 (Identités remarquables)

Pour tous vecteurs a et b de E , on a les identités

$$\begin{aligned}\|a + b\|^2 &= \|a\|^2 + 2(a | b) + \|b\|^2 \\ \|a - b\|^2 &= \|a\|^2 - 2(a | b) + \|b\|^2 \\ (a - b | a + b) &= \|a\|^2 - \|b\|^2.\end{aligned}$$

En particulier, on peut exprimer le produit scalaire à partir de la norme euclidienne par la *formule de polarisation* :

$$(a | b) = \frac{1}{2}(\|a + b\|^2 - \|a\|^2 - \|b\|^2).$$

On remarque que le vecteur nul est le seul vecteur x vérifiant $\|x\| = 0$. On remarque aussi $\|\lambda x\| = |\lambda| \|x\|$ pour tout λ réel et tout x de E . De plus, on dispose des résultats qui suivent :

Théorème 2 (Inégalité de Cauchy-Schwarz¹)

Pour tous vecteurs x et y de E , on a l'inégalité

$$|(x | y)| \leq \|x\| \|y\|$$

avec égalité si, et seulement si, la famille (x, y) est liée.

1. On peut avoir quelquefois des doutes sur la présence de carrés ou de racines lorsque l'on énonce une inégalité de Cauchy-Schwarz. Ceux-ci peuvent être levés par des considérations d'homogénéité : la norme des vecteurs se mesurent en mètre et le produit scalaire en mètre carré, les deux membres de l'inégalité doivent avoir la même unité.

Par exemple, sur \mathbb{R}^n muni du produit scalaire canonique, l'inégalité de Cauchy-Schwarz donne

$$\left| \sum_{k=1}^n x_k y_k \right| \leq \left(\sum_{k=1}^n x_k^2 \right)^{1/2} \left(\sum_{k=1}^n y_k^2 \right)^{1/2}.$$

Sur $C([a; b], \mathbb{R})$ muni du produit scalaire usuel, l'inégalité de Cauchy-Schwarz se lit

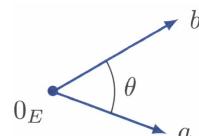
$$\left| \int_a^b f(t)g(t) dt \right| \leq \left(\int_a^b |f(t)|^2 dt \right)^{1/2} \left(\int_a^b |g(t)|^2 dt \right)^{1/2}.$$

Lorsque a et b sont deux vecteurs non nuls, l'inégalité de Cauchy-Schwarz permet d'affirmer l'existence d'un unique réel $\theta \in [0; \pi]$ pour lequel

$$(a | b) = \|a\| \|b\| \cos \theta.$$

Définition

Le réel θ est appelé *écart angulaire*¹ entre les vecteurs a et b . Selon que $\theta < \pi/2$, $\theta = \pi/2$ ou $\theta > \pi/2$, on dit que l'angle formé par a et b est *aigu*, *droit* ou *obtus*.



Théorème 3 (Inégalité triangulaire)

Pour tous vecteurs x et y de E , on a l'inégalité

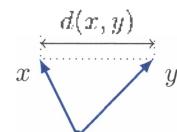
$$\|x + y\| \leq \|x\| + \|y\|$$

avec égalité si, et seulement si, x et y sont colinéaires et de produit scalaire positif (on dit que x et y sont *positivement liés*).

Tout comme la valeur absolue permet de mesurer la distance entre deux réels et le module la distance entre deux complexes, la norme euclidienne permet de définir la distance entre deux vecteurs.

Définition

Si x et y sont deux vecteurs de E , on appelle *distance* séparant les vecteurs x et y le réel $d(x, y) = \|y - x\|$.



11.1.3 Vecteurs orthogonaux

E désigne un espace préhilbertien de produit scalaire $(\cdot | \cdot)$.

Définition

Deux vecteurs x et y de E sont dits *orthogonaux* si $(x | y) = 0$.

1. Cette notion angulaire n'est pas orientée : elle se figure par un arc de cercle non fléché de longueur inférieure à π .

Le vecteur nul est le seul vecteur orthogonal à tout autre. En particulier, le vecteur nul est le seul vecteur orthogonal à lui-même :

$$(x|x) = 0 \implies x = 0_E.$$

Soit $e = (e_i)_{i \in I}$ une famille de vecteurs de E indexée par un ensemble I .

Définition

On dit que e est une *famille orthogonale* si elle est constituée de vecteurs deux à deux orthogonaux, c'est-à-dire

$$\forall (i, j) \in I^2, \quad i \neq j \implies (e_i | e_j) = 0.$$

Définition

On dit que e est une *famille orthonormale*¹ si de plus ses vecteurs sont unitaires

$$\forall (i, j) \in I^2, \quad (e_i | e_j) = \delta_{i,j} \quad \text{avec} \quad \delta_{i,j} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{sinon.} \end{cases}$$

Théorème 4 (Théorème de Pythagore)

Si $e = (e_1, \dots, e_n)$ est une famille orthogonale de vecteurs de E ,

$$\left\| \sum_{i=1}^n e_i \right\|^2 = \sum_{i=1}^n \|e_i\|^2.$$

On peut en déduire :

Théorème 5

Toute famille orthogonale ne comportant pas le vecteur nul est libre.

En particulier, les familles orthonormales sont libres.

Théorème 6 (Orthonormalisation de Schmidt)

Si (u_1, \dots, u_n) est une famille libre de vecteurs de E , il existe une unique famille orthonormale (e_1, \dots, e_n) vérifiant

$$\text{Vect}(u_1, \dots, u_k) = \text{Vect}(e_1, \dots, e_k) \quad \text{et} \quad (e_k | u_k) > 0 \quad \text{pour tout } k \in [1; n].$$

On dit que la famille (e_1, \dots, e_n) est la *famille orthonormalisée* de (u_1, \dots, u_n) par le procédé de Schmidt.

La démarche algorithmique permettant de déterminer la famille (u_1, \dots, u_n) à partir de (e_1, \dots, e_n) est présentée dans le sujet 3 p. 408

1. On parle aussi indifféremment de *famille orthonormée*.

11.1.4 Orthogonal d'une partie

E désigne un espace préhilbertien de produit scalaire $(\cdot | \cdot)$.

Définition

On appelle *orthogonal* d'une partie A de E l'ensemble noté A^\perp constitué des vecteurs de E orthogonaux à tous les vecteurs de A :

$$A^\perp = \{x \in E \mid \forall a \in A, (a | x) = 0\}.$$

Puisque le vecteur nul est orthogonal à tout vecteur, on a $\{0_E\}^\perp = E$. Inversement, seul le vecteur nul est orthogonal à tout vecteur de E et donc $E^\perp = \{0_E\}$.

Théorème 7

A^\perp est un sous-espace vectoriel de E .

Notons l'inclusion $A \subset (A^\perp)^\perp$ et l'égalité $A^\perp = (\text{Vect } A)^\perp$. En particulier, si e_1, \dots, e_p sont des vecteurs de E , on a

$$\text{Vect}(e_1, \dots, e_p)^\perp = \{e_1, \dots, e_p\}^\perp.$$

11.1.5 Sous-espaces orthogonaux

Soit F et G des sous-espaces vectoriels d'un espace préhilbertien E .

Définition

On dit que les deux sous-espaces F et G sont *orthogonaux* lorsqu'ils sont formés de vecteurs deux à deux orthogonaux, c'est-à-dire si $(x | y) = 0$ pour tout x de F et tout y de G .

L'orthogonalité des sous-espaces vectoriels signifie l'inclusion¹ $F \subset G^\perp$.

11.2 Espaces euclidiens

11.2.1 Définitions

Définition

On appelle *espace euclidien* tout espace préhilbertien réel de dimension finie.

Lorsque l'on munit \mathbb{R}^n du produit scalaire canonique, on dit que \mathbb{R}^n est muni de sa *structure euclidienne canonique*.

Définition

Une *base orthonormale* d'un espace euclidien est une famille de vecteurs qui est à la fois une base et une famille orthonormale.

La base canonique de \mathbb{R}^n est une base orthonormale pour le produit scalaire canonique.

1. Ou, et c'est équivalent $G \subset F^\perp$.

La famille $e = (e_1, \dots, e_n)$ que l'on obtient en appliquant le procédé d'orthonormalisation de Schmidt à une base $u = (u_1, \dots, u_n)$ est une base¹ orthonormale de l'espace. On en déduit :

Théorème 8

Tout espace euclidien possède une base orthonormale.

On peut aussi compléter les familles orthonormales en bases orthonormales :

Théorème 9 (Théorème de la base orthonormale incomplète)

Toute famille orthonormale d'un espace euclidien peut être complétée en une base orthonormale.

11.2.2 Coordonnées dans une base orthonormale

Soit E un espace euclidien et $e = (e_1, \dots, e_n)$ une base orthonormale de E .

On peut aisément calculer les coordonnées d'un vecteur dans la base orthonormale e à l'aide du produit scalaire.

Théorème 10

Les coordonnées x_1, \dots, x_n d'un vecteur x de E dans la base orthonormale e sont données par

$$x_k = (e_k | x) \quad \text{pour tout } k \in [1; n].$$

On dispose ainsi de la formule

$$x = \sum_{k=1}^n (e_k | x) e_k.$$

On peut aussi calculer produit scalaire et norme à l'aide des coordonnées des vecteurs dans une base orthonormale :

Théorème 11

Si x et y sont des vecteurs de E de coordonnées x_1, \dots, x_n et y_1, \dots, y_n dans la base orthonormale e ,

$$(x | y) = x_1 y_1 + \cdots + x_n y_n \quad \text{et} \quad \|x\|^2 = x_1^2 + \cdots + x_n^2.$$

Si X et Y sont les colonnes des coordonnées des vecteurs x et y , les calculs qui précèdent correspondent aux opérations

$$(x | y) = {}^t X Y \quad \text{et} \quad \|x\|^2 = {}^t X X.$$

1. Il est remarquable que la matrice de passage de e à u est triangulaire supérieure à coefficients diagonaux strictement positifs.

11.2.3 Produit mixte

Soit E un espace euclidien orienté de dimension $n \geq 1$.

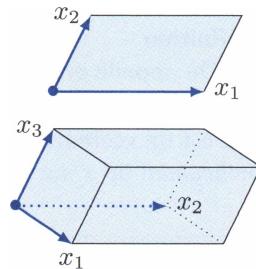
Si $e = (e_1, \dots, e_n)$ et $e' = (e'_1, \dots, e'_n)$ sont deux bases orthonormales directes de E , le déterminant de la famille e' dans la base e est égal à 1. Par la formule de changement de base (Th. 8 p. 356), le déterminant d'une famille de vecteurs est le même dans toute base orthonormale directe.

Définition

On appelle produit mixte d'une famille (x_1, \dots, x_n) de vecteurs de E , le réel $[x_1, \dots, x_n]$ égal au déterminant de cette famille dans toute base orthonormale directe de E .

Lorsque $n = 2$, $|[x_1, x_2]|$ détermine l'aire du parallélogramme défini par les vecteurs x_1 et x_2 .

Lorsque $n = 3$, $|[x_1, x_2, x_3]|$ détermine le volume du parallélépipède défini par les vecteurs x_1, x_2, x_3 .



Théorème 12

Soit (x_1, \dots, x_n) une famille de vecteurs de E .

Le produit mixte $[x_1, \dots, x_n]$ est non nul si, et seulement si, la famille (x_1, \dots, x_n) est une base.

Le produit mixte $[x_1, \dots, x_n]$ est strictement positif lorsque cette base est directe et strictement négatif sinon.

Le produit mixte $[x_1, \dots, x_n]$ se comprend alors comme un *volume orienté*.

Théorème 13

Soit (x_1, \dots, x_n) une famille de vecteurs de E . Pour tout endomorphisme u de E

$$[u(x_1), \dots, u(x_n)] = \det(u) \times [x_1, \dots, x_n].$$

Un endomorphisme multiplie donc les volumes orientés par son déterminant.

11.2.4 Supplémentaire orthogonal

Soit E un espace euclidien.

Théorème 14

Si F est un sous-espace vectoriel de E alors $F \oplus F^\perp = E$.

L'espace F^\perp est donc un supplémentaire¹ de F dans E que l'on appelle le *supplémentaire*.

1. On verra dans le sujet 30 p. 432 qu'il est possible en dimension infinie que F^\perp ne soit pas un supplémentaire de F .

orthogonal de F . Il vérifie :

$$\dim F^\perp = \dim E - \dim F \quad \text{et} \quad (F^\perp)^\perp = F.$$

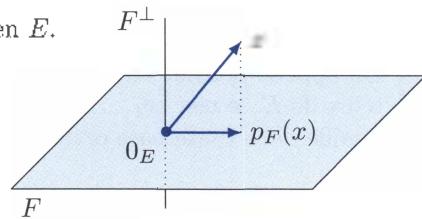
11.2.5 Projection et symétrie orthogonale

Soit F un sous-espace vectoriel d'un espace euclidien E .

Définition

On appelle *projection orthogonale* sur F la projection p_F sur F parallèlement à F^\perp .

Si x est un vecteur de E , $p_F(x)$ se nomme le *projété orthogonal* de x sur F .



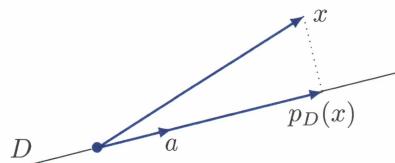
Théorème 15

Si (e_1, \dots, e_m) est une base orthonormale du sous-espace F , on a

$$p_F(x) = \sum_{k=1}^m (e_k | x) e_k \quad \text{pour tout } x \in E.$$

En particulier, si D est une droite vectorielle engendrée par un vecteur a , on a

$$p_D(x) = \frac{(a | x)}{\|a\|^2} a \quad \text{pour tout } x \in E.$$



Définition

On appelle *symétrie orthogonale* par rapport à F , la symétrie s_F par rapport à F et parallèlement à F^\perp .

Lorsque F est un hyperplan H , on parle de *réflexion* d'hyperplan H .

11.2.6 Distance à un sous-espace vectoriel

Soit F un sous-espace vectoriel d'un espace euclidien E .

Théorème 16

Soit x un vecteur de E . Pour tout vecteur y de F

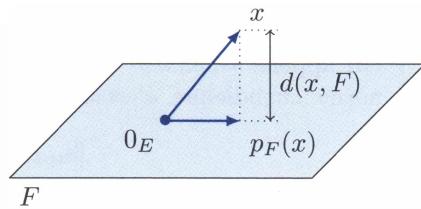
$$\|x - y\| \geq \|x - p_F(x)\|$$

avec égalité si, et seulement si, $y = p_F(x)$.

Définition

On appelle distance $d(x, F)$ d'un vecteur x de E au sous-espace vectoriel F , la distance minimale du vecteur x aux vecteurs de F .

Le résultat précédent entraîne que la distance d'un vecteur à un sous-espace vectoriel est atteinte en son projeté orthogonal.

**Théorème 17**

Pour tout vecteur x de E

$$d(x, F) = \|x - p_F(x)\|.$$

11.2.7 Hyperplans dans un espace euclidien

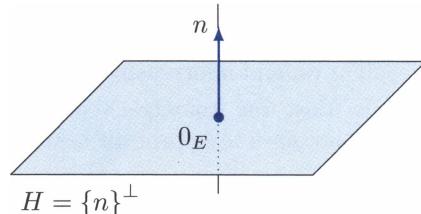
Soit H un hyperplan d'un espace euclidien E .

Définition

La droite vectorielle H^\perp se nomme la *droite normale* de H . Tout vecteur non nul de celle-ci est appelé *vecteur normal* à H .

Si n est un vecteur normal à H alors, pour tout vecteur x de E

$$x \in H \iff (n | x) = 0.$$



Ainsi, la donnée d'un vecteur normal (non nul) suffit à définir un hyperplan.

Les hyperplans affines de direction H sont alors les ensembles solutions des équations

$$(E_\lambda): (n | x) = \lambda \quad \text{avec} \quad \lambda \in \mathbb{R}.$$

En géométrie affine, le produit scalaire de deux vecteurs u et v est usuellement noté $u \cdot v$. Si A est un point d'un hyperplan affine de vecteur normal \bar{n} , cet hyperplan est formé des points M vérifiant

$$\overrightarrow{AM} \cdot \bar{n} = 0.$$

Si le vecteur \bar{n} est unitaire, le produit scalaire $\overrightarrow{AM} \cdot \bar{n}$ détermine la *distance orientée*¹ du point M à l'hyperplan affine.

11.3 Isométries vectorielles

E désigne un espace euclidien de produit scalaire $(\cdot | \cdot)$ et de dimension $n \geq 1$.

1. La valeur est positive si le vecteur \overrightarrow{AM} pointe le même demi-espace que \bar{n} , elle est négative sinon.

11.3.1 Définition

Définition

On appelle *isométrie vectorielle* de E tout endomorphisme u de E conservant la norme euclidienne, c'est-à-dire vérifiant :

$$\|u(x)\| = \|x\| \quad \text{pour tout } x \in E.$$

Le noyau d'une isométrie vectorielle est réduit au vecteur nul. Une isométrie vectorielle est donc un automorphisme de E . On parle parfois d'*automorphisme orthogonal* pour désigner une isométrie vectorielle.

Théorème 18

L'ensemble $O(E)$ des isométries vectorielles est un sous-groupe du groupe linéaire $(\mathrm{GL}(E), \circ)$. On l'appelle le *groupe orthogonal de E* .

Les endomorphismes Id_E et $-\mathrm{Id}_E$ sont des isométries vectorielles.

11.3.2 Caractérisations

Théorème 19 (Conservation du produit scalaire)

Soit u un endomorphisme de E . On a équivalence entre :

- (i) u est une isométrie vectorielle ;
- (ii) u conserve le produit scalaire, c'est-à-dire vérifie :

$$(u(x)|u(y)) = (x|y) \quad \text{pour tout } (x, y) \in E^2.$$

En particulier, les isométries vectorielles conservent l'orthogonalité des vecteurs.

Les symétries orthogonales sont des isométries vectorielles¹ et donc conservent le produit scalaire. Inversement, parmi les symétries, seules les symétries orthogonales sont des isométries.

Théorème 20 (Image d'une base orthonormale)

Soit u un endomorphisme de E et $e = (e_1, \dots, e_n)$ une base orthonormale de E . On a équivalence entre :

- (i) u est une isométrie vectorielle ;
- (ii) la famille $(u(e_1), \dots, u(e_n))$ est une base orthonormale.

11.3.3 Matrices orthogonales

Définition

On dit qu'une matrice A de $\mathcal{M}_n(\mathbb{R})$ est *orthogonale* si ${}^t A A = I_n$.

1. En revanche, les projections orthogonales autre que Id_E n'en sont pas.

Les matrices orthogonales sont inverses égales à leurs transposées. Les matrices I_n et $-I_n$ sont orthogonales.

Théorème 21

L'ensemble¹ $O_n(\mathbb{R})$ des matrices orthogonales de taille n est un sous-groupe du groupe $(GL_n(\mathbb{R}), \times)$. On l'appelle le *groupe orthogonal d'ordre n*.

En identifiant les colonnes de $\mathcal{M}_{n,1}(\mathbb{R})$ avec les vecteurs de \mathbb{R}^n constitués des mêmes coefficients, on peut définir un produit scalaire $\langle \cdot, \cdot \rangle$ sur l'espace $\mathcal{M}_{n,1}(\mathbb{R})$ à partir du produit scalaire canonique sur \mathbb{R}^n . Ce produit scalaire est donné par le calcul matriciel suivant :

$$\langle X, Y \rangle = {}^tXY \quad \text{pour tous } X \text{ et } Y \text{ de } \mathcal{M}_{n,1}(\mathbb{R}).$$

De la même manière², on introduit un produit scalaire sur l'espace $\mathcal{M}_{1,n}(\mathbb{R})$ des lignes donné par :

$$\langle X, Y \rangle = X {}^tY \quad \text{pour tous } X \text{ et } Y \text{ de } \mathcal{M}_{1,n}(\mathbb{R}).$$

On peut alors caractériser facilement qu'une matrice est orthogonale en étudiant ses rangées.

Théorème 22

Soit A une matrice de $\mathcal{M}_n(\mathbb{R})$ de colonnes C_1, \dots, C_n et de lignes L_1, \dots, L_n .

On a équivalence entre :

- (i) la matrice A est orthogonale;
- (ii) la famille (C_1, \dots, C_n) est orthonormale;
- (iii) la famille (L_1, \dots, L_n) est orthonormale.

Les notions de base orthonormale, d'isométrie et de matrice orthogonale sont liées par les deux résultats qui suivent :

Théorème 23

Soit $e = (e_1, \dots, e_n)$ une base orthonormale de l'espace E .

Une famille $e' = (e'_1, \dots, e'_n)$ de vecteurs de E est une base orthonormale si, et seulement si, la matrice P figurant la famille e' dans la base orthonormale e est orthogonale.

Une matrice de passage P entre deux bases orthonormales est donc une matrice orthogonale et par conséquent facile à inverser : $P^{-1} = {}^tP$.

Théorème 24

Soit $e = (e_1, \dots, e_n)$ une base orthonormale de l'espace E .

Un endomorphisme u de E est une isométrie vectorielle si, et seulement si, sa matrice dans la base orthonormale e est orthogonale.

1. On rencontre aussi la notation $O(n)$ pour désigner le groupe orthogonal d'ordre n .

2. Plus généralement, le sujet 2 p. 406 introduit le produit scalaire canonique sur $\mathcal{M}_{n,p}(\mathbb{R})$.

11.3.4 Isométries positives et isométries négatives

Les matrices orthogonales sont de déterminants égaux à 1 ou -1 . Il en est de même pour les isométries vectorielles.

Définition

On dit qu'une matrice orthogonale est *positive* (resp. *négative*) lorsque celle-ci est de déterminant 1 (resp. -1).

On note¹ $\mathrm{SO}_n(\mathbb{R})$ l'ensemble des matrices orthogonales positives de taille n . Il s'agit d'un sous-groupe de $\mathrm{O}_n(\mathbb{R})$ appelé *groupe spécial orthogonal*.

Les matrices de $\mathrm{SO}_n(\mathbb{R})$ correspondent aux matrices de passage entre les bases orthonormales directes d'un espace euclidien orienté.

Définition

On dit qu'une isométrie vectorielle est *positive*² (resp. *négative*) lorsque celle-ci est de déterminant 1 (resp. -1).

Les réflexions sont des isométries négatives.

On note $\mathrm{SO}(E)$ l'ensemble des isométries positives de E . C'est un sous-groupe de $\mathrm{O}(E)$ appelé *groupe spécial orthogonal* de E .

11.3.5 Isométries vectorielles positives en dimension 2

E désigne un *plan euclidien orienté*, c'est-à-dire espace euclidien orienté de dimension 2.

Théorème 25

Les matrices orthogonales positives de $\mathcal{M}_2(\mathbb{R})$ sont exactement les matrices de la forme

$$R(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \text{ avec } \theta \in \mathbb{R}.$$

De plus, ces dernières commutent entre elles.

En conséquence, le groupe $\mathrm{SO}_2(\mathbb{R})$ est commutatif ce qui entraîne le résultat suivant :

Théorème 26

Une isométrie positive d'un plan euclidien orienté a la même matrice dans toute base orthonormale directe.

Définition

Pour $\theta \in \mathbb{R}$, on appelle *rotation d'angle θ* de E l'isométrie positive notée Rot_θ dont la matrice est $R(\theta)$ dans les bases orthonormales directes de E .

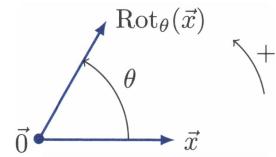
On remarque $R(\theta)R(\theta') = R(\theta + \theta')$ ce qui établit

$$\mathrm{Rot}_\theta \circ \mathrm{Rot}_{\theta'} = \mathrm{Rot}_{\theta+\theta'}.$$

1. On rencontre aussi la notation $\mathrm{SO}(n)$.

2. On parle aussi d'*isométries directes* et *indirectes*.

On peut alors définir la notion d'*angle orienté* entre deux vecteurs \vec{u} et \vec{v} non nuls du plan E : on dit que θ est une *mesure de l'angle orienté*² de \vec{u} à \vec{v} si la rotation d'angle θ transforme le vecteur unitaire $\frac{1}{\|\vec{u}\|} \cdot \vec{u}$ en $\frac{1}{\|\vec{v}\|} \cdot \vec{v}$. On note $\theta \equiv (\vec{u}; \vec{v})$ [2π]. Les rotations conservent les angles orientés.



11.3.6 Isométries vectorielles négatives en dimension 2

E désigne un espace euclidien de dimension 2 (autrement dit, un *plan euclidien*).

Théorème 27

Les matrices orthogonales négatives de $M_2(\mathbb{R})$ sont les matrices de la forme

$$S(\theta) = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix} \quad \text{avec } \theta \in \mathbb{R}.$$

Ces matrices vérifient $S(\theta)^2 = I_2$ ce qui entraîne le résultat suivant :

Théorème 28

Les isométries négatives du plan sont les réflexions, c'est-à-dire les symétries orthogonales par rapport à des droites.

Les réflexions changent les angles orientés en leur opposé.

Les isométries d'un plan euclidien orienté se limitent donc aux rotations et aux réflexions. La composée de deux réflexions est une rotation tandis que la composée d'une rotation et d'une réflexion est une réflexion.

11.4 Exercices d'apprentissage

11.4.1 Produit scalaire

Exercice 1

Pour P et Q éléments de $\mathbb{R}[X]$, on pose

$$(P | Q) = \int_0^1 P(t)Q(t) dt.$$

Montrer que $(\cdot | \cdot)$ définit un produit scalaire sur $\mathbb{R}[X]$.

- À cause de la nature géométrique de ce qui suit, on adopte une notation fléchée des vecteurs.
- Les mesures d'un même angle orienté sont égales modulo 2π et sont figurées par des arcs fléchés allant d'un vecteur à l'autre dans le sens direct ou indirect. Modulo 2π , l'angle orienté est égal à l'écart angulaire lorsque la famille est directe, il est égal à l'opposé sinon.

Solution

Il s'agit d'établir que $(\cdot | \cdot)$ est une forme bilinéaire symétrique et définie positive.

méthode

|| On vérifie que l'application $(\cdot | \cdot)$ est bien définie à valeurs dans \mathbb{R} .

Pour $P, Q \in \mathbb{R}[X]$, la fonction $t \mapsto P(t)Q(t)$ est continue sur le segment $[0 ; 1]$, on peut donc introduire son intégrale ce qui détermine un nombre réel. L'application $(\cdot | \cdot)$ est bien définie de $\mathbb{R}[X] \times \mathbb{R}[X]$ vers \mathbb{R} .

méthode

|| On montre la symétrie puis la linéarité en l'une des variables avant d'affirmer la bilinéarité.

Soit $P, Q \in \mathbb{R}[X]$. On a immédiatement par commutativité de la multiplication la propriété de symétrie

$$(Q|P) = \int_0^1 Q(t)P(t) dt = \int_0^1 P(t)Q(t) dt = (P|Q).$$

Soit $P, Q, R \in \mathbb{R}[X]$ et $\lambda, \mu \in \mathbb{R}$. On obtient par opérations et linéarité de l'intégrale

$$\begin{aligned} (P|\lambda Q + \mu R) &= \int_0^1 P(t)(\lambda Q(t) + \mu R(t)) dt = \lambda \int_0^1 P(t)Q(t) dt + \mu \int_0^1 P(t)R(t) dt \\ &= \lambda(P|Q) + \mu(P|R). \end{aligned}$$

L'application $(\cdot | \cdot)$ est linéaire en sa deuxième variable et donc bilinéaire car symétrique.

méthode

|| On vérifie que $(\cdot | \cdot)$ est définie positive en observant, pour tout $P \in \mathbb{R}[X]$,

$$(P|P) \geq 0 \quad \text{et} \quad (P|P) = 0 \implies P = 0.$$

Soit $P \in \mathbb{R}[X]$. Par intégration bien ordonnée d'une fonction positive, on a

$$(P|P) = \int_0^1 \underbrace{(P(t))^2}_{\geq 0} dt \geq 0.$$

De plus, si $(P|P) = 0$, on peut affirmer que la fonction $t \mapsto (P(t))^2$ est nulle sur $[0 ; 1]$ car continue, positive et d'intégrale nulle. On en déduit que le polynôme P possède une infinité de racines et il s'agit donc du polynôme nul¹.

Finalement, $(\cdot | \cdot)$ détermine un produit scalaire sur $\mathbb{R}[X]$.

Exercice 2

Pour $A, B \in \mathcal{M}_{n,p}(\mathbb{R})$, on pose $\langle A, B \rangle = \text{tr}({}^t A B)$.

(a) Montrer que $\langle \cdot, \cdot \rangle$ définit un produit scalaire sur $\mathcal{M}_{n,p}(\mathbb{R})$.

(b) Montrer que la base constituée des matrices élémentaires de $\mathcal{M}_{n,p}(\mathbb{R})$ est une base orthonormale pour le produit scalaire précédent.

1. On ne peut pas se contenter de conclure que la fonction $t \mapsto P(t)$ est nulle sur $[0 ; 1]$ car ici l'espace d'étude est l'espace des polynômes et non celui des fonctions définies sur $[0 ; 1]$.

Solution

(a) Pour $A, B \in \mathcal{M}_{n,p}(\mathbb{R})$, le produit ${}^t AB$ est bien défini et détermine une matrice carrée de taille p . On peut alors calculer sa trace et affirmer que l'application $\langle \cdot, \cdot \rangle$ est bien définie de $\mathcal{M}_{n,p}(\mathbb{R}) \times \mathcal{M}_{n,p}(\mathbb{R})$ vers \mathbb{R} .

Soit $A, B \in \mathcal{M}_{n,p}(\mathbb{R})$. La trace d'une matrice étant aussi celle de sa transposée, on acquiert la propriété de symétrie par le calcul suivant :

$$\langle B, A \rangle = \text{tr}({}^t BA) = \text{tr}\left({}^t ({}^t BA)\right) = \text{tr}({}^t AB) = \langle A, B \rangle.$$

De plus, pour $A, B, C \in \mathcal{M}_{n,p}(\mathbb{R})$ et $\lambda, \mu \in \mathbb{R}$, on obtient par linéarité de la trace

$$\begin{aligned} \langle A, \lambda B + \mu C \rangle &= \text{tr}({}^t A (\lambda B + \mu C)) = \text{tr}(\lambda {}^t AB + \mu {}^t AC) \\ &= \lambda \text{tr}({}^t AB) + \mu \text{tr}({}^t AC) = \lambda \langle A, B \rangle + \mu \langle A, C \rangle. \end{aligned}$$

L'application $\langle \cdot, \cdot \rangle$ est une forme bilinéaire symétrique. Reste à montrer que celle-ci est définie positive.

méthode

|| On exprime $\langle A, A \rangle$ à l'aide des coefficients de A .

Soit $A \in \mathcal{M}_{n,p}(\mathbb{R})$ de coefficient général $a_{i,j}$. La trace de ${}^t AA$ est la somme de ses coefficients diagonaux

$$\langle A, A \rangle = \text{tr}({}^t AA) = \sum_{j=1}^p [{}^t AA]_{j,j} = \sum_{j=1}^p \left(\sum_{i=1}^n \underbrace{[{}^t A]_{j,i}}_{-a_{i,j}} \underbrace{[A]_{i,j}}_{=a_{i,j}} \right) = \sum_{j=1}^p \left(\sum_{i=1}^n a_{i,j}^2 \right).$$

On en déduit immédiatement $\langle A, A \rangle \geq 0$. De plus, si $\langle A, A \rangle = 0$, on obtient la nullité d'une somme de coefficients positifs et ceux-ci sont donc tous nuls. On en déduit que la matrice A est nulle.

Finalement, $\langle \cdot, \cdot \rangle$ définit un produit scalaire¹ sur $\mathcal{M}_{n,p}(\mathbb{R})$.

(b) méthode

|| $E_{i,j} E_{k,\ell} = \delta_{j,k} E_{i,\ell}$ détermine le produit² deux matrices élémentaires.

Pour (i, j) et (k, ℓ) deux couples de $[1 : n] \times [1 : p]$,

$$\langle E_{i,j}, E_{k,\ell} \rangle = \text{tr}({}^t E_{i,j} E_{k,\ell}) = \text{tr}(E_{j,i} E_{k,\ell}) - \delta_{i,k} \text{tr}(E_{j,\ell}).$$

Or la trace $E_{j,\ell}$ vaut 1 ou 0 selon que la valeur 1 figure ou non sur la diagonale de la matrice élémentaires et donc

$$\langle E_{i,j}, E_{k,\ell} \rangle = \delta_{i,k} \delta_{j,\ell} = \begin{cases} 1 & \text{si } (i, j) = (k, \ell) \\ 0 & \text{sinon.} \end{cases}$$

1. Par un calcul analogue à celui de $\langle A, A \rangle$, on vérifie que $\langle A, B \rangle$ est la somme des produits $a_{i,j} b_{i,j}$. Le produit scalaire introduit s'assimile donc au produit scalaire canonique sur \mathbb{R}^{np} , on dit que c'est le *produit scalaire canonique* de $\mathcal{M}_{n,p}(\mathbb{R})$.

2. Voir sujet 1 p. 321, $\delta_{j,k}$ désigne le symbole de Kronecker.

La famille des matrices élémentaires est donc orthonormale et, puisqu'il s'agit d'une base, on peut parler de base orthonormale.

11.4.2 Base orthonormale

Exercice 3

On munit \mathbb{R}^3 de son produit scalaire canonique.

Orthonormaliser par le procédé de Schmidt la famille de vecteurs (u_1, u_2, u_3) avec

$$u_1 = (1, 1, 0), \quad u_2 = (1, 0, 1) \quad \text{et} \quad u_3 = (1, 1, 1).$$

Solution

méthode

Pour orthonormaliser une famille libre¹ (u_1, \dots, u_n) :

- on pose $v_1 = u_1$;
- on pose $v_2 = u_2 + \lambda v_1$ avec λ tel que $(v_2 | v_1) = 0$;
- on pose $v_3 = u_3 + \lambda v_1 + \mu v_2$ avec λ et μ tels que $(v_3 | v_1) = (v_3 | v_2) = 0$;
- etc.

Enfin², on divise chaque vecteur v_i par sa norme afin d'exprimer les vecteurs e_i .

$$e_1 = \frac{1}{\|v_1\|} v_1, \quad e_2 = \frac{1}{\|v_2\|} v_2, \quad e_3 = \frac{1}{\|v_3\|} v_3, \dots$$

On vérifie que la famille (u_1, u_2, u_3) est libre en étudiant la nullité d'une combinaison linéaire ou, plus rapidement, par le calcul du déterminant de celle-ci dans la base canonique c de \mathbb{R}^3

$$\det_c(u_1, u_2, u_3) = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{vmatrix} \stackrel{L_1 \leftarrow L_1 - L_3}{=} \begin{vmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & 1 \end{vmatrix} \stackrel{L_2 \leftarrow L_2 - L_3}{=} \begin{vmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & 1 \end{vmatrix} = 1 \neq 0.$$

On peut alors appliquer le protocole d'orthonormalisation de Schmidt. On pose

$$v_1 = u_1 = (1, 1, 0).$$

On détermine le vecteur v_2 de la forme $v_2 = u_2 + \lambda v_1$ orthogonal à v_1 :

$$(v_2 | v_1) = 0 \iff \underbrace{(u_2 | v_1)}_{=1} + \lambda \underbrace{(v_1 | v_1)}_{=2} = 0.$$

Pour $\lambda = -1/2$, on obtient

$$v_2 = u_2 - \frac{1}{2} v_1 = \left(\frac{1}{2}, -\frac{1}{2}, 1 \right).$$

1. Si la famille (u_1, \dots, u_n) n'est pas libre, le protocole échoue en déterminant un vecteur v_i nul.
 2. On peut aussi normer les vecteurs au fur et à mesure du calcul.

On détermine le vecteur v_3 de la forme $v_3 = u_3 + \lambda v_1 + \mu v_2$ orthogonal à v_1 et v_2

$$\begin{cases} (v_3 | v_1) = 0 \\ (v_3 | v_2) = 0 \end{cases} \iff \begin{cases} 2 + 2\lambda = 0 \\ 1 + \frac{1}{2}\mu = 0. \end{cases}$$

Pour $\lambda = -1$ et $\mu = -2/3$, on obtient

$$v_3 = u_3 - v_1 - \frac{2}{3}v_2 = \left(-\frac{1}{3}, \frac{1}{3}, \frac{1}{3} \right).$$

Enfin, on divise chaque vecteur v_1, v_2, v_3 par sa norme¹ pour former la famille orthonormale (e_1, e_2, e_3) cherchée

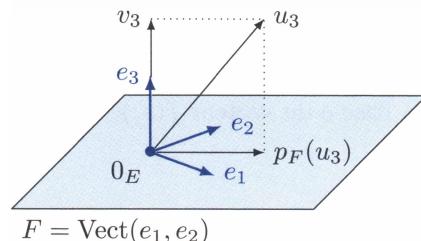
$$e_1 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \right), \quad e_2 = \left(\frac{1}{\sqrt{6}}, \frac{-1}{\sqrt{6}}, \frac{2}{\sqrt{6}} \right) \quad \text{et} \quad e_3 = \left(-\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right).$$

Pour $k \in \llbracket 1 ; n \rrbracket$, le vecteur v_k est en fait donné par

$$e_k = \frac{1}{\|v_k\|} v_k \quad \text{avec} \quad v_k = u_k - \sum_{i=1}^{k-1} (e_i | u_k) e_i$$

où la somme exprimant le vecteur v_k détermine le projeté orthogonal de u_k sur l'espace

$$\text{Vect}(e_1, \dots, e_{k-1}) = \text{Vect}(v_1, \dots, v_{k-1}).$$



Exercice 4

On munit \mathbb{R}^3 de son produit scalaire canonique. Déterminer une base orthonormale de \mathbb{R}^3 dont les deux premiers vecteurs appartiennent au plan

$$P = \{(x, y, z) \in \mathbb{R}^3 \mid x - z = 0\}.$$

Solution

méthode

- || On complète une base orthonormale du plan à l'aide d'un vecteur normal unitaire.

Les vecteurs $u_1 = (0, 1, 0)$ et $u_2 = (1, 0, 1)$ appartiennent au plan P et sont orthogonaux². Les coefficients de x, y, z dans l'équation du plan déterminent les coordonnées 1, 0 et -1 d'un vecteur $u_3 = (1, 0, -1)$ normal au plan P et donc orthogonal aux vecteurs

1. Pour calculer e_2 et e_3 , il est plus commode de diviser $2v_2 = (1, -1, 2)$ et $3v_3 = (-1, 1, 1)$ par leur norme ce qui conduit au même résultat.

2. Si l'on choisit des vecteurs u_1 et u_2 moins pertinents, on peut cependant appliquer le protocole d'orthonormalisation de Schmidt pour former une base orthonormale du plan.

précédents. On forme alors une base (e_1, e_2, e_3) telle que voulue en divisant chacun de ces vecteurs par sa norme

$$e_1 = (0, 1, 0), \quad e_2 = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right) \quad \text{et} \quad e_3 = \left(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}} \right).$$

Exercice 5

Soit f un endomorphisme d'un espace euclidien E de dimension n et $A = (a_{i,j})$ sa matrice dans une base orthonormale $e = (e_1, \dots, e_n)$.

Justifier que $a_{i,j} = (e_i | f(e_j))$ pour tous i et j dans $[1; n]$.

Solution

méthode

|| La i -ème coordonnée d'un vecteur x dans une base orthonormale s'obtient en faisant le produit scalaire avec le i -ème vecteur de base (Th. 10 p. 398).

Par définition de la matrice figurant un endomorphisme dans une base, on a

$$A = \text{Mat}_{(e_1, \dots, e_n)}(f(e_1), \dots, f(e_n)).$$

Le coefficient d'indice (i, j) de la matrice A correspond donc à la i -ème coordonnée dans la base e du vecteur $f(e_j)$. La base e étant supposée orthonormale, on obtient

$$a_{i,j} = (e_i | f(e_j)).$$

11.4.3 Orthogonal d'une partie

Exercice 6

Soit E un espace préhilbertien muni d'un produit scalaire noté $(\cdot | \cdot)$.

(a) Soit A et B deux parties de E avec $A \subset B$. Montrer $B^\perp \subset A^\perp$.

Soit F et G deux sous-espaces vectoriels de E .

(b) Montrer $(F + G)^\perp = F^\perp \cap G^\perp$.

(c) Montrer $(F \cap G)^\perp \supset F^\perp + G^\perp$.

Que devient cette inclusion si l'on suppose que l'espace E est euclidien ?

Solution

(a) Soit $x \in B^\perp$.

méthode

|| On montre que x appartient à A^\perp en observant $(a | x) = 0$ pour tout $a \in A$.

Soit a un élément arbitraire de A . Celui-ci est élément de B car on a supposé A inclus dans B . On a donc $(a | x) = 0$ puisque x est orthogonal à tout vecteur de B . Ainsi, x est élément de A^\perp et l'on peut affirmer l'inclusion $B^\perp \subset A^\perp$.

(b) On raisonne par double inclusion.

Sachant $F \subset F + G$, on a par le résultat précédent $(F + G)^\perp \subset F^\perp$. De même $G \subset F + G$ donne $(F + G)^\perp \subset G^\perp$ et donc $(F + G)^\perp \subset F^\perp \cap G^\perp$.

Inversement, soit $x \in F^\perp \cap G^\perp$. Ce vecteur x est orthogonal à tout vecteur de F et à tout vecteur de G . Il est donc orthogonal aux vecteurs de $F + G$. En effet, si y est un vecteur de $F + G$, on peut écrire $y = a + b$ avec $a \in F$ et $b \in G$ donc

$$(x|y) = (x|a+b) = (\underbrace{x|a}_{=0} + \underbrace{x|b}_{=0}) = 0.$$

Ainsi, $F^\perp \cap G^\perp \subset (F + G)^\perp$ puis on obtient l'égalité voulue par double inclusion.

(c) On sait $F \cap G \subset F$ et donc $F^\perp \subset (F \cap G)^\perp$. De même, on a $G^\perp \subset (F \cap G)^\perp$.

méthode

|| L'orthogonal d'une partie est un sous-espace vectoriel et donc stable par combinaison linéaire.

Puisque $(F \cap G)^\perp$ est un sous-espace vectoriel contenant F^\perp et G^\perp , il contient aussi leur somme : $F^\perp + G^\perp \subset (F \cap G)^\perp$.

Si l'espace E est euclidien¹, on peut transformer cette inclusion en égalité. On peut pour cela raisonner par les dimensions ou, plus efficacement, employer le résultat qui suit :

méthode

|| Lorsque F est un sous-espace vectoriel d'un espace euclidien, $(F^\perp)^\perp = F$.

On emploie l'égalité $(F + G)^\perp = F^\perp \cap G^\perp$ avec les espaces F^\perp et G^\perp au lieu de F et G et l'on obtient

$$(F^\perp + G^\perp)^\perp = (F^\perp)^\perp \cap (G^\perp)^\perp = F \cap G.$$

Il suffit alors de passer à l'orthogonal pour conclure

$$F^\perp + G^\perp = ((F^\perp + G^\perp)^\perp)^\perp = (F \cap G)^\perp.$$

Exercice 7

Soit F et G deux sous-espaces vectoriels supplémentaires et orthogonaux d'un espace préhilbertien E . Montrer $G = F^\perp$.

1. Si l'espace E n'est pas euclidien, l'inclusion peut être stricte (voir sujet 30 p. 432).

Solution

On raisonne par double inclusion. Puisque les espaces F et G sont supposés orthogonaux, les vecteurs de G sont orthogonaux à tous les vecteurs de F et donc¹ $G \subset F^\perp$.

méthode

On montre l'inclusion réciproque² en décomposant un vecteur de F^\perp en la somme d'un vecteur de F et d'un vecteur de G .

Soit x un élément de F^\perp . On peut écrire $x = a + b$ avec $a \in F$ et $b \in G$ car on suppose les espaces F et G supplémentaires dans E . Le vecteur x est orthogonal à tout vecteur de F et en particulier au vecteur a . Or

$$\underbrace{(a|x)}_{=0} = (a|a+b) = \|a\|^2 + \underbrace{(a|b)}_{=0} = \|a\|^2.$$

On a alors $\|a\| = 0$ et donc $a = 0_E$. Ainsi, $x = b$ et x est élément de G . On peut alors affirmer l'inclusion $F^\perp \subset G$ et conclure à l'égalité $G = F^\perp$.

Exercice 8

On munit \mathbb{R}^4 de son produit scalaire canonique.

(a) Former la matrice dans la base canonique de la projection orthogonale p sur l'espace

$$F = \{(x, y, z, t) \in \mathbb{R}^4 \mid x - y - z - t = x - z - 2t = 0\}.$$

(b) Calculer la distance à F du vecteur $(1, 2, 3, 4)$.

Solution**(a) méthode**

On détermine une base orthonormale de F afin de calculer la projection p par la formule du Th. 15 p. 400.

On résout les équations définissant F afin de déterminer une base de cet espace

$$\begin{cases} x - y - z - t = 0 \\ x - z - 2t = 0 \end{cases} \xrightarrow{L_2 \leftarrow L_2 - L_1} \begin{cases} x - y - z - t = 0 \\ y - t = 0 \end{cases} \xrightarrow{\quad} \begin{cases} x = z + 2t \\ y = t \end{cases}$$

On a donc

$$F = \{(z+2t, t, z, t) \mid (z, t) \in \mathbb{R}^2\} = \text{Vect}(\underbrace{(1, 0, 1, 0)}_{=u_1}, \underbrace{(2, 1, 0, 1)}_{=u_2}).$$

1. Une orthogonalité de deux sous-espaces vectoriels F et G exprime une inclusion de l'un dans l'orthogonal de l'autre : $F \subset G^\perp$ ou, et c'est équivalent, $G \subset F^\perp$. L'orthogonalité des espaces ne peut suffire à justifier l'égalité d'un espace avec l'orthogonal de l'autre.

2. Si l'on suppose l'espace E de dimension finie, une conclusion plus rapide est possible par un argument de dimension : $\dim F^\perp = \dim E - \dim F = \dim G$.

La famille (u_1, u_2) constitue une base de l'espace F . Par le procédé de Schmidt, on l'orthonormalise en (e_1, e_2) avec

$$e_1 = \frac{1}{\sqrt{2}}(1, 0, 1, 0) \quad \text{et} \quad e_2 = \frac{1}{2}(1, 1, -1, 1).$$

On peut alors calculer le projeté sur F de n'importe que vecteur v de \mathbb{R}^4 par la formule

$$p(v) = (e_1 | v)e_1 + (e_2 | v)e_2. \quad (*)$$

En particulier, on peut calculer les images des vecteurs de la base canonique et former la matrice¹ A de p dans celle-ci :

$$A = \frac{1}{4} \begin{pmatrix} 3 & 1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 3 & -1 \\ 1 & 1 & -1 & 1 \end{pmatrix}.$$

(b) méthode

|| La distance d'un vecteur à un sous-espace vectoriel est la distance de celui-ci à son projeté orthogonal (Th. 17 p. 401).

La matrice A , ou la formule $(*)$, permet de calculer le projeté orthogonal du vecteur $v = (1, 2, 3, 4)$ sur F . On obtient $p(v) = (3, 1, 1, 1)$. La distance de v à F est alors

$$d(v, F) = \|v - p(v)\| = \|(-2, 1, 2, 3)\| = \sqrt{4 + 1 + 4 + 9} = 3\sqrt{2}.$$

11.4.4 Isométrie et matrices orthogonales

Exercice 9

Vérifier que la matrice suivante est orthogonale :

$$A = \frac{1}{3} \begin{pmatrix} 1 & 2 & 2 \\ -2 & -1 & 2 \\ 2 & -2 & 1 \end{pmatrix}.$$

Solution

méthode

|| On vérifie² que les colonnes (ou les lignes) sont unitaires et deux à deux orthogonales (Th. 22 p. 403).

1. La matrice obtenue est symétrique car, lorsque p est une projection orthogonale, on peut montrer $(p(u) | v) = (u | p(v))$ pour tous vecteurs u et v . En particulier, $(e_i | p(e_j)) = (p(e_i) | e_j) = a_{j,i}$.

2. Vérifier que la matrice est de déterminant ± 1 est une condition nécessaire et mais non suffisante : calculer le déterminant de A n'est ici d'aucune utilité. En revanche, on peut vérifier ${}^t A A = I_n$ ce qui correspond à reproduire des calculs semblables à ceux menés ici.

Notons C_1, C_2 et C_3 les colonnes de A . Le produit scalaire de deux colonnes de coefficients (x, y, z) et (x', y', z') est $xx' + yy' + zz'$. Un calcul direct donne alors

$$(C_1 | C_2) = \frac{1}{9}(2+2-4) = 0 \quad \|C_1\|^2 = \frac{1}{9}(1+4+4) = 1$$

$$(C_2 | C_3) = \frac{1}{9}(4-2-2) = 0 \quad \|C_2\|^2 = \frac{1}{9}(4+1+4) = 1$$

$$(C_3 | C_1) = \frac{1}{9}(2-4+2) = 0 \quad \|C_3\|^2 = \frac{1}{9}(4+4+1) = 1$$

La matrice A est donc orthogonale.

Exercice 10

Soit f une isométrie vectorielle d'un espace euclidien E et F un sous-espace vectoriel de E . Montrer

$$f(F^\perp) = (f(F))^\perp.$$

Solution

méthode

Une isométrie vectorielle conserve le produit scalaire (Th. 19 p. 402) et donc l'orthogonalité.

Vérifions que les espaces $f(F)$ et $f(F^\perp)$ sont orthogonaux. Soit y un élément de $f(F)$ et y' un élément de $f(F^\perp)$. On peut écrire $y = f(x)$ et $y' = f(x')$ avec x et x' éléments de F et F^\perp respectivement. On a alors par conservation du produit scalaire

$$(y | y') = (f(x) | f(x')) = (x | x') = 0$$

car les vecteurs x et x' sont orthogonaux. Ainsi, on peut affirmer que les espaces $f(F)$ et $f(F^\perp)$ sont orthogonaux ce qui signifie l'inclusion¹

$$f(F^\perp) \subset (f(F))^\perp.$$

méthode

Un isomorphisme transforme une famille libre en une famille libre et par conséquent conserve la dimension².

L'isométrie f est un automorphisme donc

$$\dim f(F) = \dim F \quad \text{et} \quad \dim f(F^\perp) = \dim F^\perp.$$

Par suite

$$\begin{aligned} \dim f(F^\perp) &= \dim F^\perp = \dim E - \dim F \\ &= \dim E - \dim f(F) = \dim (f(F))^\perp. \end{aligned}$$

Par inclusion et égalité des dimensions, on conclut

$$f(F^\perp) = (f(F))^\perp.$$

1. Cela signifie aussi l'inclusion $f(F) \subset (f(F^\perp))^\perp$ mais pas encore l'égalité!

2. Voir sujet 5 p. 284.

11.5 Exercices d'entraînement

11.5.1 Produit scalaire

Exercice 11 *

Soit $(x_1, \dots, x_n) \in \mathbb{R}^n$. Montrer l'inégalité qui suit en déterminant le cas d'égalité

$$\left(\sum_{k=1}^n x_k \right)^2 \leq n \sum_{k=1}^n x_k^2.$$

Solution

méthode

On applique l'inégalité de Cauchy-Schwarz (Th. 2 p. 394) à deux vecteurs bien choisis de \mathbb{R}^n .

A partir des vecteurs $x = (x_1, \dots, x_n)$ et $u = (1, \dots, 1)$ de \mathbb{R}^n , l'inégalité de Cauchy-Schwarz $(x|u)^2 \leq \|x\|^2 \|u\|^2$ donne pour le produit scalaire canonique

$$\left(\sum_{k=1}^n x_k \right)^2 = \left(\sum_{k=1}^n x_k \times 1 \right)^2 \leq \left(\sum_{k=1}^n x_k^2 \right) \left(\sum_{k=1}^n 1^2 \right) = n \sum_{k=1}^n x_k^2.$$

De plus, il y a égalité si, et seulement si, les vecteurs $x = (x_1, \dots, x_n)$ et $u = (1, \dots, 1)$ sont colinéaires, c'est-à-dire si, et seulement si, $x_1 = \dots = x_n$.

Exercice 12 *

Soit $A, B \in \mathcal{M}_n(\mathbb{R})$. Établir

$$(\text{tr}(AB))^2 \leq \text{tr}(A^t A) \text{tr}(B^t B).$$

Solution

méthode

On applique l'inégalité de Cauchy-Schwarz en commençant par identifier le produit scalaire concerné.

On sait¹ que l'on définit un produit scalaire sur $\mathcal{M}_n(\mathbb{R})$ en posant

$$\langle A, B \rangle = \text{tr}(A^t B).$$

En appliquant l'inégalité de Cauchy-Schwarz aux éléments $A^t A$ et B , on obtient

$$|\text{tr}(AB)| \leq \sqrt{\text{tr}(A^t A)} \sqrt{\text{tr}(B^t B)}.$$

En élevant au carré et en écrivant $\text{tr}(A^t A) = \text{tr}(A^t A)$ (Th. 16 p. 319), on obtient l'inégalité voulue.

1. Voir sujet 2 p. 406.

Exercice 13 **

Pour P et Q dans $\mathbb{R}[X]$, on pose

$$\varphi(P, Q) = \frac{1}{2\pi} \int_{-\pi}^{\pi} P(e^{i\theta})Q(e^{-i\theta}) d\theta.$$

- (a) Montrer que φ définit un produit scalaire sur $\mathbb{R}[X]$.
- (b) Montrer que $(X^n)_{n \in \mathbb{N}}$ est une base orthonormale pour ce produit scalaire.

Solution(a) **méthode**

On vérifie que l'application φ est à valeurs dans \mathbb{R} puis qu'il s'agit d'une forme bilinéaire symétrique définie positive.

Soit $P, Q \in \mathbb{R}[X]$. Les coefficients des polynômes P et Q étant réels, on a

$$\overline{\varphi(P, Q)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{P(e^{i\theta})Q(e^{-i\theta})} d\theta = \frac{1}{2\pi} \int_{-\pi}^{\pi} P(e^{-i\theta})Q(e^{i\theta}) d\theta = \varphi(Q, P).$$

De plus, le changement de variable $\alpha = -\theta$ donne

$$\varphi(Q, P) = \frac{1}{2\pi} \int_{\theta=-\pi}^{\pi} Q(e^{i\theta})P(e^{-i\theta}) d\theta = \frac{1}{2\pi} \int_{\alpha=-\pi}^{\pi} Q(e^{-i\alpha})P(e^{i\alpha}) d\alpha = \varphi(P, Q).$$

On en déduit, d'une part que $\overline{\varphi(P, Q)} = \varphi(P, Q)$ et l'application φ est à valeurs réelles, d'autre part que l'application φ est symétrique.

Soit $\lambda, \mu \in \mathbb{R}$ et $P, Q, R \in \mathbb{R}[X]$. Par linéarité de l'intégrale, on a immédiatement

$$\begin{aligned} \varphi(P, \lambda Q + \mu R) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} P(e^{i\theta})(\lambda Q(e^{-i\theta}) + \mu R(e^{-i\theta})) d\theta \\ &= \frac{\lambda}{2\pi} \int_{-\pi}^{\pi} P(e^{i\theta})Q(e^{-i\theta}) d\theta + \frac{\mu}{2\pi} \int_{-\pi}^{\pi} P(e^{i\theta})R(e^{-i\theta}) d\theta \\ &= \lambda \varphi(P, Q) + \mu \varphi(P, R). \end{aligned}$$

L'application φ est linéaire en sa deuxième variable et donc bilinéaire.

Enfin, pour $P \in \mathbb{R}[X]$, on a par intégration en bon ordre d'une fonction positive

$$\varphi(P, P) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |P(e^{i\theta})|^2 d\theta \geqslant 0.$$

Aussi, si $\varphi(P, P) = 0$, on peut affirmer par nullité de l'intégrale d'une fonction continue et positive que $P(e^{i\theta}) = 0$ pour tout $\theta \in [-\pi; \pi]$. Le polynôme réel P admet alors une infinité de racines complexes, à savoir tous les complexes de module 1 : c'est le polynôme nul.

On peut alors conclure que φ est un produit scalaire sur $\mathbb{R}[X]$.

(b) La famille $(X^n)_{n \in \mathbb{N}}$ est une base de $\mathbb{R}[X]$. Elle est orthonormale car, pour tous naturels k et ℓ ,

$$\varphi(X^k, X^\ell) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(k-\ell)\theta} d\theta = \begin{cases} 1 & \text{si } k = \ell \\ 0 & \text{sinon.} \end{cases}$$

En effet,

$$\varphi(X^k, X^k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} 1 d\theta = 1 \quad \text{et} \quad \varphi(X^k, X^\ell)_{k \neq \ell} = \left[\frac{1}{i(k-\ell)} e^{i(k-\ell)\theta} \right]_{-\pi}^{\pi} = 0.$$

Exercice 14 ***

Soit $a, b, c, d \in \mathbb{R}$. Pour $u = (x, y) \in \mathbb{R}^2$ et $v = (x', y') \in \mathbb{R}^2$, on pose

$$\varphi(u, v) = axx' + bxy' + cx'y + dyy'.$$

À quelles conditions sur a, b, c, d , l'application φ est-elle un produit scalaire sur \mathbb{R}^2 ?

Solution

L'application φ est bien définie de $\mathbb{R}^2 \times \mathbb{R}^2$ vers \mathbb{R} .

méthode

Dans une analyse, on réunit des conditions sur les réels a, b, c, d en particulier le calcul du produit scalaire à des vecteurs bien choisis.

Analyse : Supposons que φ soit un produit scalaire sur \mathbb{R}^2 . En prenant $u = (1, 0)$ et $v = (0, 1)$, l'égalité de symétrie $\varphi(u, v) = \varphi(v, u)$ donne $b = c$. Aussi, pour $u = (x, y) \in \mathbb{R}^2$,

$$\varphi(u, u) = ax^2 + 2bxy + dy^2.$$

En particulier, pour $u = (1, 0) \neq (0, 0)$, on a $\varphi(u, u) > 0$ et donc $a > 0$. En considérant de nouveau $u = (x, y)$, on peut écrire sachant a non nul

$$\varphi(u, u) = ax^2 + 2bxy + dy^2 = a \underbrace{\left(x + \frac{b}{a}y \right)^2}_{\geq 0} + \underbrace{\frac{ad - b^2}{a}y^2}_{\geq 0}.$$

En prenant $u = (-b, a) \neq (0, 0)$, on a $\varphi(u, u) > 0$ et donc $ad - b^2 > 0$. Vérifions que ces conditions sont suffisantes.

Synthèse : Supposons $b = c$, $a > 0$ et ¹ $ad > b^2$. L'application φ est symétrique et l'on vérifie aisément qu'elle est linéaire en l'une de ses variables donc bilinéaire. De plus, pour tout $u = (x, y) \in \mathbb{R}^2$,

$$\varphi(u, u) = ax^2 + 2bxy + dy^2 = a \underbrace{\left(x + \frac{b}{a}y \right)^2}_{\geq 0} + \underbrace{\frac{ad - b^2}{a}y^2}_{\geq 0} \geq 0.$$

1. Ce qui entraîne que d est aussi strictement positif.

Enfin, si $\varphi(u, u) = 0$, on peut affirmer par nullité d'une somme de quantités positives

$$x + \frac{y}{a} = 0 \quad \text{et} \quad y = 0$$

et donc $u = (0, 0)$. L'application φ est alors un produit scalaire.

11.5.2 Espace euclidien

Exercice 15 **

Soit (e_1, \dots, e_n) une famille de vecteurs unitaires d'un espace euclidien E vérifiant

$$\sum_{k=1}^n (e_k | x)^2 = \|x\|^2 \quad \text{pour tout } x \in E.$$

Montrer que (e_1, \dots, e_n) est une base orthonormale de E .

Solution

Les vecteurs de la famille (e_1, \dots, e_n) sont unitaires. Vérifions qu'ils sont deux à deux orthogonaux. Soit $\ell \in \llbracket 1 ; n \rrbracket$. On applique l'hypothèse au vecteur $x = e_\ell$ et l'on isole le terme d'indice ℓ de la somme

$$\underbrace{\|e_\ell\|^2}_{=1} = \sum_{k=1}^n (e_k | e_\ell)^2 = \underbrace{\sum_{\substack{1 \leq k \leq n \\ k \neq \ell}} (e_k | e_\ell)^2}_{=0} + \underbrace{\|e_\ell\|^4}_{=1}$$

En simplifiant, on obtient

$$\sum_{\substack{1 \leq k \leq n \\ k \neq \ell}} (e_k | e_\ell)^2 = 0.$$

Ceci est une somme nulle de termes tous positifs, ses termes sont donc tous nuls

$$(e_k | e_\ell) = 0 \quad \text{pour tout } \ell \neq k.$$

Ainsi, la famille (e_1, \dots, e_n) est orthogonale donc orthonormale. Il reste à établir que celle-ci est une base.

méthode

|| Une famille orthonormale est assurément libre (Th. 5 p. 396).

Montrons que (e_1, \dots, e_n) est génératrice en étudiant l'orthogonal de $\text{Vect}(e_1, \dots, e_n)$. Soit $x \in \text{Vect}(e_1, \dots, e_n)^\perp$. Le vecteur x est orthogonal à chaque vecteur e_k donc

$$\|x\|^2 = \sum_{k=1}^n \underbrace{(e_k | x)^2}_{=0} = 0.$$

On en déduit x est nul et par conséquent¹ $\text{Vect}(e_1, \dots, e_n)^\perp = \{0_E\}$. En passant à l'orthogonal, on obtient $\text{Vect}(e_1, \dots, e_n)^\perp = E$. Par suite, la famille (e_1, \dots, e_n) est génératrice de E et c'est donc une base orthonormale de E .

Exercice 16 **

Soit f un endomorphisme d'un espace vectoriel euclidien E vérifiant $\langle f(x), x \rangle = 0$ pour tout $x \in E$. Comparer $\text{Ker}(f)$ et $\text{Im}(f)$.

Solution
méthode

- || Manipuler l'hypothèse $\langle f(x), x \rangle = 0$ avec un seul vecteur x n'est pas suffisant : on étudie $\langle f(x+y), x+y \rangle$ pour $x, y \in E$.

Soit x et y quelconques dans E . L'hypothèse d'étude donne $\langle f(x+y), x+y \rangle = 0$. En développant par bilinéarité

$$\begin{aligned} \underbrace{\langle f(x+y), x+y \rangle}_{=0} &= \langle f(x) + f(y), x+y \rangle = \langle f(x) + f(y), x \rangle + \langle f(x) + f(y), y \rangle \\ &= \underbrace{\langle f(x), x \rangle}_{=0} + \langle f(y), x \rangle + \langle f(x), y \rangle + \underbrace{\langle f(y), y \rangle}_{=0} = \langle f(y), x \rangle + \langle f(x), y \rangle. \end{aligned}$$

Si x appartient au noyau de f , ce qui précède donne $\langle f(y), x \rangle = 0$ pour tout $y \in E$ et x est donc orthogonal à tout vecteur de l'image de f . Ainsi,

$$\text{Ker}(f) \subset (\text{Im}(f))^\perp.$$

De plus, par la formule du rang, il y a égalité des dimensions et donc

$$\text{Ker}(f) = (\text{Im}(f))^\perp.$$

11.5.3 Projection orthogonale

Exercice 17 *

On considère un espace vectoriel euclidien E muni d'une base orthonormale (i, j, k) . Former la matrice dans (i, j, k) de la projection orthogonale sur le plan P d'équation

$$x - 2y + z = 0.$$

1. Précisément, on a seulement établi une inclusion $\text{Vect}(e_1, \dots, e_n)^\perp \subset \{0_E\}$ mais l'inclusion réciproque est entendue car le vecteur nul est orthogonal à tout vecteur.

Solution**méthode**

|| L'équation du plan fournit un vecteur normal à partir duquel il est facile d'exprimer la projection.

Les coefficients de x , y et z dans l'équation du plan déterminent les coordonnées 1, -2 et 1 dans (i, j, k) d'un vecteur normal n au plan : $n = i - 2j + k$. On peut alors projeter un vecteur u de E orthogonalement sur la droite normale $D = \text{Vect } n$ par la formule

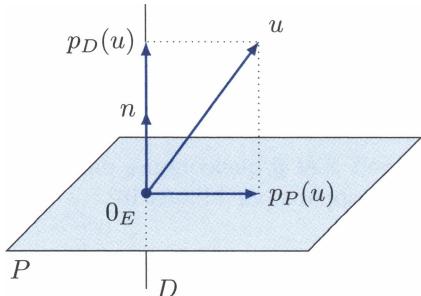
$$p_D(u) = \frac{(n|u)}{\|n\|^2} n.$$

On en déduit la projection de u sur le plan $P = D^\perp$

$$p_P(u) = u - p_D(u) = u - \frac{(n|u)}{\|n\|^2} n.$$

En calculant les images des vecteurs de bases, on forme la matrice de la projection orthogonale sur P :

$$\frac{1}{6} \begin{pmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{pmatrix}.$$

**Exercice 18 ***

On munit $E = \mathcal{M}_n(\mathbb{R})$ du produit scalaire¹ donné par $\langle A, B \rangle = \text{tr}(^t AB)$.

On introduit $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ les sous-espaces des matrices symétriques et antisymétriques de $\mathcal{M}_n(\mathbb{R})$

$$\mathcal{S}_n(\mathbb{R}) = \{M \in \mathcal{M}_n(\mathbb{R}) \mid {}^t M = M\} \quad \text{et} \quad \mathcal{A}_n(\mathbb{R}) = \{M \in \mathcal{M}_n(\mathbb{R}) \mid {}^t M = -M\}.$$

(a) Montrer que $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ sont l'orthogonal l'un de l'autre.

(b) Exprimer le projeté orthogonal sur $\mathcal{S}_n(\mathbb{R})$ d'une matrice M de $\mathcal{M}_n(\mathbb{R})$.

Solution**(a) méthode**

|| On vérifie que les espaces $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ sont orthogonaux avant d'employer un argument de complémentarité.

Soit $S \in \mathcal{S}_n(\mathbb{R})$ et $A \in \mathcal{A}_n(\mathbb{R})$. On a ${}^t S = S$ et ${}^t A = -A$. On en déduit

$$\langle S, A \rangle = \text{tr}(^t SA) = \text{tr}(SA) \quad \text{et} \quad \langle S, A \rangle = \langle A, S \rangle = \text{tr}(^t AS) = \text{tr}(-AS) = -\text{tr}(AS).$$

1. Voir sujet 2 p. 406.

Or $\text{tr}(AS) = \text{tr}(SA)$ (Th. 16 p. 319) et donc $\langle A, S \rangle = 0$.

Ainsi, les espaces $\mathcal{S}_n(\mathbb{R})$ et $\mathcal{A}_n(\mathbb{R})$ sont orthogonaux ce qui signifie que l'on a l'inclusion $\mathcal{A}_n(\mathbb{R}) \subset (\mathcal{S}_n(\mathbb{R}))^\perp$. Or ces espaces sont aussi supplémentaires¹ et donc²

$$\dim \mathcal{A}_n(\mathbb{R}) = \dim \mathcal{M}_n(\mathbb{R}) - \dim \mathcal{S}_n(\mathbb{R}) = \dim (\mathcal{S}_n(\mathbb{R}))^\perp.$$

Par inclusion et égalité des dimensions, on peut affirmer

$$\mathcal{A}_n(\mathbb{R}) = (\mathcal{S}_n(\mathbb{R}))^\perp.$$

Un raisonnement symétrique, ou l'emploi de la formule $(F^\perp)^\perp = F$ (valable pour F sous-espace vectoriel d'un espace euclidien), donne pareillement

$$\mathcal{S}_n(\mathbb{R}) = (\mathcal{A}_n(\mathbb{R}))^\perp.$$

(b) méthode

|| On détermine le projeté orthogonal a d'un vecteur x sur un sous-espace F en écrivant $x = a + b$ avec $a \in F$ et $b \in F^\perp$.

On décompose une matrice $M \in \mathcal{M}_n(\mathbb{R})$ en la somme d'une matrice symétrique et d'une matrice antisymétrique en écrivant

$$M = \underbrace{\frac{1}{2}(M + {}^t M)}_{\in \mathcal{S}_n(\mathbb{R})} + \underbrace{\frac{1}{2}(M - {}^t M)}_{\in \mathcal{A}_n(\mathbb{R}) = (\mathcal{S}_n(\mathbb{R}))^\perp}.$$

Le projeté orthogonal de M sur $\mathcal{S}_n(\mathbb{R})$ est donc $\frac{1}{2}(M + {}^t M)$.

Exercice 19 **

Soit x et y deux vecteurs distincts d'un espace euclidien de dimension supérieure à 2.

(a) On suppose $(x|y) = \|y\|^2$. Montrer qu'il existe un unique hyperplan H de E tel que $y = p_H(x)$.

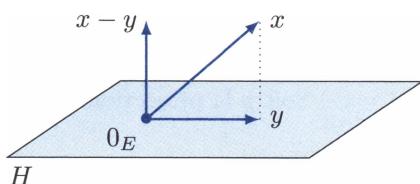
(b) À quelle condition existe-t-il une réflexion échangeant x et y ?

Solution

méthode

|| On caractérise un hyperplan par l'obtention d'un vecteur normal à celui-ci.

(a) Analyse : Supposons que H soit un hyperplan solution. Puisque le vecteur y est le projeté de x sur y , le vecteur différence $x - y$ est orthogonal à H . Or celui-ci est non nul car on a supposé les vecteurs x et y distincts. Le vecteur $x - y$ est donc un vecteur normal à H . On en déduit $H = \{x - y\}^\perp$.



1. Voir sujet 3 p. 322.
2. Voir aussi le sujet 7 p. 411.

Synthèse : Soit H l'hyperplan de vecteur normal $y - x$. L'hypothèse $(x | y) = (y | y)$ donne $(x - y | y) = 0$ ce qui assure que $y \in H$. De plus, on peut écrire

$$x = \underbrace{y}_{\in H} + \underbrace{(x - y)}_{\in H^\perp}$$

et l'on a donc $p_H(x) = y$.

(b) *Analyse :* Supposons qu'il existe une réflexion¹ σ d'hyperplan H échangeant x et y . Puisqu'une réflexion conserve la norme, on a nécessairement $\|x\| = \|y\|$. De plus, on a par linéarité $\sigma(x - y) = \sigma(x) - \sigma(y) = y - x$ et le vecteur $x - y$ est changé en son opposé. Or seuls les vecteurs de la droite normale à l'hyperplan de réflexion H sont changés en leur opposé. On a donc $H = \{x - y\}^\perp$.

Synthèse : Supposons $\|x\| = \|y\|$ et considérons la réflexion σ d'hyperplan H de vecteur normal $x - y$. Le vecteur $x + y$ appartient à H car

$$(x + y | x - y) = \|x\|^2 - \|y\|^2 = 0.$$

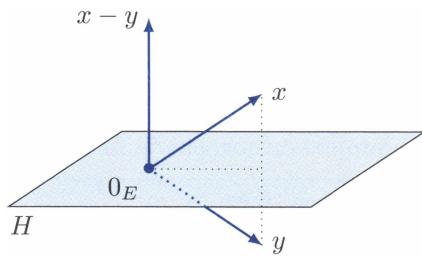
On peut alors écrire

$$x = \underbrace{\frac{1}{2}(x + y)}_{\in H} + \underbrace{\frac{1}{2}(x - y)}_{\in H^\perp}$$

et donc

$$\sigma(x) = \frac{1}{2}(x + y) - \frac{1}{2}(x - y) = y$$

puis $\sigma(y) = \sigma^2(x) = x$.



En résumé, il existe une réflexion échangeant x et y si, et seulement si, $\|x\| = \|y\|$. De plus, si tel est le cas, celle-ci est unique.

Exercice 20 ***

Soit p une projection vectorielle d'un espace euclidien E .

Montrer que la projection p est orthogonale si, et seulement si,

$$\|p(x)\| \leq \|x\| \quad \text{pour tout } x \in E.$$

Solution

(\Rightarrow) Soit p la projection orthogonale sur un sous-espace vectoriel F .

méthode

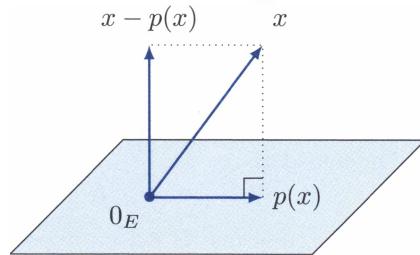
|| On utilise le théorème de Pythagore.

1. Rappelons qu'une réflexion est par définition une symétrie orthogonale par rapport à un hyperplan.

Soit $x \in E$. On écrit $x = p(x) + (x - p(x))$ avec les vecteurs $p(x)$ et $x - p(x)$ orthogonaux car $x \in F$ et $x - p(x) \in F^\perp$. Le théorème de Pythagore donne alors

$$\|x\|^2 = \|p(x)\|^2 + \|x - p(x)\|^2 \geq \|p(x)\|^2$$

et donc $\|p(x)\| \leq \|x\|$.



(\Leftarrow) Soit p une projection vérifiant $\|p(x)\| \leq \|x\|$ pour tout $x \in E$.

méthode

On montre l'orthogonalité de l'espace F sur lequel p projette avec l'espace G parallèlement auquel p opère.

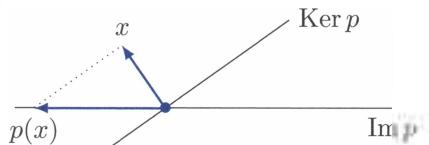
Puisque p est une projection vectorielle, les espaces $F = \text{Im}(p)$ et $G = \text{Ker}(p)$ sont supplémentaires et p est la projection sur F parallèlement à G . Soit $a \in F$ et $b \in G$. Pour tout réel λ , considérons le vecteur¹ $x = a + \lambda b$. On a $p(x) = a$ et l'hypothèse $\|p(x)\| \leq \|x\|$ donne

$$\|a\|^2 \leq \|a + \lambda b\|^2 = \|a\|^2 + 2\lambda(a|b) + \lambda^2\|b\|^2.$$

Après simplification, on obtient l'inéquation $2\lambda(a|b) + \lambda^2\|b\|^2 \geq 0$ valable pour tout réel λ . Si $(a|b) \neq 0$, on peut écrire l'absurdité²

$$\underbrace{2\lambda(a|b) + \lambda^2\|b\|^2}_{\geq 0} \underset{\lambda \rightarrow 0}{\sim} \underbrace{2\lambda(a|b)}_{\text{change de signe}}.$$

On en déduit $(a|b) = 0$. Ainsi, les espaces F et G sont orthogonaux. Or ceux-ci sont aussi supplémentaires, ils sont donc l'orthogonal³ l'un de l'autre. On peut alors conclure que p est une projection orthogonale.



Lorsqu'une projection n'est pas orthogonale, le projeté d'un vecteur peut être plus long que le vecteur initial.

1. On peut aussi introduire le vecteur $y = \lambda a + b$ pour lequel $2\lambda(a|b) + \|b\|^2 \geq 0$ pour tout réel λ . Ceci oblige $(a|b) = 0$ car la fonction affine $\lambda \mapsto 2\lambda(a|b) + \|b\|^2$ change de signe dans le cas contraire.

2. On peut aussi affirmer $(a|b) = 0$ en observant que le trinôme $\|b\|^2\lambda^2 + 2(a|b)\lambda$ est de discriminant négatif car ne peut pas posséder deux racines réelles distinctes.

3. Voir sujet 7 p. 411.

11.5.4 Calcul de distance à un sous-espace vectoriel

Exercice 21 *

On munit $E = \mathcal{M}_n(\mathbb{R})$ du produit scalaire canonique¹ donné par $\langle A, B \rangle = \text{tr}({}^t AB)$.

On considère

$$H = \{M \in \mathcal{M}_n(\mathbb{R}) \mid \text{tr}(M) = 0\}.$$

- (a) Justifier que H est un hyperplan et déterminer un vecteur normal de H .
- (b) Exprimer simplement la distance à H d'une matrice M de $\mathcal{M}_n(\mathbb{R})$.

Solution

(a) méthode

|| On exprime la condition $\text{tr}(M) = 0$ sous la forme $\langle A, M \rangle = 0$.

Pour tout $M \in \mathcal{M}_n(\mathbb{R})$, on remarque $\text{tr}(M) = \text{tr}({}^t \mathbf{I}_n M) = \langle \mathbf{I}_n, M \rangle$. On en déduit

$$H = \{M \in \mathcal{M}_n(\mathbb{R}) \mid \langle \mathbf{I}_n, M \rangle = 0\} = \{\mathbf{I}_n\}^\perp \quad \text{avec} \quad \mathbf{I}_n \neq \mathbf{O}_n.$$

Ainsi, H est l'hyperplan défini par le vecteur normal \mathbf{I}_n .

(b) méthode

|| La distance d'un vecteur à un sous-espace vectoriel est la distance de celui-ci à son projeté orthogonal (Th. 17 p. 401).

Le projeté orthogonal de M sur l'hyperplan H se déduit du projeté sur la droite normale $D = \text{Vect } \mathbf{I}_n$:

$$p_H(M) = M - p_D(M) \quad \text{avec} \quad p_D(M) = \frac{\langle \mathbf{I}_n, M \rangle}{\|\mathbf{I}_n\|^2} \mathbf{I}_n$$

et donc²

$$d(M, H) = \|M - p_H(M)\| = \|p_D(M)\| = \frac{|\langle \mathbf{I}_n, M \rangle|}{\|\mathbf{I}_n\|} = \frac{1}{\sqrt{n}} |\text{tr}(M)|.$$

Exercice 22 **

Calculer

$$m = \inf_{(a,b) \in \mathbb{R}^2} \int_0^1 (t^2 - (at + b))^2 dt.$$

1. Voir sujet 2 p. 406.

2. Plus généralement, les mêmes calculs montrent que la distance d'un vecteur x à un hyperplan de vecteur normal n est donnée par la formule $d(x, H) = \frac{|\langle n, x \rangle|}{\|n\|}$.

Solution**méthode**

On interprète la borne inférieure m en fonction de la distance d'un vecteur à un sous-espace vectoriel pour un produit scalaire à déterminer.

On introduit le produit scalaire¹ sur $\mathbb{R}_2[X]$ donné par

$$(P|Q) = \int_0^1 P(t)Q(t) dt.$$

En introduisant la norme euclidienne associée

$$\int_0^1 (t^2 - (at + b))^2 dt = \|X^2 - (aX + b)\|^2$$

et donc

$$m = \inf_{(a,b) \in \mathbb{R}^2} \|X^2 - (aX + b)\|^2 = \left(\inf_{(a,b) \in \mathbb{R}^2} \|X^2 - (aX + b)\| \right)^2 = \left(d(X^2, \mathbb{R}_1[X]) \right)^2.$$

Il reste à déterminer le projeté orthogonal $aX + b$ de X^2 sur $\mathbb{R}_1[X]$. Plutôt que d'orthonormaliser par le procédé de Schmidt² la base $(1, X)$, on détermine a et b de sorte que

$$X^2 - (aX + b) \in (\mathbb{R}_1[X])^\perp.$$

Ceci revient à chercher a et b tels que $X^2 - (aX + b)$ est orthogonal³ à 1 et X :

$$\begin{cases} (X^2 - (aX + b)|1) = 0 \\ (X^2 - (aX + b)|X) = 0 \end{cases} \iff \begin{cases} \frac{1}{2}a + b = \frac{1}{3} \\ \frac{1}{3}a + \frac{1}{2}b = \frac{1}{6} \end{cases}$$

On obtient $a = 1$ et $b = -1/6$. Le projeté orthogonal de X^2 sur $\mathbb{R}_1[X]$ est donc $X - 1/6$ et l'on peut conclure

$$\begin{aligned} m &= \left\| X^2 - X + \frac{1}{6} \right\|^2 \\ &= \left(X^2 - X + \frac{1}{6} | X^2 \right) + \left(X^2 - X + \frac{1}{6} | - X + \underbrace{\frac{1}{6}}_{\in \mathbb{R}_1[X]} \right) \\ &= \left(\frac{1}{5} - \frac{1}{4} + \frac{1}{18} \right) + 0 = \frac{1}{180}. \end{aligned}$$

1. Dans le sujet 1 p. 405 on a vu que l'expression proposée définit un produit scalaire sur $\mathbb{R}[X]$, elle définit alors a fortiori un produit scalaire sur le sous-espace vectoriel $\mathbb{R}_2[X]$.

2. Cette orthonormalisation produit les polynômes $P_0 = 1$ et $P_1 = \sqrt{3}(2X - 1)$.

3. Car $\text{Vect}(1, X)^\perp = \{1, X\}^\perp$.

11.5.5 Isométries

Exercice 23 *

Soit s_1 et s_2 deux réflexions de droites D_1 et D_2 d'un plan euclidien orienté E . On introduit \vec{u}_1 et \vec{u}_2 des vecteurs directeurs des droites D_1 et D_2 et l'on pose θ une mesure de l'angle orienté de \vec{u}_1 à \vec{u}_2 . Préciser la composée $s_2 \circ s_1$.

Solution

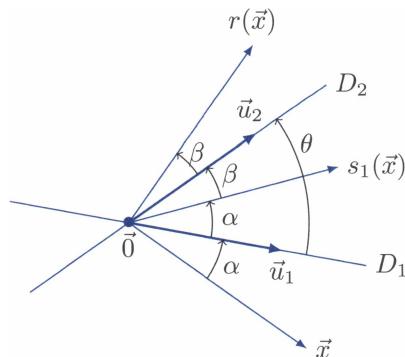
méthode

|| Les isométries de E sont exclusivement les rotations (isométries positives) et les réflexions (isométries négatives).

Par composition de deux isométries négatives, $r = s_2 \circ s_1$ est une isométrie positive, c'est-à-dire une rotation (Th. 26 p. 404). Déterminons l'angle de celle-ci en mesurant l'angle orienté de \vec{u}_1 à $r(\vec{u}_1)$.

Le vecteur \vec{u}_1 appartient à la droite de réflexion D_1 . On a donc $s_1(\vec{u}_1) = \vec{u}_1$ puis $r(\vec{u}_1) = s_2(\vec{u}_1)$. Par la relation de Chasles

$$\begin{aligned} (\vec{u}_1; r(\vec{u}_1)) &\equiv (\vec{u}_1; s_2(\vec{u}_1)) \\ &\equiv (\vec{u}_1; \vec{u}_2) + (\vec{u}_2; s_2(\vec{u}_1)) \quad [2\pi]. \end{aligned}$$



Sachant qu'une réflexion change les angles orientés en leur opposé, on écrit

$$(\vec{u}_1; r(\vec{u}_1)) \equiv (s_2(\vec{u}_2); s_2(\vec{u}_1)) \equiv -(\vec{u}_2; \vec{u}_1) \equiv (\vec{u}_1; \vec{u}_2) \quad [2\pi]$$

donc

$$(\vec{u}_1; r(\vec{u}_1)) \equiv 2(\vec{u}_1; \vec{u}_2) \equiv 2\theta \quad [2\pi].$$

On peut conclure que r est la rotation d'angle 2θ .

Exercice 24 *

Soit r une rotation et s une réflexion d'un plan euclidien orienté E .

Simplifier les composées $s \circ r \circ s$ et $r \circ s \circ r$.

Solution

méthode

|| Les composées $s \circ r$ et $r \circ s$ sont des isométries négatives.

La composée $s \circ r$ est une réflexion (Th. 28 p. 405) donc $(s \circ r)^2 = \text{Id}_E$ ce qui donne $s \circ r \circ s \circ r = \text{Id}_E$. En composant par r^{-1} à droite, on obtient $s \circ r \circ s = r^{-1}$. En composant par s^{-1} à gauche, il vient $r \circ s \circ r = s^{-1} = s$.

Exercice 25 **

Soit f une isométrie d'un espace euclidien E et $g = f - \text{Id}_E$.

(a) Montrer l'égalité $\text{Im}(g) = (\text{Ker}(g))^\perp$.

On note p la projection orthogonale sur $\text{Ker}(g)$ et l'on introduit pour tout $n \in \mathbb{N}^*$

$$p_n = \frac{1}{n} (\text{Id}_E + f + f^2 + \cdots + f^{n-1}).$$

(b) Démontrer que, pour tout $x \in E$, $\lim_{n \rightarrow +\infty} \|(p_n - p)(x)\| = 0$.

Solution**(a) méthode**

À l'aide d'un argument de dimension, il suffit d'établir que les espaces $\text{Ker}(g)$ et $\text{Im}(g)$ sont orthogonaux.

Soit x un élément de $\text{Ker}(g)$ et y dans $\text{Im}(g)$. On a $g(x) = 0_E$ et donc $f(x) = x$. Aussi, on peut écrire $y = g(a) = f(a) - a$ avec $a \in E$. Par bilinéarité du produit scalaire

$$(x|y) = (x|f(a) - a) = (x|f(a)) - (x|a) = (f(x)|f(a)) - (x|a).$$

Or f est une isométrie et donc conserve le produit scalaire (Th. 19 p. 402)

$$(f(x)|f(a)) = (x|a).$$

On a alors $(x|y) = 0$. Les espaces $\text{Ker}(g)$ et $\text{Im}(g)$ sont donc orthogonaux ce qui produit l'inclusion $\text{Im}(g) \subset (\text{Ker}(g))^\perp$. Enfin, par la formule du rang,

$$\text{rg}(g) = \dim E - \dim \text{Ker}(g) = \dim(\text{Ker}(g))^\perp$$

et l'on peut conclure $\text{Im}(g) = (\text{Ker}(g))^\perp$ par inclusion et égalité des dimensions.

(b) Soit $x \in E$. La projection p est la projection sur $\text{Ker}(g)$ parallèlement à son orthogonal $(\text{Ker}(g))^\perp = \text{Im}(g)$.

méthode

On exprime x comme la somme d'un vecteur de $\text{Ker}(g)$ et d'un vecteur de $\text{Im}(g)$.

On écrit $x = a + b$ avec $a = p(x) \in \text{Ker}(g)$ et $b \in \text{Im}(g)$.

D'une part, on a $g(a) = 0_E$ donc $f(a) = a$ ce qui donne

$$p_n(a) = \underbrace{\frac{1}{n} (a + f(a) + \cdots + f^{n-1}(a))}_{=a} = a.$$

D'autre part, il existe un vecteur c tel que $b = g(c)$, c'est-à-dire $b = f(c) - c$, et l'on observe alors un télescopage

$$p_n(b) = \frac{1}{n}(\text{Id} + f + f^2 + \cdots + f^{n-1}) \circ (f - \text{Id}_E)(c) = \frac{1}{n}(f^n(c) - c).$$

Par linéarité $p_n(x) = p_n(a) + p_n(b)$ et donc

$$\|(p_n - p)(x)\| = \frac{1}{n}\|f^n(c) - c\|.$$

Enfin, puisque l'isométrie f conserve la norme, on a $\|f^n(c)\| = \|c\|$ donc

$$\|(p_n - p)(x)\| \underset{n \rightarrow +\infty}{\longrightarrow} 0.$$

Exercice 26 **

Soit f un endomorphisme d'un espace euclidien E conservant l'orthogonalité, c'est-à-dire vérifiant, pour tous x et y de E ,

$$(x|y) = 0 \implies (f(x)|f(y)) = 0.$$

Montrer qu'il existe $\lambda \in \mathbb{R}_+$ et $\varphi \in O(E)$ tels que $f = \lambda\varphi$.

Solution

méthode

On montre que les vecteurs unitaires sont envoyés sur des vecteurs de normes égales à l'aide de l'identité remarquable $(a+b|a-b) = \|a\|^2 - \|b\|^2$.

Soit x et y deux vecteurs unitaires de E . Les vecteurs $x+y$ et $x-y$ sont orthogonaux car

$$(x+y|x-y) = \|x\|^2 - \|y\|^2 = 0$$

Les vecteurs images $f(x+y)$ et $f(x-y)$ sont donc orthogonaux et l'on obtient par linéarité de f

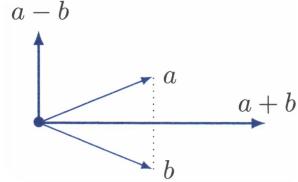
$$\underbrace{(f(x+y)|f(x-y))}_{=0} = (f(x)+f(y)|f(x)-f(y)) = \|f(x)\|^2 - \|f(y)\|^2.$$

On en déduit que les images par f des vecteurs unitaires de E sont tous de même norme. Notons $\lambda \in \mathbb{R}_+$ cette valeur commune et montrons alors $\|f(x)\| = \lambda \|x\|$ pour tout $x \in E$.

méthode

Lorsqu'il n'est pas nul, on divise x par sa norme pour le transformer en un vecteur unitaire.

Si le vecteur x est nul, la propriété $\|f(x)\| = \lambda \|x\|$ est immédiate.



Si le vecteur x est non nul, on peut introduire le vecteur unitaire

$$y = \frac{1}{\|x\|} x.$$

Par linéarité de f , on a

$$\|f(y)\| = \left\| f\left(\frac{1}{\|x\|}x\right) \right\| = \left\| \frac{1}{\|x\|}f(x) \right\| = \frac{1}{\|x\|}\|f(x)\|.$$

L'égalité $\|f(y)\| = \lambda$ donne alors $\|f(x)\| = \lambda \|x\|$.

On peut ensuite conclure en discutant selon l'éventuelle nullité de λ :

Cas : $\lambda = 0$. L'endomorphisme f est nul et l'on a $f = \lambda\varphi$ avec $\varphi = \text{Id}_E \in O(E)$.

Cas : $\lambda \neq 0$. On peut introduire

$$\varphi = \frac{1}{\lambda}f \quad \text{vérifiant} \quad \|\varphi(x)\| = \|x\| \quad \text{pour tout } x \in E.$$

L'application φ est linéaire et conserve la norme, c'est une isométrie vectorielle. On peut donc à nouveau écrire $f = \lambda\varphi$ avec $\varphi \in O(E)$.

Exercice 27 **

Soit E un espace euclidien de dimension $n \geq 1$ et f une fonction de E vers E vérifiant

$$f(0_E) = 0_E \quad \text{et} \quad \forall (x, y) \in E^2, \quad \|f(x) - f(y)\| = \|x - y\|.$$

- (a) Montrer que $\|f(x)\| = \|x\|$ pour tout x de E .
- (b) Etablir $(f(x)|f(y)) = (x|y)$ pour tous x et y de E .
- (c) En introduisant une base orthonormale de E , établir que l'application f est linéaire.

Solution

(a) L'égalité $\|f(x) - f(y)\| = \|x - y\|$ avec $y = 0_E$ donne directement $\|f(x)\| = \|x\|$.

(b) méthode

On exploite l'identité remarquable $\|a - b\|^2 = \|a\|^2 - 2(a|b) + \|b\|^2$.

Soit x et y deux éléments de E . En développant les deux membres de l'égalité

$$\|f(x) - f(y)\|^2 = \|x - y\|^2$$

on obtient

$$\|f(x)\|^2 - 2(f(x)|f(y)) + \|f(y)\|^2 = \|x\|^2 - 2(x|y) + \|y\|^2.$$

On simplifie cette relation en exploitant $\|f(x)\| = \|x\|$ et $\|f(y)\| = \|y\|$ pour obtenir l'égalité voulue $(f(x)|f(y)) = (x|y)$.

(c) Considérons $e = (e_1, \dots, e_n)$ une base orthonormale de E . En vertu de la conservation du produit scalaire obtenue ci-dessus, la famille $e' = (e'_1, \dots, e'_n)$ constituée des vecteurs $e'_i = f(e_i)$ est aussi une base orthonormale de E .

méthode

|| On peut exprimer les coordonnées d'un vecteur dans une base orthonormale à l'aide du produit scalaire (Th. 10 p. 398).

Soit x un vecteur de E . On peut écrire $f(x)$ à l'aide de ses coordonnées dans la base orthonormale e'

$$f(x) = \sum_{i=1}^n (e_i | f(x)) e'_i = \sum_{i=1}^n (f(e_i) | f(x)) e'_i = \sum_{i=1}^n (e_i | x) e'_i.$$

Par linéarité du produit scalaire en la deuxième variable, l'expression ci-dessus assure la linéarité de f . Finalement, f est une isométrie vectorielle¹.

11.5.6 Matrices orthogonales

Exercice 28 **

Soit $A = (a_{i,j}) \in O_n(\mathbb{R})$. Montrer

$$\sum_{1 \leq i, j \leq n} |a_{i,j}| \leq n\sqrt{n} \quad \text{et} \quad \left| \sum_{1 \leq i, j \leq n} a_{i,j} \right| \leq n.$$

Solution

méthode

|| La présence d'une racine peut faire penser à l'inégalité de Cauchy-Schwarz.

Soit $i \in \llbracket 1 ; n \rrbracket$. Par l'inégalité de Cauchy-Schwarz

$$\sum_{j=1}^n |a_{i,j}| = \sum_{j=1}^n |a_{i,j}| \times 1 \leq \left(\sum_{j=1}^n a_{i,j}^2 \right)^{1/2} \left(\sum_{j=1}^n 1^2 \right)^{1/2}. \quad (*)$$

La matrice A étant orthogonale, ses lignes sont unitaires et donc

$$\sum_{j=1}^n a_{i,j}^2 = 1 \quad \text{pour tout } i \in \llbracket 1 ; n \rrbracket.$$

La relation (*) se relit alors

$$\sum_{j=1}^n |a_{i,j}| \leq \sqrt{n}.$$

¹. En exploitant une translation pour ramener l'image de 0_E en 0_E , cette étude assure qu'une application $f: E \rightarrow E$ qui conserve la distance euclidienne est la composée d'une isométrie et d'une translation.

Il suffit ensuite de sommer ces comparaisons pour i allant de 1 à n pour affirmer la première inégalité demandée

$$\sum_{1 \leq i \leq n} |a_{i,j}| \leq n\sqrt{n}.$$

La seconde inégalité est plus délicate à obtenir.

méthode

|| La somme des $a_{i,j}$ est égale à ${}^t X A X$ pour X la colonne dont tous les coefficients sont égaux à 1.

Rappelons que le produit scalaire canonique sur l'espace $\mathcal{M}_{n,n}(\mathbb{R})$ des colonnes est donné par la formule

$$\langle X, Y \rangle = {}^t X Y.$$

On peut donc écrire

$$\sum_{1 \leq i \leq n} a_{i,j} = {}^t X A X = \langle X, AX \rangle.$$

Par l'inégalité de Cauchy-Schwarz,

$$|\langle X, AX \rangle| \leq \|X\| \|AX\|.$$

Or

$$\|AX\|^2 = \langle AX, AX \rangle = {}^t (AX) AX = {}^t X {}^t A A X - {}^t X X = \|X\|^2.$$

On peut alors conclure

$$\left| \sum_{1 \leq i \leq n} a_{i,j} \right| \leq \|X\| = n.$$

Exercice 29 ***

Soit M une matrice orthogonale de taille $n = p + q$ que l'on écrit par blocs

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \quad \text{avec} \quad A \in \mathcal{M}_p(\mathbb{R}) \text{ et } D \in \mathcal{M}_q(\mathbb{R}).$$

Montrer

$$(\det(A))^2 = (\det(D))^2.$$

Solution

La matrice M étant orthogonale, on a ${}^t M M = I_n$, c'est-à-dire

$$\begin{pmatrix} {}^t A A + {}^t C C & {}^t A B + {}^t C D \\ {}^t B A + {}^t D C & {}^t B B + {}^t D D \end{pmatrix} = \begin{pmatrix} I_p & 0 \\ 0 & I_q \end{pmatrix}.$$

méthode

|| On introduit une matrice triangulaire par blocs N tel que le produit $N M$ soit aussi triangulaire par blocs en vertu de l'égalité qui précède.

Considérons la matrice¹

$$N = \begin{pmatrix} {}^t A & {}^t C \\ 0 & I_q \end{pmatrix}.$$

On calcule le produit NM par blocs ce qui donne

$$NM = \begin{pmatrix} {}^t AA + {}^t CC & {}^t AB + {}^t CD \\ C & D \end{pmatrix} = \begin{pmatrix} I_p & 0 \\ C & D \end{pmatrix}.$$

En passant cette relation au déterminant, on obtient

$$\det(A) \times \det(M) = \det(D).$$

En effet, le déterminant d'une matrice triangulaire par blocs est le produit des déterminants des blocs diagonaux (Th. 16 p. 359) ce qui donne $\det(N) = \det({}^t A) = \det(A)$ et de même $\det(NM) = \det(D)$.

Puisque le déterminant d'une matrice orthogonale vaut 1 ou -1 , on peut conclure

$$\det(A) = \pm \det(D).$$

11.6 Exercices d'approfondissement

Exercice 30 **

On munit l'espace $E = C([-1; 1], \mathbb{R})$ du produit scalaire défini par

$$\langle f, g \rangle = \int_{-1}^1 f(t)g(t) dt.$$

On pose

$$F = \{f \in E \mid \forall t \in [-1; 0], f(t) = 0\} \quad \text{et} \quad G = \{g \in E \mid \forall t \in [0; 1], g(t) = 0\}.$$

- (a) Montrer que $F^\perp = G$.
- (b) Les sous-espaces vectoriels F et F^\perp sont-ils supplémentaires ?
- (c) Comparer $F^\perp + G^\perp$ et $(F \cap G)^\perp$.

Solution

- (a) Soit $f \in F$ et $g \in G$. Par la relation de Chasles

$$\langle f, g \rangle = \int_{-1}^0 \underbrace{f(t)}_{=0} g(t) dt + \int_0^1 f(t) \underbrace{g(t)}_{=0} dt = 0.$$

1. La matrice $N' = \begin{pmatrix} I_p & 0 \\ 0 & D \end{pmatrix}$ convient aussi.

Les sous-espaces vectoriels F et G sont donc orthogonaux ce qui permet d'affirmer l'inclusion $G \subset F^\perp$.

Inversement, soit $g \in F^\perp$. Montrons que $g(x) = 0$ pour tout $x \in [0; 1]$. Par l'absurde, on suppose $g(x_0) \neq 0$ pour un certain $x_0 \in]0; 1[$.

méthode

On exploite la continuité de g en x_0 pour définir une fonction f de F telle que $\langle f, g \rangle \neq 0$ en tant qu'intégrale d'une fonction continue positive qui n'est pas la fonction nulle.

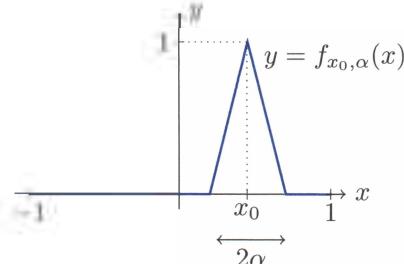
Quitte à considérer la fonction $-g$, on peut supposer $g(x_0) > 0$. Par continuité de g au point x_0 , il existe $\alpha > 0$ tel que

$$[x - \alpha; x + \alpha] \subset]0; 1[\quad \text{et} \quad g(t) > 0 \quad \text{pour tout } [x - \alpha; x + \alpha].$$

Considérons alors la fonction $f = f_{x_0, \alpha}$ définie par le schéma ci-contre.

La fonction f appartient à F et le produit fg est une fonction continue, positive mais qui n'est pas la fonction nulle. On a donc

$$\int_{-1}^1 f(t)g(t) dt > 0.$$



C'est absurde car on a supposé g élément de F^\perp . On a donc $g(x) = 0$ pour tout x de $]0; 1[$ puis aussi pour $x = 0$ et $x = 1$ par continuité de g en ces points.

Finalement, $g \in G$ et l'on peut affirmer $F^\perp \subset G$ puis l'égalité par double inclusion.

(b) Les sous-espaces vectoriels F et $F^\perp = G$ ne peuvent pas être supplémentaires car toute fonction de la somme $F + G$ s'annule en 0.

(c) Comme au-dessus, on montre $G^\perp = F$ et donc $F^\perp + G^\perp = G + F$ est un espace constitué de fonctions s'annulant toutes en 0. Cependant, $(F \cap G)^\perp = \{0\}^\perp = E$. On a donc l'inclusion $F^\perp + G^\perp \subset (F \cap G)^\perp$ et celle-ci est stricte¹.

Exercice 31 ** (Famille équiangulaire)

Soit u_1, \dots, u_n des vecteurs unitaires d'un espace euclidien E de dimension $n \geq 2$ dont le produit scalaire est noté $\langle \cdot, \cdot \rangle$. On suppose qu'il existe un réel c tel que $\langle u_i, u_j \rangle = c$ pour tous les indices i et j distincts dans $[1; n]$.

Pour quelles valeurs de c peut-on affirmer que la famille (u_1, \dots, u_n) est liée ?

1. Celle-ci est toujours vraie, voir sujet 6 p. 410.

2. On peut aussi remarquer que, pour $H = F + G$, $H^\perp = F^\perp \cap G^\perp = \{0\}$ et donc H est strictement inclus dans $(H^\perp)^\perp$.

Solution**méthode**

|| On étudie la liberté de la famille (u_1, \dots, u_n) .

Soit $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$ tel que $\lambda_1 u_1 + \dots + \lambda_n u_n = 0_E$. En faisant le produit scalaire des deux membres de cette équation avec le vecteur u_i , on obtient pour tout $i \in [1; n]$

$$c\lambda_1 + \dots + c\lambda_{i-1} + \lambda_i + c\lambda_{i+1} + \dots + c\lambda_n = 0.$$

Considérons alors la matrice

$$A = \begin{pmatrix} 1 & & & (c) \\ & \ddots & & \\ (c) & & \ddots & \\ & & & 1 \end{pmatrix} \in \mathcal{M}_n(\mathbb{R})$$

et la colonne X dont les coefficients sont les $\lambda_1, \dots, \lambda_n$. Les équations précédentes fournissent le système $AX = 0$. Si la matrice A est inversible, l'équation $AX = 0$ entraîne $X = 0$ et la famille (u_1, \dots, u_n) est alors libre.

Inversement, supposons la matrice A non inversible. Il existe une relation linéaire sur ses colonnes ce qui permet d'écrire $\mu_1 C_1 + \dots + \mu_n C_n = 0$ avec μ_1, \dots, μ_n des réels non tous nuls. Considérons alors le vecteur $v = \mu_1 u_1 + \dots + \mu_n u_n$ et montrons que celui-ci est nul.

méthode

|| Le vecteur v est nul car combinaison linéaire des vecteurs u_i mais aussi orthogonal à chacun des u_i .

Pour tout $i \in [1; n]$, le vecteur v est orthogonal à u_i car la valeur du produit scalaire $\langle u_i, v \rangle$ est déterminée par le coefficient en i -ème ligne de la colonne $\mu_1 C_1 + \dots + \mu_n C_n$ et ce dernier est nul. On a donc

$$v \in \text{Vect}(u_1, \dots, u_n) \cap \text{Vect}(u_1, \dots, u_n)^\perp = \{0_E\}.$$

Ainsi, on a la relation linéaire $\mu_1 u_1 + \dots + \mu_n u_n = 0_E$ et l'on peut affirmer que la famille (u_1, \dots, u_n) est liée.

En résumé, la famille (u_1, \dots, u_n) est libre si, et seulement si, la matrice A est inversible.

Sachant¹ $\det(A) = (1 + (n - 1)c)(1 - c)^{n-1}$, on peut conclure que (u_1, \dots, u_n) est liée si, et seulement si²,

$$c = 1 \quad \text{ou} \quad c = -\frac{1}{n-1}$$

Exercice 32 ** (Décomposition QR d'une matrice inversible)

Soit $A \in \text{GL}_n(\mathbb{R})$.

(a) Montrer qu'il existe une matrice Q orthogonale et une matrice R triangulaire supérieure à coefficients diagonaux strictement positifs telles que $A = QR$.

(b) Montrer l'unicité de cette écriture.

1. Voir sujet 5 p. 367.

2. Le cas $c = 1$ correspond au cas où les vecteurs unitaires u_1, \dots, u_n sont tous égaux.

Solution(a) **méthode**

On interprète A comme la matrice de passage d'une base à une autre que l'on orthonormalise par le procédé de Schmidt.

On munit l'espace \mathbb{R}^n de sa structure euclidienne canonique et l'on introduit sa base canonique $c = (c_1, \dots, c_n)$. La matrice A figure dans la base c une famille de vecteurs $u = (u_1, \dots, u_n)$ de \mathbb{R}^n :

$$A = \text{Mat}_c(u_1, \dots, u_n).$$

Puisque la matrice A est inversible, la famille u est une base de \mathbb{R}^n et A se comprend comme la matrice de passage de c à u . On orthonormalise alors la base u par le procédé de Schmidt ce qui forme une base orthonormale $e = (e_1, \dots, e_n)$.

La matrice de passage Q de c à e est orthogonale car il s'agit d'une matrice de passage entre deux bases orthonormales.

La matrice de passage R de e à u est triangulaire supérieure car, pour tout $j \in \llbracket 1; n \rrbracket$,

$$u_j \in \text{Vect}(e_1, \dots, e_j).$$

De plus, les coefficients diagonaux de R sont strictement positifs. En effet, ceux-ci sont donnés par les coordonnées des vecteurs u_j dans la base orthonormale e . Le j -ème coefficient diagonal est donc $(e_j | u_j)$ et celui-ci est strictement positif¹.

Enfin, les trois matrices de passages A , Q et R sont liées par la formule $A = QR$ ce qui résout le problème. En effet, si P est la matrice de passage d'une base e à une base e' et Q celle de la base e' à une base e'' alors les colonnes X , X' et X'' formées des coordonnées d'un même vecteur x dans les bases e , e' et e'' sont liées par les formules de changement de base (Th. 12 p. 317)

$$X = PX' \quad \text{et} \quad X' = QX''.$$

On en déduit $X = PQX''$ et la matrice PQ est donc la matrice de passage de e à e'' .

(b) Supposons $A = QR = Q'R'$ avec Q, Q' matrices orthogonales et R, R' matrices triangulaires supérieures à coefficients diagonaux strictement positifs. En organisant les membres de cette égalité, on dispose de l'identité $Q'^{-1}Q = R'R^{-1}$.

La matrice $M = Q'^{-1}Q$ est orthogonale par opérations dans le groupe $O_n(\mathbb{R})$ et est égale à la matrice $R'R^{-1}$ qui est triangulaire supérieure à coefficients diagonaux positifs. Montrons que cette matrice ne peut être que la matrice I_n .

La première colonne de M est unitaire car M est orthogonale mais c'est aussi la première colonne d'une matrice triangulaire supérieure à coefficients diagonaux positifs. La première colonne de M est donc égale à la première colonne de I_n . La deuxième colonne de M est unitaire et orthogonale à la première. C'est de plus la deuxième colonne d'une matrice triangulaire supérieure à coefficients diagonaux positifs. Cette deuxième colonne n'est autre que la seconde colonne de la matrice I_n . On répète ce raisonnement pour toutes les colonnes et l'on conclut $M = I_n$ puis $Q = Q'$ et $R = R'$.

1. Voir Th. 6 p. 396.

Exercice 33 ** (Déterminant de Gram)

Soit (x_1, \dots, x_n) une famille de vecteurs d'un espace vectoriel euclidien E . La matrice de $\mathcal{M}_n(\mathbb{R})$ de coefficient général $\langle x_i, x_j \rangle$ est notée $G(x_1, \dots, x_n)$, on l'appelle *matrice de Gram* associée à la famille (x_1, \dots, x_n) .

(a) Montrer que, si la famille (x_1, \dots, x_n) est liée, $\det(G(x_1, \dots, x_n)) = 0$.

On suppose désormais la famille (x_1, \dots, x_n) libre et l'on introduit $e = (e_1, \dots, e_n)$ une base orthonormale de l'espace F qu'elle engendre. On note $A = (a_{i,j})$ la matrice figurant la famille (x_1, \dots, x_n) dans la base e .

(b) Exprimer $G(x_1, \dots, x_n)$ en fonction de A et de ${}^t A$. En déduire

$$\det(G(x_1, \dots, x_n)) > 0.$$

(c) On introduit $u \in E$. Montrer

$$d(u, F) = \sqrt{\frac{\det(G(u, x_1, \dots, x_n))}{\det(G(x_1, \dots, x_n))}}$$

Solution**(a) méthode**

Une relation linéaire sur les vecteurs de la famille (x_1, \dots, x_n) se retrouve à l'identique sur les rangées de $G(x_1, \dots, x_n)$.

Supposons la famille (x_1, \dots, x_n) liée. Il existe des réels $\lambda_1, \dots, \lambda_n$ non tous nuls tels que $\lambda_1 x_1 + \dots + \lambda_n x_n = 0_E$. En faisant le produit scalaire des deux membres de cette égalité avec le vecteur x_i , on obtient pour tout $i \in [1 ; n]$

$$\lambda_1 \langle x_i, x_1 \rangle + \dots + \lambda_n \langle x_i, x_n \rangle = 0.$$

En notant C_1, \dots, C_n les colonnes de $G(x_1, \dots, x_n)$, les égalités ci-dessus donne la relation linéaire

$$\lambda_1 C_1 + \dots + \lambda_n C_n = 0.$$

On en déduit que la matrice $G(x_1, \dots, x_n)$ n'est pas inversible et son déterminant est nul.

(b) méthode

On sait exprimer le produit scalaire de deux vecteurs à partir de leurs coordonnées dans une base orthonormale.

Les coefficients de la matrice A déterminent les coordonnées des vecteurs x_1, \dots, x_n dans la base orthonormale e . Pour $i, j \in [1 ; n]$, les coefficients $a_{1,i}, \dots, a_{n,i}$ d'une part, et $a_{1,j}, \dots, a_{n,j}$ d'autre part, correspondent aux coordonnées des vecteurs x_i et x_j et donc

$$\langle x_i, x_j \rangle = \sum_{k=1}^n a_{k,i} a_{k,j} = \sum_{k=1}^n [A]_{k,i} [A]_{k,j} = \sum_{k=1}^n [A]_{i,k} [A]_{k,j} = [AA]_{i,j}.$$

Ainsi, $G(x_1, \dots, x_n) = {}^t A A$. Enfin, la matrice A est inversible car de rang n égal à celui de la famille libre (x_1, \dots, x_n) et donc

$$\det(G(x_1, \dots, x_n)) = (\det(A))^2 > 0.$$

(c) On décompose le vecteur u en $u = a + b$ avec $a \in F$ et $b \in F^\perp$. Le vecteur a est le projeté orthogonal de u sur F et donc

$$d(u, F) = \|u - p_F(u)\| = \|u - a\| = \|b\|.$$

Parallèlement, étudions le déterminant de la matrice $G(u, x_1, \dots, x_n)$. Pour tout i de $\{1, \dots, n\}$, on a $\langle x_i, u \rangle = \langle x_i, a \rangle$ car le vecteur b est orthogonal aux vecteurs x_i . La matrice $G(u, x_1, \dots, x_n)$ peut donc s'écrire par blocs

$$\begin{pmatrix} \|u\|^2 & \langle u, x_1 \rangle & \cdots & \langle u, x_n \rangle \\ \langle x_1, u \rangle & G(x_1, \dots, x_n) & & \\ \vdots & & & \\ \langle x_n, u \rangle & & & \end{pmatrix} = \begin{pmatrix} \|a\|^2 + \|b\|^2 & \langle a, x_1 \rangle & \cdots & \langle a, x_n \rangle \\ \langle x_1, a \rangle & G(x_1, \dots, x_n) & & \\ \vdots & & & \\ \langle x_n, a \rangle & & & \end{pmatrix}.$$

On décompose la première colonne en la somme de deux colonnes afin d'isoler le vecteur b . Par multilinéarité du déterminant en la famille des colonnes, il vient

$$\det(G(u, x_1, \dots, x_n)) = \underbrace{\begin{vmatrix} \|a\|^2 & \langle a, x_1 \rangle & \cdots & \langle a, x_n \rangle \\ \langle x_1, a \rangle & G(x_1, \dots, x_n) & & \\ \vdots & & & \\ \langle x_n, a \rangle & & & \end{vmatrix}}_{=0 \text{ car la famille est liée}} + \underbrace{\begin{vmatrix} \|b\|^2 & \langle a, x_1 \rangle & \cdots & \langle a, x_n \rangle \\ 0 & G(x_1, \dots, x_n) & & \\ \vdots & & & \\ 0 & & & \end{vmatrix}}_{\neq 0}.$$

On développe ensuite le second déterminant en sa première colonne et l'on obtient

$$\det(G(u, x_1, \dots, x_n)) = \underbrace{\det(G(a, x_1, \dots, x_n))}_{=0 \text{ car la famille est liée}} + \|b\|^2 \underbrace{\det(G(x_1, \dots, x_n))}_{\neq 0}.$$

Finalement¹,

$$d(u, F) = \sqrt{\frac{\det(G(u, x_1, \dots, x_n))}{\det(G(x_1, \dots, x_n))}}$$

1. Si l'on oriente le sous-espace vectoriel F et si l'on considère e une base orthonormale directe, la formule $\det(G(x_1, \dots, x_n)) = (\det(A))^2$ donne $\det(G(x_1, \dots, x_n)) = [x_1, \dots, x_n]^2$. Le produit mixte mesurant une aire en dimension 2 et un volume en dimension 3, on peut interpréter géométriquement le résultat en cours.

CHAPITRE 12

Probabilités

12.1 Probabilité sur un univers fini

12.1.1 Expérience aléatoire

Définition

|| L'ensemble des résultats possibles d'une expérience aléatoire est appelé *univers*, il est généralement noté Ω .

Les éléments ω de Ω constituent les *issues* (ou *réalisations*) de l'expérience aléatoire.

Dans la suite, l'univers Ω est supposé être un ensemble fini.

12.1.2 Événements

Définition

|| On appelle *événement* de l'univers Ω toute partie A de Ω .

Lorsque la partie A est un singleton, on dit qu'il s'agit d'un *événement élémentaire*. L'événement \emptyset est appelé *événement impossible* tandis que l'événement Ω est appelé *événement certain*.

Les opérations ensemblistes se traduisent en langage événementiel. Si A et B sont deux événements de l'univers Ω , on définit l'*intersection* des événements A et B par l'ensemble $A \cap B$. Cet événement est noté « A et B ». L'*union* des événements A et B est l'ensemble $A \cup B$, cet événement est noté « A ou B ». Plus généralement, on définit aussi l'*intersection* et l'*union* de plusieurs événements. On définit encore l'*événement contraire* d'un événement A comme étant son complémentaire dans Ω , on le note \bar{A} .

Lorsque $A \subset B$, on dit que l'événement A entraîne B ou encore que A implique B .

Définition

On dit que les deux événements A et B sont *incompatibles* lorsque leur intersection est l'événement impossible, c'est-à-dire lorsque $A \cap B = \emptyset$.

Définition

On appelle *système complet d'événements* de l'univers Ω toute famille finie $(A_i)_{1 \leq i \leq p}$ d'événements de Ω deux à deux incompatibles et d'union égale à Ω .

A et \bar{A} forment un système complet d'événements.

12.1.3 Probabilité

Définition

On appelle *probabilité* sur l'univers fini Ω toute application P de¹ $\wp(\Omega)$ vers $[0; 1]$ vérifiant $P(\Omega) = 1$ et, pour tous événements A et B de Ω ,

$$A \cap B = \emptyset \implies P(A \cup B) = P(A) + P(B) \quad (\text{propriété d'additivité}).$$

Un *espace probabilisé* est un couple (Ω, P) constitué d'un univers fini Ω et d'une probabilité P sur Ω .

Si l'univers fini Ω n'est pas vide, on définit une probabilité sur Ω en posant

$$P(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)} \quad \text{pour tout } A \subset \Omega.$$

C'est la *probabilité uniforme* sur Ω .

Le choix d'un univers pour modéliser une expérience aléatoire est souvent contraint par la détermination d'une probabilité sur celui-ci. Dans la pratique, on privilégie les espaces munis d'une probabilité uniforme mais ce ne sont pas les seuls manipulés.

12.1.4 Propriétés calculatoires

Soit (Ω, P) un espace probabilisé fini.

Théorème 1

Pour tout événement A de Ω , $P(\bar{A}) = 1 - P(A)$.

En particulier, on a $P(\emptyset) = 1 - P(\Omega) = 0$.

Théorème 2

Pour tous événements A et B de Ω , $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

En particulier,

$$P(A \cup B) \leq P(A) + P(B).$$

1. Rappelons que $\wp(\Omega)$ désigne l'ensemble des parties de Ω .

Cette dernière propriété se généralise par récurrence à l'union de plusieurs événements A_1, \dots, A_n pour former l'*inégalité de Boole*

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mathbb{P}(A_i).$$

Aussi :

Théorème 3

Si $(A_i)_{1 \leq i \leq n}$ est une famille d'événements deux à deux incompatibles,

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mathbb{P}(A_i).$$

12.1.5 Construction d'une probabilité

On suppose l'univers Ω de cardinal $N \in \mathbb{N}^*$. On peut énumérer ses éléments et écrire

$$\Omega = \{\omega_1, \dots, \omega_N\} \quad \text{avec } \omega_1, \dots, \omega_N \text{ deux à deux distincts.}$$

Pour k allant de 1 à N , les *probabilités élémentaires* $p_k = \mathbb{P}(\{\omega_k\})$ constituent une famille (p_1, \dots, p_N) de réels positifs de somme égale à 1. La connaissance de celle-ci suffit à calculer la probabilité de n'importe quel événement A puisque

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcup_{\omega_k \in A} \{\omega_k\}\right) = \sum_{\omega_k \in A} \mathbb{P}(\{\omega_k\}) = \sum_{\omega_k \in A} p_k.$$

Inversement, on peut définir une probabilité sur Ω à partir des probabilités élémentaires :

Théorème 4

Si (p_1, \dots, p_N) est une famille de réels positifs de somme égale à 1, il existe une unique probabilité sur $\Omega = \{\omega_1, \dots, \omega_N\}$ vérifiant

$$\mathbb{P}(\{\omega_k\}) = p_k \quad \text{pour tout } k \in [1; N].$$

La probabilité uniforme est obtenue pour $p_1 = \dots = p_N = \frac{1}{N}$.

12.2 Probabilités conditionnelles

(Ω, \mathbb{P}) désigne un espace probabilisé fini.

12.2.1 Définition

Définition

Soit A un événement de Ω de probabilité non nulle. Pour tout événement B de Ω , la *probabilité de B sachant A* est définie par¹ :

$$P(B|A) \stackrel{\text{def}}{=} \frac{P(A \cap B)}{P(A)}.$$

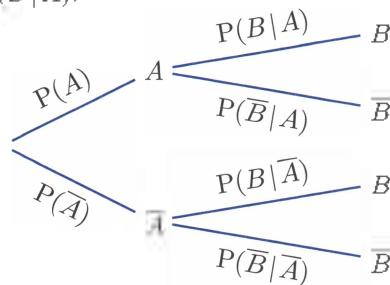
On note aussi $P_A(B)$ la probabilité de B sachant A ce qui introduit l'application

$$P_A : \begin{cases} \wp(\Omega) \rightarrow [0;1] \\ B \mapsto P_A(B) = P(B|A). \end{cases}$$

Celle-ci définit une probabilité sur Ω ce qui autorise l'usage des formules vues précédemment. Par exemple,

$$P_A(\bar{B}) = 1 - P_A(B).$$

Une expérience aléatoire est quelquefois visualisée par un *arbre de probabilités* figurant la compréhension du déroulement de l'expérience. Les probabilités conditionnelles pondèrent un tel arbre.



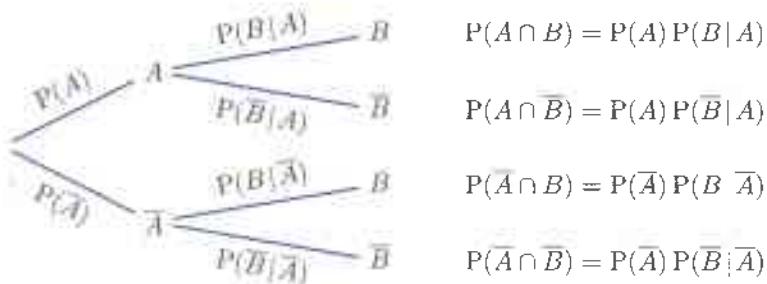
12.2.2 Formule des probabilités composées

Théorème 5

Si A et B sont deux événements de Ω avec A de probabilité non nulle,

$$P(A \cap B) = P(A) P(B|A).$$

Lorsqu'une expérience aléatoire est visualisée par un arbre, cette formule permet le calcul de la probabilité de réalisation d'une feuille.



1. La notation $P(B|A)$ peut prêter à confusion : $(B|A)$ ne désigne pas un événement dont nous calculons la probabilité par P !

On peut généraliser cette formule à plusieurs événements : si A_1, \dots, A_n sont des événements de Ω avec $A_1 \cap \dots \cap A_{n-1}$ de probabilité non nulle¹,

$$P(A_1 \cap \dots \cap A_n) = P(A_1) P(A_2 | A_1) \dots P(A_n | A_1 \cap \dots \cap A_{n-1}).$$

12.2.3 Formule des probabilités totales

Théorème 6

Si $(A_i)_{1 \leq i \leq n}$ est un système complet d'événements de Ω tous de probabilités non nulles alors, pour tout événement B ,

$$P(B) = \sum_{i=1}^n P(A_i) P(B | A_i).$$

Sur un arbre, cette formule permet le calcul de la probabilité d'un événement qui se retrouve sur plusieurs feuilles. En reprenant la figure précédente :

$$P(B) = P(A) P(B | A) + P(\bar{A}) P(B | \bar{A}).$$

12.2.4 Formule de Bayes

Théorème 7

Si A et B sont deux événements de probabilités non nulles,

$$P(A | B) = \frac{P(B | A) P(A)}{P(B)}.$$

La formule de Bayes permet d'étudier la probabilité d'un événement A non directement observable mais dont une conséquence B est réalisée.

12.2.5 Événements indépendants

Définition

On dit que deux événements A et B de l'espace probabilisé (Ω, P) sont *indépendants* si $P(A \cap B) = P(A) P(B)$.

Si A et B sont indépendants et si $P(B) > 0$ alors $P(A | B) = P(A)$.

Définition

On dit que des événements A_1, \dots, A_n de l'espace probabilisé (Ω, P) sont *mutuellement indépendants* si, pour tout $m \in \llbracket 1 ; n \rrbracket$ et pour tous $i_1 < \dots < i_m$ choisis dans $\llbracket 1 ; n \rrbracket$,

$$P\left(\bigcap_{k=1}^m A_{i_k}\right) = \prod_{k=1}^m P(A_{i_k}).$$

1. Par inclusion, A_1 , $A_1 \cap A_2$, etc. sont aussi de probabilités non nulles.

L'indépendance mutuelle ne doit pas être confondue avec l'indépendance deux à deux.

On modélise les résultats des lancers successifs d'une pièce, qu'elle soit équilibrée ou non, par des événements mutuellement indépendants. Il en est de même pour les lancers d'un dé ou les tirages avec remise dans une urne. Ce n'est pas le cas des tirages sans remise pour lesquels il sera utile d'employer la formule des probabilités composées pour tenir compte de la modification de la composition de l'urne.

12.3 Exercices d'apprentissage

Exercice 1

Pour chacune des expériences qui suit, proposer un espace probabilisé (Ω, P) permettant de l'étudier.

- (a) On tire successivement et sans remise six boules dans une urne contenant des boules numérotées de 1 à 49.
- (b) On lance deux dés équilibrés.
- (c) Dix individus prennent place sur dix chaises réparties autour d'une table.
- (d) On lance une pièce équilibrée. Si celle-ci tombe côté 'pile', on tire une boule dans une urne contenant une boule blanche et deux boules rouges. Sinon, on tire une boule dans une urne contenant trois boules blanches et une boule rouge.

Solution

(a) méthode

Lors des expériences par tirage, il convient de distinguer les tirages avec remise de ceux sans remise.

Le tirage ayant lieu sans remise, les valeurs des boules constituant une réalisation de l'expérience sont deux à deux distinctes. Un ensemble Ω convenable est donc l'ensemble des arrangements de 6 éléments de $[1 ; 49]$.

méthode

En l'absence d'hypothèses supplémentaires, on présume équiprobabilité et indépendance.

En supposant que les tirages sont équiprobables et que le tirage d'une boule n'a pas plus de conséquence pour les tirages suivants que l'absence de la boule dans l'urne, on munit l'univers Ω de la probabilité uniforme.

(b) Le lancer des deux dés est *a priori* simultané et les dés identiques. Le résultat de l'expérience est donc un ensemble constitué de deux valeurs avec possibilités de répétitions. On peut ainsi obtenir les valeurs '1' et '3' ou les valeurs '6' et '6'. La première situation est cependant deux fois plus probable que la seconde ce qui complique (un peu) la définition d'une probabilité sur l'ensemble de ces issues.

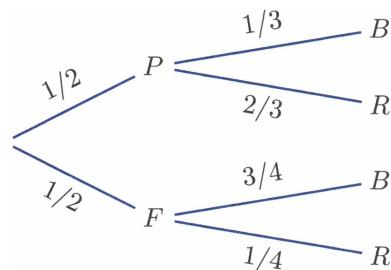
méthode

Il est parfois utile de distinguer des objets qui ne le sont pas *a priori*. On peut par exemple attribuer une couleur à un dé ou numérotter des boules visuellement identiques.

Distinguons les dés lancés en supposant que l'un est blanc et l'autre rouge¹. Les réalisations de l'expérience peuvent s'apparenter à des couples de valeurs comprises entre 1 et 6, le premier élément du couple étant déterminé par la valeur du dé blanc, le second par la valeur du dé rouge. L'univers Ω est alors $\llbracket 1 ; 6 \rrbracket \times \llbracket 1 ; 6 \rrbracket$ et on le munit de la probabilité uniforme.

(c) On attribue un numéro de 1 à 10 à chaque individu et un numéro de 1 à 10 à chaque chaise. Une issue de l'expérience définit une correspondance qui associe à chaque individu la chaise qu'il occupe, c'est-à-dire une fonction de $\llbracket 1 ; 10 \rrbracket$ vers $\llbracket 1 ; 10 \rrbracket$. En présumant qu'aucune chaise n'est occupée par deux individus, cette correspondance est bijective et l'on peut donc choisir $\Omega = S_{10}$ ensemble des permutations de $\llbracket 1 ; 10 \rrbracket$. Encore une fois, on munit cet univers de la probabilité uniforme.

(d) Le résultat de l'expérience apparaît comme un couple² formé du résultat du lancer de la pièce que l'on note P ou F et la couleur de la boule tirée que l'on note B ou R . L'ensemble Ω est donc la réunion des couples $(P, B), (P, R), (F, B), (F, R)$. Par soucis de concision, on préfère écrire PB plutôt que (P, B) et donc $\Omega = \{PB, PR, FB, FR\}$. La probabilité sur Ω de déduit de l'arbre ci-contre et est résumée dans le tableau suivant :



PB	PF	FB	FR
$1/6$	$1/3$	$3/8$	$1/8$

Exercice 2

Soit A , B et C trois événements d'un univers fini Ω .

(a) Exprimer en langage naturel les événements

(i) $A \cup B \cup C$ (ii) $(A \cup B \cup C) \setminus (A \cap B \cap C)$ (iii) $(A \cap B) \cup (B \cap C) \cup (C \cap A)$.

(b) Exprimer par les opérations ensemblistes les événements :

(i) « Seul un des trois événements A , B ou C est réalisé³ »;

(ii) « Au plus deux des trois événements A , B ou C sont réalisés ».

(c) Exprimer les propriétés :

(i) « Si A est réalisé, aucun des événements B ou C ne l'est »;

(ii) « Lorsque A est réalisé, un et un seul des événements B ou C a lieu ».

1. On peut aussi lancer les deux dés successivement en supposant que le résultat d'un lancer est sans conséquence sur le suivant.

2. Il ne faut pas limiter l'étude de l'expérience au simple tirage de la boule.

Solution

(a) (i) Par les lois de Morgan $\overline{A \cup B \cup C} = \overline{A} \cap \overline{B} \cap \overline{C}$. L'événement $\overline{A \cup B \cup C}$ signifie qu'aucun des événements A , B ou C n'est réalisé.

(ii) L'événement $A \cup B \cup C$ traduit la réalisation d'au moins l'un des trois événements tandis que $A \cap B \cap C$ signifie la réalisation de chacun des trois événements. L'événement $(A \cup B \cup C) \setminus (A \cap B \cap C)$ signifie la réalisation d'au moins un des événements mais pas des trois.

(iii) L'événement $A \cap B$ correspond à la réalisation conjointe de A et B . L'événement $(A \cap B) \cup (B \cap C) \cup (C \cap A)$ apparaît alors comme la réalisation d'au moins 2 des trois événements A , B et C .

(b) (i) La réalisation de A sans celle de B ni celle de C s'écrit $A \cap \overline{B} \cap \overline{C}$. L'événement étudié s'exprime donc $(A \cap B \cap C) \cup (\overline{A} \cap B \cap C) \cup (\overline{A} \cap \overline{B} \cap C)$.

(ii) La réalisation de A ou B sans celle de C s'exprime $(A \cup B) \cap \overline{C}$ et l'on peut proposer une solution analogue à la précédente sous réserve de ne pas oublier la situation où aucun des événements n'est réalisé. Cependant,

méthode

|| Il est quelquefois plus commode d'exprimer l'événement contraire.

Dire qu'au plus deux des trois événements A , B ou C sont réalisés signifie aussi que les trois événements ne sont pas tous réalisés, c'est-à-dire $\overline{A \cap B \cap C}$.

(c) (i) On exprime ici une incompatibilité de A et de l'événement $B \cup C$. La propriété étudiée s'écrit $A \cap (B \cup C) = \emptyset$. On peut aussi écrire $A \subset \overline{B \cup C}$.

(ii) On exprime ici une implication donc une inclusion : $A \subset (B \cup C) \setminus (B \cap C)$.

Exercice 3

Soit $n \in \mathbb{N}^*$. Déterminer une probabilité sur l'univers $\Omega = \{1, 2, \dots, n\}$ telle que la probabilité de l'événement $\{1, 2, \dots, k\}$ soit proportionnelle à k^2 .

Solution**méthode**

|| Une probabilité est entièrement déterminée par ses probabilités élémentaires qui sont des réels positifs de somme égale à 1 (Th. 4 p. 441).

Analyse : Soit P une probabilité solution. Il existe un réel a tel que, pour tout $k \in [1; n]$,

$$P(\{1, 2, \dots, k\}) = ak^2.$$

Puisque $P(\Omega) = 1$, on obtient immédiatement la valeur de a : $a = 1/n^2$. Aussi,

$$\{1, 2, \dots, k\} = \{1, 2, \dots, k-1\} \cup \{k\}.$$

3. On dit qu'un événement est *réalisé* si l'issue de l'expérience aléatoire lui appartient.

Par additivité, c'est-à-dire par calcul d'une probabilité d'une réunion d'événements incompatibles¹,

$$P(\{1, 2, \dots, k\}) = P(\{1, 2, \dots, k-1\}) + P(\{k\})$$

et donc

$$P(\{k\}) = P(\{1, \dots, k\}) - P(\{1, \dots, k-1\}) = \frac{2k-1}{n^2}.$$

Ces probabilités élémentaires suffisent à déterminer entièrement la probabilité P.

Synthèse : Considérons les réels

$$p_k = \frac{2k-1}{n^2} \quad \text{avec } k \in [1; n].$$

Ceux-ci sont positifs et de somme égale à 1 car²

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

On peut donc définir une probabilité P sur Ω en posant les probabilités élémentaires

$$P(\{k\}) = \frac{2k-1}{n^2} \quad \text{pour tout } k \in \{1, \dots, n\}.$$

Il reste à vérifier que celle-ci convient. Pour tout $k \in \{1, \dots, n\}$, on obtient par additivité

$$P(\{1, 2, \dots, k\}) = P\left(\bigcup_{i=1}^k \{i\}\right) = \sum_{i=1}^k P(\{i\}) = \sum_{i=1}^k \frac{2i-1}{n^2} = \frac{1}{n^2} \left(2 \sum_{i=1}^k i - k \right) = \frac{k^2}{n^2}.$$

Finalement, la probabilité P ainsi déterminée est (l'unique) solution du problème posé.

Exercice 4

On tire successivement deux boules dans une urne contenant 7 boules blanches et 3 boules rouges.

- (a) Quelle est la probabilité que la première boule tirée soit rouge ?
- (b) Quelle est la probabilité d'avoir tiré deux boules rouges ?
- (c) Quelle est la probabilité que la seconde boule tirée soit rouge ?
- (d) Quelle est la probabilité qu'au moins l'une des deux boules soit rouge ?
- (e) La seconde boule tirée est rouge. Quelle est la probabilité que la première boule tirée le soit aussi ?

1. On parle aussi quelquefois de *réunion disjointe* et l'on utilise les symboles \sqcup ou \uplus pour écrire cette opération.

2. La somme $1 + 2 + \dots + n$ vaut $\frac{n(n+1)}{2}$.

Solution**méthode**

Pour s'exprimer correctement lors de la résolution d'un exercice de probabilité, il est important de dénommer les événements utiles.

Introduisons les événements :

- R_1 = « La première boule tirée est rouge »,
 R_2 = « La seconde boule tirée est rouge ».

On peut aussi nommer les événements contraires $B_1 = \overline{R_1}$ et $B_2 = \overline{R_2}$. L'expérience peut être visualisée par l'arbre ci-contre où les branches sont pondérées par des probabilités conditionnelles : lors du deuxième tirage la composition de l'urne n'est plus la même que lors du premier tirage.

(a) La probabilité pour la première boule tirée d'être rouge dépend simplement de la composition initiale de l'urne

$$P(R_1) = \frac{3}{10} = 0,3.$$

(b) La probabilité de tirer deux boules rouges est celle de l'événement $R_1 \cap R_2$. La composition de l'urne au deuxième tirage est fonction du résultat du premier. Sachant que R_1 est de probabilité non nulle, on utilise la formule des probabilités composées (Th. 5 p. 442)

$$P(R_1 \cap R_2) = P(R_1) P(R_2 | R_1) = \frac{3}{10} \cdot \frac{2}{9} = \frac{1}{15} = 0,067 \text{ à } 10^{-3} \text{ près.}$$

(c) L'événement R_2 apparaît sur plusieurs feuilles de l'arbre figurant l'expérience, sa probabilité se déduit de la formule des probabilités totales (Th. 6 p. 443).

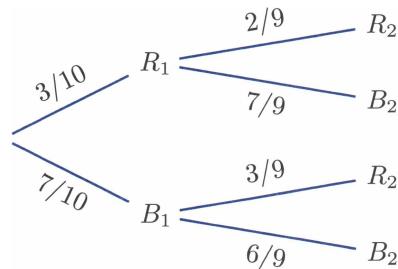
méthode

Pour appliquer la formule des probabilités totales, on identifie un système complet d'événements.

Les événements R_1 et B_1 constituent un système complet d'événements de probabilités non nulles car B_1 est simplement l'événement contraire de R_1 . Par la formule des probabilités totales

$$P(R_2) = P(R_1) P(R_2 | R_1) + P(B_1) P(R_2 | B_1) = \frac{3}{10} \cdot \frac{2}{9} + \frac{7}{10} \cdot \frac{3}{9} = \frac{27}{90} = \frac{3}{10}.$$

Cette valeur correspond à la probabilité que la première boule tirée soit rouge. On peut l'expliquer par un argument de symétrie. On numérote de 1 à 10 les boules constituant l'urne. Un tirage correspond au choix équiprobable d'un couple de deux valeurs distinctes comprises entre 1 et 10. Or en échangeant les deux éléments du couple, on peut affirmer



qu'il y autant de tirages dont la première boule est rouge que de tirages dont la seconde boule est rouge !

(d) méthode

|| Les événements s'exprimant par « au moins un(e) » invite à l'introduction de l'événement contraire qui s'exprime par « aucun(e) ».

L'événement étudié est simplement $\overline{B_1 \cap B_2}$. Par la formule des probabilités composées

$$P(\overline{B_1 \cap B_2}) = 1 - P(B_1 \cap B_2) = 1 - P(B_1)P(B_2 | B_1) = \frac{8}{15} \simeq 0,533 \text{ à } 10^{-3} \text{ près.}$$

(e) méthode

|| Il s'agit de remonter dans l'arbre à partir d'une feuille : on utilise la formule de Bayes (Th. 7 p. 443).

On étudie la probabilité de R_1 sachant R_2 . Ces deux événements sont de probabilités non nulles et la formule de Bayes donne

$$P(R_1 | R_2) = \frac{P(R_2 | R_1)P(R_1)}{P(R_2)} = \frac{\frac{2}{9} \cdot \frac{9}{10}}{\frac{3}{10}} = \frac{2}{9} = 0,222 \text{ à } 10^{-3} \text{ près.}$$

Cette valeur est égale la probabilité de R_2 sachant R_1 ce qui s'explique à nouveau par l'argument de symétrie précédent. Il est notable que cette valeur est inférieure à la probabilité de R_1 ce qui peut être attendu. Si l'on tire initialement une boule rouge, la probabilité que la boule suivante soit rouge est inférieure à ce que l'on obtient si l'on tire initialement une boule blanche. Savoir que l'on a obtenu une boule rouge au second tirage réduit la probabilité que le premier tirage soit aussi celui d'une boule rouge.

Exercice 5

On lance successivement deux dés équilibrés et l'on considère les événements

$A = \text{« La valeur du premier dé lancé est paire »},$

$B = \text{« La valeur du second dé lancé est paire »},$

$C = \text{« La somme des valeurs des deux dés est paire »}.$

Vérifier que les événements A , B et C sont deux à deux indépendants sans être mutuellement indépendants.

Solution

méthode

|| On justifie l'indépendance de deux événements A et B en vérifiant l'égalité $P(A \cap B) = P(A)P(B)$.

Chacun des deux dés étant équilibré, on peut affirmer $P(A) = P(B) = 1/2$.

méthode

|| L'indépendance de deux événements est souvent une hypothèse de la modélisation de l'expérience.

C'est une hypothèse implicite de ce sujet, on suppose l'indépendance des valeurs obtenues par chacun des deux dés. On peut alors modéliser l'expérience en considérant l'univers $\Omega = [1 ; 6] \times [1 ; 6]$ muni de la probabilité uniforme. À partir de ce modèle, on retrouve les probabilités des événements A et B déjà affirmées ci-dessus mais on peut aussi calculer celle de $A \cap B$:

$$P(A \cap B) = \frac{\text{Card}(A \cap B)}{\text{Card}(\Omega)} = \frac{\text{Card}\{(2, 2), (2, 4), \dots, (6, 6)\}}{36} = \frac{9}{36} = \frac{1}{4}.$$

Les événements A et B sont donc indépendants car $P(A \cap B) = P(A) P(B)$.

En remarquant $P(C) = 1/2$ et $A \cap B = A \cap C = B \cap C$, on établit aussi que A et C d'une part, B et C d'autre part, sont indépendants.

Cependant, les événements A , B et C ne sont pas mutuellement indépendants puisque

$$P(A \cap B \cap C) = P(A \cap B) = \frac{1}{4} \neq \frac{1}{8} = P(A) P(B) P(C).$$

12.4 Exercices d'entraînement

12.4.1 Définition d'une probabilité

Exercice 6 *

Soit A et B deux événements d'un espace probabilisé (Ω, P) .

- (a) On suppose $A \subset B$. Montrer $P(A) \leq P(B)$.
- (b) On suppose $A \cap B = \emptyset$. Montrer $P(A) \leq 1 - P(B)$.
- (c) On suppose $P(A) = 0,3$, $P(B) = 0,5$ et $P(A \cup B) = 0,6$. Calculer $P(\bar{A} \cup B)$.

Solution
méthode

|| Une probabilité sur l'univers Ω est une application P de $\wp(\Omega)$ vers $[0 ; 1]$ vérifiant $P(\Omega) = 1$ et la propriété d'additivité

$$A \cap B = \emptyset \implies P(A \cup B) = P(A) + P(B).$$

- (a) Si A est inclus dans B , l'événement B est la réunion des événements incompatibles A et $B \setminus A$. Par additivité

$$P(B) = P(A) + P(B \setminus A) \quad \text{avec} \quad P(B \setminus A) \geq 0.$$

La probabilité de B est donc supérieure à celle de A .

(b) Si A et B sont incompatibles, on a par additivité $P(A \cup B) = P(A) + P(B)$. Or la probabilité $P(A \cup B)$ est inférieure à 1 et donc $P(A) + P(B) \leq 1$ ce qui donne directement l'inégalité voulue¹.

(c) Par passage à l'événement contraire, $P(\bar{A} \cup \bar{B}) = 1 - P(A \cap B)$. Les événements $A \cap B$ et $\bar{A} \cap \bar{B}$ sont incompatibles et de réunion égale à A , on a donc par additivité

$$P(A) = P(\bar{A} \cap \bar{B}) + P(A \cap B).$$

La probabilité d'une union et de l'intersection de deux événements sont liées par la formule (Th. 2 p. 440)

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

On en déduit $P(A \cap B) = 0,3 + 0,5 - 0,6 = 0,2$ puis $P(\bar{A} \cap \bar{B}) = 0,3 - 0,2 = 0,1$ et enfin $P(\bar{A} \cup \bar{B}) = 0,9$.

Exercice 7 *

À quelle(s) condition(s) sur les réels x et y existe-t-il une probabilité P sur l'ensemble à 3 éléments $\Omega = \{a, b, c\}$ vérifiant

$$P(\{a, b\}) = x \quad \text{et} \quad P(\{b, c\}) = y ?$$

Solution

méthode

|| On caractérise une probabilité par ses valeurs sur les événements élémentaires.

Analyse : Soit P une probabilité solution. Par événement contraire (Th. 1 p. 440)

$$P(\{c\}) = 1 - P(\{a, b\}) = 1 - x \quad \text{et} \quad P(\{a\}) = 1 - P(\{b, c\}) = 1 - y$$

puis

$$P(\{b\}) = 1 - P(\{a, c\}) = x + y - 1$$

car on a par additivité

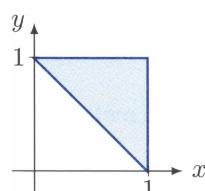
$$P(\{a, c\}) = P(\{a\} \cup \{c\}) = P(\{a\}) + P(\{c\}) = 2 - x - y.$$

Les trois probabilités élémentaires devant être positives, on obtient les conditions

$$x \leq 1, \quad y \leq 1 \quad \text{et} \quad x + y \geq 1.$$

Synthèse : Supposons les trois conditions précédentes remplies. Puisque

$$(1 - y) + (x + y - 1) + (1 - x) = 1.$$



1. On peut aussi écrire $\bar{A} \subset \bar{B}$ avec $P(\bar{B}) = 1 - P(B)$.

on peut définir une probabilité sur Ω par les conditions (Th. 4 p. 441)

$$P(\{a\}) = 1 - y, \quad P(\{b\}) = x + y - 1 \quad \text{et} \quad P(\{c\}) = 1 - y.$$

Cette probabilité est solution car

$$\begin{aligned} P(\{a, b\}) &= P(\{a\} \cup \{b\}) = P(\{a\}) + P(\{b\}) = x \\ P(\{b, c\}) &= P(\{b\} \cup \{c\}) = P(\{b\}) + P(\{c\}) = y. \end{aligned}$$

Exercice 8 **

Soit P une probabilité sur un ensemble Ω et A, B deux événements de Ω . On pose

$$x = P(A \cap B), \quad y = P(A \cap \bar{B}), \quad z = P(\bar{A} \cap B) \quad \text{et} \quad t = P(\bar{A} \cap \bar{B}).$$

(a) Vérifier

$$P(A)P(B) - P(A \cap B) = yz - xt.$$

(b) En déduire

$$|P(A)P(B) - P(A \cap B)| \leq \frac{1}{4}$$

Solution

(a) En écrivant $\Omega = B \cup \bar{B}$, on a

$$P(A) = P(A \cap \Omega) = P((A \cap B) \cup (A \cap \bar{B})).$$

Les événements $A \cap B$ et $A \cap \bar{B}$ sont incompatibles et l'on a donc par additivité

$$P(A) = P(A \cap B) + P(A \cap \bar{B}) = x + y.$$

De même, on obtient $P(B) = x + z$ donc

$$P(A)P(B) - P(A \cap B) = (x + y)(x + z) - x = x(x + y + z) + yz - x. \quad (*)$$

Aussi $\Omega = (A \cup \bar{A}) \cap (B \cup \bar{B})$ et l'on observe par développement que Ω est la réunion des événements deux à deux incompatibles $A \cap B$, $A \cap \bar{B}$, $\bar{A} \cap B$ et $\bar{A} \cap \bar{B}$. On en déduit $x + y + z + t = 1$ et l'égalité $(*)$ devient

$$P(A)P(B) - P(A \cap B) = x(1 - t) + yz - x = yz - xt.$$

(b) méthode

¶ Pour tout x réel, on sait¹ $x(1 - x) \leq \frac{1}{4}$.

¹. La fonction $x \mapsto x(1 - x)$ est représentée par une parabole tournée vers le bas dont le sommet est en $x = 1/2$.

Les réels x et t étant positifs, on a

$$P(A)P(B) - P(A \cap B) \leq yz.$$

Or $z = 1 - x - y - t \leq 1 - y$ et donc

$$P(A)P(B) - P(A \cap B) \leq y(1 - y) \leq \frac{1}{4}.$$

De même, on obtient

$$P(A \cap B) - P(A)P(B) \leq xt \leq x(1 - x) \leq \frac{1}{4}$$

et on peut alors conclure à l'inégalité proposée.

12.4.2 Événements indépendants

Exercice 9 *

Soit A et B deux événements incompatibles d'un espace probabilisé (Ω, P) . À quelle condition les événements A et B sont-ils indépendants ?

Solution

méthode

Il ne faut confondre indépendance et incompatibilité !

Les événements A et B étant incompatibles, on a $A \cap B = \emptyset$ et donc $P(A \cap B) = 0$. Si les événements A et B sont indépendants, on a aussi $P(A \cap B) = P(A)P(B)$ et donc $P(A)P(B) = 0$. On en déduit $P(A) = 0$ ou $P(B) = 0$.

Inversement, si $P(A) = 0$, on a aussi $P(A \cap B) = 0 = P(A)P(B)$ car $A \cap B \subset A$. Les événements A et B sont alors indépendants. On conclut de la même façon si $P(B) = 0$.

En résumé, deux événements incompatibles sont indépendants si, et seulement si, l'un d'eux est de probabilité nulle.

Exercice 10 *

Soit A, B, C trois événements d'un espace probabilisé (Ω, P) .

(a) On suppose les événements A, B et C mutuellement indépendants. Montrer que les événements A et $B \cup C$ sont indépendants.

(b) On suppose A et B d'une part, A et C d'autre part, indépendants. Les événements A et $B \cup C$ sont-ils indépendants ?

Solution(a) **méthode**

L'indépendance mutuelle de A , B et C donne $P(A \cap B \cap C) = P(A)P(B)P(C)$ mais aussi $P(A \cap B) = P(A)P(B)$ et des égalités analogues pour $P(A \cap C)$ et $P(B \cap C)$.

Par distributivité de l'intersection sur l'union puis développement de la probabilité d'une union, on écrit

$$P(A \cap (B \cup C)) = P((A \cap B) \cup (A \cap C)) = P(A \cap B) + P(A \cap C) - P(A \cap B \cap C).$$

Par indépendance puis factorisation

$$\begin{aligned} P(A \cap (B \cup C)) &= P(A)P(B) + P(A)P(C) - P(A)P(B)P(C) \\ &= P(A)(P(B) + P(C) - P(B)P(C)). \end{aligned}$$

Enfin, on a aussi

$$P(B \cup C) = P(B) + P(C) - P(B \cap C) = P(B) + P(C) - P(B)P(C)$$

et donc $P(A \cap (B \cup C)) = P(A)P(B \cup C)$. Les événements A et $B \cup C$ sont indépendants¹.

(b) Par un exemple, on observe que les indépendances de A et B et de A et C n'entraînent pas celle de A et $B \cup C$. Considérons pour cela l'univers $\Omega = \llbracket 1 ; 6 \rrbracket$ muni de la probabilité uniforme et les événements

$$A := \{2, 4, 6\}, \quad B = \{1, 2\} \quad \text{et} \quad C = \{2, 3\}.$$

On a

$$P(A) = \frac{1}{3}, \quad P(B) = P(C) = \frac{1}{3} \quad \text{et} \quad P(A \cap B) = P(A \cap C) = \frac{1}{6}.$$

On vérifie donc $P(A \cap B) = P(A)P(B)$ et $P(A \cap C) = P(A)P(C)$. En revanche

$$P(B \cup C) = \frac{1}{2} \quad \text{et} \quad P(A \cap (B \cup C)) = \frac{1}{6} \neq \frac{1}{3} = P(A)P(B \cup C).$$

Les événements A et $B \cup C$ ne sont pas indépendants².

Exercice 11 *

Deux urnes contiennent des boules blanches et rouges. Les proportions de boules blanches dans ces urnes sont respectivement égales à p et q avec $p, q \in]0 ; 1[$.

De façon équiprobable on choisit l'une des deux urnes et l'on tire avec remise deux boules dans celle-ci. À quelle condition a-t-on l'indépendance des deux événements « La première boule tirée est blanche » et « La seconde boule tirée est blanche » ?

1. Aussi, et plus immédiatement, les événements A et $B \cap C$ sont indépendants. Plus généralement, si A_1, \dots, A_n sont des événements mutuellement indépendants, tout événement construit par opérations d'union, d'intersection et de passage au complémentaire à partir de A_1, \dots, A_p est indépendant d'un événement qui serait construit à partir de A_{p+1}, \dots, A_n .

2. Les événements A et $B \cap C$ ne sont pas non plus indépendants.

Solution

Lorsque l'on opère un tirage avec remise il est d'usage de supposer les tirages indépendants. Cependant, connaître le résultat du premier tirage dans le protocole en cours apporte une information sur l'urne dans laquelle on opère le second tirage.

Introduisons les événements :

$A = \text{« La première boule tirée est blanche »},$

$B = \text{« La deuxième boule tirée est blanche »},$

$U = \text{« L'urne contenant les boules blanches en proportion } p \text{ a été choisie »}.$

méthode

|| On étudie l'indépendance de A et B en comparant $P(A \cap B)$ et $P(A)P(B)$.

Le couple (U, \bar{U}) est un système complet d'événements de probabilités non nulles et la formule des probabilités totales donne

$$P(A) = P(A|U)P(U) + P(A|\bar{U})P(\bar{U}) = \frac{1}{2}(p+q).$$

La probabilité de B est identique à celle de A . Aussi, les événements A et B étant indépendants pour les probabilités conditionnelles P_U et $P_{\bar{U}}$.

$$\begin{aligned} P(A \cap B) &= P(A \cap B|U)P(U) + P(A \cap B|\bar{U})P(\bar{U}) \\ &= P(A|U)P(B|U)P(U) + P(A|\bar{U})P(B|\bar{U})P(\bar{U}) = \frac{1}{2}(p^2 + q^2). \end{aligned}$$

On peut alors étudier l'indépendance des événements A et B :

$$\begin{aligned} P(A \cap B) - P(A)P(B) &\iff \frac{1}{2}(p^2 + q^2) = \frac{1}{4}(p+q)^2 \\ &\iff p^2 - 2pq + q^2 = 0 \\ &\iff p = q. \end{aligned}$$

Finalement, les événements A et B sont indépendants si, et seulement si, les proportions p et q sont identiques.

Exercice 12 **

Soit A, B, C trois événements d'un espace probabilisé (Ω, P) .

On suppose A indépendant de $B \cap C$, B indépendant de $A \cap C$ et C indépendant de $A \cap B$. On suppose aussi A indépendant de $B \cup C$ et $P(A) > 0$.

Etablir que les événements A, B, C sont mutuellement indépendants.

Solution**méthode**

|| On vérifie que les probabilités de $A \cap B$, $B \cap C$, $C \cap A$ et $A \cap B \cap C$ sont les produits des probabilités des événements intersectés.

Les premières hypothèses d'indépendance donnent les égalités

$$P(A \cap B \cap C) = P(A) P(B \cap C) = P(B) P(A \cap C) = P(C) P(A \cap B). \quad (*)$$

Aussi, par l'indépendance de A avec $B \cup C$, on a $P(A \cap (B \cup C)) = P(A) P(B \cup C)$ avec

$$P(A \cap (B \cup C)) = P((A \cap B) \cup (A \cap C)) = P(A \cap B) + P(A \cap C) - P(A \cap B \cap C).$$

On obtient ainsi l'identité

$$P(A) P(B \cup C) = P(A \cap B) + P(A \cap C) - P(A \cap B \cap C).$$

Par la première égalité de (*), on transforme celle-ci en

$$P(A) P(B \cup C) = P(A \cap B) + P(A \cap C) - P(A) P(B \cap C)$$

puis en réorganisant les membres

$$P(A) P(B \cup C) + P(A) P(B \cap C) = P(A \cap B) + P(A \cap C).$$

Sachant $P(B \cup C) + P(B \cap C) = P(B) + P(C)$, on obtient

$$P(A)(P(B) + P(C)) = P(A \cap B) + P(A \cap C).$$

On multiplie les deux membres par $P(C)$

$$P(A) P(C)(P(B) + P(C)) = P(C) P(A \cap B) + P(C) P(A \cap C).$$

La dernière égalité de (*) donne alors

$$P(A) P(C)(P(B) + P(C)) = (P(B) + P(C)) P(A \cap C). \quad (**)$$

On poursuit en discutant selon que la somme $P(B) + P(C)$ est nulle ou non.

Cas : $P(B) + P(C) > 0$. On simplifie (**) par $P(B) + P(C)$ pour obtenir

$$P(A \cap C) = P(A) P(C).$$

Un raisonnement symétrique donne

$$P(A \cap B) = P(A) P(B).$$

Les égalités (*) donnent alors

$$P(A \cap B \cap C) = P(A) P(B) P(C) \quad \text{et} \quad P(A) P(B \cap C) = P(A) P(B) P(C).$$

En simplifiant la dernière égalité par $P(A)$ qui est supposée non nulle, on obtient

$$P(B \cap C) = P(B) P(C).$$

Finalement, la probabilité d'une intersection quelconque d'événements parmi A , B et C est le produit des probabilités des événements considérés : ces trois événements sont mutuellement indépendants.

Cas : $P(B) + P(C) = 0$. On a $P(B) = P(C) = 0$ et les probabilités des événements $A \cap B$, $A \cap C$, $B \cap C$ et $A \cap B \cap C$ sont aussi nulles car ces événements sont inclus dans des événements de probabilités nulles. Les événements A , B et C sont alors mutuellement indépendants.

Exercice 13 ***

Soit (Ω, P) un espace probabilisé.

(a) Montrer que si A et B sont deux événements indépendants alors A et \bar{B} sont indépendants.

Pour $\varepsilon = 1$ ou -1 et A un événement de Ω , on note

$$A^\varepsilon = \begin{cases} A & \text{si } \varepsilon = 1 \\ \bar{A} & \text{si } \varepsilon = -1. \end{cases}$$

(b) Soit $n \geq 2$ et A_1, \dots, A_n des événements mutuellement indépendants de (Ω, P) . Montrer que, pour tout $(\varepsilon_1, \dots, \varepsilon_n) \in \{-1, 1\}^n$, les événements $A_1^{\varepsilon_1}, \dots, A_n^{\varepsilon_n}$ sont mutuellement indépendants.

Solution

(a) La famille (B, \bar{B}) est un système complet d'événements permettant d'écrire

$$P(A) = P(A \cap B) + P(A \cap \bar{B}).$$

Par indépendance de A et B , on a $P(A \cap B) = P(A)P(B)$ et donc

$$P(A) = P(A)P(B) + P(A \cap \bar{B}).$$

On en déduit

$$P(A \cap \bar{B}) = P(A) - P(A)P(B) = P(A)(1 - P(B)) = P(A)P(\bar{B}).$$

Les événements A et \bar{B} sont indépendants.

(b) méthode

On commence par vérifier que $A_1, \dots, A_{n-1}, \bar{A}_n$ sont mutuellement indépendants.

Il s'agit de vérifier que la probabilité de n'importe quelle intersection d'événements parmi $A_1, \dots, A_{n-1}, \bar{A}_n$ est le produit des probabilités des événements concernés.

Soit $m \in [2; n]$ et $i_1 < \dots < i_m$ choisis dans $[1; n]$.

Cas : $i_m < n$. On a immédiatement

$$P(A_{i_1} \cap \dots \cap A_{i_m}) = P(A_{i_1}) \times \dots \times P(A_{i_m})$$

car les événements A_1, \dots, A_{m-1} sont mutuellement indépendants.

Cas : $i_m = n$. On écrit

$$\mathrm{P}(A_{i_1} \cap \dots \cap A_{i_{m-1}} \cap A_n) = \mathrm{P}(A \cap \overline{B})$$

avec $A = A_{i_1} \cap \dots \cap A_{i_{m-1}}$ et $B = A_n$. Les événements A et B sont indépendants car la mutuelle indépendance des A_1, \dots, A_n permet d'écrire

$$\begin{aligned}\mathrm{P}(A \cap B) &= \mathrm{P}(A_{i_1} \cap \dots \cap A_{i_{m-1}} \cap B) \\ &= \mathrm{P}(A_{i_1}) \times \dots \times \mathrm{P}(A_{i_{m-1}}) \times \mathrm{P}(B) \\ &= \mathrm{P}(A_{i_1} \cap \dots \cap A_{i_{m-1}}) \mathrm{P}(B) = \mathrm{P}(A) \mathrm{P}(B).\end{aligned}$$

L'étude de la question précédente donne alors $\mathrm{P}(A \cap \overline{B}) = \mathrm{P}(A) \mathrm{P}(\overline{B})$, autrement dit,

$$\mathrm{P}(A_{i_1} \cap \dots \cap A_{i_{m-1}} \cap \overline{A_n}) = \mathrm{P}(A_{i_1}) \times \dots \times \mathrm{P}(A_{i_{m-1}}) \times \mathrm{P}(\overline{A_n}).$$

Ainsi, les événements $A_1, \dots, A_{n-1}, \overline{A_n}$ sont mutuellement indépendants.

A l'aide d'une permutation des événements, on établit aussi que, pour tout $j \in \llbracket 1 ; n \rrbracket$,

$A_1, \dots, A_j, \dots, A_n$ mutuellement indépendants \implies

$A_1, \dots, \overline{A_j}, \dots, A_n$ mutuellement indépendants.

En passant à l'événement contraire plusieurs des événements A_1, \dots, A_n tant que nécessaire, on établit que les événements $A_1^{\varepsilon_1}, \dots, A_n^{\varepsilon_n}$ sont mutuellement indépendants.

12.4.3 Calcul de probabilités

Exercice 14 *

On considère des dés équilibrés. Lequel des événements qui suivent est le plus probable ?

- (a) $A = \text{« Ne pas obtenir de ‘un’ ni de ‘six’ en 2 lancers »}.$
- (b) $B = \text{« Obtenir un ‘six’ en moins de 4 lancers »}.$
- (c) $C = \text{« Obtenir un ‘double six’ en moins de 24 lancers de deux dés »}.$
- (d) $D = \text{« Obtenir toutes les valeurs de ‘un’ à ‘six’ en moins de 8 lancers »}.$

Solution

(a) Chacun des deux lancers détermine une valeur élément de $\llbracket 1 ; 6 \rrbracket$. On modélise l'expérience par l'univers $\Omega = \llbracket 1 ; 6 \rrbracket^2$ muni de la probabilité uniforme. L'événement étudié se traduit par l'ensemble $A = \llbracket 2 ; 5 \rrbracket^2$. Sa probabilité est

$$\mathrm{P}(A) = \frac{\mathrm{Card}(\llbracket 2 ; 5 \rrbracket^2)}{\mathrm{Card}(\llbracket 1 ; 6 \rrbracket^2)} = \left(\frac{2}{3}\right)^2 \approx 0,444 \text{ à } 10^{-3} \text{ près.}$$

(b) Si on lance une première fois le dé et que l'on obtient immédiatement un ‘six’, l’événement B est réalisé et il n'est plus utile de continuer de lancer le dé. Cependant, afin de simplifier l’expression de l'espace probabilisé modélisant l’expérience, on suppose que le dé est lancé exactement quatre fois et l'on s'intéresse à la probabilité qu'au moins un ‘six’ figure parmi ces quatre lancers. On considère donc l'univers $\Omega = [[1; 6]]^4$ muni de la probabilité uniforme.

méthode

|| On raisonne par l’événement contraire.

L’événement contraire de l’événement B est l’événement ne pas obtenir de ‘six’ en 4 lancers et par un calcul semblable au précédent

$$P(B) = 1 - \left(\frac{5}{6}\right)^4 \simeq 0,518 \text{ à } 10^{-3} \text{ près.}$$

(c) En distinguant les deux dés, on considère l'univers $\Omega = ([[1; 6]] \times [[1; 6]])^{24}$ muni de la probabilité uniforme. On raisonne à nouveau par l’événement contraire

$$P(C) = 1 - \left(\frac{35}{36}\right)^{24} \simeq 0,491 \text{ à } 10^{-3} \text{ près.}$$

(d) On modélise l’expérience par l’univers $\Omega = [[1; 6]]^8$ muni de la probabilité uniforme. Dénombrons, sans répétitions, les éléments¹ constituant l’événement D . Ceux-ci sont deux types :

Cas : une valeur dans $[1; 6]$ est prise trois fois. On choisit cette valeur (6 possibilités) puis les positions qu’elle occupe ($\binom{8}{3}$ possibilités) et enfin on complète les 5 positions restantes par les permutations des autres valeurs ($5!$ possibilités). Au total, il y a

$$6 \times \binom{8}{3} \times 5! = 40\,320 \text{ éléments de ce type.}$$

Cas : deux valeurs dans $[1; 6]$ sont prises deux fois. On choisit ces valeurs ($\binom{6}{2}$ possibilités). Pour la plus petite des deux, on choisit les deux positions qu’elle occupe ($\binom{8}{2}$ possibilités). Pour la plus grande, on choisit ses deux positions parmi les six restantes ($\binom{6}{2}$ possibilités). Enfin, on complète les 4 positions vacantes afin d’obtenir les autres valeurs ce qui offre $4!$ possibilités. Au total, il y a

$$\binom{6}{2} \times \binom{8}{2} \times \binom{6}{2} \times 4! = 151\,200 \text{ éléments de ce type.}$$

On a alors

$$P(D) = \frac{40\,320 + 151\,200}{6^8} \simeq 0,114 \text{ à } 10^{-3} \text{ près.}$$

L’événement B est donc le plus probable.

1. Le problème s’apparente au calcul du nombre de surjections d’un ensemble de $[1; 8]$ vers $[1; 6]$ (voir sujet 19 p. 223).

Exercice 15 *

On lance deux fois un dé. Montrer que la probabilité d'obtenir deux fois la même valeur est minimale lorsque le dé est équilibré.

Solution

Pour $k \in [1; 6]$, introduisons les événements :

$$\begin{aligned} A_k &= \text{« On obtient la valeur } k \text{ au premier lancer »,} \\ B_k &= \text{« On obtient la valeur } k \text{ au second lancer ».} \end{aligned}$$

Notons aussi p_k la probabilité commune des événements A_k et B_k . Les réels p_1, \dots, p_6 sont positifs et de somme égale à 1.

Dans ce sujet, on étudie la probabilité de l'événement

$$A = (A_1 \cap B_1) \cup \dots \cup (A_6 \cap B_6).$$

Les événements $A_k \cap B_k$ étant deux à deux incompatibles, on a par additivité

$$P(A) = P(A_1 \cap B_1) + \dots + P(A_6 \cap B_6).$$

Le cadre hypothétique de l'expérience laisse supposer que les deux lancers sont indépendants et donc $P(A_k \cap B_k) = P(A_k)P(B_k)$ pour tout $k \in [1; 6]$. On obtient alors

$$P(A) = p_1^2 + \dots + p_6^2.$$

méthode

Par l'inégalité de Cauchy-Schwarz, on sait¹ que, pour tous x_1, \dots, x_n réels,

$$\sum_{k=1}^n x_k \leq n \sum_{k=1}^n x_k^2$$

avec égalité si, et seulement si, $x_1 = \dots = x_n$.

En appliquant ce résultat, on a

$$P(A) \geq \frac{1}{6} \sum_{k=1}^6 p_k = \frac{1}{6}$$

avec égalité si, et seulement si, $p_1 = \dots = p_6$ ce qui correspond à la situation d'un dé équilibré.

1. Voir sujet 11 p. 415.

Exercice 16 **

Un archer a la probabilité $p \in [0; 1]$ d'atteindre une cible à chaque essai. Soit $n \in \mathbb{N}^*$ et $k \in [1; n]$.

- Quelle est la probabilité qu'il atteigne au moins une fois la cible en n tentatives ?
- Quelle est la probabilité qu'il touche sa cible pour la première fois lors du n -ième essai ?
- Quelle est la probabilité qu'il touche exactement k cibles en n essais ?
- Quelle est la probabilité qu'il touche sa k -ième cible lors du n -ième essai ?

Solution

Pour $k \in [1; n]$, on introduit les événements :

$$A_k = \text{« L'archer atteint sa cible lors du } k\text{-ième essai ».}$$

Le cadre hypothétique donne $P(A_k) = p$ et laisse présumer que les événements A_k sont mutuellement indépendants.

méthode

On exprime les différents événements étudiés à l'aide des événements A_k et l'on exploite l'indépendance de ceux-ci pour le calcul de probabilité¹.

(a) Dans cette question, on calcule la probabilité de l'événement $B_n = A_1 \cup \dots \cup A_n$: on raisonne par l'événement contraire.

$$P(\overline{B_n}) = P(\overline{A_1 \cup \dots \cup A_n}) = P(\overline{A_1} \cap \dots \cap \overline{A_n}).$$

Par indépendance² des événements $\overline{A_1}, \dots, \overline{A_n}$

$$P(\overline{A_1} \cap \dots \cap \overline{A_n}) = P(\overline{A_1}) \times \dots \times P(\overline{A_n}) = (1-p)^n.$$

Ainsi,

$$P(B_n) = 1 - (1-p)^n.$$

(b) Dans cette question, on étudie l'événement $C_n = \overline{A_1} \cap \dots \cap \overline{A_{n-1}} \cap A_n$ traduisant un série de $n-1$ échecs suivis d'un succès. Par indépendance, on obtient

$$P(\overline{A_1} \cap \dots \cap \overline{A_{n-1}} \cap A_n) = P(\overline{A_1}) \times \dots \times P(\overline{A_{n-1}}) \times P(A_n) = (1-p)^{n-1}p.$$

(c) Introduisons l'événement étudié

$$D_{k,n} = \text{« L'archer touche } k \text{ cibles en } n \text{ essais ».}$$

1. On peut modéliser l'expérience en introduisant un espace probabilisé (Ω, P) avec $\Omega = \{0, 1\}^n$ exprimant l'échec ou la réussite de l'archer à chacune de ses n tentatives. Cependant, il est assez fréquent de mener des calculs de probabilités sans préciser l'espace (Ω, P) sous-jacent à l'étude.

2. Voir sujet 13 p. 457.

Pour $\varepsilon \in \{-1, 1\}$, adoptons la notation

$$A^\varepsilon = \begin{cases} A & \text{si } \varepsilon = 1 \\ \bar{A} & \text{si } \varepsilon = -1. \end{cases}$$

L'événement $D_{k,n}$ est la réunion des événements incompatibles

$$A(\varepsilon) = A_1^{\varepsilon_1} \cap \dots \cap A_n^{\varepsilon_n}$$

avec $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n) \in \{-1, 1\}^n$ famille comportant exactement k fois la valeur 1. Par indépendance, on obtient la probabilité d'un tel événement $A(\varepsilon)$

$$P(A(\varepsilon)) = \prod_{k=1}^n P(A_k^{\varepsilon_k}) = p^k(1-p)^{n-k}.$$

Cette probabilité ne dépend pas du choix de la famille $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$ comportant k fois la valeur 1. Au surplus, il y a exactement $\binom{n}{k}$ familles ε possibles car la détermination d'une telle famille se comprend comme le choix de k éléments dans l'ensemble $[1; n]$: ces éléments sont les positions des ε_i égaux à 1.

On peut alors calculer la probabilité de $D_{k,n}$:

$$P(D_{k,n}) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Dans le chapitre 13 sur les variables aléatoires, on dira que le nombre de succès dans une série de n tentatives indépendantes ayant une probabilité de réussite individuelle égale à p suit une loi binomiale de paramètres n et p .

(d) Notons l'événement étudié

$$E_{k,n} = \text{« L'archer touche sa } k\text{-ième cible lors du } n\text{-ième essai ».}$$

Cet événement signifie que l'archer a touché $k-1$ cibles en $n-1$ essais et touche aussi la cible lors du n -ième essai. Par indépendance¹ de ces deux événements

$$P(E_{k,n}) = P(D_{k-1,n-1}) \times P(A_n) = \binom{n-1}{k-1} p^k (1-p)^{n-k}.$$

Exercice 17 ***

(a) Soit A_1, \dots, A_n des événements d'un espace probabilisé (Ω, P) . Montrer

$$P\left(\bigcup_{i=1}^n A_i\right) \geq \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j).$$

(b) Montrer que la probabilité qu'un Valet côtoie une Dame dans un jeu mélangé de cinquante-deux cartes est supérieure à 0,47.

1. L'événement $D_{k-1,n-1}$ s'exprime à l'aide des événements A_1, \dots, A_{n-1} et est pour cette raison indépendant de A_n .

Solution

(a) On raisonne par récurrence sur $n \in \mathbb{N}^*$.

Pour $n = 1$, la propriété est immédiate (la deuxième en somme en second membre est vide donc nulle).

Pour $n = 2$, la propriété voulue se relit $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$: elle est vraie car il y a égalité.

Supposons la propriété établie au rang $n \geq 1$ et considérons A_1, \dots, A_n et A_{n+1} des événements de (Ω, \mathcal{P}) . On peut écrire

$$P\left(\bigcup_{i=1}^{n+1} A_i\right) = P(A \cup B) \quad \text{avec} \quad A = \bigcup_{i=1}^n A_i \text{ et } B = A_{n+1}.$$

L'égalité $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ donne alors

$$P\left(\bigcup_{i=1}^{n+1} A_i\right) = P\left(\bigcup_{i=1}^n A_i\right) + P(A_{n+1}) - P\left(\left(\bigcup_{i=1}^n A_i\right) \cap A_{n+1}\right).$$

Grâce à l'hypothèse de récurrence

$$P\left(\bigcup_{i=1}^n A_i\right) + P(A_{n+1}) \geq \sum_{i=1}^{n+1} P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j). \quad (*)$$

Aussi, par l'inégalité de Boole¹

$$P\left(\left(\bigcup_{i=1}^n A_i\right) \cap A_{n+1}\right) = P\left(\bigcup_{i=1}^n (A_i \cap A_{n+1})\right) \leq \sum_{i=1}^n P(A_i \cap A_{n+1}). \quad (**)$$

En combinant (*) et (**), on obtient

$$P\left(\bigcup_{i=1}^{n+1} A_i\right) \geq \sum_{i=1}^{n+1} P(A_i) - \sum_{1 \leq i < j \leq n+1} P(A_i \cap A_j).$$

La récurrence est établie.

(b) Un mélange des cinquante deux cartes du jeu s'apparente à une permutation de l'ensemble des entiers allant de 1 à 52. On modélise l'expérience par l'univers $\Omega = \mathcal{S}_{52}$ à $52!$ éléments muni de la probabilité uniforme. Dans cet univers, on étudie la probabilité de l'événement :

$E = \langle\!\langle \text{Un Valet côtoie une Dame à l'intérieur du jeu de cartes} \rangle\!\rangle$.

1. Voir p. 441.

méthode

L'événement étudié est la réunion des événements

$A_i = \text{« Un Valet côtoie une Dame en les positions } i \text{ et } i+1 \text{ ».}$

L'événement E est la réunion des événements A_i pour $i \in [1; 51]$. On applique la formule précédente à cette union ce qui nécessite d'évaluer $P(A_i)$ et $P(A_i \cap A_j)$.

Soit $i \in [1; 51]$. Pour former un élément de A_i , on choisit la Dame, le Valet ainsi que leur position relative ($4 \times 4 \times 2$ possibilités). Ceci définit les cartes en position i et $i+1$ et l'on complète les 50 autres positions par les $50!$ permutations possibles des cartes restantes. On a donc

$$P(A_i) = \frac{\text{Card}(A_i)}{\text{Card}(\Omega)} = \frac{4 \times 4 \times 2 \times 50!}{52!} = \frac{32}{52 \times 51}.$$

Soit $i < j$ deux éléments de $[1; 51]$. Pour déterminer $P(A_i \cap A_j)$ on distingue deux cas.

Cas : $j = i+1$. L'intersection $A_i \cap A_{i+1}$ se comprend comme l'encadrement d'un Valet en position $i+1$ par deux Dames ou l'inverse. Pour former cet événement, on choisit le Valet et les deux Dames ou la Dame et les deux Valets ($4 \times 4 \times 3 + 4 \times 4 \times 3$). On complète ensuite les 49 autres positions avec les permutations des cartes restantes :

$$P(A_i \cap A_{i+1}) = \frac{2 \times 4 \times 4 \times 3 \times 49!}{52!} = \frac{96}{52 \times 51 \times 50}.$$

Cas : $j > i+1$. L'intersection $A_i \cap A_j$ est obtenue par le choix des deux Dames, des deux Valets et de leurs positions relatives complétées des permutations des 48 cartes restantes :

$$P(A_i \cap A_j) = \frac{\underbrace{2 \times 4 \times 4}_{\text{première paire}} \times \underbrace{2 \times 3 \times 3}_{\text{seconde paire}} \times 48!}{52!} = \frac{576}{52 \times 51 \times 50 \times 49}.$$

Enfin, il y a 50 couples (i, j) avec $j = i+1$ et $\binom{50}{2}$ couples¹ (i, j) avec $j > i+1$. L'application de la formule de la première question donne alors

$$P(E) \geq \frac{8}{17} \approx 0,470 \text{ à } 10^{-3} \text{ près.}$$

12.4.4 Probabilités conditionnelles

Exercice 18 *

Dans une commode à 7 tiroirs figure un billet de 1 dollar avec la probabilité p . Céline a exploré sans succès les six premiers tiroirs. Quelle est la probabilité qu'elle découvre le billet dans le septième tiroir ?

1. Pour former un tel couple, on choisit deux éléments parmi les entiers de 1 à 50 et l'on ajoute 1 au plus grand.

Solution

On introduit les événements

- A_i = « Le billet se trouve dans le i -ème tiroir »,
- B = « Le billet ne figure pas dans la commode ».

Les événements A_1, \dots, A_7 et B constituent un système complet et le cadre hypothétique donne¹ $P(A_1) = \dots = P(A_7) = p/7$ et $P(B) = 1 - p$.

méthode

La probabilité cherchée est une probabilité conditionnelle. On détermine celle-ci en reprenant la formule de définition.

La probabilité de A_7 sachant $\overline{A_1}, \dots, \overline{A_6}$ est

$$P(A_7 | \overline{A_1} \cap \dots \cap \overline{A_6}) = \frac{P(\overline{A_1} \cap \dots \cap \overline{A_6} \cap A_7)}{P(\overline{A_1} \cap \dots \cap \overline{A_6})}.$$

D'une part, l'événement A_7 entraîne $\overline{A_1} \cap \dots \cap \overline{A_6}$ et donc

$$P(\overline{A_1} \cap \dots \cap \overline{A_6} \cap A_7) = P(A_7) = \frac{p}{7}.$$

D'autre part, par passage² à l'événement contraire puis calcul d'une probabilité d'événements incompatibles

$$P(\overline{A_1} \cap \dots \cap \overline{A_6}) = P(\overline{A_1 \cup \dots \cup A_6}) = 1 - P(A_1 \cup \dots \cup A_6) = 1 - \frac{6p}{7}.$$

Finalement,

$$P(A_7 | \overline{A_1} \cap \dots \cap \overline{A_6}) = \frac{\frac{p}{7}}{1 - \frac{6p}{7}} = \frac{p}{7 - 6p}.$$

Exercice 19 ** (Paradoxe des deux enfants)

Une famille a deux enfants.

- Quelle est la probabilité que les deux soient des garçons ?
- Quelle est cette probabilité sachant que l'aîné est un garçon ?
- On sait qu'au moins l'un des enfants est un garçon, quelle est la probabilité que les deux le soient ?
- On sait que l'un des deux enfants est un garçon et qu'il est né un 29 février, quelle est la probabilité que l'autre enfant soit aussi un garçon ?

1. À défaut de précisions, on peut supposer l'équiprobabilité.

2. Ici, on ne peut pas calculer la probabilité de l'intersection par le produit des probabilités car les événements A_1, \dots, A_6 ne sont pas indépendants.

Solution

Pour $i = 1, 2$ introduisons l'événement :

$$G_i = \text{« Le } i\text{-ème enfant de la famille est un garçon ».}$$

On considère¹ les événements G_1 et G_2 indépendants et l'on suppose² $P(G_i) = 0,5$. Dans chaque question, on s'intéresse à l'événement $A = G_1 \cap G_2$.

(a) Par indépendance des événements G_1 et G_2 , on a $P(A) = P(G_1) \times P(G_2) = 1/4$.

(b) Par définition d'une probabilité conditionnelle

$$P(A|G_1) = \frac{P(G_1 \cap G_2)}{P(G_1)} = P(G_2) = \frac{1}{2}.$$

(c) Par définition d'une probabilité conditionnelle et en calculant la probabilité d'une union à l'aide de celle d'une intersection

$$P(A|G_1 \cup G_2) = \frac{P(G_1 \cap G_2)}{P(G_1) + P(G_2) - P(G_1 \cap G_2)} = \frac{1}{3}.$$

(d) **méthode**

|| On introduit des événements³ liés aux dates de naissance des deux enfants.

Pour $i = 1, 2$, introduisons l'événement :

$$D_i = \text{« Le } i\text{-ème enfant de la famille est né le 29 février ».}$$

Les événements G_1 , G_2 , D_1 et D_2 sont considérés mutuellement indépendants avec, en première approximation⁴,

$$P(D_1) = P(D_2) = \frac{1}{366 + 3 \times 365} = p.$$

Savoir qu'un garçon est né le 29 février correspond à l'événement

$$B = (G_1 \cap D_1) \cup (G_2 \cap D_2).$$

1. Dans un premier temps, l'univers Ω est l'ensemble des quatre possibilités GG , GF , FG et FF exprimant le sexe des enfants selon le rang de naissance. Cet univers est muni de la probabilité uniforme et les trois premières questions de ce sujet peuvent être résolues par simples calculs dans cet espace probabilisé.

2. En première approximation seulement ! D'après l'institut national d'études démographiques, il est né ces dix dernières entre 104,5 et 105 garçons pour 100 filles en France métropolitaine.

3. Pour exprimer ces événements, il est indispensable d'agrandir l'univers Ω dans lequel on mène l'étude. On peut par exemple adjoindre le jour de naissance au sexe des enfants. Exprimer précisément cet univers Ω n'est pas utile au calcul de probabilité réalisé ici.

4. Les années bissextiles ont lieu tous les 4 ans mais les années séculaires ne sont pas bissextiles, à moins que l'année soit un multiple de 400...

On veut ici calculer $P(A|B)$. On commence par déterminer la probabilité de B qui est celle d'une union

$$P(B) = P(G_1 \cap D_1) + P(G_2 \cap D_2) - P(G_1 \cap D_1 \cap G_2 \cap D_2) = \frac{1}{2}p + \frac{1}{2}p - \frac{1}{4}p^2 = \frac{4p - p^2}{4}.$$

On poursuit en calculant la probabilité de $A \cap B$.

$$\begin{aligned} P(A \cap B) &= P((G_1 \cap G_2 \cap D_1) \cup (G_1 \cap G_2 \cap D_2)) \\ &= P(G_1 \cap G_2 \cap D_1) + P(G_1 \cap G_2 \cap D_2) - P((G_1 \cap G_2 \cap D_1 \cap D_2)) \\ &= \frac{1}{4}p + \frac{1}{4}p - \frac{1}{4}p^2 = \frac{2p - p^2}{4}. \end{aligned}$$

Finalement,

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{2p - p^2}{4p} \approx 0,5.$$

Exercice 20 **

Chaque jour du lundi au vendredi, le professeur Zinzin a la probabilité $p \in]0; 1[$ d'égarter ses notes de cours en salle de classe. Peu lui importe car il improvise à chaque fois, mais ce vendredi soir il ne les retrouve plus et cela le contrarie car il s'y trouvait un dessin de sa fille Libi. Il est cependant certain de les avoir eues en sa possession le lundi matin. Quelle est la probabilité pour le professeur d'avoir perdu ses notes de cours le mercredi ?

Solution

On introduit les événements Lu , Ma , Mc , Je , Ve associés à la perte des notes de cours les jours correspondants. Le cadre hypothétique donne

$$P(Lu) = p, \quad P(Ma|Lu) = p, \quad P(Mc|Lu \cap Ma) = p, \text{ etc.}$$

On souhaite calculer $P(Mc|Lu \cup Ma \cup Me \cup Je \cup Ve)$.

méthode

Par la formule des probabilités composées (Th. 5 p. 442), on détermine la probabilité de ne pas perdre les notes de cours durant les premiers jours de la semaine.

On a immédiatement $P(\overline{Lu}) = 1-p$. Cette probabilité est non nulle et on peut appliquer la formule des probabilités composées pour écrire

$$P(\overline{Lu} \cap \overline{Ma}) = P(\overline{Ma}|\overline{Lu}) P(\overline{Lu})$$

avec $P(\overline{Ma}|\overline{Lu}) = 1 - P(Ma|\overline{Lu}) = 1 - p$. On a donc

$$P(\overline{Lu} \cap \overline{Ma}) = (1-p)^2.$$

De même, on obtient

$$P(\overline{Lu} \cap \overline{Ma} \cap \overline{Me}) = P(\overline{Mc}|\overline{Lu} \cap \overline{Ma}) P(\overline{Lu} \cap \overline{Ma}) = (1-p)^3.$$

Plus généralement

$$P(\overline{Lu} \cap \overline{Ma} \cap \overline{Me} \cap \overline{Je}) = (1-p)^4 \quad \text{et} \quad P(\overline{Lu} \cap \overline{Ma} \cap \overline{Me} \cap \overline{Je} \cap \overline{Vc}) = (1-p)^5.$$

L'événement Me se confond avec $Me \cap \overline{Lu} \cap \overline{Ma}$ car est inclus dans $\overline{Lu} \cap \overline{Ma}$. On a donc

$$P(Me) = P(\overline{Lu} \cap \overline{Ma}) - P(\overline{Lu} \cap \overline{Ma} \cap \overline{Mc}) = (1-p)^2 - (1-p)^3 = p(1-p)^2.$$

Aussi, par passage à l'événement contraire

$$P(Lu \cup Ma \cup Mc \cup Je \cup Vc) = 1 - P(\overline{Lu} \cap \overline{Ma} \cap \overline{Me} \cap \overline{Je} \cap \overline{Vc}) = 1 - (1-p)^5.$$

On peut alors conclure

$$P(Me | Lu \cup Ma \cup Mc \cup Je \cup Vc) = \frac{p(1-p)^2}{1 - (1-p)^5}.$$

Exercice 21 **

Une lampe est éteinte dans une pièce lorsque survient une coupure d'électricité. Des individus pénètrent dans cette pièce et basculent plusieurs fois l'interrupteur en espérant que la lumière s'allume, sans succès... Quand l'électricité revient, quelle est la probabilité que la lumière soit allumée sachant que n individus sont entrés dans la pièce et que chacun a la probabilité $p \in]0; 1[$ d'avoir repositionné l'interrupteur dans l'état où celui-ci figurait lorsqu'il est entré.

Solution

(a) Introduisons les événements :

A_0 = « L'interrupteur est en position ouverte au début de l'expérience »

et, pour $k \in [1; n]$,

A_k = « L'interrupteur est en position ouverte lorsque ressort le k -ième individu »,

B_k = « L'individu k repositionne l'interrupteur dans l'état où il était en entrant ».

Le cadre hypothétique donne $P(A_0) = 1$, $P(B_k) = p \in]0; 1[$ et l'on veut déterminer $P(\overline{A_n})$.

méthode

On établit une formule de récurrence sur les $P(A_k)$ par la formule des probabilités totales (Th. 6 p. 443).

Le couple $(B_k, \overline{B_k})$ est un système complet d'événements et donc

$$P(A_{k+1}) = P(A_{k+1} | B_k) P(B_k) + P(A_{k+1} | \overline{B_k}) P(\overline{B_k}).$$

Or

$$P(A_{k+1} | B_k) = \frac{P(A_{k+1} \cap B_k)}{P(B_k)} = \frac{P(A_k \cap B_k)}{P(B_k)} = P(A_k)$$

car les événements A_k et B_k peuvent être considérés indépendants dans cette expérience.
Mutatis mutandis¹

$$P(A_{k+1} | \overline{B_k}) = P(\overline{A_k}) = 1 - P(A_k).$$

On obtient donc

$$P(A_{k+1}) = p P(A_k) + (1-p)(1 - P(A_k)) = (2p-1) P(A_k) + (1-p).$$

La suite $(P(A_k))_{0 \leq k \leq n}$ est arithmético-géométrique² et l'on obtient au terme des calculs

$$P(A_n) = \frac{1 + (2p-1)^n}{2} \quad \text{et} \quad P(\overline{A_n}) = \frac{1 - (2p-1)^n}{2}.$$

Exercice 22 **

Dans une usine, 2 % des articles produits sont défectueux. Un contrôle qualité permet d'écartier 99 % des articles lorsqu'ils sont défectueux mais aussi 5 % des articles qui ne le sont pas !

- (a) Quelle est la probabilité qu'il y ait une erreur de contrôle ?
- (b) Quelle est la probabilité qu'un article écarté par le contrôle soit défectueux ?
- (c) Quelle est la probabilité qu'un article en sortie d'usine soit défectueux ?

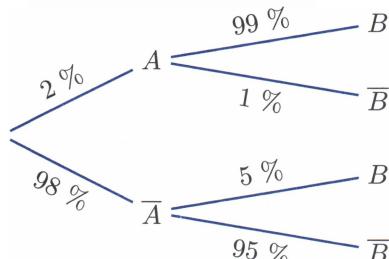
Solution

On introduit les événements :

- $A = \langle\!\langle \text{L'article produit est défectueux} \rangle\!\rangle$,
 $B = \langle\!\langle \text{Le contrôle qualité refuse l'article} \rangle\!\rangle$.

Le cadre hypothétique détermine

$$P(A) = 0,02 \quad P(B|A) = 0,99 \quad P(B|\overline{A}) = 0,05.$$



(a) Une erreur de contrôle correspond à l'événement $C = (A \cap \overline{B}) \cup (\overline{A} \cup B)$. Par union de deux événements incompatibles, on a

$$P(C) = P(A \cap \overline{B}) + P(\overline{A} \cap B).$$

La formule des probabilités composées donne alors

$$P(C) = P(\overline{B}|A) P(A) + P(B|\overline{A}) P(\overline{A})$$

avec $P(B|A) = 1 - P(\overline{B}|A)$. On obtient donc

$$P(C) = 0,01 \times 0,02 + 0,05 \times 0,98 \simeq 0,049 \text{ à } 10^{-3} \text{ près.}$$

1. Locution latine signifiant « En modifiant ce qui doit être changé ».

2. Voir section 6.7.1 du chapitre 6 de l'ouvrage *Exercices d'analyse MPSI*.

(b) On veut ici déterminer $P(A|B)$

méthode

On étudie la probabilité que l'événement A soit la cause de la réalisation de l'événement B : on emploie la formule de Bayes (Th. 7 p. 443).

Les événements A et \bar{A} forment un système complet d'événements permettant de calculer la probabilité de B par la formule des probabilités totales

$$P(B) = P(B|A)P(A) + P(B|\bar{A})P(\bar{A}) = 0,99 \times 0,02 + 0,05 \times 0,98.$$

Par la formule de Bayes

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{0,99 \times 0,02}{0,99 \times 0,02 + 0,05 \times 0,98} \simeq 0,29 \text{ à } 10^{-2} \text{ près.}$$

(c) On veut ici calculer $P(A|\bar{B})$. Il s'agit encore d'une application de la formule de Bayes

$$P(A|\bar{B}) = \frac{P(\bar{B}|A)P(A)}{P(\bar{B})} = \frac{0,01 \times 0,02}{1 - (0,99 \times 0,02 + 0,05 \times 0,98)} \simeq 5,4 \cdot 10^{-4} \text{ à } 10^{-5} \text{ près.}$$

Exercice 23 *** (Urne de Polya)

Une urne contient initialement b boules blanches et r boules rouges. On tire dans celle-ci une boule, on note sa couleur et on la remet accompagnée de d boules de la même couleur. On répète l'expérience n fois (avec $n \in \mathbb{N}^*$).

Déterminer la probabilité que la boule obtenue soit blanche lors du n -ième tirage.

Solution

Pour $k = 1, 2, \dots$ introduisons les événements :

$$B_k = \text{« Lors du } k\text{-ième tirage la boule tirée est blanche ».}$$

On veut calculer $P(B_n)$.

Au premier tirage, la composition de l'urne est connue et la probabilité que la boule tirée soit blanche est

$$P(B_1) = \frac{b}{b+r}.$$

Au second tirage, il faut tenir compte du résultat du précédent tirage. Puisque les événements B_1 et \bar{B}_1 constituent un système complet, on peut écrire par la formule des probabilités totales

$$P(B_2) = P(B_2|B_1)P(B_1) + P(B_2|\bar{B}_1)P(\bar{B}_1)$$

Le cadre de l'expérience donne

$$P(B_2 | B_1) = \frac{b+d}{b+r+d} \quad \text{et} \quad P(B_2 | \bar{B}_1) = \frac{b}{b+r+d}$$

et l'on obtient après calculs

$$P(B_2) = \frac{b+d}{b+r+d} \cdot \frac{b}{b+r} + \frac{b}{b+r+d} \cdot \frac{r}{b+r} = \frac{b}{b+r} = P(B_1). \quad (*)$$

méthode

|| On montre par récurrence que la probabilité de B_n est toujours égale à $\frac{b}{b+r}$.

Précisément, montrons par récurrence sur $n \in \mathbb{N}^*$ que la probabilité que la boule soit blanche lors du n -ième tirage vaut $\frac{b}{b+r}$ en notant b et r les nombres de boules blanches et rouges constituant *initiallement* l'urne.

La propriété est immédiatement vérifiée pour $n = 1$ et les calculs ci-dessus montrent qu'elle est aussi valable quand $n = 2$.

Supposons la propriété vraie au rang $n \geq 1$ et étudions le résultat du $(n+1)$ -ième tirage en fonction du résultat du *premier* tirage. Par la formule des probabilités totales

$$P(B_{n+1}) = P(B_{n+1} | B_1)P(B_1) + P(B_{n+1} | \bar{B}_1)P(\bar{B}_1).$$

méthode

|| Un $(n+1)$ -ième tirage sachant B_1 peut s'interpréter comme un n -ième tirage.

Si la première boule obtenue est blanche, le $(n+1)$ -ième tirage se comprend comme un n -ième tirage à partir d'une urne initialement constituée de $b+d$ boules blanches et de r boules rouges. L'hypothèse de récurrence donne alors

$$P(B_{n+1} | B_1) = \frac{b+d}{b+r+d}.$$

Mutatis mutandis

$$P(B_{n+1} | \bar{B}_1) = \frac{b}{b+r+d}.$$

Le calcul de $P(B_{n+1})$ conduit alors à répéter ceux déjà menés dans (*) et l'on obtient

$$P(B_{n+1}) = \frac{b}{b+r}.$$

La récurrence est établie.

12.5 Exercices d'approfondissement

Exercice 24 ** (Fonction indicatrice d'Euler¹)

Soit n un entier naturel supérieur à 2. On munit l'univers $\Omega = \llbracket 1 ; n \rrbracket$ de la probabilité uniforme et, pour tout entier d divisant n , on introduit l'événement

$$A_d = \{1 \leq k \leq n \mid d \text{ divise } k\}.$$

(a) Calculer $P(A_d)$.

Soit p_1, \dots, p_r les facteurs premiers de n .

(b) Montrer que les événements A_{p_1}, \dots, A_{p_r} sont mutuellement indépendants.

On note

$$B = \{1 \leq k \leq n \mid k \text{ et } n \text{ sont premiers entre eux}\}.$$

(c) Montrer

$$P(B) = \prod_{i=1}^r \left(1 - \frac{1}{p_i}\right).$$

Solution

(a) Les multiples de d dans $\llbracket 1 ; n \rrbracket$ sont $d, 2d, \dots, n$. Il y en a exactement n/d et donc

$$P(A_d) = \frac{\text{Card}(A_d)}{\text{Card}(\Omega)} = \frac{n/d}{n} = \frac{1}{d}.$$

(b) méthode

Si p et q sont deux entiers premiers entre eux, on a

$$p \mid k \text{ et } q \mid k \iff pq \mid k. \quad (*)$$

Soit $m \in \llbracket 1 ; r \rrbracket$ et $i_1 < \dots < i_m$ choisis dans $\llbracket 1 ; r \rrbracket$. Les nombres premiers p_{i_1}, \dots, p_{i_r} étant deux à deux distincts, ils sont deux à deux premiers entre eux et la propriété (*) donne pour tout $k \in \llbracket 1 ; n \rrbracket$

$$(p_{i_1} \mid k \text{ et } \dots \text{ et } p_{i_m} \mid k) \iff p_{i_1} \times \dots \times p_{i_m} \mid k.$$

On a donc

$$A_{p_{i_1}} \cap \dots \cap A_{p_{i_m}} = A_d \quad \text{avec} \quad d = p_{i_1} \times \dots \times p_{i_m}.$$

On en déduit

$$P(A_{p_{i_1}} \cap \dots \cap A_{p_{i_m}}) = \frac{1}{d} = \frac{1}{p_{i_1} \times \dots \times p_{i_m}} = \prod_{k=1}^m P(A_{p_{i_k}}).$$

Les événements A_{p_1}, \dots, A_{p_r} sont donc mutuellement indépendants.

1. Cette étude propose un calcul probabiliste des valeurs de la fonction indicatrice d'Euler présentée dans le chapitre 2 de l'ouvrage *Exercices d'algèbre et de probabilités MP* dans la même collection.

(c) Soit $k \in \llbracket 1 ; n \rrbracket$. Les entiers k et n sont premiers entre eux si, et seulement si, ils n'ont pas de diviseurs premiers en commun. Ainsi, $B = \overline{A_{p_1}} \cap \dots \cap \overline{A_{p_r}}$. Les événements A_{p_1}, \dots, A_{p_r} étant mutuellement indépendants, $\overline{A_{p_1}}, \dots, \overline{A_{p_r}}$ le sont aussi¹ et l'on a directement

$$P(B) = \prod_{i=1}^r P(\overline{A_{p_i}}) = \prod_{i=1}^r \left(1 - \frac{1}{p_i}\right).$$

Exercice 25 ** (Loi des successions de Laplace)

On dispose de $N + 1$ urnes numérotées de 0 à N . L'urne de numéro k contient k boules blanches et $N - k$ boules rouges. On choisit une urne au hasard, chaque choix étant équitable. Dans l'urne choisie, on tire des boules avec remise.

(a) Soit $n \in \mathbb{N}$. Quelle est la probabilité que la $(n + 1)$ -ième boule tirée soit blanche sachant que les n précédentes le sont toutes ?

(b) Que devient cette probabilité lorsque N tend vers l'infini ?

Solution

Pour $k \in \llbracket 0 ; N \rrbracket$ et $\ell \in \llbracket 1 ; n + 1 \rrbracket$, on introduit les événements suivants :

$A_k = \text{« L'urne choisie est celle de numéro } k \text{ »}$

$B_\ell = \text{« La } \ell\text{-ième boule tirée est blanche »}.$

Le cadre hypothétique fournit

$$P(A_k) = \frac{1}{N+1} \quad \text{et} \quad P(B_\ell | A_k) = \frac{k}{N}.$$

(a) On souhaite calculer $P(B_{n+1} | B_1 \cap \dots \cap B_n)$. La famille (A_0, A_1, \dots, A_N) constitue un système complet d'événements et la formule des probabilités totales donne

$$P(B_1 \cap \dots \cap B_n) = \sum_{k=0}^N P(B_1 \cap \dots \cap B_n | A_k) P(A_k).$$

Les tirages successifs ayant lieu avec remise dans une même urne, on a la propriété d'indépendance² qui permet d'écrire

$$P(B_1 \cap \dots \cap B_n | A_k) = \prod_{i=1}^n P(B_i | A_k) = \left(\frac{k}{N}\right)^n.$$

Ainsi,

$$P(B_1 \cap \dots \cap B_n) = \frac{1}{N+1} \sum_{k=1}^N \left(\frac{k}{N}\right)^n$$

1. Voir sujet 13 p. 457.

2. Plus précisément, le cadre hypothétique assure que les événements B_1, \dots, B_n sont indépendants dans l'univers muni de la probabilité conditionnelle P_{A_k} .

et, finalement,

$$P(B_{n+1} | B_1 \cap \dots \cap B_n) = \frac{P(B_1 \cap \dots \cap B_n \cap B_{n+1})}{P(B_1 \cap \dots \cap B_n)} = \frac{1}{N} \cdot \frac{\sum_{k=1}^N k^{n+1}}{\sum_{k=1}^N k^n}$$

(b) **méthode**

On détermine un équivalent des sommes du quotient précédent en faisant apparaître une somme de Riemann.

On écrit

$$\sum_{k=1}^N k^n = \frac{1}{N} \sum_{k=1}^N \left(\frac{k}{N}\right)^n \times N^{n+1}.$$

La fonction $f: t \mapsto t^n$ est définie et continue sur $[0; 1]$, le théorème sur les sommes de Riemann¹ donne

$$\frac{1}{N} \sum_{k=1}^N f\left(\frac{k}{N}\right) \xrightarrow[N \rightarrow +\infty]{} \int_0^1 f(t) dt = \left[\frac{1}{n+1} t^{n+1} \right]_0^1 = \frac{1}{n+1}$$

Ainsi,

$$\sum_{k=1}^N k^n \underset{N \rightarrow +\infty}{\sim} \frac{N^{n+1}}{n+1}.$$

En substituant $n+1$ à n , on détermine aussi un équivalent de la somme des k^{n+1} et l'on conclut

$$P(B_{n+1} | B_1 \cap \dots \cap B_n) \xrightarrow[N \rightarrow +\infty]{} \frac{n+1}{n+2}.$$

Exercice 26 ***

Soit $n \in \mathbb{N}^*$. À chaque suite finie $x = (x_1, \dots, x_n)$ élément de $\{-1, 1\}^n$ on associe la suite $s = (s_0, \dots, s_n)$ avec

$$s_0 \in \mathbb{Z} \quad \text{et} \quad s_k = s_{k-1} + x_k \text{ pour } k \in \llbracket 1; n \rrbracket.$$

La suite s détermine une ligne brisée articulée autour des points de coordonnées (k, s_k) avec k allant de 0 à n . On dit que cette ligne brisée détermine un chemin allant de s_0 à s_n en n étapes.

(a) Soit $\ell \in \mathbb{Z}$ et $m \in \mathbb{N}$. Combien existe-t-il de chemins allant de ℓ à m en n étapes ?

(b) On suppose $\ell, m \in \mathbb{N}$. Expliquer pourquoi il y a autant de chemins joignant $-\ell$ à m en n étapes que de chemins joignant ℓ à m en n étapes et coupant l'axe des abscisses.

(c) À une élection opposant deux candidats, l'un l'emporte avec 42 voix contre 24 pour l'autre. Quelle est la probabilité que le candidat vainqueur ait été majoritaire (au sens large) tout au long du dépouillement ?

1. Voir Th 9 du chapitre 10 de l'ouvrage *Exercices d'analyse MPSI*.

Solution

(a) On souhaite dénombrer les chemins vérifiant $s_0 = \ell$ et $s_n = m$.

Soit $x = (x_1, \dots, x_n)$ dans l'ensemble $X = \{-1, 1\}^n$. On remarque que si p désigne le nombre de 1 figurant dans x , $n - p$ est le nombre de -1 et l'on a

$$s_n = s_0 + p - (n - p) = s_0 + 2p - n.$$

On en déduit que si $m - \ell + n$ n'est pas un nombre pair, il n'y a pas de chemins solutions. Sinon, on peut introduire un entier p pour lequel $m = \ell + 2p - n$.

Cas : $p < 0$ ou $p > n$. À nouveau, il n'y a pas de chemins solutions.

Cas : $0 \leq p \leq n$. Les chemins solutions¹ correspondent aux suites x contenant p termes égaux 1 et les autres égaux à -1 . Il y a $\binom{n}{p}$ positions possibles pour les termes égaux à 1 et autant de chemins solutions.

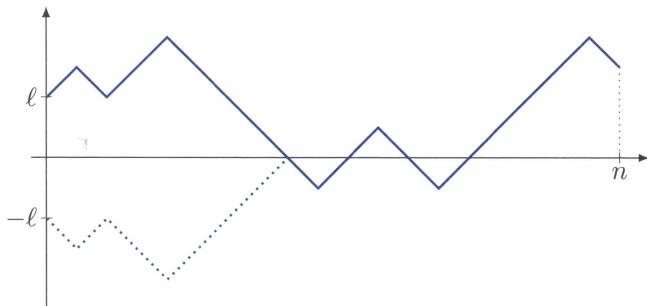
(b) méthode

Tout chemin allant de ℓ à m en n étapes et coupant l'axe des abscisses peut être associé de façon bijective à un chemin joignant $-\ell$ à m .

Si $\ell = 0$, la propriété est immédiate. Supposons pour la suite $\ell \in \mathbb{N}^*$ et considérons un chemin allant de ℓ à m en n étapes et coupant l'axe des abscisses. Ce chemin est défini à partir d'une suite $x' = (x_1, \dots, x_n)$ de X et, puisqu'il coupe l'axe des abscisses, il existe un indice k dans $\llbracket 1 ; n \rrbracket$ tel que $\ell + x_1 + \dots + x_k = 0$. Si plusieurs entiers k sont possibles considérons le plus petit de ceux-ci et introduisons la suite $x' = (x'_1, \dots, x'_n)$ de X donnée par

$$(x'_1, \dots, x'_k) = (-x_1, \dots, -x_k) \quad \text{et} \quad (x'_{k+1}, \dots, x'_n) = (x_{k+1}, \dots, x_n).$$

La suite x' permet de définir un chemin joignant $-\ell$ à m en n étapes comme illustrée ci-dessous



Ainsi, à chaque suite x de X définissant un chemin de ℓ à m coupant l'axe des abscisses on peut associer une suite x' de X définissant un chemin de $-\ell$ à m . Inversement, une suite x' déterminant un chemin de $-\ell$ à m coupe nécessairement l'axe des abscisses. En suivant une démarche symétrique à celle qui précède, on peut lui associer une suite x

1. Dans ce contexte, l'entier m figure parmi $\ell - n, \ell - n + 2, \dots, \ell + n$.

de X définissant un chemin de ℓ à m qui coupe l'axe des abscisses. Ces deux associations étant réciproques l'une de l'autre, elles sont bijectives et il y a autant¹ de chemins de ℓ à m coupant l'axe des abscisses que de chemins de $-\ell$ à m .

(c) Le dépouillement des $n = 66$ bulletins peut s'apparenter à une suite (x_1, \dots, x_n) où x_i vaut 1 si le i -ème bulletin est favorable au candidat vainqueur et -1 sinon. Sachant que le vainqueur l'emporte avec $m = 18$ voix de plus, la suite détermine un chemin joignant 0 à m en n étapes. Toutes ces suites sont équiprobables et définissent les éléments de l'univers Ω dans lequel on veut calculer la probabilité de l'événement :

$A = \langle\text{Le candidat vainqueur est majoritaire tout au long du dépouillement}\rangle.$

Les suites (x_1, \dots, x_n) éléments de l'événement A déterminent un chemin allant de 0 à m qui peut toucher l'axe des abscisses mais pas le traverser. En ajoutant une unité à l'ordonnée de chaque point du chemin, on définit un chemin allant de 1 à $m+1$ qui doit être toujours strictement au-dessus de l'axe des abscisses. Par les résultats des questions précédentes, on sait :

- il y a exactement $\binom{n}{p}$ chemins possibles avec $p = \frac{m+1}{2}$;
- il y a autant de chemins de $(0, 1)$ à $(n, m+1)$ qui touchent voire traversent l'axe des abscisses que de chemins de $(0, -1)$ à $(n, m+1)$ à savoir $\binom{n}{q}$ avec $q = \frac{m+n+2}{2}$.

La probabilité étant uniforme, on conclut

$$P(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)} = 1 - \frac{\binom{n}{q}}{\binom{n}{p}} = 1 - \frac{\binom{a+b}{a+1}}{\binom{a+b}{a}} = \frac{a-b+1}{a+1} \approx 0,442 \text{ à } 10^{-3} \text{ près}$$

avec $a = 42$ et $b = 24$ les nombres de voix recueillies par chaque candidat.

Si l'on souhaite connaître la probabilité de l'événement B que le candidat vainqueur ait été majoritaire au sens strict tout au long du dépouillement, le premier vote doit lui être attribué (probabilité $a/(a+b)$) et le dépouillement des $a+b-1$ autres bulletins ($a-1$ pour lui, b pour l'adversaire) doit satisfaire la propriété de dépouillement majoritaire calculée ci-dessus. On obtient

$$P(B) = \frac{a}{a+b} \cdot \frac{a-b}{a} = \frac{a-b}{a+b} \approx 0,273 \text{ à } 10^{-3} \text{ près.}$$

1. Ce résultat est connu sous l'appellation de *principe de réflexion*.

CHAPITRE 13

Variables aléatoires

(Ω, P) désigne un espace probabilisé fini.

13.1 Variables aléatoires sur un espace probabilisé fini

13.1.1 Variables aléatoires

Définition

Une *variable aléatoire* sur l'espace probabilisé (Ω, P) est une application X définie sur l'univers Ω et à valeurs dans un ensemble¹ E .

$$X : \begin{cases} \Omega \rightarrow E \\ \omega \mapsto X(\omega). \end{cases}$$

Lorsque la variable X prend ses valeurs dans \mathbb{R} , on dit que X est une *variable aléatoire réelle*.

Définition

Lorsque X est une variable aléatoire sur Ω à valeurs dans E et A une partie de E , on note $(X \in A)$ (ou $\{X \in A\}$) l'événement $X^{-1}(A)$.

L'événement $(X \in A)$ réunit les issues ω pour lesquelles $X(\omega) \in A$. Il s'agit d'une partie de l'univers Ω et l'on peut donc en mesurer la probabilité.

Lorsque A désigne un singleton $\{x\}$, l'événement $(X \in A)$ s'écrit² $(X = x)$.

1. Sans perte de généralité on peut supposer si besoin l'ensemble E fini car l'univers Ω l'est.
2. Si X est une variable réelle et $x \in \mathbb{R}$, l'événement $(X = x)$ désigne $(X \in]-\infty ; x])$, etc.

13.1.2 Loi d'une variable aléatoire

Soit X une variable aléatoire définie sur l'espace probabilisé (Ω, P) et E un ensemble fini contenant¹ l'ensemble $X(\Omega)$ des valeurs prises par X .

Définition

On appelle *loi sur*² E de la variable aléatoire X l'application

$$P_X : \wp(E) \rightarrow [0; 1]$$

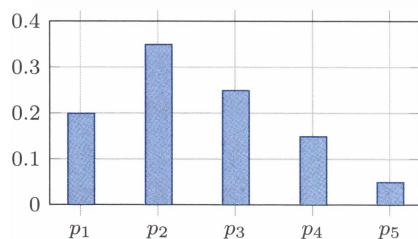
définie par $P_X(A) \stackrel{\text{def}}{=} P(X \in A)$ pour toute partie A de E .

Théorème 1

La loi P_X définit une probabilité sur E et celle-ci est entièrement déterminée par les valeurs de $P(X = x)$ pour x dans E .

Si x_1, \dots, x_n constituent une énumération des éléments de E , les $p_i = P(X = x_i)$ déterminent une famille (p_1, \dots, p_n) de réels positifs de somme égale à 1. La loi de X est souvent confondues avec cette famille qui peut être figurée par un tableau à une ligne ou par un diagramme en bâton comme ci-contre.

x_1	x_2	x_3	x_4	x_5
0,20	0,35	0,25	0,15	0,05



13.1.3 Lois usuelles

Définition

On dit qu'une variable aléatoire X suit une *loi uniforme* sur un ensemble E fini à $n \in \mathbb{N}^*$ éléments lorsque $X(\Omega) = E$ et $P(X = x) = 1/n$ pour tout x de E .

Définition

On dit qu'une variable aléatoire X suit une *loi de Bernoulli* de paramètre $p \in [0; 1]$ (et l'on note $X \sim \mathcal{B}(p)$) lorsque $X(\Omega) \subset \{0, 1\}$ et que l'on a

$$P(X = 0) = 1 - p \quad \text{et} \quad P(X = 1) = p.$$

Les lois de Bernoulli servent à modéliser les épreuves à deux issues : succès (valeur 1, probabilité p) ou échec (valeur 0, probabilité $q = 1 - p$).

Si A est un événement de Ω , la fonction indicatrice $\mathbf{1}_A$ suit une loi de Bernoulli de paramètre $p = P(A)$ pour laquelle l'appartenance à A s'apparente à un succès.

1. En pratique, E est souvent l'ensemble $X(\Omega)$ mais ce peut aussi être un ensemble plus grand.
2. Très couramment, on parle simplement de la *loi* de X .

Définition

On dit qu'une variable aléatoire X suit une *loi binomiale* de paramètres¹ $n \in \mathbb{N}$ et $p \in [0; 1]$ (et note $X \sim B(n, p)$) lorsque $X(\Omega) \subset \llbracket 0; n \rrbracket$ et

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \text{ pour tout } k \in \llbracket 0; n \rrbracket.$$

Dans la suite, on énoncera que le nombre de succès lors de la répétition de n épreuves de Bernoulli indépendantes de même paramètre $p \in [0; 1]$ suit une loi binomiale de paramètres n et p .

13.1.4 Variable aléatoire composée

Soit X une variable aléatoire sur (Ω, P) et f une application définie sur l'ensemble des valeurs prises par X .

Définition

|| On note $f(X)$ la variable aléatoire obtenue par la composition $f \circ X$.

Lorsque la fonction f présente une notation usuelle, on adapte celle-ci à la description de la variable composée. C'est ainsi que pour une variable aléatoire réelle X , on peut écrire

$$X^2, |X|, e^X, \dots$$

Théorème 2

La loi de la variable aléatoire $Y = f(X)$ est entièrement déterminée par celle de X .

13.1.5 Opérations sur les variables aléatoires réelles

Les variables aléatoires réelles sur Ω correspondent aux applications de Ω vers une partie de \mathbb{R} . Elles peuvent se comprendre comme des fonctions de Ω vers \mathbb{R} et, à ce titre, il est possible d'opérer par addition, multiplication, etc. sur les variables aléatoires réelles comme on le fait avec les fonctions réelles. En particulier, l'ensemble des variables aléatoires réelles sur Ω est un espace vectoriel réel.

13.2 Vecteurs aléatoires**13.2.1 Loi conjointe**

Soit X et Y des variables aléatoires sur (Ω, P) à valeurs dans des ensembles finis E et F respectivement.

1. On dit aussi que X suit une loi binomiale de *taille* n et de *probabilité de réussite* p .

Définition

On appelle *couple de variables aléatoires* défini par X et Y la variable aléatoire¹ $Z = (X, Y)$ déterminée par

$$Z: \begin{cases} \Omega \rightarrow E \times F \\ \omega \mapsto (X(\omega), Y(\omega)) \end{cases}$$

La loi de la variable Z est appelée la *loi conjointe* de X et Y .

La loi de Z est souvent figurée par un tableau dont les entrées sont les valeurs possibles de X et Y .

13.2.2 Lois marginales

Soit Z une variable aléatoire sur (Ω, P) à valeurs dans un ensemble fini $E \times F$. Pour chaque ω de Ω , $Z(\omega)$ désigne un couple $(X(\omega), Y(\omega))$ avec $X(\omega) \in E$ et $Y(\omega) \in F$ ce qui introduit deux variables aléatoires X et Y dont Z est le couple.

Définition

Les lois des variables aléatoires X et Y sont appelées les *lois marginales* de Z .

Théorème 3

La loi de Z détermine entièrement ses lois marginales car

$$P(X = x) = \sum_{y \in F} P((X, Y) = (x, y)) \quad \text{pour tout } x \in E \text{ et}$$

$$P(Y = y) = \sum_{x \in E} P((X, Y) = (x, y)) \quad \text{pour tout } y \in F.$$

Lorsque l'on visualise la loi conjointe par un tableau, les lois marginales s'obtiennent en sommant sur les rangées².

13.2.3 Lois conditionnelles

Soit X et Y des variables aléatoires sur (Ω, P) à valeurs dans des ensembles finis E et F respectivement et Z le couple (X, Y) .

Définition

Pour $x \in E$ tel que $P(X = x) > 0$, on appelle *loi conditionnelle* de Y sachant $(X = x)$ la loi de Y pour la probabilité conditionnelle $P_{(X=x)}$. Cette loi conditionnelle est déterminée par

$$P(Y = y | X = x) \stackrel{\text{def}}{=} \frac{P(X = x; Y = y)}{P(X = x)} \quad \text{pour tout } y \in F.$$

1. On parle quelquefois de *vecteur aléatoire* lorsque l'on étudie une variable aléatoire à valeurs dans un espace produit.

2. Voir sujet 2 p. 486.

Théorème 4

La loi de X et les lois conditionnelles de Y connaissant les valeurs prises par X déterminent entièrement la loi conjointe de X et Y et donc la loi de Y :

$$P(Y = y) = \sum_x P(Y = y | X = x) P(X = x) \quad \text{pour tout } y \in F$$

où la somme porte¹ sur les $x \in E$ vérifiant $P(X = x) > 0$.

13.2.4 Couples de variables indépendantes

Soit X et Y des variables aléatoires sur (Ω, P) à valeurs dans des ensembles E et F .

Définition

On dit que les deux variables aléatoires X et Y sont *indépendantes* si, pour toute partie A de E et toute partie B de F , les événements $(X \in A)$ et $(Y \in B)$ sont indépendants :

$$P((X, Y) \in A \times B) = P(X \in A) P(Y \in B).$$

Théorème 5

Les variables aléatoires X et Y sont indépendantes si, et seulement si,

$$P(X = x, Y = y) = P(X = x) P(Y = y) \quad \text{pour tout } (x, y) \in X(\Omega) \times Y(\Omega).$$

Deux événements A et B sont indépendants si, et seulement si, les fonctions indicatrices $\mathbf{1}_A$ et $\mathbf{1}_B$ définissent des variables aléatoires indépendantes.

Théorème 6

Si les variables X et Y sont indépendantes alors, pour toutes fonctions f et g définies respectivement sur E et F , les variables composées $f(X)$ et $g(Y)$ sont indépendantes.

13.2.5 Variables mutuellement indépendantes

Soit X_1, \dots, X_n des variables aléatoires sur l'espace (Ω, P) à valeurs dans des ensembles finis E_1, \dots, E_n .

Définition

On dit que les variables X_1, \dots, X_n sont *mutuellement indépendantes* si, pour toutes parties A_1, \dots, A_n de E_1, \dots, E_n , les événements $(X_i \in A_i)$ sont mutuellement indépendants.

1. On pourrait limiter la somme aux x éléments de $X(\Omega)$ mais cela ne suffit pas pour pouvoir affirmer $P(X = x) > 0$: il est possible qu'une valeur prise par la variable X soit associée à des issues ω de probabilités nulles.

Théorème 7

Les variables X_1, \dots, X_n sont mutuellement indépendantes si, et seulement si,

$$P(X_1 = x_1, \dots, X_n = x_n) = P(X_1 = x_1) \times \dots \times P(X_n = x_n)$$

pour tout $(x_1, \dots, x_n) \in X_1(\Omega) \times \dots \times X_n(\Omega)$.

13.2.6 Schéma de Bernoulli

Définition

Un schéma de Bernoulli de paramètre $p \in [0; 1]$ est une suite X_1, \dots, X_n de variables aléatoires mutuellement indépendantes suivant chacune une même loi de Bernoulli de paramètre p .

Un schéma de Bernoulli se rencontre lorsque l'on étudie la répétition d'épreuves de Bernoulli indépendantes. La variable $S = X_1 + \dots + X_n$ donne alors le nombre de succès rencontrés.

Théorème 8

Si X_1, \dots, X_n sont des variables définissant un schéma de Bernoulli de paramètre p , la variable $S = X_1 + \dots + X_n$ suit une loi binomiale de paramètres n et p .

13.3 Espérance et variance d'une variable aléatoire réelle

13.3.1 Espérance

Soit X une variable aléatoire réelle sur (Ω, P) et E un ensemble fini contenant les valeurs prises par X .

Définition

L'espérance de la variable aléatoire réelle X est le réel

$$E(X) = \sum_{x \in E} x P(X = x)$$

L'espérance d'une variable aléatoire réelle est entièrement déterminée par sa loi, elle définit la moyenne probabiliste de la variable X .

On dit qu'une variable aléatoire est *centrée* lorsque son espérance est nulle.

Si X est une variable aléatoire constante égale à un réel C , $E(X) = C$.

Si X est la fonction indicatrice d'un événement A de Ω , $E(\mathbf{1}_A) = P(A)$.

Si X suit une loi de Bernoulli de paramètre p , $E(X) = p$.

Si X suit une loi binomiale de paramètres n et p , $E(X) = np$.

L'espérance de la variable X peut aussi être calculée par la formule

$$E(X) = \sum_{\omega \in \Omega} P(\{\omega\}) X(\omega).$$

13.3.2 Propriétés

Théorème 9 (Linéarité de l'espérance)

Si X et Y sont des variables aléatoires réelles sur Ω alors, pour tout $\lambda \in \mathbb{R}$,

$$E(\lambda X) = \lambda E(X) \quad \text{et} \quad E(X + Y) = E(X) + E(Y).$$

Si a et b sont deux réels, $E(aX + b) = aE(X) + b$. En particulier, si X est une variable aléatoire d'espérance μ , la variable $X - \mu$ est centrée.

Théorème 10

Si X est une variable aléatoire sur Ω à valeurs dans \mathbb{R}_+ alors $E(X) \geq 0$.

On en déduit la propriété de croissance de l'espérance :

$$X \leq Y \implies E(X) \leq E(Y).$$

Théorème 11 (Formule de transfert)

Si X est une variable aléatoire réelle sur Ω et f une fonction réelle définie sur un ensemble E contenant les valeurs prises par X

$$E(f(X)) = \sum_{x \in E} f(x) P(X = x).$$

En particulier, $m_k = E(X^k) = \sum_{x \in E} x^k P(X = x)$ pour tout $k \in \mathbb{N}$.

Théorème 12

Si X et Y sont deux variables aléatoires réelles indépendantes

$$E(XY) = E(X) E(Y).$$

La réciproque est fausse.

13.3.3 Variance

Soit X une variable aléatoire réelle sur (Ω, P) .

Définition

On appelle *moment d'ordre* $k \in \mathbb{N}$ de la variable X le réel $m_k = E(X^k)$.

Le moment d'ordre 0 vaut 1, le moment d'ordre 1 est l'espérance.

Définition

On appelle *variance* de la variable X le réel positif

$$V(X) \stackrel{\text{def}}{=} E((X - E(X))^2).$$

On définit aussi l'*écart-type*¹ de X par $\sigma(X) = \sqrt{V(X)}$. Variance et écart-type quantifient « l'écart d'une variable aléatoire à sa moyenne ».

Théorème 13

Si a et b sont deux réels, $V(aX + b) = a^2 V(X)$.

Si X est une variable aléatoire d'espérance μ et de variance $\sigma^2 > 0$, la variable $\frac{X-\mu}{\sigma}$ est d'espérance nulle et de variance 1, on dit qu'elle est *centrée réduite*.

Théorème 14 (Formule de Huygens)

$$V(X) = E(X^2) - E(X)^2.$$

Si X est une variable aléatoire constante, $V(X) = 0$.

Si X suit une loi de Bernoulli de paramètre p , $V(X) = p(1-p)$.

Si X suit une loi binomiale de paramètres n et p , $V(X) = np(1-p)$.

13.3.4 Covariance

Soit X et Y deux variables aléatoires réelles sur (Ω, P) .

Définition

On appelle *covariance* des variables X et Y le réel

$$\text{Cov}(X, Y) \stackrel{\text{def}}{=} E((X - E(X))(Y - E(Y))).$$

En particulier, $V(X) = \text{Cov}(X, X)$.

Théorème 15 (Formule de Huygens)

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y).$$

Lorsque les variables X et Y sont indépendantes, $\text{Cov}(X, Y) = 0$.

Par linéarité de l'espérance, l'application covariance est bilinéaire et par conséquent :

Théorème 16 (Variance d'une somme)

$$V(X + Y) = V(X) + 2 \text{Cov}(X, Y) + V(Y).$$

En particulier, la variance d'une somme de variables aléatoires réelles deux à deux indépendantes est égale à la somme des variances.

1. L'espérance et l'écart-type s'expriment dans la même unité que les valeurs de la variable X .

13.3.5 Inégalités de concentration

Théorème 17 (Inégalité de Markov)

Si X est une variable aléatoire réelle à valeurs positives,

$$a P(X \geq a) \leq E(X) \quad \text{pour tout } a \in \mathbb{R}_+.$$

Théorème 18 (Inégalité de Bienaymé-Tchebychev)

Si X est une variable aléatoire réelle d'espérance μ et de variance σ^2 alors, pour tout réel α strictement positif,

$$P(|X - \mu| \geq \alpha) \leq \frac{\sigma^2}{\alpha^2}.$$

Cette inégalité permet d'étudier l'écart d'une variable à son espérance.

13.4 Exercices d'apprentissage

Exercice 1

On tire successivement $n \in \mathbb{N}^*$ boules dans une urne contenant b boules blanches et r boules rouges et l'on pose X le nombre de boules blanches obtenues.

- (a) Déterminer la loi de X lorsque le tirage a lieu avec remise.
- (b) Déterminer la loi de X lorsque le tirage a lieu sans remise et que $n \leq b + r$.

Solution

(a) méthode

|| Un tirage avec remise s'apparente à une succession d'épreuves de Bernoulli indépendantes : la variable X suit une loi binomiale.

Si l'on interprète le tirage d'une boule blanche comme un succès, la variable X donne le nombre de succès pour n épreuves indépendantes (car le tirage a lieu avec remise) ayant chacune la probabilité $p = b/(b + r)$ de réussite. La variable X suit donc une loi binomiale de paramètres n et p :

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} \frac{b^k r^{n-k}}{(b+r)^n} \quad \text{pour tout } k \in \llbracket 0 ; n \rrbracket.$$

(b) Distinguons les différentes boules contenues dans l'urne, par exemple en les numérotant de 1 à $N = b + r$. Le tirage ayant lieu sans remise, on peut choisir pour univers Ω associé à l'expérience l'ensemble des arrangements de longueur n d'éléments de $\llbracket 1 ; N \rrbracket$ muni de la probabilité uniforme. Le cardinal de Ω est $N(N-1) \times \cdots \times (N-n+1)$.

Pour $k \in \llbracket 0 ; n \rrbracket$, l'événement $(X = k)$ correspond à l'ensemble des arrangements contenant exactement k boules blanches. Dénombrons cet ensemble. Pour construire un arrangement élément de l'événement $(X = k)$, on commence par choisir les k positions occupées

par les boules blanches ($\binom{n}{k}$ possibilités). On choisit ensuite les k boules blanches occupant ces positions successives parmi les b boules disponibles ($b \times (b-1) \times \cdots \times (b-k+1)$ possibilités¹). Enfin, on complète les $\ell = n - k$ positions encore libres par le choix de $n - k$ boules distinctes parmi les r boules rouges ($r \times (r-1) \times \cdots \times (r-\ell+1)$ possibilités). La probabilité étant uniforme, on obtient

$$P(X = k) = \begin{cases} \frac{(N-n)!}{N!} \cdot \frac{n!}{k!\ell!} \cdot \frac{b!}{(b-k)!} \cdot \frac{r!}{(r-\ell)!} & \text{si } k \leq b \text{ et } \ell = n - k \leq r \\ 0 & \text{sinon.} \end{cases}$$

De façon simplifiée, on peut² écrire³ pour tout $k \in \llbracket 0 ; n \rrbracket$

$$P(X = k) = \frac{\binom{b}{k} \binom{r}{\ell}}{\binom{N}{n}} \quad \text{avec } N = a + b \text{ et } n = k + \ell.$$

Exercice 2

On lance deux dés équilibrés et l'on note Y et Z les plus petite et plus grande valeurs obtenues.

- (a) Par un tableau, déterminer la loi conjointe de Y et Z .
- (b) En déduire les lois de Y et Z .
- (c) Déterminer les lois de Z sachant Y .

Solution

Supposons pouvoir distinguer les deux dés et notons X_1 et X_2 les valeurs de chacun⁴. Les variables aléatoires X_1 et X_2 sont indépendantes et suivent une loi uniforme sur $\llbracket 1 ; 6 \rrbracket$. Les variables Y et Z correspondent alors

$$Y = \min(X_1, X_2) \quad \text{et} \quad Z = \max(X_1, X_2).$$

(a) Le vecteur aléatoire (Y, Z) est à valeurs dans $E = \llbracket 1 ; 6 \rrbracket^2$. Soit (i, j) un élément de E . Étudions la probabilité de l'événement $A_{i,j} = \{(Y, Z) = (i, j)\}$.

Cas : $i > j$. L'événement $A_{i,j}$ est impossible et donc de probabilité nulle.

Cas : $i = j$. L'événement $A_{i,i}$ se confond avec $(X_1 = i) \cap (X_2 = i)$. Par indépendance

$$P(A_{i,i}) = P(X_1 = i) P(X_2 = i) = \frac{1}{36}.$$

Cas : $i < j$. L'événement $A_{i,j}$ est la réunion des événements $(X_1 = i) \cap (X_2 = j)$ et $(X_1 = j) \cap (X_2 = i)$ qui sont incompatibles donc

$$P(A_{i,j}) = P(X_1 = i) P(X_2 = j) + P(X_1 = j) P(X_2 = i) = \frac{1}{18}.$$

1. Il peut figurer un zéro dans ce produit s'il n'y a pas suffisamment de boules blanches pour constituer un tel arrangement.

2. Rappelons que $\binom{n}{p} = 0$ lorsque $p > n$.

3. On dit que la variable X suit une *loi hypergéométrique* de paramètres n , b et N .

4. On peut modéliser l'expérience par l'univers $\Omega = \llbracket 1 ; 6 \rrbracket^2$ muni de la probabilité uniforme mais il n'est pas nécessaire de l'exprimer pour la suite.

On résume ces valeurs dans le tableau¹ ci-dessous :

	$Z = 1$	$Z = 2$	$Z = 3$	$Z = 4$	$Z = 5$	$Z = 6$
$Y = 1$	1/36	1/18	1/18	1/18	1/18	1/18
$Y = 2$	0	1/36	1/18	1/18	1/18	1/18
$Y = 3$	0	0	1/36	1/18	1/18	1/18
$Y = 4$	0	0	0	1/36	1/18	1/18
$Y = 5$	0	0	0	0	1/36	1/18
$Y = 6$	0	0	0	0	0	1/36

(b) **méthode**

|| Les lois de Y et Z s'obtiennent en sommant sur les rangées (Th. 3 p. 480).

Pour $k \in [1; 6]$, on a

$$P(Y = k) = \sum_{j=1}^6 P(Y = k, Z = j) \quad \text{et} \quad P(Z = k) = \sum_{i=1}^6 P(Y = i, Z = k).$$

On peut résumer le résultat de ces calculs² dans le tableau ci-dessous³ :

	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
$P(Y = k)$	11/36	9/36	7/36	5/36	3/36	1/36
$P(Z = k)$	1/36	3/36	5/36	7/36	9/36	11/36

(c) Dans le tableau figurant la loi conjointe, la loi de Z sachant ($Y = i$) est déterminée par les proportions des valeurs $P(Y = i, Z = j)$ vis-à-vis de $P(Y = i)$, c'est-à-dire vis-à-vis de la somme des valeurs de la rangée. On peut résumer le calcul de ces lois par le tableau suivant donnant $P(Z = j | Y = i)$ selon les valeurs de i et j :

	$j = 1$	$j = 2$	$j = 3$	$j = 4$	$j = 5$	$j = 6$
$i = 1$	1/11	2/11	2/11	2/11	2/11	2/11
$i = 2$	0	1/9	2/9	2/9	2/9	2/9
$i = 3$	0	0	1/7	2/7	2/7	2/7
$i = 4$	0	0	0	1/5	2/5	2/5
$i = 5$	0	0	0	0	1/3	2/3
$i = 6$	0	0	0	0	0	1

1. On pourra observer que la somme des valeurs du tableau est égale à 1.

2. Un calcul de direct des lois de Y et Z est aussi possible. Par exemple, pour déterminer la loi de Z on commence par évaluer $P(Z \leq k) = (k/6)^2$ pour tout $k \in [1; 6]$ puis on détermine $P(Z = k)$ comme la différence $P(Z \leq k) - P(Z \leq k - 1) = (2k - 1)/36$.

3. On pourra remarquer que les variables Y et $7 - Z$ suivent la même loi ce qui ne signifie pour autant que ces variables sont égales : la loi ne caractérise pas la variable aléatoire.

Exercice 3

Deux variables aléatoires¹ indépendantes X et Y suivent des lois binomiales de tailles n et m et de même probabilité de réussite $p \in]0; 1[$.

- (a) Identifier la loi suivie par la variable aléatoire $S = X + Y$.
- (b) Soit $s \in [0; n + m]$. Identifier la loi de X sachant $(S = s)$.

Solution

(a) Les variables X et Y prennent leurs valeurs dans $[0; n]$ et $[0; m]$ respectivement et leur somme S prend donc ses valeurs dans $[0; n + m]$.

Soit $s \in [0; n + m]$.

méthode

|| On exprime l'événement $(S = s)$ selon les valeurs possibles prises par X et Y .

L'événement $(S = s)$ se décompose² en la réunion des événements deux à deux incompatibles $(X = i) \cap (Y = j)$ pour i et j naturels vérifiant $i + j = s$. Par additivité

$$\mathbb{P}(S = s) = \sum_{i+j=s} \mathbb{P}(X = i, Y = j).$$

Par indépendance des variables aléatoires X et Y ,

$$\mathbb{P}(X = i, Y = j) = \mathbb{P}(X = i) \times \mathbb{P}(Y = j) = \binom{n}{i} p^i (1-p)^{n-i} \times \binom{m}{j} p^j (1-p)^{m-j}$$

et donc

$$\mathbb{P}(S = s) = \sum_{i+j=s} \binom{n}{i} \binom{m}{j} p^s (1-p)^{n+m-s}.$$

méthode

|| Le coefficient de X^s dans $(1+X)^n(1+X)^m$ détermine la valeur³ de la somme

$$\sum_{i+j=s} \binom{n}{i} \binom{m}{j}.$$

Dans l'égalité $(1+X)^n(1+X)^m = (1+X)^{n+m}$, l'identification des coefficients de X^s donne

$$\sum_{i+j=s} \binom{n}{i} \binom{m}{j} = \binom{n+m}{s}$$

1. À défaut de précision, les variables aléatoires introduites sont toujours supposées définies sur un même espace probabilisé fini (Ω, \mathcal{P}) .

2. Les événements $(Y = j)$ avec j allant de 0 à m constituent un système complet d'événements et $(S = s)$ est alors la réunion des événements incompatibles $(S = s) \cap (Y = j)$. Or ce dernier se confond avec $(X = i) \cap (Y = j)$ pour $i = s - j$.

3. Voir sujet 22 p. 76.

et donc

$$P(S = s) = \binom{n+m}{s} p^s (1-p)^{n+m-s}$$

Finalement, la variable aléatoire S suit une loi binomiale¹ de taille $n+m$ et de probabilité de réussite p .

(b) Sachant ($S = s$), la variable aléatoire X prend ses valeurs dans $\llbracket 0; s \rrbracket$ et, par définition d'une probabilité conditionnelle, on a pour tout $i \in \llbracket 0; s \rrbracket$

$$P(X = i | S = s) = \frac{P(X = i, S = s)}{P(S = s)} = \frac{P(X = i, Y = j)}{P(S = s)} \quad \text{avec } i + j = s.$$

On poursuit le calcul à l'aide de l'indépendance des variables X et Y

$$P(X = i | S = s) = \frac{P(X = i) P(Y = j)}{P(S = s)} = \frac{\binom{n}{i} p^i (1-p)^{n-i} \binom{m}{j} p^j (1-p)^{m-j}}{\binom{n+m}{s} p^s (1-p)^{n+m-s}}.$$

Après simplifications, on conclut²

$$P(X = i | S = s) = \frac{\binom{n}{i} \binom{m}{j}}{\binom{n+m}{s}}.$$

Exercice 4

Soit X une variable aléatoire à valeurs dans $\llbracket 0; n \rrbracket$. On suppose qu'il existe un réel a tel que

$$P(X = k) = \frac{a}{k!(n-k)!} \quad \text{pour tout } k \in \llbracket 0; n \rrbracket.$$

Calculer l'espérance et la variance de X .

Solution

On commence par déterminer la valeur de a .

méthode

Si une variable aléatoire X prend ses valeurs dans un ensemble fini E , la somme des $P(X = x)$ pour x parcourant E est égale à 1.

La valeur de a se déduit de l'égalité

$$\sum_{k=0}^n P(X = k) = 1.$$

1. Les variables X et Y peuvent être déterminées par la somme de n et m variables de Bernoulli mutuellement indépendantes et de même paramètre p (Th. 8 p. 482). La variable $S = X + Y$ se comprend alors comme la somme de $n+m$ variables de Bernoulli indépendantes de paramètre p .

2. On peut reconnaître une loi hypergéométrique de paramètres s , n et $n+m$ (voir sujet 1 p. 485).

Sachant

$$\sum_{k=0}^n \frac{1}{k!(n-k)!} = \frac{1}{n!} \sum_{k=0}^n \binom{n}{k} = \frac{(1+1)^n}{n!} = \frac{2^n}{n!}$$

on obtient $a = n!/2^n$ et donc, pour tout $k \in [0; n]$,

$$P(X = k) = \binom{n}{k} \frac{1}{2^k} \cdot \frac{1}{2^{n-k}}.$$

La variable aléatoire X suit une loi binomiale de paramètres n et $p = 1/2$.

méthode

Lorsque X suit une loi binomiale de paramètres n et p

$$E(X) = np \quad \text{et} \quad V(X) = np(1-p).$$

Ici, on a donc

$$E(X) = \frac{n}{2} \quad \text{et} \quad V(X) = \frac{n}{4}.$$

Exercice 5

Soit $a < b$ deux entiers. Calculer l'espérance et la variance d'une variable X suivant une loi uniforme sur $[a; b]$.

Solution

La variable X prend ses valeurs dans l'ensemble $[a; b]$ à $n = b - a + 1$ éléments et

$$P(X = k) = \frac{1}{n} \quad \text{pour tout } k \in [a; b].$$

méthode

Afin de faciliter les calculs, on écrit $X = a - 1 + Y$ avec Y variable aléatoire suivant une loi uniforme sur $[1; n]$.

L'espérance et la variance de X se déduisent des espérance et variance de Y

$$E(X) = a - 1 + E(Y) \quad \text{et} \quad V(X) = V(Y).$$

Un calcul direct détermine l'espérance de Y

$$E(Y) = \sum_{k=1}^n k P(X = k) = \sum_{k=1}^n k \cdot \frac{1}{n} = \frac{n+1}{2} \quad \text{car} \quad \sum_{k=1}^n k = \frac{n(n+1)}{2}.$$

La variance de Y peut être obtenue par la formule de Huygens (Th. 14 p. 484).

$$E(Y^2) = \sum_{k=1}^n k^2 P(X = k) = \sum_{k=1}^n \frac{k^2}{n} = \frac{(n+1)(2n+1)}{6} \quad \text{car} \quad \sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$$

et donc

$$V(Y) = E(Y^2) - E(Y)^2 = \frac{n^2 - 1}{12}.$$

Finalement,

$$E(X) = \frac{a+b}{2} \quad \text{et} \quad V(X) = \frac{(b-a+1)^2 - 1}{12}.$$

13.5 Exercices d'entraînement

13.5.1 Loi binomiale

Exercice 6 *

Une variable aléatoire X suit une loi binomiale paramètres $n \in \mathbb{N}^*$ et $p \in]0; 1[$.

Pour quelle valeur de l'entier $k \in [0; n]$, la probabilité $P(X = k)$ est-elle maximale ?

Solution

Par définition d'une loi binomiale, on a

$$u_k = P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{pour tout } k \in [0; n].$$

méthode

|| On étudie la monotonie de la suite finie (u_0, u_1, \dots, u_n) .

Etudions le quotient de deux termes successifs de la suite $(u_k)_{0 \leq k \leq n}$. Pour $k \in [1; n]$

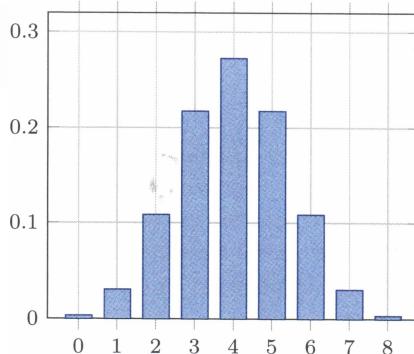
$$\frac{u_k}{u_{k-1}} = \frac{\binom{n}{k} p^k (1-p)^{n-k}}{\binom{n}{k-1} p^{k-1} (1-p)^{n-k+1}} = \frac{n-k+1}{k} \cdot \frac{p}{1-p}.$$

On en déduit

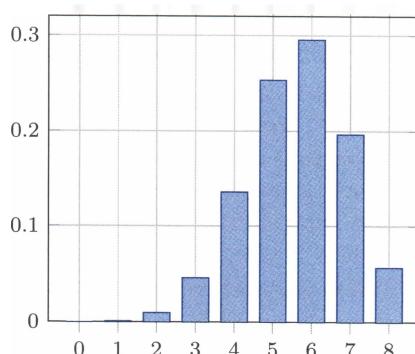
$$\begin{aligned} u_{k-1} < u_k &\iff 1 < \frac{u_k}{u_{k-1}} \\ &\iff k \leq (n+1)p \\ &\iff k \leq k_0 \quad \text{avec } k_0 \text{ la partie entière de } (n+1)p. \end{aligned}$$

La suite $(u_k)_{0 \leq k \leq k_0}$ est croissante tandis que la suite $(u_k)_{k_0 \leq k \leq n}$ est strictement décroissante. Le maximum de la suite $(u_k)_{0 \leq k \leq n}$ est donc atteint en $k = k_0$.

Loi binomiale $n = 8$, $p = 0,5$ et $k_0 = 4$



Loi binomiale $n = 8$, $p = 0,7$ et $k_0 = 6$



Exercice 7 *

Deux archers tirent indépendamment sur n cibles. À chaque tir, le premier archer a la probabilité p de toucher et le second la probabilité q .

Quelle est la loi suivie par le nombre de cibles touchées au moins une fois ?

Solution**méthode**

|| On introduit des variables de Bernoulli modélisant les résultats des archers sur chaque cible.

Numérotons les cibles de 1 à n . Pour $i \in [1 ; n]$, définissons la variable aléatoire X_i égale à 1 lorsque la cible i est touchée par le premier archer et 0 sinon. Définissons de même la variable Y_i associée au succès du second archer. Les variables X_1, \dots, X_n et Y_1, \dots, Y_n sont mutuellement indépendantes, chaque X_i suit une loi de Bernoulli de paramètre p tandis que Y_i suit une loi de paramètre q . La variable $Z_i = \max(X_i, Y_i)$ détermine si une cible a été atteinte au moins une fois. Le nombre de cibles touchées au moins une fois est donc

$$N = \sum_{i=1}^n Z_i.$$

méthode

|| Les variables Z_i suivent une même loi de Bernoulli.

La variable Z_i prend pour valeurs 0 ou 1 et l'on a $Z_i = 0$ si, et seulement si, $X_i = Y_i = 0$. Par indépendance

$$\mathrm{P}(Z_i = 0) = \mathrm{P}(X_i = 0, Y_i = 0) = \mathrm{P}(X_i = 0) \mathrm{P}(Y_i = 0) = (1-p)(1-q).$$

La variable Z_i suit donc une loi de Bernoulli de paramètre

$$r = 1 - (1-p)(1-q) = p + q - pq.$$

Par somme de variables de Bernoulli indépendantes et de même paramètre r (Th. 8 p. 482), la variable N suit une loi binomiale de paramètres n et r .

Exercice 8 **

Le standardiste d'un centre d'appel téléphonique d'un service après-vente a la probabilité p d'apporter une solution à l'appel d'un client. Lorsqu'il n'y parvient pas, il transmet l'appel à un spécialiste qui a la probabilité q de résoudre le problème posé. Si ce dernier n'y parvient pas, un réparateur est envoyé au domicile du client.

On suppose que le centre d'appel a été contacté n fois. Déterminer la loi de la variable aléatoire N donnant le nombre d'interventions du réparateur.

Solution

On note X et Y les variables aléatoires donnant le nombre d'appels résolus par respectivement le standardiste et le spécialiste. On veut déterminer $N = n - (X + Y)$.

Les appels sont indépendants et chacun détermine une variable de Bernoulli de paramètre p prenant la valeur 1 lorsque le standardiste apporte une solution. La variable X suit alors une loi binomiale de paramètres n et p :

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{pour tout } k \in [0; n].$$

De plus, pour $k \in [0; n]$, lorsque l'événement $(X = k)$ est réalisé, il y a $n - k$ appels qui sont transmis au spécialiste et chacun a la probabilité q d'être résolu. Par la même argumentation qu'au-dessus

$$P(Y = \ell | X = k) = \binom{n-k}{\ell} q^\ell (1-q)^{n-k-\ell} \quad \text{pour tout } \ell \in [0; n-k].$$

méthode

La connaissance de la loi de X et des lois de Y connaissant la valeur de X suffisent à déterminer la loi conjointe de X et Y et l'on peut en déduire la loi de toutes les variables qui sont fonction de X et Y .

La variable $X + Y$ prend ses valeurs dans $[0; n]$ et, pour tout $m \in [0; n]$,

$$P(X + Y = m) = \sum_{k+\ell=m} P(X = k, Y = \ell) = \sum_{k=0}^m P(X = k, Y = m - k)$$

avec

$$P(X = k, Y = m - k) = P(X = k) P(Y = m - k | X = k).$$

Ainsi,

$$P(X + Y = m) = \sum_{k=0}^m \left(\binom{n}{k} p^k (1-p)^{n-k} \times \binom{n-k}{m-k} q^{m-k} (1-q)^{n-m} \right).$$

Or

$$\binom{n}{k} \binom{n-k}{m-k} = \frac{n!}{k!(m-k)!(n-m)!} = \binom{n}{m} \binom{m}{k}$$

et l'on peut donc écrire

$$\begin{aligned} P(X + Y = m) &= \binom{n}{m} \sum_{k=0}^m \binom{m}{k} p^k q^{m-k} (1-p)^{n-k} (1-q)^{n-m} \\ &= \binom{n}{m} (1-p)^{n-m} (1-q)^{n-m} \sum_{k=0}^m \binom{m}{k} p^k (1-p)^{m-k} q^{m-k}. \end{aligned}$$

Enfin, par la formule du binôme,

$$\begin{aligned} \mathrm{P}(X + Y = m) &= \binom{n}{m} (1-p)^{n-m} (1-q)^{m-n} (p + (1-p)q)^m \\ &= \binom{n}{m} r^m (1-r)^{n-m} \quad \text{avec } r = p + q - pq. \end{aligned}$$

Finalement, la variable $N = n - (X + Y)$ suit une loi binomiale¹ de paramètres n et $1-r = (1-p)(1-q)$.

13.5.2 Espérances et variances

Exercice 9 *

Soit X une variable aléatoire prenant ses valeurs dans $\llbracket 0 ; n \rrbracket$. Établir

$$\mathrm{E}(X) = \sum_{k=1}^n \mathrm{P}(X \geq k).$$

Solution

méthode

On exprime $\mathrm{P}(X \geq k)$ comme une somme de probabilités $\mathrm{P}(X = \ell)$ puis on échange les deux sommes.

L'événement $(X \geq k)$ est la réunion des événements incompatibles $(X = \ell)$ pour l'entier ℓ allant de k à n . On a donc par additivité

$$\sum_{k=1}^n \mathrm{P}(X \geq k) = \sum_{k=1}^n \left(\sum_{j=k}^n \mathrm{P}(X = j) \right).$$

La double somme correspond à une somme triangulaire portant sur les couples (j, k) de $\llbracket 1 ; n \rrbracket^2$ vérifiant $j \leq k$. En échangeant les deux sommes, on obtient

$$\sum_{k=1}^n \mathrm{P}(X \geq k) = \sum_{j=1}^n \left(\sum_{k=1}^j \underbrace{\mathrm{P}(X = j)}_{\text{constante}} \right) = \sum_{j=1}^n j \mathrm{P}(X = j).$$

Enfin, en adjoignant un terme nul d'indice $j = 0$, on reconnaît l'espérance de X

$$\sum_{k=1}^n \mathrm{P}(X \geq k) = \sum_{j=0}^n j \mathrm{P}(X = j) = \mathrm{E}(X).$$

1. On peut trouver ce résultat plus rapidement en modifiant le protocole en un autre *équivalent*. Standardiste et spécialiste réceptionnent tous les appels et la probabilité qu'un appel ne trouve pas de solutions est alors $(1-p)(1-q)$. Le nombre total N d'appels sans solutions suit alors une loi de Bernoulli de paramètres n et $(1-p)(1-q)$.

Exercice 10 *

Soit X_1, \dots, X_n des variables aléatoires réelles mutuellement indépendantes suivant une même loi d'espérance μ et de variance σ^2 .

(a) Calculer l'espérance de la variable

$$M_n = \frac{1}{n} \sum_{i=1}^n X_i.$$

On dit que M_n est un *estimateur* de l'espérance μ .

(b) Calculer la variance de M_n .

(c) Pour quelle valeur du réel λ , la variable aléatoire

$$V_n = \lambda \sum_{i=1}^n (X_i - M_n)^2$$

peut-elle être considérée comme un estimateur de la variance σ^2 ?

Solution

(a) Par linéarité de l'espérance, on a immédiatement

$$\mathbb{E}(M_n) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i) = \mu.$$

(b) méthode

|| Lorsque les variables sont deux à deux indépendantes, la variance d'une somme est la somme des variances (Th. 16 p. 484).

Les variables aléatoires X_i étant deux à deux indépendantes

$$\mathbb{V}(M_n) = \frac{1}{n^2} \mathbb{V}\left(\sum_{i=1}^n X_i\right) = \frac{1}{n^2} \sum_{i=1}^n \mathbb{V}(X_i) = \frac{1}{n} \sigma^2.$$

(c) En développant le carré

$$V_n = \lambda \left(\sum_{i=1}^n X_i^2 - 2M_n \underbrace{\sum_{i=1}^n X_i}_{=nM_n} + nM_n^2 \right) = \lambda \sum_{i=1}^n X_i^2 - n\lambda M_n^2.$$

Par linéarité de l'espérance

$$\mathbb{E}(V_n) = \lambda \sum_{i=1}^n \mathbb{E}(X_i^2) - n\lambda \mathbb{E}(M_n^2).$$

Par la formule de Huygens (Th. 14 p. 484)

$$\mathrm{E}(X_i^2) = \mathrm{V}(X_i) + \mathrm{E}(X_i)^2 \quad \text{et} \quad \mathrm{E}(M_n^2) = \mathrm{V}(M_n) + \mathrm{E}(M_n)^2.$$

On obtient donc

$$\mathrm{E}(V_n) = \lambda \sum_{i=1}^n (\sigma^2 + \mu^2) - n\lambda \left(\frac{\sigma^2}{n} + \mu^2 \right).$$

Ainsi,

$$\mathrm{E}(V_n) = \sigma^2 \iff \lambda = \frac{1}{n-1}.$$

Exercice 11 **

Deux variables aléatoires réelles indépendantes X et Y prennent leurs valeurs dans un ensemble E et suivent une même loi.

(a) Soit f et g deux fonctions réelles croissantes définies sur E . En étudiant l'espérance de $Z = (f(Y) - f(X))(g(Y) - g(X))$, établir

$$\mathrm{E}(f(X)g(X)) \geq \mathrm{E}(f(X)) \mathrm{E}(g(X)).$$

(b) Application : On suppose que X prend ses valeurs dans $E \subset \mathbb{R}_+$. Montrer

$$\mathrm{E}(X) \mathrm{E}\left(\frac{1}{X}\right) \geq 1.$$

Solution

(a) Les deux fonctions réelles f et g étant croissantes, les facteurs $f(Y) - f(X)$ et $g(Y) - g(X)$ ont tous deux le signe de $Y - X$. La variable aléatoire Z est donc à valeurs dans \mathbb{R}_+ et son espérance est positive. On peut donc écrire

$$\mathrm{E}\left((f(Y) - f(X))(g(Y) - g(X))\right) \geq 0. \quad (*)$$

méthode

|| L'espérance d'un produit de deux variables indépendantes est le produit des espérances (Th. 12 p. 483).

Par linéarité, on peut développer le calcul de l'espérance de Z

$$\mathrm{E}(Z) = \mathrm{E}(f(Y)g(Y)) - \mathrm{E}(f(X)g(Y)) - \mathrm{E}(f(Y)g(X)) + \mathrm{E}(f(X)g(X)).$$

Les variables X et Y étant indépendantes, les variables $f(X)$ et $g(Y)$ d'une part, et $f(Y)$ et $g(X)$ d'autre part, le sont aussi (Th. 6 p. 481). On a donc

$$\mathrm{E}(f(X)g(Y)) = \mathrm{E}(f(X)) \mathrm{E}(g(Y)) \quad \text{et} \quad \mathrm{E}(f(Y)g(X)) = \mathrm{E}(f(Y)) \mathrm{E}(g(X)).$$

Enfin, les variables $f(X)$ et $f(Y)$ suivent la même loi, il en de même pour $g(X)$ et $g(Y)$ et aussi pour $f(X)g(X)$ et $f(Y)g(Y)$. On obtient donc

$$\mathrm{E}(Z) = 2\mathrm{E}(f(X)g(X)) - 2\mathrm{E}(f(X))\mathrm{E}(g(X)).$$

L'inégalité (*) donne alors directement celle voulue¹.

(b) Il suffit d'appliquer l'inégalité qui précède aux fonctions définies et croissantes sur E

$$f: x \mapsto x \quad \text{et} \quad g: x \mapsto -\frac{1}{x}.$$

13.5.3 Calcul d'espérances et de variances

Exercice 12 *

Une population de N individus est infectée par un virus dans une faible proportion p . Des analyses sanguines permettent de détecter la présence du virus dans n'importe quel échantillon. Afin de réduire le nombre d'analyses, on se propose de déterminer les individus malades en réunissant ceux-ci par groupes de n et, pour simplifier l'étude, on suppose que n divise N . On rassemble une partie des échantillons sanguins des individus de chaque groupe et l'on teste l'échantillon obtenu. Si le résultat du groupe est positif, on analyse individuellement les échantillons des individus du groupe.

- (a) Déterminer la loi de la variable aléatoire X donnant le nombre de groupes positifs.
- (b) On note Y la variable aléatoire donnant le nombre d'analyses effectuées. Calculer l'espérance et la variance de Y en fonction de n , N et p .
- (c) Que vaut cette espérance pour les valeurs $N = 1\,000$, $n = 10$ et $p = 0,01$?

Solution

(a) Dans chaque groupe, le nombre Z d'individus infectés suit une loi binomiale de paramètres n et p . L'échantillon du groupe est positif dès que $Z \geq 1$. La probabilité qu'un échantillon regroupé soit positif est donc²

$$q = P(Z \geq 1) = 1 - P(Z = 0) = 1 - (1 - p)^n.$$

Il y a en tout $m = N/n$ groupes et chacun a, indépendamment des autres, la probabilité q d'être infecté, le nombre X de groupes infectés suit donc une loi binomiale de paramètres m et q .

(b) méthode

|| Y est une fonction de X et son espérance s'en déduit.

On analyse chaque échantillon des m groupes et, pour chaque groupe infecté, on analyse les n échantillons des individus du groupe. Le nombre total d'analyses est donc donné par

$$Y = m + nX.$$

1. Cette inégalité entre en résonance et généralise celle du sujet 27 du chapitre 1 de l'ouvrage *Exercices d'analyse MPSI*.

2. C'est la probabilité de l'événement contraire de « Tous les individus du groupe sont sains ».

Par linéarité, l'espérance de Y se déduit de celle de X :

$$\mathbb{E}(Y) = m + n\mathbb{E}(X) = m + nmq = m + N(1 - (1 - p)^n).$$

(c) Pour les valeurs proposées, on procède en moyenne à 196 analyses.

Exercice 13 **

Une urne contient b boules blanches et r boules rouges. On tire simultanément n boules dans cette urne et l'on note X le nombre de boules blanches obtenues.

Déterminer l'espérance et la variance de la variable X .

Solution

Le tirage simultané des n boules s'apparente à un tirage sans remise : la loi de X a déjà été calculée dans le sujet 1 p. 485. La variable X prend ses valeurs dans¹ $\llbracket 0 ; n \rrbracket$ et, pour tout $k \in \llbracket 0 ; n \rrbracket$,

$$\mathbb{P}(X = k) = \frac{\binom{b}{k} \binom{r}{\ell}}{\binom{N}{n}} \quad \text{avec} \quad N = b + r \text{ et } n = k + \ell.$$

L'espérance de X est

$$\mathbb{E}(X) = \sum_{k=0}^n k \mathbb{P}(X = k) = \binom{N}{n}^{-1} \sum_{k=0}^n k \binom{b}{k} \binom{r}{n-k}.$$

On peut retirer le premier terme de la somme car celui-ci est nul.

méthode

|| La somme s'interprète comme le coefficient d'une puissance de X dans le produit de deux polynômes.

Soit k un entier compris entre 1 et n . D'une part, $\binom{b}{k}$ est le coefficient de X^{k-1} dans le polynôme dérivé de $(1 + X)^b$. D'autre part, $\binom{r}{n-k}$ est le coefficient de X^{n-k} dans le polynôme $(1 + X)^r$. La somme étudiée peut donc se comprendre comme le calcul du coefficient de X^{n-1} dans le produit de ces deux polynômes :

$$((1 + X)^b)' \times (1 + X)^r = b(1 + X)^{N-1}.$$

On a donc

$$\mathbb{E}(X) = \binom{N}{n}^{-1} \times b \binom{N-1}{n-1} = \frac{nb}{N}.$$

On calcule la variance par la formule de Huygens (Th. 14 p. 484).

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \mathbb{E}(X(X-1)) + \mathbb{E}(X) - \mathbb{E}(X)^2.$$

1. Plus précisément, X prend ses valeurs dans $\llbracket m ; M \rrbracket$ avec $m = \max(0, n - r)$ et $M = \min(n, b)$. Pour les valeurs de k en dehors que $\llbracket m ; M \rrbracket$, la formule donnant $\mathbb{P}(X = k)$ est valable par nullité de l'un des coefficients binomiaux du numérateur.

Par la formule de transfert

$$\mathbb{E}(X(X-1)) = \binom{N}{n}^{-1} \sum_{k=0}^n k(k-1) \binom{b}{k} \binom{r}{n-k}.$$

Comme au-dessus, on peut simplifier les deux premiers termes de la somme car ils sont nuls puis comprendre celle-ci comme menant le calcul du coefficient de X^{n-2} dans

$$((1+X)^b)'' \times (1+X)^r = b(b-1)(1+X)^{N-2}.$$

On obtient alors

$$\mathbb{E}(X(X-1)) = \binom{N}{n}^{-1} \times b(b-1) \binom{N-2}{n-2} = \frac{n(n-1)b(b-1)}{N(N-1)}.$$

Au terme des calculs, on conclut

$$\mathbb{V}(X) = \frac{nbn(N-n)}{N^2(N-1)}.$$

Exercice 14 ***

Dans une urne contenant $n \in \mathbb{N}^*$ boules blanches et n boules rouges, on prélève successivement et sans remise des boules jusqu'à l'obtention de toutes les boules blanches. On note X le nombre total de boules alors sorties de l'urne.

- (a) Proposer un espace probabilisé (Ω, \mathcal{P}) permettant d'étudier l'expérience.
- (b) Déterminer la loi de X .
- (c) Calculer son espérance et sa variance.

Solution

(a) On pourrait proposer pour univers l'ensemble des suites finies codant la nature des boules successives tirées par les lettres 'B' et 'R'. Une telle suite devrait alors comporter n fois la lettre 'B' et $k \in [0; n]$ fois la lettre 'R' en se terminant par la lettre 'B'. Il est cependant délicat de définir une probabilité sur cet univers. Afin de proposer un univers plus simple, nous supposons que les tirages se poursuivent au delà de l'épuisement des boules blanches. On considère alors pour univers Ω l'ensemble des suites finies de longueur $2n$ comportant n fois la lettre 'B' et n fois la lettre 'R'. Cet ensemble possède $\binom{2n}{n}$ éléments et on le munit de la probabilité uniforme car les tirages sont équiprobables.

(b) La variable aléatoire X prend ses valeurs dans $[n; 2n]$. Soit $p \in [n; 2n]$.

méthode

|| On évalue $\mathbb{P}(X \leq p)$ afin d'en déduire $\mathbb{P}(X = p)$.

Pour $p \in [n; 2n]$, on a $(X \leq p)$ lorsque toutes les boules blanches figurent dans les p premiers éléments du tirage. Pour former un tel tirage, il suffit de déterminer les

places des n boules blanches parmi les p premières positions, les autres places étant alors occupées par des boules rouges. On a donc

$$\mathrm{P}(X \leq p) = \frac{1}{\mathrm{Card}(\Omega)} \binom{p}{n} = \binom{2n}{n}^{-1} \binom{p}{n}.$$

Pour $p \in [n+1; 2n]$, on en déduit¹ par la formule du triangle de Pascal

$$\begin{aligned} \mathrm{P}(X = p) &= \mathrm{P}(X \leq p) - \mathrm{P}(X \leq p-1) \\ &= \binom{2n}{n}^{-1} \left(\binom{p}{n} - \binom{p-1}{n} \right) = \binom{2n}{n}^{-1} \binom{p-1}{n-1} \end{aligned}$$

et cette formule est aussi valable si $p = n$.

(c) L'espérance de X est donnée par la formule

$$\mathrm{E}(X) = \sum_{p=n}^{2n} p \mathrm{P}(X = p) = \binom{2n}{n}^{-1} \sum_{p=n}^{2n} p \binom{p-1}{n-1}.$$

Or, pour tout $p \in [n; 2n]$,

$$p \binom{p-1}{n-1} = \frac{p!}{(p-n)!(n-1)!} = n \binom{p}{n}$$

et donc

$$\mathrm{E}(X) = n \binom{2n}{n}^{-1} \sum_{p=n}^{2n} \binom{p}{n}.$$

On achève le calcul à l'aide de la formule du triangle de Pascal²

$$\sum_{p=0}^n \binom{n+p}{n} = \binom{n}{n} + \binom{n+1}{n} + \cdots + \binom{2n}{n} = \binom{2n+1}{n+1}$$

et l'on obtient

$$\mathrm{E}(X) = \frac{n(2n+1)}{n+1}$$

méthode

|| La variance de X se déduit de l'espérance de $X(X+1)$.

Par la formule de transfert

$$\mathrm{E}(X(X+1)) = \sum_{p=n}^{2n} p(p+1) \mathrm{P}(X = p) = \binom{2n}{n}^{-1} \sum_{p=n}^{2n} p(p+1) \binom{p-1}{n-1}$$

1. La variable X donne directement accès au nombre de boules rouges sorties de l'urne ou au nombre de boules rouges restant dans l'urne à la fin du tirage. Par symétrie des tirages, on peut aussi en déduire le nombre de boules rouges tirées avant d'obtenir une première boule blanche.

2. Voir sujet 20 p. 75.

avec, pour tout $p \in \llbracket n ; 2n \rrbracket$,

$$p(p+1) \binom{p-1}{n-1} = (p+1) \times n \binom{p}{n} = n(n+1) \binom{p+1}{n+1}$$

et donc

$$\mathbb{E}(X(X+1)) = n(n+1) \binom{2n}{n}^{-1} \sum_{p=n}^{2n} \binom{p+1}{n+1} = n(n+1) \binom{2n}{n}^{-1} \binom{2n+2}{n+2}.$$

On obtient alors

$$\mathbb{E}(X(X+1)) = \frac{2n(n+1)(2n+1)}{(n+2)}$$

puis

$$\text{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \mathbb{E}(X(X+1)) - \mathbb{E}(X) - \mathbb{E}(X)^2 = \frac{n^2(2n+1)}{(n+1)^2(n+2)}.$$

Exercice 15 ***

Un groupe de $n \geq 2$ individus échange des cadeaux. Chaque individu apporte un cadeau qu'il dispose dans une urne. Chacun leur tour, les individus retirent de l'urne un cadeau avec le risque de piocher leur propre cadeau. On note X la variable aléatoire donnant le nombre d'individus ayant pioché les cadeaux qu'ils ont apportés.

- (a) Proposer un espace probabilisé (Ω, \mathcal{P}) permettant d'étudier l'expérience.
- (b) Calculer l'espérance et la variance de X .

Solution

(a) On attribue aux individus un numéro allant de 1 à n et l'on numérote de la même façon les cadeaux que ceux-ci ont apportés. Une répartition par tirage des cadeaux s'apparente à une fonction de $\llbracket 1 ; n \rrbracket$ vers $\llbracket 1 ; n \rrbracket$ qui associe à chaque individu le cadeau que celui-ci a tiré. Dans le protocole, cette association est bijective et l'on peut choisir pour univers Ω l'ensemble \mathcal{S}_n à $n!$ éléments constitué des permutations de $\llbracket 1 ; n \rrbracket$. Les différentes répartitions étant équiprobables, l'univers Ω est muni de la probabilité uniforme.

- (b) Par définition de l'espérance

$$\mathbb{E}(X) = \sum_{k=0}^n k \mathbb{P}(X = k).$$

Ce calcul est lié au dénombrement des individus qui ont reçu leur propre cadeau.

méthode

On approfondit l'expérience en poursuivant celle-ci par le choix arbitraire et uniforme d'un individu. On étudie ensuite si cet individu a reçu son propre cadeau.

On « agrandit » l'espace probabilisé en considérant l'univers $\Omega' = \Omega \times \llbracket 1 ; n \rrbracket$ muni de la probabilité uniforme. Dans l'univers Ω' , un couple (σ, i) traduit une répartition σ des cadeaux et le choix d'un individu i . On étudie alors l'événement

$$A = \text{« L'individu choisi a reçu son propre cadeau ».}$$

Pour chaque individu, la variable aléatoire déterminant le numéro du cadeau qu'il reçoit suit une loi uniforme sur $\llbracket 1 ; n \rrbracket$. On a donc un premier calcul immédiat

$$\mathbb{P}(A) = \frac{1}{n}$$

Parallèlement, on peut retrouver cette probabilité en introduisant le système complet constitué des événements $(X = k)$ pour $k \in \llbracket 0 ; n \rrbracket$

$$\mathbb{P}(A) = \sum_{k=0}^n \mathbb{P}(A | X = k) \mathbb{P}(X = k).$$

Or lorsque k individus ont reçu leur propre cadeau, la probabilité que l'individu choisi figure parmi ceux-ci est égale à k/n et donc $\mathbb{P}(A | X = k) = k/n$. On en déduit

$$\sum_{k=0}^n \frac{k}{n} \mathbb{P}(X = k) = \frac{1}{n} \quad \text{puis} \quad \mathbb{E}(X) = 1.$$

Pour calculer la variance, on utilise la formule de Huygens en commençant par calculer

$$\mathbb{E}(X(X - 1)) = \sum_{k=0}^n k(k - 1) \mathbb{P}(X = k).$$

méthode

|| Comme au-dessus, on approfondit l'expérience afin de pouvoir calculer cette somme.

On poursuit l'expérience initiale en choisissant un premier individu puis un second distinct du premier et en étudiant l'événement

$$B = \text{« Les deux individus choisis ont reçu leur propre cadeau ».}$$

D'une part,

$$\mathbb{P}(B) = \frac{1}{n(n-1)}$$

En effet, pour chaque couple d'individus distincts, la variable aléatoire donnant le couple formé par les cadeaux qui leurs sont attribués suit une loi uniforme sur l'ensemble à $n(n - 1)$ éléments des couples d'entiers distincts choisis dans $\llbracket 1 ; n \rrbracket$

D'autre part,

$$\mathbb{P}(B) = \sum_{k=0}^n \mathbb{P}(B | X = k) \mathbb{P}(X = k).$$

Or, il y a une probabilité égale à $\frac{k}{n}$ que le premier individu choisi s'est vu attribuer son cadeau puis une probabilité égale à $\frac{n-k}{n-1}$ pour le second individu sachant que le premier a reçu le sien. On a donc

$$P(B|X=k) = \frac{k}{n} \cdot \frac{n-k}{n-1}.$$

On en déduit

$$\sum_{k=0}^n \frac{k}{n} \cdot \frac{n-k}{n-1} P(X=k) = \frac{1}{n(n-1)} \quad \text{puis} \quad E(X(X-1)) = 1,$$

Finalement,

$$V(X) = E(X(X-1)) + E(X) - E(X)^2 = 1.$$

13.5.4 Couples de variables aléatoires

Exercice 16 *

Deux joueurs lancent chacun n fois une pièce équilibrée. On note X le nombre de côtés ‘faces’ obtenus par le premier joueur et Y celui du second.

- (a) Calculer $P(X = Y)$.
- (b) En déduire $P(X \leq Y)$.

Solution

(a) À défaut de précision, les lancers sont supposés indépendants et les variables X et Y se comprennent comme donnant le nombre de succès dans la répétition de n épreuves de Bernoulli indépendantes de paramètre $1/2$. Les variables X et Y suivent donc chacune des lois binomiales de paramètres n et $1/2$ indépendantes :

$$P(X=k) = P(Y=k) = \binom{n}{k} \frac{1}{2^n} \quad \text{pour tout } k \in [0; n].$$

L’événement $(X = Y)$ est la réunion des événements incompatibles $(X = k, Y = k)$ pour k allant de 0 à n . On a donc

$$P(X = Y) = \sum_{k=0}^n P(X = k, Y = k).$$

Par indépendance

$$P(X = Y) = \sum_{k=0}^n P(X = k) P(Y = k) = \sum_{k=0}^n \binom{n}{k}^2 \frac{1}{4^n} = \frac{1}{4^n} \sum_{k=0}^n \binom{n}{k}^2.$$

Cette dernière somme a déjà été calculée dans le sujet 23 p. 77 et l’on obtient

$$P(X = Y) = \frac{1}{4^n} \binom{2n}{n}.$$

(b) **méthode**

|| Les événements $(X = Y)$, $(X < Y)$ et $(X > Y)$ constituent un système complet.

Par symétrie de l'expérience, on a $P(X < Y) = P(X > Y)$ et donc

$$\begin{aligned} 1 &= P(X = Y) + P(X < Y) + P(X > Y) \\ &= P(X = Y) + 2P(X < Y) \\ &= 2P(X \leq Y) - P(X = Y). \end{aligned}$$

On peut alors conclure

$$P(X \leq Y) = \frac{1}{2} \left(1 + \frac{1}{4^n} \binom{2n}{n} \right).$$

Exercice 17 **

Soit X une variable aléatoire à valeurs dans $\llbracket 1 ; n \rrbracket$ et telle que chaque élément de $\llbracket 1 ; n \rrbracket$ est une valeur pris par X avec une probabilité non nulle. Soit aussi Y une variable aléatoire telle que, pour tout $k \in \llbracket 1 ; n \rrbracket$, la loi de Y sachant $(X = k)$ est uniforme sur $\llbracket 1 ; k \rrbracket$

- (a) Exprimer la loi de Y en fonction de celle de X .
- (b) En déduire l'espérance de Y en fonction de celle de X .
- (c) Retrouver ce résultat en considérant la variable $Z = X + 1 - Y$.

Solution

(a) Soit $k \in \llbracket 1 ; n \rrbracket$. Lorsque l'événement $(X = k)$ a lieu, les hypothèses assurent que Y prend ses valeurs dans $\llbracket 1 ; k \rrbracket$ et

$$P(Y = j | X = k) = \frac{1}{k} \quad \text{pour tout } j \in \llbracket 1 ; k \rrbracket.$$

méthode

|| La connaissance des probabilités de Y sachant X permet de d'exprimer la loi de Y en fonction de celle de X (Th. 4 p. 481).

La variable Y prend globalement ses valeurs dans $\llbracket 1 ; n \rrbracket$. Introduisons $j \in \llbracket 1 ; n \rrbracket$ et calculons $P(Y = j)$. Les événements $(X = k)$ pour k allant de 1 à n constituent un système complet d'événements de probabilités non nulles et la formule des probabilités totales donne

$$P(Y = j) = \sum_{k=1}^n P(Y = j | X = k) P(X = k).$$

Lorsque $k < j$, la probabilité de $(Y = j)$ sachant $(X = k)$ est nulle et l'on peut simplifier les termes correspondants dans la somme puis conclure

$$P(Y = j) = \sum_{k=j}^n P(Y = j | X = k) P(X = k) = \sum_{k=j}^n \frac{1}{k} P(X = k).$$

(b) L'espérance de Y s'exprime

$$\mathbb{E}(Y) = \sum_{j=1}^n j \mathbb{P}(Y=j) = \sum_{j=1}^n \left(\sum_{k=j}^n \frac{j}{k} \mathbb{P}(X=k) \right).$$

méthode

|| Pour progresser dans le calcul, on échange les deux sommes.

La double somme correspond à une somme triangulaire portant sur les couples (j, k) de $\llbracket 1 ; n \rrbracket^2$ vérifiant $j \leq k$. En échangeant les deux sommes, on obtient

$$\mathbb{E}(Y) = \sum_{k=1}^n \left(\sum_{j=1}^k \frac{j}{k} \mathbb{P}(X=k) \right) = \sum_{k=1}^n \frac{k+1}{2} \mathbb{P}(X=k) \text{ car } \sum_{j=1}^k j = \frac{k(k+1)}{2}.$$

Par la formule de transfert,

$$\mathbb{E}(Y) = \mathbb{E}\left(\frac{X+1}{2}\right) = \frac{1}{2} \mathbb{E}(X) + \frac{1}{2}.$$

(c) Soit $k \in \llbracket 1 ; n \rrbracket$. La loi conditionnelle de Y sachant $(X=k)$ est uniforme sur $\llbracket 1 ; k \rrbracket$. La loi conditionnelle de $Z = X+1-Y$ sachant $(X=k)$ se confond avec celle de $k+1-Y$ qui est aussi uniforme sur $\llbracket 1 ; k \rrbracket$. Les deux variables Y et Z satisfont les mêmes conditions qui déterminent leurs lois, elles suivent donc la même loi¹ et ont par conséquent la même espérance. Or $\mathbb{E}(Y) + \mathbb{E}(Z) = \mathbb{E}(X+1) = \mathbb{E}(X) + 1$ et l'on retrouve

$$\mathbb{E}(Y) = \frac{1}{2} \mathbb{E}(X) + \frac{1}{2}$$

13.5.5 Covariance

Exercice 18 *

Soit U et V deux variables de Bernoulli indépendantes de paramètres p et $q \in]0 ; 1[$. On pose $X = U + V$ et $Y = U - V$.

- (a) Calculer la covariance de X et Y .
- (b) Les variables X et Y sont-elles indépendantes ?

Solution

(a) méthode

|| On peut calculer la covariance de deux variables par la formule de Huygens (Th. 15 p. 484).

1. Sans pour autant être égales...

On a $\text{Cov}(X, Y) = \text{E}(XY) - \text{E}(X)\text{E}(Y)$ avec, par linéarité,

$$\text{E}(X)\text{E}(Y) = (\text{E}(U) + \text{E}(V))(\text{E}(U) - \text{E}(V)) = (p+q)(p-q)$$

et

$$\text{E}(XY) = \text{E}(U^2 - V^2) = \text{E}(U^2) - \text{E}(V^2) = \text{E}(U) - \text{E}(V) = p - q$$

car $U^2 = U$ et $V^2 = V$ puisque U et V prennent leurs valeurs dans $\{0, 1\}$.

On obtient finalement $\text{Cov}(X, Y) = (p-q)(1-(p+q))$.

(b) Lorsque $p \neq q$ et $p+q \neq 1$, la covariance de X et Y est non nulle et on peut directement affirmer que ces variables ne sont pas indépendantes.

méthode

|| La nullité de la covariance¹ de deux variables U et V n'est qu'une condition nécessaire à leur indépendance.

Lorsque $p = q$ ou $p+q = 1$, la covariance de X et Y est nulle et l'on ne peut pas conclure immédiatement.

méthode

|| On vérifie l'indépendance de deux variables aléatoires X et Y en comparant $\text{P}(X = x, Y = y)$ et $\text{P}(X = x)\text{P}(Y = y)$ pour tout couple (x, y) formé de valeurs prises par X et Y (Th. 5 p. 481).

On remarque que l'événement $(X = 2, Y = 1)$ est impossible alors que, par indépendance,

$$\text{P}(X = 2) = \text{P}(U = 1, V = 1) = pq \neq 0$$

et

$$\text{P}(Y = 1) = \text{P}(U = 1, V = 0) + \text{P}(U = 0, V = 1) = p(1-q) + q(1-p) \neq 0.$$

La propriété $\text{P}(X = 2, Y = 1) = \text{P}(X = 2)\text{P}(Y = 1)$ n'est donc pas satisfaite, les variables X et Y ne sont pas indépendantes².

Exercice 19 **

Soit X_1, \dots, X_n, X_{n+1} des variables aléatoires mutuellement indépendantes suivant chacune des lois de Bernoulli de même paramètre $p \in [0; 1]$. Pour tout $k \in [1; n]$ on pose $Y_k = X_k X_{k+1}$. Calculer la variance de $S_n = Y_1 + \dots + Y_n$.

1. Il en est de même pour la vérification de l'égalité $\text{E}(XY) = \text{E}(X)\text{E}(Y)$.

2. Cette situation illustre qu'une covariance peut être nulle sans que les variables soient indépendantes.

Solution

Chaque variable Y_k pour k allant de 1 à n suit une loi de Bernoulli de paramètre p^2 car elle prend ses valeurs dans $\{0, 1\}$ et car on a par indépendance de X_k et X_{k+1}

$$\mathrm{P}(Y_k = 1) = \mathrm{P}(X_k = 1, X_{k+1} = 1) = \mathrm{P}(X_k = 1)\mathrm{P}(X_{k+1} = 1) = p^2.$$

Cependant, les variables Y_1, \dots, Y_n ne sont pas indépendantes.

méthode

|| La variance de la somme $Y_1 + \dots + Y_n$ se déduit du calcul des covariances des variables Y_1, \dots, Y_n (Th. 16 p. 484).

Soit i et j deux indices distincts dans $\llbracket 1; n \rrbracket$. Calculons la covariance de Y_i et Y_j :

$$\begin{aligned}\mathrm{Cov}(Y_i, Y_j) &= \mathrm{E}(Y_i Y_j) - \mathrm{E}(Y_i) \mathrm{E}(Y_j) \\ &= \mathrm{E}(X_i X_{i+1} X_j X_{j+1}) - \mathrm{E}(X_i) \mathrm{E}(X_{i+1}) \mathrm{E}(X_j) \mathrm{E}(X_{j+1}).\end{aligned}$$

Quitte à échanger, on peut supposer $i < j$ car on sait $\mathrm{Cov}(Y_i, Y_j) = \mathrm{Cov}(Y_j, Y_i)$. On est alors amené à distinguer deux cas.

Cas : $i+1 < j$. Les variables X_i, X_{i+1}, X_j et X_{j+1} sont deux à deux distinctes et l'on a par indépendance

$$\mathrm{E}(X_i X_{i+1} X_j X_{j+1}) = \mathrm{E}(X_i) \mathrm{E}(X_{i+1}) \mathrm{E}(X_j) \mathrm{E}(X_{j+1}).$$

La covariance de Y_i et Y_j est alors nulle.

Cas : $i+1 = j$. Les variables X_{i+1} et X_j sont identiques donc $X_{i+1} X_j = X_{i+1}^2 = X_{i+1}$ puis

$$\mathrm{E}(X_i X_{i+1} X_j X_{j+1}) = \mathrm{E}(X_i X_{i+1} X_{i+2}) = \mathrm{E}(X_i) \mathrm{E}(X_{i+1}) \mathrm{E}(X_{i+2}).$$

La covariance de Y_i et Y_{i+1} vaut alors

$$\mathrm{Cov}(Y_i, Y_{i+1}) = p^3 - p^4.$$

Notons qu'il y a exactement $(n-1)$ couples (i, j) avec $j = i+1$.

Il reste alors à développer la variance de la somme par bilinéarité de la covariance :

$$\begin{aligned}\mathrm{V}(S_n) &= \mathrm{Cov}(S_n, S_n) = \sum_{i=1}^n \left(\sum_{j=1}^n \mathrm{Cov}(Y_i, Y_j) \right) = \sum_{i=1}^n \mathrm{V}(Y_i) + 2 \sum_{i=1}^{n-1} \left(\sum_{j=i+1}^n \mathrm{Cov}(Y_i, Y_j) \right) \\ &= np^2(1-p^2) + 2(n-1)(p^3 - p^4) \\ &= p^2(1-p)(n(3p+1) - 2p).\end{aligned}$$

Exercice 20 **

Soit X et Y deux variables aléatoires réelles avec $\mathrm{V}(X) > 0$. Déterminer $(a, b) \in \mathbb{R}^2$ minimisant la quantité

$$\mathrm{E}\left(\left(Y - (aX + b)\right)^2\right).$$

Solution

Par la formule de Huygens

$$\begin{aligned} \mathrm{E}\left(\left(Y - (aX + b)\right)^2\right) &= \mathrm{V}(Y - (aX + b)) + \mathrm{E}(Y - (aX + b))^2 \\ &= \mathrm{V}(Y - aX) + (\mathrm{E}(Y) - a \mathrm{E}(X) - b)^2. \end{aligned}$$

méthode

Il s'agit de la somme de deux termes positifs. Le premier peut être rendu minimal pour une bonne valeur de a et, lorsque celle-ci est connue, le second peut être rendu nul pour un bon choix de b .

En développant le calcul de la variance

$$\mathrm{V}(Y - aX) = a^2 \mathrm{V}(X) - 2a \mathrm{Cov}(X, Y) + \mathrm{V}(Y).$$

Cette quantité est une expression du second degré en la variable a qui est minimale pour

$$a = \frac{\mathrm{Cov}(X, Y)}{\mathrm{V}(X)}.$$

Parallèlement, le terme $(\mathrm{E}(Y) - a \mathrm{E}(X) - b)^2$ est nul pour $b = \mathrm{E}(Y) - a \mathrm{E}(X)$.

Finalement, l'espérance étudiée est minimale pour

$$a = \frac{\mathrm{Cov}(X, Y)}{\mathrm{V}(X)} \quad \text{et} \quad b = \frac{\mathrm{E}(Y) \mathrm{V}(X) - \mathrm{E}(X) \mathrm{Cov}(X, Y)}{\mathrm{V}(X)}.$$

Ces valeurs de a et b réalisent une *régression linéaire* : elles donnent la meilleure approximation linéaire de Y en fonction de la variable X .

13.5.6 Inégalités de concentration

Exercice 21 *

Soit X une variable aléatoire réelle et $g: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ une fonction croissante. Montrer que, pour tout réel a positif,

$$g(a) \mathrm{P}(|X| \geq a) \leq \mathrm{E}\left(g(|X|)\right).$$

Solution**méthode**

On peut appliquer l'inégalité de Markov (Th. 17 p. 485) à la variable composée $Y = g(|X|)$ ou reproduire la démonstration de celle-ci.

On privilégie ici la deuxième démarche¹. Soit $a \in \mathbb{R}_+$. On a

$$\mathrm{P}(|X| \geq a) = \mathrm{E}(\mathbf{1}_A) \quad \text{avec} \quad A = (|X| \geq a).$$

1. La fonction g étant seulement supposée croissante et non strictement croissante, on sera attentif à écrire l'inclusion $(|X| \geq a) \subset (g(|X|) \geq g(a))$ et non l'égalité lors de l'application de l'inégalité de Markov à $Y = g(|X|)$.

Or

$$g(|X|) \geq g(a)\mathbf{1}_A.$$

En effet, si $|X| < a$, le premier membre est positif alors que le second est nul et, si $|X| \geq a$, le premier membre est supérieur au second par croissance de g .

Par croissance et linéarité de l'espérance, on obtient alors

$$\mathbb{E}(g(|X|)) \geq \mathbb{E}(g(a)\mathbf{1}_A) = g(a)\mathbb{P}(A).$$

Exercice 22 *

Une population présente une propriété dans une proportion inconnue $p \in]0 ; 1[$ que l'on souhaite estimer. On choisit un échantillon de n personnes et l'on pose $X_i = 1$ si le i -ème individu présente la propriété étudiée, 0 sinon. On considère que les variables aléatoires X_i ainsi définies sont indépendantes et suivent chacune une loi de Bernoulli de paramètre p . Enfin, on pose $S_n = X_1 + \dots + X_n$.

(a) Soit $\varepsilon > 0$. Établir

$$\mathbb{P}\left(\left|\frac{S_n}{n} - p\right| > \varepsilon\right) \leq \frac{1}{4n\varepsilon^2}.$$

(b) Pour $\varepsilon = 0,05$, quelle valeur de n peut-on choisir pour que S_n/n constitue une valeur approchée de p à ε près avec une probabilité supérieure à 95 % ?

Solution

(a) **méthode**

|| On applique l'inégalité de Bienaymé-Tchebychev (Th. 18 p. 485).

Par somme de variables de Bernoulli indépendantes, la variable S_n suit une loi binomiale de paramètres n et p . On sait donc

$$\mathbb{E}(S_n) = np \quad \text{et} \quad \mathbb{V}(S_n) = np(1-p).$$

En appliquant l'inégalité de Bienaymé-Tchebychev à la variable S_n

$$\mathbb{P}\left(\left|\frac{S_n}{n} - p\right| > \varepsilon\right) = \mathbb{P}(|S_n - np| > n\varepsilon) \leq \frac{p(1-p)}{n\varepsilon^2}.$$

Enfin, l'inégalité classique¹ $p(1-p) \leq 1/4$ permet de conclure.

(b) Il suffit de choisir n de sorte que $1/4n\varepsilon^2 \leq 0,05$. La valeur $n = 2\,000$ convient.

1. Voir sujet 4 du chapitre 1 de l'ouvrage *Exercices d'analyse MPSI*.

Exercice 23 ***

Soit X_1, \dots, X_n des variables aléatoires indépendantes de loi uniforme sur $\{-1, 1\}$. On pose $S_n = X_1 + \dots + X_n$.

(a) Montrer que $\ln(\cosh \lambda) \leq e^{\lambda^2/2}$ pour tout réel λ .

(b) Établir que, pour tout $\alpha > 0$,

$$\Pr\left(\frac{S_n}{n} > \alpha\right) \leq e^{-n\alpha^2/2}.$$

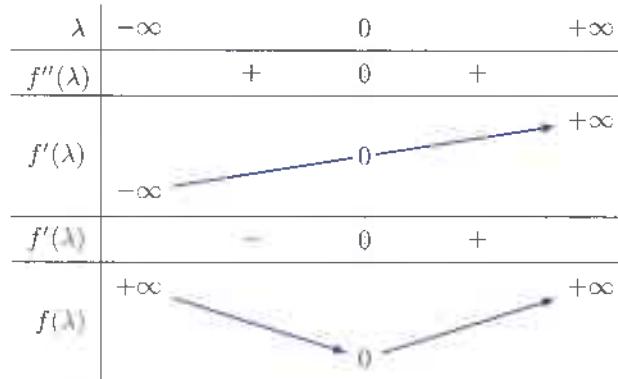
Solution(a) **méthode**

|| On étudie l'inégalité équivalente obtenue par passage au logarithme.

Considérons la fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ définie par $f(\lambda) = \frac{1}{2}\lambda^2 - \ln(\cosh \lambda)$. La fonction f est indéfiniment dérivable avec, pour tout $\lambda \in \mathbb{R}$,

$$f'(\lambda) = \lambda - \tanh \lambda \quad \text{et} \quad f''(\lambda) = \tanh^2 \lambda.$$

On en déduit les variations puis le signe de f :



La fonction f est donc positive et par conséquent $\ln(\cosh \lambda) \leq \frac{1}{2}\lambda^2$ pour tout λ réel. On conclut à l'inégalité voulue par croissance de l'exponentielle.

(b) **méthode**

|| On applique l'inégalité de Markov à une variable aléatoire bien choisie en fonction de S_n .

Soit α et λ des réels strictement positifs. Par stricte croissance de la fonction $t \mapsto e^{\lambda t}$, on a l'égalité d'événements

$$\left(S_n > n\alpha\right) = \left(S_n > \lambda S_n \alpha / \lambda\right) = \left(e^{\lambda S_n} > e^{\lambda n \alpha}\right).$$

Par l'inégalité de Markov appliquée à la variable $Y = e^{\lambda S_n}$, il vient

$$\mathrm{P}\left(\frac{S_n}{n} > \alpha\right) = \mathrm{P}(e^{\lambda S_n} > e^{\lambda n \alpha}) \leq e^{-\lambda n \alpha} \mathrm{E}(e^{\lambda S_n}).$$

On peut calculer l'espérance figurant en second membre par l'indépendance des variables X_1, \dots, X_n qui entraîne¹ celle des variables $e^{\lambda X_1}, \dots, e^{\lambda X_n}$.

$$\mathrm{E}(e^{\lambda S_n}) = \mathrm{E}(e^{\lambda X_1} \times \dots \times e^{\lambda X_n}) = \mathrm{E}(e^{\lambda X_1}) \times \dots \times \mathrm{E}(e^{\lambda X_n}) = \left(\frac{e^\lambda + e^{-\lambda}}{2}\right)^n.$$

En exploitant l'inégalité de la question précédente, on poursuit

$$\mathrm{P}\left(\frac{S_n}{n} > \alpha\right) \leq e^{-\lambda n \alpha} (\mathrm{ch} \lambda)^n \leq e^{-\lambda n \alpha} e^{n \lambda^2 / 2}.$$

Il reste à choisir convenablement λ . Pour la valeur² $\lambda = \alpha$, on conclut³

$$\mathrm{P}\left(\frac{S_n}{n} > \alpha\right) \leq e^{-n \alpha^2} e^{n \alpha^2 / 2} = e^{-n \alpha^2 / 2}.$$

13.6 Exercices d'approfondissement

Exercice 24 *

Soit X_1, \dots, X_n des variables aléatoires indépendantes suivant toutes une même loi de Bernoulli de paramètre $p \in [0; 1]$. On forme U la colonne de $\mathcal{M}_{n,1}(\mathbb{R})$ dont les éléments sont les valeurs respectives des variables X_1, \dots, X_n .

Donner la probabilité que $M = U^t U$ soit une matrice de projection.

Solution

La matrice U est une colonne de hauteur n , sa transposée U^t est une ligne de longueur n et le produit $M = U^t U$ est une matrice carrée de taille n . Celle-ci est une matrice de projection si, et seulement si, $M^2 = M$.

méthode

¶ Lorsque X et Y sont deux colonnes, le calcul ${}^t X Y$ détermine un scalaire⁴.

On a

$$M^2 = (U^t U)(U^t U) = U({}^t \underbrace{U U}_{\text{colonne}})^t U = \lambda M \quad \text{avec} \quad \lambda = {}^t U U.$$

1. Cette affirmation se justifie aisément par le Th. 7 p. 482.
2. Le polynôme du second degré $\lambda \mapsto \frac{n}{2} \lambda^2 - n \alpha \lambda$ est minimal en α .
3. Ce résultat est un cas particulier de l'inégalité de Chernoff que l'on retrouvera dans le sujet 25 du chapitre 9 de l'ouvrage *Exercices d'algèbre et de probabilités MP*.
4. Lorsque X et Y sont des colonnes réelles de hauteur n , ${}^t X Y$ correspond au calcul du produit scalaire canonique sur $\mathcal{M}_{n,1}(\mathbb{R})$.

La matrice M est donc une matrice de projection si, et seulement si, $\lambda = 1$. Or

$${}^t UU = \sum_{i=1}^n X_i^2 = \sum_{i=1}^n X_i$$

car X_1, \dots, X_n prennent leurs valeurs dans $\{0, 1\}$. Enfin, les variables X_1, \dots, X_n étant indépendantes et suivant une même loi de Bernoulli de paramètre p , la variable λ suit une loi binomiale de paramètres n et p . La probabilité que M soit une matrice de projection est alors

$$\mathbb{P}(\lambda = 1) = \binom{n}{1} p(1-p)^{n-1} = np(1-p)^{n-1}.$$

Exercice 25 *

Soit X_1 et X_2 deux variables aléatoires réelles. On suppose que $E(X_1^k) = E(X_2^k)$ pour tout $k \in \mathbb{N}$. Montrer que les variables X_1 et X_2 suivent les mêmes lois.

Solution

méthode

|| On observe $E(L(X_1)) = E(L(X_2))$ pour tout polynôme réel L .

Soit $L = a_0 + a_1 X + \dots + a_N X^N$ un polynôme réel. Par linéarité de l'espérance, on peut écrire

$$E(L(X_1)) = \sum_{k=0}^N a_k E(X_1^k) = \sum_{k=0}^N a_k E(X_2^k) = E(L(X_2)).$$

Pour montrer que les variables X_1 et X_2 suivent les mêmes lois, on vérifie l'égalité $P(X_1 = x) = P(X_2 = x)$ pour toutes les valeurs x prises par ces variables. Introduisons E l'ensemble des valeurs prises par les variables X_1 et X_2 . L'ensemble E est fini et l'on peut énumérer ses éléments

$$E = \{x_1, \dots, x_n\} \quad \text{avec } x_1, \dots, x_n \text{ deux à deux distincts.}$$

Soit $i \in \llbracket 1; n \rrbracket$. Montrons $P(X_1 = x_i) = P(X_2 = x_i)$ en introduisant un polynôme¹ L_i prenant la valeur 1 en x_i et la valeur 0 en les x_j pour $j \neq i$.

Par la formule de transfert,

$$E(L_i(X_1)) = \sum_{j=1}^n L_i(x_j) P(X_1 = x_j) = P(X_1 = x_i)$$

et de même $E(L_i(X_2)) = P(X_2 = x_i)$. On a donc $P(X_1 = x_i) = P(X_2 = x_i)$ pour tout i de $\llbracket 1; n \rrbracket$ et l'on peut affirmer que les variables X_1 et X_2 suivent la même loi.

1. Un tel polynôme existe, voir sujet 6 p. 285.

Exercice 26 **

On considère deux dés à 6 faces non nécessairement équilibrés, non nécessairement identiques. On note X_1 et X_2 les variables aléatoires indépendantes déterminant les valeurs de ces deux dés.

(a) Montrer que le polynôme réel $P = 1 + X + X^2 + \cdots + X^{10}$ ne peut pas se factoriser dans $\mathbb{R}[X]$ comme un produit de deux polynômes réels de degré 5.

(b) Montrer qu'il est impossible que $X_1 + X_2$ suive une loi uniforme sur $[2 ; 12]$.

Solution

(a) On observe $(X - 1)P = X^{11} - 1$. Hormis le nombre 1 qui en est racine simple, le polynôme $X^{11} - 1$ ne possède pas de racines réelles¹. Par l'absurde, si le polynôme P peut être factorisé en un produit de deux polynômes réels de degrés impairs, chaque facteur admet une racine réelle qui est alors racine de $X^{11} - 1$. C'est absurde.

(b) Introduisons les probabilités des valeurs des différentes faces des deux dés

$$p_i = P(X_1 = i) \quad \text{et} \quad q_i = P(X_2 = i) \quad \text{pour tout } i \in [1 ; 6].$$

méthode

On introduit les polynômes réels

$$P_1 = p_1 + p_2X + \cdots + p_6X^5 \quad \text{et} \quad Q_1 = q_1 + q_2X + \cdots + q_6X^5.$$

Les lois de X_1 et X_2 permettent de déterminer celle de $X_1 + X_2$. Cette variable prend ses valeurs dans $[2 ; 12]$ et

$$P(X_1 + X_2 = 2) = P(X_1 = 1, X_2 = 1) = P(X_1 = 1)P(X_2 = 1) = p_1q_1,$$

$$P(X_1 + X_2 = 3) = P(X_1 = 1)P(X_2 = 2) + P(X_1 = 2)P(X_2 = 1) = p_1q_2 + p_2q_1, \text{ etc.}$$

De façon générale, on observe que

$$P(X_1 + X_2 = k) = \sum_{i+j=k} p_i q_j \quad \text{pour tout } k \in [2 ; 12]$$

où la somme s'étend sur les couples $(i, j) \in [1 ; 6]^2$ vérifiant la condition $i + j = k$.

Ce calcul est aussi celui que l'on mène lorsque l'on étudie les coefficients du polynôme $Q = P_1 P_2$. On peut alors affirmer que la loi de $X_1 + X_2$ est donnée par les coefficients de Q .

Supposons par l'absurde que la variable $X_1 + X_2$ suit une loi uniforme sur $[2 ; 12]$. Tous les coefficients du polynôme Q sont alors égaux à $1/11$ ce qui donne l'égalité

$$(p_1 + p_2X + \cdots + p_6X^5)(q_1 + q_2X + \cdots + q_6X^5) = \frac{1}{11}(1 + X + \cdots + X^{10}).$$

Les coefficients p_6 et q_6 sont nécessairement non nuls et ce qui précède est alors absurde car contredit le résultat de la première question.

1. Pour $n \in \mathbb{N}^*$, les racines complexes de $X^n - 1$ sont les racines n -ièmes de l'unité, il y en a n parmi lesquelles les racines réelles sont 1 et -1 , cette dernière uniquement lorsque l'entier n est pair.

Exercice 27 * (Théorème de Weierstrass)**

Soit $f: [0; 1] \rightarrow \mathbb{R}$ une fonction continue. Pour $n \in \mathbb{N}^*$ et $x \in [0; 1]$, on introduit une variable aléatoire $X_n = S_n/n$ où S_n suit une loi binomiale de paramètres n et x .

(a) Vérifier que $B_n(f): x \mapsto E(f(X_n))$ est une fonction polynomiale sur $[0; 1]$.

Soit $\varepsilon > 0$.

(b) Justifier l'existence d'un réel $\alpha > 0$ et d'un réel $M \in \mathbb{R}_+$ vérifiant

$$\forall (x, y) \in [0; 1]^2, |y - x| \leq \alpha \implies |f(y) - f(x)| \leq \varepsilon \quad \text{et} \quad \forall x \in [0; 1], |f(x)| \leq M.$$

(c) On considère l'événement $A_n = (|X_n - x| \leq \alpha)$. Montrer

$$E(|f(X_n) - f(x)| \mathbf{1}_{A_n}) \leq \varepsilon \quad \text{et} \quad E(|f(X_n) - f(x)| \mathbf{1}_{A_n^c}) \leq \frac{M}{2n\alpha^2}.$$

(d) Conclure que, pour n assez grand,

$$\sup_{x \in [0; 1]} |B_n(f)(x) - f(x)| \leq 2\varepsilon.$$

Solution

(a) Soit $x \in [0; 1]$. La variable S_n prend ses valeurs dans $[0; n]$ et la variable X_n prend les siennes dans $[0; 1]$. La variable composée $f(X_n)$ est donc bien définie et par la formule de transfert

$$B_n(f)(x) = E(f(X_n)) = \sum_{k=0}^n f\left(\frac{k}{n}\right) P(S_n = k) = \sum_{k=0}^n \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k}.$$

La fonction $x \mapsto B_n(f)(x)$ est donc bien polynomiale sur $[0; 1]$.

(b) La fonction f est continue sur le segment $[0; 1]$, le théorème de Heine¹ assure que f est uniformément continue ce qui donne l'existence du réel $\alpha > 0$ tel que voulu.

La fonction f est continue sur le segment $[0; 1]$, le théorème des bornes atteintes² affirme que f est bornée ce qui permet d'introduire $M \in \mathbb{R}_+$ tel que souhaité.

(c) On a $|f(X_n) - f(x)| \mathbf{1}_{A_n} \leq \mathbf{1}_{A_n}$ car, si la valeur de la fonction indicatrice n'est pas nulle, on a $|X_n - x| \leq \alpha$ et donc $|f(X_n) - f(x)| \leq \varepsilon$. Par croissance de l'espérance

$$E(|f(X_n) - f(x)| \mathbf{1}_{A_n}) \leq \varepsilon E(\mathbf{1}_{A_n}) = \varepsilon P(A_n) \leq \varepsilon.$$

Aussi, on a

$$|f(X_n) - f(x)| \leq |f(X_n)| + |f(x)| \leq 2M$$

1. Voir Th 1 du chapitre 10 de l'ouvrage *Exercices d'analyse MPSI*.

2. Voir Th 16 du chapitre 7 de l'ouvrage *Exercices d'analyse MPSI*.

ce qui entraîne la comparaison

$$|f(X_n) - f(x)| \mathbf{1}_{\overline{A_n}} \leq 2M \mathbf{1}_{\overline{A_n}}.$$

Par croissance de l'espérance, il vient

$$\mathbb{E}(|f(X_n) - f(x)| \mathbf{1}_{\overline{A_n}}) \leq 2M \mathbb{E}(\mathbf{1}_{\overline{A_n}}) = 2M \mathbb{P}(\overline{A_n}).$$

méthode

On majore la probabilité de l'écart à la moyenne par l'inégalité de Bienaymé-Tchebychev.

L'espérance de X_n est égale à x et sa variance égale à $\frac{x(1-x)}{n}$. L'inégalité de Bienaymé-Tchebychev donne alors

$$\mathbb{P}(\overline{A_n}) = \mathbb{P}(|X_n - x| > \alpha) \leq \frac{x(1-x)}{n\alpha^2}.$$

Enfin, par l'inégalité $x(1-x) \leq 1/4$, on conclut

$$\mathbb{E}(|f(X_n) - f(x)| \mathbf{1}_{\overline{A_n}}) \leq \frac{M}{2n\alpha^2}.$$

(d) La borne supérieure introduite existe car elle porte sur une fonction continue sur un segment donc bornée. Pour obtenir l'inégalité demandée, il suffit de constater

$$|B_n(f)(x) - f(x)| \leq 2\varepsilon \quad \text{pour tout } x \in [0; 1].$$

Les événements A_n et $\overline{A_n}$ constituent un système complet d'événements ce qui permet d'écrire $1 = \mathbf{1}_{A_n} + \mathbf{1}_{\overline{A_n}}$. Ce qui précède donne alors

$$\mathbb{E}(|f(X_n) - f(x)|) = \mathbb{E}(|f(X_n) - f(x)| \mathbf{1}_{A_n}) + \mathbb{E}(|f(X_n) - f(x)| \mathbf{1}_{\overline{A_n}}) \leq \varepsilon + \frac{M}{2n\alpha^2}.$$

Pour n assez grand, on a $\frac{M}{2n\alpha^2} \leq \varepsilon$ et donc $\mathbb{E}(|f(X_n) - f(x)|) \leq 2\varepsilon$.

Enfin, l'espérance d'une constante étant la valeur de celle-ci, on a pour tout $x \in [0; 1]$

$$|B_n(f)(x) - f(x)| = |\mathbb{E}(f(X_n)) - \mathbb{E}(f(x))| = |\mathbb{E}(f(X_n) - f(x))|.$$

Par croissance de l'espérance, on peut affirmer $|\mathbb{E}(Y)| \leq \mathbb{E}(|Y|)$ pour toute variable réelle Y et donc conclure

$$|B_n(f)(x) - f(x)| \leq \mathbb{E}(|f(X_n) - f(x)|) \leq 2\varepsilon.$$

Ainsi, la fonction f continue sur $[0; 1]$ peut être approchée *uniformément* par une fonction polynomiale. Ce résultat sera de nouveau présenté dans le cours de seconde année (Th 7 du chapitre 7 de l'ouvrage *Exercices d'analyse MP* dans la même collection).

Table des matières

1 Ensembles et applications	3
1.1 Rudiments de logique	3
1.2 Ensembles	6
1.3 Applications	8
1.4 Relations binaires	12
1.5 Exercices d'apprentissage	13
Phrases quantifiées	14
Raisonnements	17
Opérations dans $\wp(E)$	21
Applications	23
Relations binaires	27
1.6 Exercices d'entraînement	29
Opérations dans $\wp(E)$	29
Injection, surjection, bijection	30
Images directes et images réciproques	36
Relations binaires	40
1.7 Exercices d'approfondissement	42
2 Calculs algébriques	47
2.1 Les entiers naturels et le principe de récurrence	47
2.2 Sommes et produits	48
2.3 Formules du binôme et de factorisation	52
2.4 Systèmes d'équations linéaires	54
2.5 Exercices d'apprentissage	56
Principe de récurrence	56
Sommes et produits	57

Coefficients binomiaux	62
Systèmes d'équations linéaires	64
2.6 Exercices d'entraînement	66
Principe de récurrence	66
Sommes numériques	69
Produits numériques	72
Coefficients binomiaux	74
2.7 Exercices d'approfondissement	80
3 Arithmétique des entiers	87
3.1 Divisibilité	87
3.2 PGCD et PPCM	89
3.3 Entiers premiers entre eux	91
3.4 Nombres premiers	92
3.5 Exercices d'apprentissage	94
Divisibilité	94
PGCD et PPCM	95
Entiers premiers entre eux	97
Nombres premiers	99
3.6 Exercices d'entraînement	101
Etudes arithmétiques	101
Nombres premiers	108
L'infinité des nombres premiers	109
Décomposition en facteurs premiers	111
3.7 Exercices d'approfondissement	112
4 Structures algébriques usuelles	119
4.1 Loi de composition interne	119
4.2 Structure de groupe	122
4.3 Structure d'anneau	123
4.4 Exercices d'apprentissage	125
Loi de composition interne	125
Groupes	127
4.5 Exercices d'entraînement	130
Loi de composition interne	130
Groupes	132
Sous-groupes	134
Groupes finis	137
Anneaux et corps	138
4.6 Exercices d'approfondissement	141

5 Polynômes et fractions rationnelles	149
5.1 L'anneau des polynômes	149
5.2 Racines d'un polynôme	152
5.3 Dérivation	156
5.4 Arithmétique des polynômes	157
5.5 Le corps des fractions rationnelles	160
5.6 Exercices d'apprentissage	163
Polynômes	163
Racines	165
Arithmétique des polynômes	168
Décompositions en éléments simples	171
5.7 Exercices d'entraînement	176
Généralités	176
Racines	179
Relations coefficients-racines d'un polynôme scindé	181
Arithmétiques des polynômes	184
Familles de polynômes classiques	188
Les fractions rationnelles	193
5.8 Exercices d'approfondissement	198
6 Dénombrement	205
6.1 Cardinal d'un ensemble fini	205
6.2 Cardinaux usuels	207
6.3 Listes, arrangements, combinaisons	208
6.4 Exercices d'apprentissage	209
Généralités	209
Dénombrements	210
6.5 Exercices d'entraînement	213
Démonstrations combinatoires	213
Dénombrements	214
Compositions d'un entier	218
Dénombrements ensemblistes	221
Dénombrements d'applications	223
6.6 Exercices d'approfondissement	227
7 Espaces vectoriels	235
7.1 Espaces vectoriels	235
7.2 Sous-espaces vectoriels	237
7.3 Famille de vecteurs	240
7.4 Espaces de dimension finie	241
7.5 Sous-espaces affines	244
7.6 Exercices d'apprentissage	245
Sous-espaces vectoriels	245
Liberté d'une famille de vecteurs	249
7.7 Exercices d'entraînement	253

Généralités sur les espaces vectoriels	253
Liberté	259
Supplémentarité	260
Rang d'une famille de vecteurs	264
L'espace des polynômes	265
Sous-espaces affines	267
7.8 Exercices d'approfondissement	268
8 Les applications linéaires	273
8.1 Applications linéaires	273
8.2 Endomorphismes	275
8.3 Détermination d'une application linéaire	276
8.4 Théorème du rang	277
8.5 Exercices d'apprentissage	279
Généralités	279
Applications linéaires en dimension finie	283
8.6 Exercices d'entraînement	286
Généralités	286
Projections vectorielles	291
Applications linéaires en dimension finie	296
Rang d'une application linéaire	298
Formes linéaires et hyperplans	303
8.7 Exercices d'approfondissement	305
9 Matrices	309
9.1 Calcul matriciel	309
9.2 Représentations matricielles	313
9.3 Changements de bases	317
9.4 Opérations élémentaires et systèmes linéaires	319
9.5 Exercices d'apprentissage	321
Calculs matriciels	321
Matrices et applications linéaires	326
9.6 Exercices d'entraînement	331
Les matrices carrées	331
Matrices carrées inversibles	335
Matrices et applications linéaires	337
Trace d'une matrice carrée	344
Rang	345
9.7 Exercices d'approfondissement	348
10 Déterminants	353
10.1 Groupe symétrique	353
10.2 Déterminants	355
10.3 Calculs de déterminants	358
10.4 Applications	360

10.5 Exercices d'apprentissage	361
Permutation de $\{1, \dots, n\}$	361
Calculs de déterminants	363
10.6 Exercices d'entraînement	366
Calculs de déterminants	366
Déterminant d'une matrice	375
Déterminant par blocs	377
Déterminant d'un endomorphisme	379
Applications	383
10.7 Exercices d'approfondissement	387
11 Espaces préhilbertiens réels	393
11.1 Produit scalaire	393
11.2 Espaces euclidiens	397
11.3 Isométries vectorielles	401
11.4 Exercices d'apprentissage	405
Produit scalaire	405
Base orthonormale	408
Orthogonal d'une partie	410
Isométric et matrices orthogonales	413
11.5 Exercices d'entraînement	415
Produit scalaire	415
Espace euclidien	418
Projection orthogonale	419
Calcul de distance à un sous-espace vectoriel	424
Isométries	426
Matrices orthogonales	430
11.6 Exercices d'approfondissement	432
12 Probabilités	439
12.1 Probabilité sur un univers fini	439
12.2 Probabilités conditionnelles	441
12.3 Exercices d'apprentissage	444
12.4 Exercices d'entraînement	450
Définition d'une probabilité	450
Événements indépendants	453
Calcul de probabilités	458
Probabilités conditionnelles	464
12.5 Exercices d'approfondissement	472
13 Variables aléatoires	477
13.1 Variables aléatoires sur un espace probabilisé fini	477
13.2 Vecteurs aléatoires	479
13.3 Espérance et variance d'une variable aléatoire réelle	482
13.4 Exercices d'apprentissage	485

13.5 Exercices d'entraînement	491
Loi binomiale	491
Espérances et variances	494
Calcul d'espérances et de variances	497
Couples de variables aléatoires	503
Covariance	505
Inégalités de concentration	508
13.6 Exercices d'approfondissement	511

Cet ouvrage propose 401 exercices d'algèbre et de probabilités regroupés par chapitre et accompagnés de résumés de cours. Il est destiné aux élèves de CPGE scientifiques de première année en filière MPSI. Il pourra aussi intéresser les étudiants préparant le CAPES de mathématiques.

Les **résumés de cours** présentent de façon synthétique les définitions et les théorèmes conformément au programme de la filière. Ils seront utiles pour une **révision rapide et efficace** et pourront servir de formulaire.

Les **exercices** proposés sont de niveaux variés et regroupés en trois catégories :

- les **exercices d'apprentissage** permettent l'acquisition des fondamentaux du cours ;
- les **exercices d'entraînement** conduisent à la maîtrise des concepts du chapitre ;
- les **exercices d'approfondissement** invitent les étudiants à une recherche plus fouillée par la mise en résonance de notions présentées dans différents chapitres.

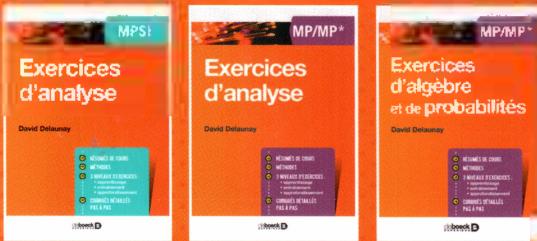
Les corrections des exercices sont **détaillées** pas à pas et accompagnées de **méthodes** mettant en lumière les démarches suivies et les idées récurrentes.



- des résumés de cours
- des méthodes
- 401 exercices de niveaux variés
- des corrigés très détaillés
- strictement conforme au programme officiel

David DELAUNAY, ancien élève de l'École normale supérieure de Cachan, est professeur agrégé de mathématiques en classes préparatoires au lycée Dupuy de Lôme de Lorient.

Collection dirigée par Olivier RODOT



ISBN : 978-2-8073-0613-4



deboeck SUPÉRIEUR

www.deboecksuperieur.com