



MARIE-CÉCILE DARRACQ • JEAN-ÉTIENNE ROMBALDI

Algèbre et géométrie pour la Licence

**LICENCES 1 & 2
MATHÉMATIQUES**

- Cours complet
- Près de 200 exercices
- Tous les corrigés détaillés



MARIE-CÉCILE DARRACQ • JEAN-ÉTIENNE ROMBALDI

Algèbre et géométrie pour la Licence

Des mêmes auteurs

DARRACQ M.-C., ROMBALDI J.-É., *Probabilités pour la Licence* (nouveau 2021)

DARRACQ M.-C., ROMBALDI J.-É., *Analyse pour la Licence*

DARRACQ M.-C., ROMBALDI J.-É., *Mathématiques pour le Capes. Probabilités* (nouveau 2021)

DARRACQ M.-C., ROMBALDI J.-É., *Mathématiques pour le Capes. Algèbre et géométrie* (nouveau 2021)

DARRACQ M.-C., ROMBALDI J.-É., *Mathématiques pour le Capes. Analyse*

ROMBALDI J.-É., *Mathématiques pour l'agrégation. Algèbre et géométrie* – 2^e édition (nouveau 2021)

ROMBALDI J.-É., *Exercices et problèmes corrigés pour l'agrégation de mathématiques*

ROMBALDI J.-É., *Leçons d'oral pour l'agrégation de mathématiques. Première épreuve : les exposés*

ROMBALDI J.-É., *Leçons d'oral pour l'agrégation de mathématiques. Seconde épreuve : les exercices*

Pour toute information sur notre fonds et les nouveautés dans votre domaine
de spécialisation, consultez notre site web :

www.deboecksuperieur.com

En couverture : Coupe d'un nautille © AdrianHancu/Istockphoto

Maquette intérieure : Hervé Soulard/Nexeme

Mise en pages de l'auteur

Maquette de couverture : Primo&Primo

Couverture : SCM, Toulouse

Dépôt légal :

Bibliothèque royale de Belgique : 2021/13647/105

Bibliothèque nationale, Paris : juillet 2021

ISBN : 978-2-8073-3221-8

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 ou de quelque manière que ce soit.*

© De Boeck Supérieur SA, 2021 - Rue du Bosquet 7, B1348 Louvain-la-Neuve
De Boeck Supérieur - 5 allée de la 2^e DB, 75015 Paris

Sommaire

Avant-propos	ix
1 Éléments de logique et de théorie des ensembles	1
2 Structure de groupe	31
3 Structures d'anneau et de corps	57
4 Division euclidienne dans $\mathbb{Z}$	71
5 Le corps $\mathbb{C}$ des nombres complexes	109
6 Espaces vectoriels réels ou complexes	139
7 Espaces vectoriels réels ou complexes de dimension finie	169
8 Opérations élémentaires et déterminants	191
9 Polynômes à coefficients réels ou complexes	211
10 Réduction des endomorphismes	245
11 Formes bilinéaires et quadratiques réelles ou complexes	263
12 Espaces préhilbertiens	305
Bibliographie	337
Index	339

Table des matières

Avant-propos	ix
1 Éléments de logique et de théorie des ensembles	1
1.1 Quelques notions de logique	1
1.2 Les connecteurs logiques de base	2
1.3 Quelques méthodes de raisonnement	4
1.4 Notions de base sur les ensembles. Quantificateurs	5
1.5 Les symboles $\sum$ et $\prod$	7
1.6 Les théorèmes de récurrence	8
1.7 L'algèbre des parties d'un ensemble	9
1.8 Applications. Notions d'injectivité, surjectivité et bijectivité	11
1.9 Relations d'ordre et d'équivalence	16
1.10 Exercices	18
2 Structure de groupe	31
2.1 Loi de composition interne	31
2.2 Groupes	33
2.3 Sous-groupes	35
2.4 Sous-groupe engendré par une partie	36
2.5 Groupes monogènes	37
2.6 Groupes finis. Théorème de Lagrange	39
2.7 Morphismes de groupes	40
2.8 Ordre d'un élément dans un groupe	43
2.9 Exercices	46
3 Structures d'anneau et de corps	57
3.1 Anneaux	57
3.2 Éléments inversibles dans un anneau unitaire	60
3.3 Sous-anneaux	61
3.4 Morphismes d'anneaux	62
3.5 Corps	62
3.6 Morphismes de corps	64
3.7 Exercices	64
4 Division euclidienne dans $\mathbb{Z}$	71
4.1 Divisibilité et congruences	71
4.2 Le théorème de division euclidienne dans $\mathbb{Z}$	72

4.3	Les systèmes de numération	73
4.4	Plus grand commun diviseur et plus petit commun multiple	74
4.5	L'algorithme d'Euclide	80
4.6	Equations diophantiennes $ax + by = c$	82
4.7	Equations $ax \equiv b \pmod{n}$	83
4.8	Le théorème Chinois	84
4.9	L'ensemble $\mathcal{P}$ des nombres premiers	85
4.10	Décomposition en facteurs premiers	87
4.11	Les théorèmes de Fermat et de Wilson	89
4.12	Les anneaux $\frac{\mathbb{Z}}{n\mathbb{Z}}$ et la fonction indicatrice d'Euler	92
4.13	Exercices	97
5	Le corps $\mathbb{C}$ des nombres complexes	109
5.1	Conditions nécessaires à la construction de $\mathbb{C}$	110
5.2	Construction de $\mathbb{C}$	110
5.3	Conjugué et module d'un nombre complexe	113
5.4	Les équations de degré 2	116
5.5	Les équations de degré 3 et 4	119
5.6	Arguments d'un nombre complexe	120
5.7	Racines n -ièmes d'un nombre complexe	125
5.8	Exercices	127
6	Espaces vectoriels réels ou complexes	139
6.1	L'espace vectoriel $\mathbb{K}^n$	139
6.2	Définition d'un espace vectoriel réel ou complexe	140
6.3	Sous-espaces vectoriels	142
6.4	Applications linéaires	146
6.5	Base canonique de $\mathbb{K}^n$ et expression matricielle des applications linéaires de $\mathbb{K}^n$ dans $\mathbb{K}^m$	149
6.6	Matrices réelles ou complexes	151
6.7	Systèmes d'équations linéaires	159
6.8	Sommes et sommes directes de sous-espaces vectoriels	160
6.9	Exercices	162
7	Espaces vectoriels réels ou complexes de dimension finie	169
7.1	Systèmes libres, systèmes générateurs et bases	169
7.2	Espaces vectoriels de dimension finie	172
7.3	Rang d'un système de vecteurs ou d'une application linéaire	179
7.4	Expression matricielle des applications linéaires	181
7.5	Formules de changement de base	184
7.6	Exercices	186
8	Opérations élémentaires et déterminants	191
8.1	Opérations élémentaires. Matrices de dilatation et de transvection	192
8.2	Déterminants des matrices carrées	195
8.3	Déterminant d'une famille de vecteurs	203
8.4	Déterminant d'un endomorphisme	204
8.5	Exercices	205

9	Polynômes à coefficients réels ou complexes	211
9.1	L'algèbre $\mathbb{K}[X]$ des polynômes à coefficients dans $\mathbb{K}$	211
9.2	Division euclidienne dans $\mathbb{K}[X]$	215
9.3	Polynômes premiers entre eux	220
9.4	Fonctions polynomiales, racines des polynômes	222
9.5	Dérivation des polynômes. Formule de Taylor	225
9.6	Polynômes irréductibles	227
9.7	Décomposition des fractions rationnelles en éléments simples	231
9.8	Polynômes d'interpolation de Lagrange	236
9.9	Exercices	238
10	Réduction des endomorphismes	245
10.1	Polynômes d'endomorphismes, polynômes annulateurs, polynôme minimal	245
10.2	Le théorème de décomposition des noyaux	246
10.3	Valeurs et vecteurs propres	248
10.4	Le théorème de Cayley-Hamilton	251
10.5	Diagonalisation	253
10.6	Endomorphismes trigonalisables	255
10.7	Exercices	257
11	Formes bilinéaires et quadratiques réelles ou complexes	263
11.1	Formes linéaires	263
11.2	Formes bilinéaires	266
11.3	Expression matricielle des formes bilinéaires (en dimension finie)	267
11.4	Formes quadratiques	269
11.5	Théorème de réduction de Gauss	271
11.6	Orthogonalité, noyau et rang	278
11.7	Signature d'une forme quadratique réelle en dimension finie	282
11.8	Quadriques dans $\mathbb{R}^n$ ou $\mathbb{C}^n$	285
11.9	Exercices	289
12	Espaces préhilbertiens	305
12.1	Produit scalaire	305
12.2	Orthogonalité	307
12.3	Procédé d'orthogonalisation de Gram-Schmidt	309
12.4	Projection orthogonale sur un sous-espace de dimension finie	311
12.5	Les endomorphismes symétriques réels	317
12.6	Espaces vectoriels hermitiens	319
12.7	Réduction des matrices normales	321
12.8	Exercices	322
	Bibliographie	337
	Index	339

Avant-propos

Ce cours d'algèbre s'adresse aux étudiants de première et deuxième année d'université. C'est le deuxième volume d'une série qui en comporte 3, le premier volume étant consacré à l'analyse et le troisième à la théorie des probabilités. Il ne s'agit pas de manuels de « méthodes » où l'on sacrifie la notion de rigueur qui est l'essence même des mathématiques. Les notions étudiées le sont de façon rigoureuse en démontrant tous (ou presque) les résultats énoncés. Chaque chapitre se termine par une série d'exercices tous corrigés en détails.

Ce deuxième volume est consacré aux notions d'algèbre habituellement enseignées en première et deux année de licence (L1 et L2), à savoir l'étude de quelques notions de logique et de théorie des ensembles, des structures de groupe, d'anneaux et de corps, en se concentrant sur l'anneau des entiers relatifs, le corps des nombres complexes, l'anneau des polynômes à coefficients réels ou complexes et les principales notions d'algèbre linéaire et bilinéaire avec la réduction des endomorphismes et des formes quadratiques. On s'intéresse également à quelques notions d'arithmétique.

Les élèves en classes préparatoires aux grandes écoles pourront aussi tirer profit de cet ouvrage.

Ces livres seront également utiles aux candidats au Capes. Une deuxième version de ce volume d'algèbre avec un dernier chapitre consacré à quelques épreuves choisies de ce concours sera également publiée à destination de ces candidats, le niveau d'exigence pour ces épreuves ne dépassant pas le niveau L1-L2. Nous espérons que ce travail leur sera utile.

Pour conclure, nous tenons à remercier les éditions De Boeck et en particulier Alain Luguët pour la confiance qu'ils nous accordent en publiant ce travail.

Chapitre 1

Éléments de logique et de théorie des ensembles

Pour les exemples et exercices traités dans ce chapitre les ensembles usuels de nombres entiers, rationnels réels et complexes sont supposés connus, au moins de manière intuitive comme cela se passe au Lycée.

1.1 Quelques notions de logique

Nous allons préciser à un premier niveau quelques notions mathématiques qui sont relativement intuitives mais nécessitent quand même des définitions rigoureuses. L'idée étant de préciser schématiquement comment se présente une théorie mathématique ainsi que la notion essentielle de démonstration.

La première notion est celle d'assertion. De manière intuitive, une assertion est un énoncé mathématique aussi rigoureux que possible qui ne peut prendre que deux valeurs de vérité à savoir « vrai » ou « faux » mais jamais entre les deux comme dans le langage courant.

Une assertion qui est toujours vraie est une tautologie.

Par exemple les énoncés suivants sont des assertions : $2 < 15$ (elle est vraie), $\sqrt{2}$ est un nombre rationnel (elle est fausse), $\cos(n\pi) = (-1)^n$ (vraie), ...

Deux assertions sont dites logiquement équivalentes, ou plus simplement équivalentes, si elles sont toutes deux vraies ou toutes deux fausses.

Il y a ensuite les énoncés qui se démontrent. Pour ce faire, on se donne des règles précises (que nous verrons par la pratique) qui permettent de construire de nouvelles assertions à partir d'assertions données.

Il ne faut pas croire que dans une théorie donnée toute assertion P soit obligatoirement démontrable. En 1931 Kurt Gödel a démontré qu'il y a des assertions non démontrables (on dit aussi qu'elles sont indécidables) : il n'est pas possible de démontrer que P est vraie ni que P est fausse.

À la base de toute théorie mathématique, on dispose d'un petit nombre d'assertions qui sont supposés vraies *a priori* (c'est-à-dire avant toute expérience) et que l'on nomme axiomes ou postulats. Ces axiomes sont élaborés par abstraction à partir de l'intuition et ne sont pas déduits d'autres relations.

Par exemple, la géométrie euclidienne est basée sur une quinzaine d'axiomes. L'un de ces axiomes est le postulat numéro 15 qui affirme que par un point donné passe une et une seule droite parallèle à une droite donnée.

Une autre exemple important est donné par la construction de l'ensemble noté $\mathbb{N}$ des entiers naturels. Cette construction peut se faire en utilisant les axiomes de Peano suivants :

- 0 est un entier naturel ;
- tout entier naturel n a un unique successeur noté $n + 1$;
- deux entiers naturels ayant même successeur sont égaux ;
- une partie P de $\mathbb{N}$ qui contient 0 et telle que si n est dans P alors le successeur de n y est aussi, est égale à $\mathbb{N}$ (axiome de récurrence).

Nous reviendrons au paragraphe 1.6 sur l'ensemble $\mathbb{N}$ en partant sur une autre base.

La théorie des ensemble est basée sur le système d'axiomes de Zermelo-Fränkel.

La notion de définition nous permet de décrire un objet ou une situation précise à l'aide du langage courant.

Les énoncés qui se démontrent sont classés en fonction de leur importance dans une théorie comme suit :

- un théorème est une assertion vraie déduite d'autres assertions, il s'agit en général d'un résultat important à retenir ;
- un lemme est un résultat préliminaire utilisé pour démontrer un théorème ;
- un corollaire est une conséquence importante d'un théorème ;
- une proposition est de manière générale un résultat auquel on peut attribuer la valeur vraie ou fausse sans ambiguïté.

Pour rédiger un énoncé mathématique, on utilise le langage courant et les objets manipulés sont représentés en général par des lettres de l'alphabet latin ou grec. Usuellement, on utilise :

- les lettres minuscules $a, b, c, \dots$ pour des objets fixés ;
- les lettres minuscules $x, y, z, t, \dots$ pour des objets inconnus à déterminer ;
- les lettres majuscules $E, F, G, H, \dots$ pour des ensembles ;
- des lettres de l'alphabet grecques minuscules ou majuscules $\alpha, \beta, \varepsilon, \delta, \dots, \Lambda, \Gamma, \Omega, \dots$

1.2 Les connecteurs logiques de base

L'élaboration de nouvelles assertions à partir d'autres se fait en utilisant les connecteurs logiques de négation, de conjonction, de disjonction, d'implication et d'équivalence définis comme suit, où P et Q désignent des assertions.

- La négation de P , notée $\neg P$, ou non P ou $\overline{P}$, est l'assertion qui est vraie si P est fausse et fausse si P est vraie. Par exemple la négation de l'assertion : « x est strictement positif » est « x est négatif ou nul ». En théorie des ensembles on admet qu'il n'existe pas d'assertion P telle que P et $\overline{P}$ soient toutes deux vraies. On dit que cette théorie est non contradictoire.

- La conjonction de P et Q , notée $P \wedge Q$ (lire P et Q), est l'assertion qui est vraie uniquement si P et Q sont toutes deux vraies (et donc fausse dans les trois autres cas). Par exemple $P \wedge \overline{P}$ est toujours faux (on se place dans des théories non contradictoires).
- La disjonction de P et Q , notée $P \vee Q$ (lire P ou Q), est l'assertion qui est vraie uniquement si l'une des deux assertions P ou Q est vraie (donc fausse si P et Q sont toutes deux fausses). Par exemple $P \vee \overline{P}$ est toujours vraie (c'est une tautologie). Il faut remarquer que le « ou » pour « ou bien » est inclusif, c'est-à-dire que P et Q peuvent être toutes deux vrais dans le cas où $P \vee Q$ est vraie. On peut aussi introduire le « ou exclusif », noté W , qui est vrai uniquement lorsque l'une des deux assertions, mais pas les deux simultanément, est vraie.
- L'implication, notée $P \rightarrow Q$, est l'assertion qui est fausse uniquement si P est vraie et Q fausse (donc vraie dans les trois autres cas). On peut remarquer que si P est fausse, alors $P \rightarrow Q$ est vraie indépendamment de la valeur de vérité de Q . L'implication est à la base du raisonnement mathématique. En partant d'une assertion P (ou de plusieurs), une démonstration aboutit à un résultat Q . Si cette démonstration est faite sans erreur, alors $P \rightarrow Q$ est vraie et on notera $P \Rightarrow Q$ (ce qui signifie que si P est vraie, alors Q est vraie). Dans ce cas, on dit que P est une condition suffisante et Q une condition nécessaire. On peut remarquer que l'implication est transitive, c'est-à-dire que si P implique Q et Q implique R , alors P implique R .
- L'équivalence de P et Q , notée $P \leftrightarrow Q$, est l'assertion qui est vraie uniquement si $P \rightarrow Q$ et $Q \rightarrow P$ sont toutes deux vraies. Dans le cas où $P \leftrightarrow Q$ est vraie on dit que P et Q sont équivalentes et on note $P \Leftrightarrow Q$ (ce qui signifie que P et Q sont, soit toutes deux vraies, soit toutes deux fausses). Dans ce cas, on dit que Q est une condition nécessaire et suffisante de P .

On peut résumer ce qui précède, en utilisant la table de vérité suivante :

P	Q	$\overline{P}$	$P \wedge Q$	$P \vee Q$	$P \rightarrow Q$	$P \leftrightarrow Q$
V	V	F	V	V	V	V
V	F	F	F	V	F	F
F	V	V	F	V	V	F
F	F	V	F	F	V	V

Les tables de vérité peuvent être utilisées pour faire certaines démonstrations. Deux assertions qui ont même table de vérité sont équivalentes.

Avec le théorème qui suit, on résume quelques règles de calcul.

Théorème 1.1.

Soient P, Q, R des propositions. On a les équivalences :

1. commutativité :

$$(P \wedge Q) \Leftrightarrow (Q \wedge P) \text{ et } (P \vee Q) \Leftrightarrow (Q \vee P)$$

2. associativité

$$(P \wedge (Q \wedge R)) \Leftrightarrow ((P \wedge Q) \wedge R) \text{ et } (P \vee (Q \vee R)) \Leftrightarrow ((P \vee Q) \vee R)$$

3. distributivité :

$$(P \wedge (Q \vee R)) \Leftrightarrow ((P \wedge Q) \vee (P \wedge R))$$

$$(P \vee (Q \wedge R)) \Leftrightarrow ((P \vee Q) \wedge (P \vee R))$$

4. négations :

$$(\overline{\overline{P}}) \Leftrightarrow (P), (\overline{P \wedge Q}) \Leftrightarrow (\overline{P} \vee \overline{Q}), (\overline{P \vee Q}) \Leftrightarrow (\overline{P} \wedge \overline{Q})$$

$$(P \rightarrow Q) \Leftrightarrow (\overline{Q} \rightarrow \overline{P}), (P \rightarrow Q) \Leftrightarrow (\overline{P} \vee Q), (\overline{P \rightarrow Q}) \Leftrightarrow (P \wedge \overline{Q})$$

Preuve. On utilise les tables de vérité (exercice laissé au lecteur). $\square$

Les équivalences $(P \wedge Q) \Leftrightarrow (\overline{P} \vee \overline{Q})$ et $(P \vee Q) \Leftrightarrow (\overline{P} \wedge \overline{Q})$ sont appelées lois de Morgan.

1.3 Quelques méthodes de raisonnement

En général l'énoncé d'une proposition à démontrer est formé d'une ou plusieurs hypothèses qui constituent l'assertion H et d'une ou plusieurs conclusions qui constituent l'assertion C . Il s'agit donc de montrer l'implication $H \Rightarrow C$. Si de plus, on peut montrer que $C \Rightarrow H$, on dira alors que la réciproque de la proposition est vraie. Les idées de base que l'on peut utiliser sont les suivantes.

- Une assertion peut toujours être remplacée par n'importe quelle assertion qui lui est équivalente.
- On peut effectuer une démonstration directe, c'est-à-dire de déduire logiquement C de H .
- L'implication étant transitive, on peut essayer de montrer que $C \Rightarrow C'$ sachant par ailleurs que $C' \Rightarrow H$.
- Dans le cas où une démonstration directe semble difficile, on peut essayer une démonstration par l'absurde qui consiste à étudier l'assertion $H \wedge \overline{C}$ équivalente à $\overline{H \rightarrow C}$ et on montre qu'on aboutit à une impossibilité si cette dernière assertion est vraie (pratiquement, on suppose que la conclusion est fausse avec les hypothèses et on aboutit à une absurdité). Il en résulte alors que $\overline{H \rightarrow C}$ est fausse, c'est-à-dire que $H \rightarrow C$ est vraie, soit $H \Rightarrow C$.
- On peut aussi essayer de montrer la contraposée $\overline{C} \Rightarrow \overline{H}$ puisque les implications $H \rightarrow C$ et $\overline{C} \rightarrow \overline{H}$ sont équivalentes.
- La démonstration par contre-exemple permet de montrer qu'une implication $H \rightarrow C$, où H et C sont des propriétés portant sur des variables x , est fausse. Pour ce faire on cherche une ou des valeurs de x pour lesquels $H(x)$ est vraie et $C(x)$ est fausse.

- La démonstration par récurrence permet de montrer qu’une propriété portant sur des entiers naturels est toujours vraie. Cette méthode de démonstration est décrite au paragraphe 1.6, où elle apparaît comme un théorème basé sur le fait que l’ensemble des entiers naturels est bien ordonné. Si on accepte l’axiome de Péano, le principe de récurrence en est une conséquence immédiate.

1.4 Notions de base sur les ensembles. Quantificateurs

Nous nous contenterons d’une définition intuitive de la notion d’ensemble.

Un ensemble est une collection d’objets possédant des propriétés communes, ces objets sont les éléments de l’ensemble. On admet l’existence d’un ensemble qui ne contient aucun élément. Cet ensemble est noté $\emptyset$ et on dit que c’est l’ensemble vide.

On utilisera les notations suivantes, pour les ensembles de nombres usuels :

- $\mathbb{N}$ est l’ensemble des entiers naturels ;
- $\mathbb{Z}$ est l’ensemble des entiers relatifs ;
- $\mathbb{Q}$ est l’ensemble des nombres rationnels
- $\mathbb{R}$ est l’ensemble des nombres réels ;
- $\mathbb{C}$ est l’ensemble des nombres complexes.

Nous serons souvent amenés à décrire un ensemble en précisant les propriétés que doivent vérifier tous ses éléments, ce que nous noterons de la façon suivante :

$$E = \{\text{description des propriétés des éléments de } E\}$$

(on dit que l’ensemble E est défini en compréhension).

Cette notion d’ensemble défini en compréhension peut conduire à des paradoxes liés au problème de « l’ensemble de tous les ensembles », mais à un premier niveau, on se contente de ce point de vue intuitif. Une étude approfondie de la théorie des ensembles peut mener assez loin. Le lecteur intéressé peut consulter le volume de Bourbaki sur les ensembles, ou tout autre ouvrage spécialisé.

On peut aussi décrire un ensemble en donnant la liste finie ou infinie de tous ces éléments, quand cela est possible, ce qui se note $E = \{x_1, x_2, \dots, x_n\}$ s’il s’agit d’un ensemble fini ou $E = \{x_1, x_2, \dots, x_n, \dots\}$ s’il s’agit d’un ensemble infini pour lequel on peut numéroter les éléments (un tel ensemble est dit dénombrable). On dit alors que l’ensemble E est défini en extension.

Un singleton est un ensemble qui ne contient qu’un élément, soit $E = \{a\}$.

Si n, m sont deux entiers relatifs, l’ensemble des entiers relatifs compris entre n et m sera noté $\{n, \dots, m\}$. Dans le cas où $m < n$, il ne peut y avoir d’entiers entre n et m et cet ensemble est tout simplement l’ensemble vide. Dans le cas où $n = m$, cet ensemble est le singleton $\{n\}$. Pour $n < m$, on notera aussi $\{n, n+1, \dots, m\}$ cet ensemble.

Si E est un ensemble, on notera alors $a \in E$ pour signifier que a est un élément de E , ce qui se lit « a appartient à E ». La négation de cette assertion est « a n’appartient pas à E » et se notera $a \notin E$.

Pour signifier qu'un ensemble F est contenu dans un ensemble E , ce qui signifie que tout élément de F est dans E , nous noterons $F \subset E$ qui se lit « F est contenu dans E ». On peut écrire de manière équivalent que $E \supset F$ pour dire que E contient F . La négation de cette assertion est notée $F \not\subset E$.

Deux ensembles E et F sont égaux si, et seulement si, ils ont les mêmes éléments, ce qui se traduit par $E \subset F$ et $F \subset E$.

On admet que si E est un ensemble, il existe un ensemble dont tous les éléments sont formés de tous les sous-ensembles (ou parties) de E . On note $\mathcal{P}(E)$ cet ensemble et on dit que c'est l'ensemble des parties de E . Ainsi $F \subset E$ est équivalent à $F \in \mathcal{P}(E)$. L'ensemble vide et E sont des éléments de $\mathcal{P}(E)$.

Par exemple pour $E = \{1, 2, 3\}$, on a :

$$\mathcal{P}(E) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$$

Pour décrire des ensembles, ou faire des raisonnements, nous utiliseront les deux quantificateurs suivants.

- Le quantificateur universel « quel que soit » ou « pour tout » noté $\forall$ utilisé pour signifier que tout élément x d'un ensemble E vérifie une propriété $P(x)$, la syntaxe étant :

$$(\forall x \in E) (P(x)) \quad (1.1)$$

- Le quantificateur existentiel « il existe » noté $\exists$ pour signifier qu'il existe au moins un élément x de E vérifiant la propriété $P(x)$, la syntaxe étant :

$$(\exists x \in E) | (P(x)) \quad (1.2)$$

Pour signifier qu'il existe un et un seul x dans E vérifiant la propriété $P(x)$, on utilisera la syntaxe $(\exists! x \in E) | (P(x))$.

La négation de l'assertion 1.1 est $(\exists x \in E) | (\overline{P(x)})$ en utilisant le symbole $|$ qui se lit « tel que » utilisé pour traduire le fait que x est tel que la propriété $\overline{P(x)}$ est vérifiée et la négation de 1.2 est $(\forall x \in E) (\overline{P(x)})$.

Nous verrons qu'il n'est pas toujours facile de traduire la négation d'une assertion en utilisant les quantificateurs. Par exemple pour traduire le fait qu'une suite $(u_n)_{n \in \mathbb{N}}$ de nombres réels est convergente vers un réel ℓ nous écrirons :

$$(\exists \ell \in \mathbb{R}) | (\forall \varepsilon > 0, \exists n_0 \in \mathbb{N} | \forall n \geq n_0, |u_n - \ell| < \varepsilon)$$

ce qui signifie qu'il existe un réel ℓ tel que quel que soit la précision $\varepsilon > 0$ que l'on choisisse l'écart entre u_n et ℓ (soit $|u_n - \ell|$) est inférieur à ε à partir d'un certain rang n_0 . La négation de cette assertion s'écrit :

$$(\forall \ell \in \mathbb{R}), (\exists \varepsilon > 0, \forall n_0 \in \mathbb{N}, \exists n \geq n_0 | |u_n - \ell| \geq \varepsilon)$$

En utilisant les quantificateurs, il faudra faire attention à l'ordre d'apparition de ces derniers. Par exemple les assertions suivantes, où f est une fonction à valeurs réelles définie sur un ensemble E :

$$\forall x \in E, \exists M > 0 | f(x) < M$$

$$\exists M > 0 \mid \forall x \in E, f(x) < M$$

ne sont pas équivalentes. La première assertion signifie que pour tout élément x de E il existe un réel $M > 0$ qui dépend à priori de x (il faudrait donc le noter $M(x)$) tel que $f(x) < M$ (par exemple $M(x) = f(x) + 1$ convient), alors que la seconde signifie qu'il existe un réel $M > 0$, indépendant de x dans E , tel que $f(x) < M$, ce qui n'est pas la même chose.

1.5 Les symboles $\sum$ et $\prod$

Si n est un entier naturel non nul et $x_1, x_2, \dots, x_n$ des entiers, rationnels, réels ou complexes, on notera :

$$\sum_{k=1}^n x_k = x_1 + x_2 + \dots + x_n \text{ et } \prod_{k=1}^n x_k = x_1 \cdot x_2 \cdot \dots \cdot x_n$$

la somme et le produit des x_k . Dans une telle somme ou produit l'indice est muet, ce qui signifie que $\sum_{k=1}^n x_k = \sum_{i=1}^n x_i$ et $\prod_{k=1}^n x_k = \prod_{i=1}^n x_i$.

La manipulation d'un produit de réels strictement positifs se ramène à une somme en utilisant la fonction logarithme $\ln \left(\prod_{k=1}^n x_k \right) = \sum_{k=1}^n \ln(x_k)$.

On peut également effectuer des changements d'indice. Par exemple, en posant $i = k + 1$, on a $\sum_{k=1}^n x_k = \sum_{i=2}^{n+1} x_{i-1} = \sum_{k=2}^{n+1} x_{k-1}$.

On peut ajouter ou multiplier de telles sommes (ou produits). Par exemple, on a :

$$\begin{aligned} \sum_{k=1}^n x_k + \sum_{k=1}^n y_k &= \sum_{k=1}^n (x_k + y_k), \quad \lambda \sum_{k=1}^n x_k = \sum_{k=1}^n \lambda x_k \\ \left(\sum_{k=1}^n x_k \right) \left(\sum_{k=1}^m y_k \right) &= \left(\sum_{j=1}^n x_j \right) \left(\sum_{k=1}^m y_k \right) = \sum_{\substack{1 \leq j \leq n \\ 1 \leq k \leq m}} x_j y_k \end{aligned}$$

Pour vérifier ce dernier résultat, on écrit que :

$$\begin{aligned} S &= \left(\sum_{k=1}^n x_k \right) \left(\sum_{k=1}^m y_k \right) = (x_1 + x_2 + \dots + x_n) \left(\sum_{k=1}^m y_k \right) \\ &= x_1 \sum_{k=1}^m y_k + \dots + x_n \sum_{k=1}^m y_k = \sum_{j=1}^n x_j \left(\sum_{k=1}^m y_k \right) \\ &= \sum_{j=1}^n \left(\sum_{k=1}^m x_j y_k \right) = \sum_{\substack{1 \leq j \leq n \\ 1 \leq k \leq m}} x_j y_k \end{aligned}$$

1.6 Les théorèmes de récurrence

On désigne par $\mathbb{N}$ l'ensemble des entiers naturels, soit $\mathbb{N} = \{0, 1, 2, \dots, n, \dots\}$. La construction de cet ensemble avec les opérations usuelles d'addition et de multiplication est admise. On note $\mathbb{N}^*$ l'ensemble $\mathbb{N}$ privé de 0.

Notre point de départ est l'axiome du bon ordre suivant : toute partie non vide de $\mathbb{N}$ admet un plus petit élément, ce qui signifie que si A est une partie non vide de $\mathbb{N}$, il existe alors un entier $m \in \mathbb{N}$ tel que $m \leq n$ pour tout $n \in A$. De cet axiome du bon ordre, on déduit les deux théorèmes fondamentaux qui suivent. Le premier résultat est souvent appelé théorème de récurrence faible et le second théorème de récurrence forte.

Théorème 1.2.

Soient $n_0 \in \mathbb{N}$ et $\mathcal{P}(n)$ une propriété portant sur les entiers $n \geq n_0$. La propriété $\mathcal{P}(n)$ est vraie pour tout entier $n \geq n_0$ si, et seulement si :

- (i) $\mathcal{P}(n_0)$ est vraie ;
- (ii) pour tout $n \geq n_0$ si $\mathcal{P}(n)$ est vrai alors $\mathcal{P}(n+1)$ est vraie.

Preuve. La condition nécessaire est évidente. En supposant les conditions (i) et (ii) vérifiées, on note A l'ensemble des entiers $n \geq n_0$ pour lesquels $\mathcal{P}(n)$ est faux. Si A est non vide, il admet alors un plus petit élément $n > n_0$ (puisque $\mathcal{P}(n_0)$ est vraie). Mais alors $\mathcal{P}(n-1)$ est vraie ce qui implique, d'après (ii), que $\mathcal{P}(n)$ est vraie, soit une contradiction. En définitive A est vide et la propriété est vraie pour tout entier $n \geq n_0$. $\square$

Théorème 1.3.

Soient $n_0 \in \mathbb{N}$ et $\mathcal{P}(n)$ une propriété portant sur les entiers $n \geq n_0$. La propriété $\mathcal{P}(n)$ est vraie pour tout entier $n \geq n_0$ si, et seulement si :

- (i) $\mathcal{P}(n_0)$ est vraie ;
- (ii) pour tout $n \geq n_0$ si $\mathcal{P}(k)$ est vrai pour tout entier k compris entre n_0 et n , alors $\mathcal{P}(n+1)$ est vraie.

Preuve. La condition nécessaire est évidente. En supposant les conditions (i) et (ii) vérifiées, on note A l'ensemble des entiers $n \geq n_0$ pour lesquels $\mathcal{P}(n)$ est faux. Si A est non vide, il admet alors un plus petit élément $n > n_0$ et $\mathcal{P}(k)$ est vraie pour tout k compris entre n_0 et $n-1$, ce qui implique que $\mathcal{P}(n)$ est vraie, soit une contradiction. En définitive A est vide et la propriété est vraie pour tout entier $n \geq n_0$. $\square$

Le théorème de récurrence nous permet de définir la fonction factorielle sur l'ensemble des entiers naturels de la façon suivante :

$$\begin{cases} 0! = 1 \\ \forall n \in \mathbb{N}, (n+1)! = (n+1)n! \end{cases}$$

De manière plus générale, c'est le théorème de récurrence qui nous assure de l'existence et de l'unicité d'une suite (réelle ou complexe) définie par :

$$\begin{cases} u_0 \text{ est donné} \\ \forall n \in \mathbb{N}, u_{n+1} = f(u_n) \end{cases}$$

où f est une fonction définie sur un ensemble I et à valeurs dans le même ensemble I . Une telle suite est dite définie par une relation de récurrence (d'ordre 1). Une telle suite peut aussi se définir en donnant les premières valeurs $u_0, u_1, \dots, u_p$ et une relation $u_{n+1} = f(u_n, \dots, u_{n-(p-1)})$ pour $n \geq p-1$. Une telle suite est dite définie par une relation de récurrence d'ordre p .

1.7 L'algèbre des parties d'un ensemble

Nous allons définir sur l'ensemble $\mathcal{P}(E)$ des parties d'un ensemble E des opérations qui vont traduire les idées intuitives de partie complémentaire, d'intersection et de réunion.

L'ensemble E étant donné et $A, B, C, \dots$ désignant des parties de E (donc des éléments de $\mathcal{P}(E)$), on définit les ensembles suivant.

- le complémentaire de A dans E est l'ensemble noté $C_E A$, ou $E \setminus A$ (lire E moins A) ou $\overline{A}$ des éléments de E qui ne sont pas dans A , ce qui peut se traduire par :

$$(x \in \overline{A}) \Leftrightarrow ((x \in E) \wedge (x \notin A))$$

ou encore par $\overline{A} = \{x \in E \mid x \notin A\}$.

- L'intersection de A et B , notée $A \cap B$, est l'ensemble des éléments de E qui sont dans A et dans B , soit :

$$(x \in A \cap B) \Leftrightarrow ((x \in A) \wedge (x \in B))$$

ou encore $A \cap B = \{x \in E \mid x \in A \text{ et } x \in B\}$. Si $A \cap B = \emptyset$, on dit alors que A et B sont disjointes. Par exemple A et $\overline{A}$ sont disjointes.

- La réunion de A et B , notée $A \cup B$, est l'ensemble des éléments de E qui sont soit dans A , soit dans B (éventuellement dans A et B) soit :

$$(x \in A \cup B) \Leftrightarrow ((x \in A) \vee (x \in B))$$

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

- La différence de A et B , notée $A \setminus B$, est l'ensemble des éléments de E qui sont dans A et qui ne sont pas dans B , soit :

$$(x \in A \setminus B) \Leftrightarrow ((x \in A) \wedge (x \notin B))$$

ou encore $A \setminus B = \{x \in A \mid x \notin B\}$. Ainsi $\overline{A} = E \setminus A$.

- La différence symétrique de A et B , notée $A \Delta B$, est l'ensemble des éléments de E qui sont soit dans A et pas dans B soit dans B et pas dans A (c'est-à-dire dans A ou exclusif dans B), soit :

$$(x \in A \Delta B) \Leftrightarrow ((x \in A) \wedge (x \notin B)) \vee ((x \in B) \wedge (x \notin A))$$

Par exemple, on a $A \Delta \emptyset = A$, $A \Delta E = \overline{A}$.

Ces opérateurs de complémentarité, intersection, réunion et différence symétrique sont décrits à l'aide des connecteurs logiques non de négation, $\wedge$ de conjonction, $\vee$ de disjonction et Δ de disjonction exclusive.

Avec le théorème qui suit, on résume les résultats essentiels relatifs à ces opérateurs ensemblistes.

Théorème 1.4.

Soient E un ensemble et $A, B, C, \dots$ des sous-ensembles de E . On a :

1. commutativité : $A \cap B = B \cap A$, $A \cup B = B \cup A$, $A \Delta B = B \Delta A$;
2. associativité :

$$A \cap (B \cap C) = (A \cap B) \cap C, \quad A \cup (B \cup C) = (A \cup B) \cup C$$

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

3. distributivité :

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C), \quad A \cap (B \Delta C) = (A \cap B) \Delta (A \cap C)$$

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

4. différence symétrique : $A \Delta A = \emptyset$, $A \Delta B = (A \setminus B) \cup (B \setminus A)$,
 $A \Delta B = (A \cap \overline{B}) \cup (B \cap \overline{A})$, $A \Delta B = (A \cup B) \setminus (A \cap B)$;
5. négations : $\overline{\overline{A}} = A$, $(A \subset B) \Leftrightarrow (\overline{B} \subset \overline{A})$, $\overline{A \cap B} = \overline{A} \cup \overline{B}$,
 $\overline{A \cup B} = \overline{A} \cap \overline{B}$.

Preuve. Laisée au lecteur. □

On notera l'analogie entre ce théorème et le théorème 1.1 sur les règles de calculs avec les connecteurs logiques.

La propriété d'associativité de l'intersection et de la réunion nous permet d'écrire $A \cap B \cap C$ et $A \cup B \cup C$ l'intersection et la réunion de trois ensembles sans se soucier de parenthèses. De manière plus générale, grâce à cette associativité, on peut définir l'intersection ou la réunion de n sous-ensembles $A_1, A_2, \dots, A_n$ de E par :

$$(x \in A_1 \cap A_2 \cap \dots \cap A_n) \Leftrightarrow ((x \in A_1) \wedge (x \in A_2) \wedge \dots \wedge (x \in A_n))$$

et :

$$(x \in A_1 \cup A_2 \cup \dots \cup A_n) \Leftrightarrow ((x \in A_1) \vee (x \in A_2) \vee \dots \vee (x \in A_n))$$

De façon condensée, on écrira $(A_k)_{1 \leq k \leq n}$ une telle famille de sous ensembles de E , $\bigcap_{k=1}^n A_k = A_1 \cap A_2 \cap \dots \cap A_n$ l'intersection et $\bigcup_{k=1}^n A_k = A_1 \cup A_2 \cup \dots \cup A_n$ la réunion.

On vérifie facilement que $E \setminus \bigcap_{k=1}^n A_k = \bigcup_{k=1}^n (E \setminus A_k)$, $E \setminus \bigcup_{k=1}^n A_k = \bigcap_{k=1}^n (E \setminus A_k)$ et

que pour tout entier j compris entre 1 et n , on a $\bigcap_{k=1}^n A_k \subset A_j \subset \bigcup_{k=1}^n A_k$.

Définition 1.1. On dit qu'une famille $(A_k)_{1 \leq k \leq n}$ de parties d'un ensemble E forme une partition de E , si les A_k sont deux à deux disjoints, c'est-à-dire que $A_k \cap A_j = \emptyset$ pour $1 \leq k \neq j \leq n$ et de réunion égale à E , soit

$$\bigcup_{k=1}^n A_k = E.$$

Dans le cas où (A_1, A_2) forme une partition de E , on a nécessairement $A_2 = \overline{A_1}$.

La notion de produit cartésien de deux ensembles sera très souvent utilisée. Elle correspond à l'idée de couples et se généralise pour aboutir à la notion de liste.

Définition 1.2. Étant donné deux ensembles E et F , on appelle produit cartésien de E par F l'ensemble $E \times F$ des couples (x, y) formés d'un élément x de E et d'un élément y de F .

Il est à noter que les couples sont ordonnés, c'est-à-dire que $(x, y) = (y, x)$ $E \times F$ si, et seulement si, $x = y$. De manière plus générale, on a $(x, y) = (x', y')$ dans $E \times F$ si, et seulement si, $x = x'$ et $y = y'$.

Dans le cas où $F = E$, on note E^2 pour $E \times E$.

On peut itérer le procédé et définir le produit cartésien $E_1 \times E_2 \times \cdots \times E_n$ de n ensembles comme l'ensemble des listes (ordonnées) $(x_1, x_2, \dots, x_n)$ formées d'un élément x_1 de E_1 suivi d'un élément x_2 de E_2 , $\dots$, suivi d'un élément x_n de E_n .

On notera de façon condensée $\prod_{k=1}^n E_k = E_1 \times \cdots \times E_n$.

Là encore, on a $(x_1, x_2, \dots, x_n) = (x'_1, x'_2, \dots, x'_n)$ dans $E \times F$ si, et seulement si, $x_k = x'_k$ pour tout k compris entre 1 et n .

Dans le cas où tous les E_k sont égaux à un même ensemble E , on notera E^n pour $E \times E \times \cdots \times E$ (n fois).

1.8 Applications. Notions d'injectivité, surjectivité et bijectivité

Les notations E, F, G désignent des ensembles.

Définition 1.3. On appelle application, ou fonction, de E dans F (ou de E vers F) toute partie Γ du produit cartésien $E \times F$ telle que :

$$\forall x \in E, \exists ! y \in F \mid (x, y) \in \Gamma$$

En notant f une application de E dans F (c'est en réalité le triplet (E, F, Γ) avec la propriété énoncée ci-dessus), on note pour tout $x \in E$, $f(x)$ l'unique élément de F tel que $(x, f(x)) \in \Gamma$ et on dit que $f(x)$ est l'image de x par f et x est un antécédent de y par f . Un antécédent de y par f n'est pas unique a priori. On dit aussi que E est l'ensemble de départ (ou l'ensemble de définition), F l'ensemble d'arrivée et Γ le graphe de l'application f .

Deux applications f et g sont égales si, et seulement si, elles ont même ensemble de départ E , même ensemble d'arrivée F et même graphe Γ , c'est-à-dire que :

$$\forall x \in E, g(x) = f(x)$$

On a tout simplement précisé l'idée d'un procédé qui associe à tout élément de E un unique élément de F . On note :

$$\begin{array}{ccc} f : & E & \rightarrow & F \\ & x & \mapsto & f(x) \end{array}$$

une telle application (ou fonction). On utilisera aussi les notation $f : E \rightarrow F$ ou $f : x \mapsto f(x)$.

Nous ne faisons pas la distinction ici entre fonction et application. Usuellement, on distingue ces notions en disant qu'une fonction de E dans F toute partie Γ du produit cartésien $E \times F$ telle que pour tout élément x de E , il existe au plus un élément y de F tel que $(x, y) \in \Gamma$. Le sous-ensemble D de E pour lequel il existe un unique élément y de F tel que $(x, y) \in \Gamma$ est appelé l'ensemble de définition de la fonction. Une application est donc une fonction pour laquelle tout élément de l'ensemble de départ E a une image dans F .

On note $\mathcal{F}(E, F)$ ou F^E l'ensemble de toutes les applications de E dans F (la deuxième notation sera justifiée plus loin).

L'application qui associe à tout x d'un ensemble E le même x est l'application identique notée Id_E , où Id si l'ensemble E est fixé.

Si f est une fonction de E dans F et D un sous-ensemble non vide de E , on définit une application g de D dans F en posant :

$$\forall x \in D, g(x) = f(x)$$

et on dit que g est la restriction de f à D , ce qui se note $g = f|_D$.

Définition 1.4. Soit f une application de E dans F . Pour toute partie A de E , l'image de A par f est le sous ensemble de F noté $f(A)$ et défini par $f(A) = \{f(x) \mid x \in A\}$. Pour toute partie B de F , l'image réciproque de B par f est le sous ensemble de E noté $f^{-1}(B)$ et défini par $f^{-1}(B) = \{x \in E \mid f(x) \in B\}$.

On a donc, pour tout $y \in F$:

$$(y \in f(A)) \Leftrightarrow (\exists x \in A \mid y = f(x))$$

et pour tout $x \in E$:

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

L'ensemble $f(E)$ est appelé l'image de f .

Exemple 1.1 On a $f(\emptyset) = \emptyset$, $f(\{x\}) = \{f(x)\}$ pour tout $x \in E$, $f^{-1}(\emptyset) = \emptyset$ et $f^{-1}(F) = E$.

Pour tout $y \in F$, $f^{-1}\{y\}$ est l'ensemble des $x \in E$ tels que $f(x) = y$ et cet ensemble peut être vide ou formé de un ou plusieurs éléments. En fait $f^{-1}\{y\}$ est l'ensemble des solutions dans E de l'équation $f(x) = y$, où y est donné dans F et x l'inconnue dans E . Cette équation peut avoir 0 ou plusieurs solutions.

On vérifie facilement le résultat suivant.

Théorème 1.5.

Soit f une application de E dans F . Pour toutes parties A, B de E et C, D de F , on a :

1. $A \subset B \Rightarrow f(A) \subset f(B)$
2. $f(A \cup B) = f(A) \cup f(B)$
3. $f(A \cap B) \subset f(A) \cap f(B)$
4. $C \subset D \Rightarrow f^{-1}(C) \subset f^{-1}(D)$
5. $f^{-1}(C \cup D) = f^{-1}(C) \cup f^{-1}(D)$
6. $f^{-1}(C \cap D) = f^{-1}(C) \cap f^{-1}(D)$
7. $f^{-1}(\overline{C}) = \overline{f^{-1}(C)}$

Preuve. Vérification immédiate. Par exemple, pour le point **2**, on peut écrire que y est dans $f(A \cup B)$ si, et seulement si, il existe x dans $A \cup B$ tel que $y = f(x)$, ce qui implique que $y \in f(A)$ dans le cas où $x \in A$ ou $y \in f(B)$ dans le cas où $x \in B$, soit $y \in f(A) \cup f(B)$ dans tous les cas. Réciproquement si $y \in f(A) \cup f(B)$, il est dans $f(A)$ ou $f(B)$ et s'écrit donc $y = f(x)$ avec x dans A ou B , ce qui signifie que $y \in f(A \cup B)$. On a donc les inclusions $f(A \cup B) \subset f(A) \cup f(B)$ et $f(A) \cup f(B) \subset f(A \cup B)$, c'est-à-dire l'égalité souhaitée.

Pour le point **3**, on a seulement une inclusion. Dire que $y \in f(A \cap B)$ équivaut à dire qu'il existe $x \in A \cap B$ tel que $y = f(x)$ et $y \in f(A) \cap f(B)$. Réciproquement, si $y \in f(A) \cap f(B)$, il existe $x_1 \in A$ et $x_2 \in B$ tels que $y = f(x_1) = f(x_2)$ et, a priori, il n'y a aucune raison pour que $x_1 = x_2$. $\square$

On dispose d'une opération importante sur les fonctions, c'est la composition des fonctions qui permet de construire de nouvelles fonctions à partir de fonctions données.

Définition 1.5. Soient f une application de E dans F et g une application de F dans G . La composée de f par g est la fonction de E dans G notée $g \circ f$ et définie par :

$$\forall x \in E, g \circ f(x) = g(f(x))$$

Ce qui peut se schématiser par :

$$\begin{array}{ccccc} E & \xrightarrow{f} & F & \xrightarrow{g} & G \\ x & \mapsto & f(x) & \mapsto & g(f(x)) \end{array}$$

On remarquera que $f \circ g$ n'est pas définie *a priori* (dans la situation de la définition).

Dans le cas où f est définie de E dans F et g de F dans E , on peut définir les applications $f \circ g$ (de F dans F) et $g \circ f$ (de E dans E) et il n'y a aucune raison pour que ces applications soient égales, même dans le cas où $F = E$.

Dans le cas où $E = F$, on dit que les applications f et g (définies de E dans E) commutent, si $f \circ g = g \circ f$.

On vérifie facilement que la loi de composition est associative, c'est-à-dire que $f \circ (g \circ h) = (f \circ g) \circ h$, quand toutes ces composées ont un sens. Cette propriété d'associativité permet de définir la composée de n applications $f_1 \circ f_2 \circ \dots \circ f_n$ sans se soucier de parenthèses.

Si f est une application de E dans E , on peut définir la suite de ses itérées par la relation de récurrence suivante :

$$\begin{cases} f^1 = f \\ \forall n \in \mathbb{N}^*, f^{n+1} = f^n \circ f \end{cases}$$

On convient que $f^0 = Id_E$. On vérifie facilement que $f^p \circ f^q = f^q \circ f^p = f^{p+q}$ pour tous entiers naturels p, q .

Les notions suivantes d'injectivité et de surjectivité sont aussi très importantes.

Définition 1.6. Soient E, F deux ensembles et f une application de E dans F . On dit que f est :

1. *injective (ou que c'est une injection)* si deux éléments distincts de E ont deux images distinctes dans F , soit :

$$x_1 \neq x_2 \text{ dans } E \Rightarrow f(x_1) \neq f(x_2) \text{ dans } F \quad (1.3)$$

2. *surjective (ou que c'est une surjection)* si tout élément de F a au moins un antécédent dans E , soit :

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

3. *bijective (ou que c'est une bijection)* si elle est à la fois injective ou surjective.

Une injection peut aussi se caractériser en disant que tout élément de F a au plus un antécédent par f , encore équivalent à dire que pour tout $y \in F$ l'équation $y = f(x)$ a au plus une solution x dans E , ce qui revient à dire que si x_1 et x_2 sont deux éléments de E tels que $f(x_1) = f(x_2)$, alors $x_1 = x_2$ (contraposée de (1.3)).

Une surjection peut se caractériser en disant que pour tout $y \in F$ l'équation $y = f(x)$ a au moins une solution x dans E , encore équivalent à dire que $f(E) = F$.

Si f est une surjection de E dans F , on dit parfois que f est une surjection de E sur (pour surjection) F .

Une bijection peut se caractériser en disant que tout élément de F a un unique antécédent par f , encore équivalent à dire que pour tout $y \in F$ l'équation $y = f(x)$ a une et une seule solution x dans E , ce qui permet de définir l'application réciproque de f , notée f^{-1} , de F dans E par :

$$(y \in F \text{ et } x = f^{-1}(y)) \Leftrightarrow (x \in E \text{ et } y = f(x))$$

Cette application f^{-1} est une bijection de F dans E . L'application $f \circ f^{-1}$ est alors l'application identité $y \mapsto y$ de F dans F et l'application $f^{-1} \circ f$ est l'application identité $x \mapsto x$ de E dans E , ce qui se note $f \circ f^{-1} = Id_F$ et $f^{-1} \circ f = Id_E$.

Définition 1.7. On appelle permutation d'un ensemble E toute bijection de E dans lui même.

On note en général $\mathfrak{S}(E)$ l'ensemble des permutations de E .

Exemple 1.2 L'application $x \mapsto x^2$ est surjective de $\mathbb{R}$ dans $\mathbb{R}^+$, mais non injective. Elle est bijective de $\mathbb{R}^+$ dans $\mathbb{R}^+$.

Dans le cas où f est une application de E dans F , on a noté pour toute partie B de F , $f^{-1}(B)$ l'image réciproque de B par f , sans aucune hypothèse de bijectivité pour f . Dans le cas où f est bijective, $f^{-1}(B)$ est aussi l'image directe de B par f^{-1} , mais dans le cas général, il faut bien prendre garde, malgré la notation, que f n'a aucune raison d'être bijective. Il faudrait en réalité utiliser un autre symbole que f^{-1} (par exemple $f^*(B)$, $f^{(-1)}(B)$, ou $f^{\zeta\zeta\boxtimes}(f)$), mais nous préférons utiliser la notation $f^{-1}(B)$ rencontrée le plus souvent. Si l'on sait de quoi l'on parle il n'y a pas de véritable problème, il s'agit seulement d'une notation.

On vérifie facilement le résultat suivant.

Théorème 1.6.

Soient E, F, G des ensembles, f une application de E dans F et g une application de F dans G .

1. Si f et g sont injectives, alors $g \circ f$ est injective (la composée de deux injections est une injection).
2. Si f et g sont surjectives, alors $g \circ f$ est surjective (la composée de deux surjections est une surjection).
3. Si f et g sont bijectives, alors $g \circ f$ est bijective (la composée de deux injections est une bijection) et $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Preuve.

1. Supposons f et g injectives. Si $g \circ f(x_1) = g \circ f(x_2)$, alors $g(f(x_1)) = g(f(x_2))$, donc $f(x_1) = f(x_2)$ puisque g est injective et $x_1 = x_2$ puisque f est injective.
2. Supposons f et g surjectives. Pour tout $z \in G$, il existe $y \in F$ tel que $z = g(y)$ puisque g est surjective et $y \in F$ s'écrit $y = f(x)$ avec $x \in E$ puisque f est surjective. On a donc $z = g \circ f(x)$ avec $x \in E$. L'application $g \circ f$ est donc surjective. De manière plus compacte, on peut écrire que :

$$(g \circ f)(E) = g(f(E)) = g(F) = G.$$

3. Les deux premiers points nous disent que $g \circ f$ est bijective si f et g le sont. Puis avec $(f^{-1} \circ g^{-1}) \circ g \circ f = f^{-1} \circ Id_F \circ f = f^{-1} \circ f = Id_E$, on déduit que $f^{-1} \circ g^{-1}$ est l'inverse de $g \circ f$.

□

Le résultat qui suit peut parfois être utile pour montrer l'injectivité, la surjectivité ou la bijectivité d'une application.

Théorème 1.7.

Soient E, F deux ensembles et f une application de E dans F .

1. *S'il existe une application g de F dans E telle que $g \circ f = Id_E$, alors f est injective.*
2. *S'il existe une application h de F dans E telle que $f \circ h = Id_F$, alors f est surjective.*
3. *S'il existe deux applications g et h de F dans E telles que $g \circ f = Id_E$ et $f \circ h = Id_F$, alors f est bijective et $g = h = f^{-1}$.*

Preuve.

1. Si x, x' dans E sont tels que $f(x) = f(x')$, alors $x = g \circ f(x) = g \circ f(x') = x'$ et f est injective.
2. Pour tout $y \in F$, on a $y = (f \circ h)(y) = f(h(y))$ avec $x = h(y) \in E$, donc f est surjective.
3. Les deux premiers points nous disent que f est bijective et de $g \circ f = Id_E$, on déduit que $f^{-1} = (g \circ f) \circ f^{-1} = g$. De même $h = g^{-1}$.

□

1.9 Relations d'ordre et d'équivalence

On se donne un ensemble non vide E et une propriété $\mathcal{P}$ sur les éléments (x, y) de E^2 . On dit que x et y dans E sont en relation, ce que l'on notera $x\mathcal{R}y$, si le couple (x, y) vérifie la propriété $\mathcal{P}$. On dit que $\mathcal{R}$ est une relation sur E .

Définition 1.8. On dit que $\mathcal{R}$ est une relation d'ordre sur E , si elle vérifie les propriétés suivantes :

- *réflexivité* : pour tout $x \in E$, on a $x\mathcal{R}x$;
- *anti-symétrie* : si x, y dans E sont tels que $x\mathcal{R}y$ et $y\mathcal{R}x$, on a alors $x = y$;
- *transitivité* : si x, y, z dans E sont tels que $x\mathcal{R}y$ et $y\mathcal{R}z$, on a alors $x\mathcal{R}z$.

Définition 1.9. Un ensemble E muni d'une relation d'ordre est dit ordonné.

Définition 1.10. Si $\mathcal{R}$ est une relation d'ordre sur E telle que deux éléments quelconques de E sont comparables, c'est-à-dire que pour tous x, y

dans E , on a $x\mathcal{R}y$ ou $y\mathcal{R}x$, on dit alors que l'ensemble E est totalement ordonné ou que $\mathcal{R}$ est une relation d'ordre total sur E .

Exemples 1.1

1. Les ensembles $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ et $\mathbb{R}$ sont totalement ordonnés par la relation $\leq$.
2. La relation de divisibilité est une relation d'ordre total sur $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$, mais ce n'est pas une relation d'ordre sur $\mathbb{Z}^* = \mathbb{Z} \setminus \{0\}$.
3. La relation d'inclusion $\subset$ est une relation d'ordre sur l'ensemble $\mathcal{P}(E)$ des parties de E .

Définition 1.11. On dit que $\mathcal{R}$ est une relation d'équivalence sur E , si elle vérifie les propriétés suivantes :

- réflexivité : pour tout $x \in E$, on a $x\mathcal{R}x$;
- symétrie : si x, y dans E sont tels que $x\mathcal{R}y$, on a alors $y\mathcal{R}x$;
- transitivité : si x, y, z dans E sont tels que $x\mathcal{R}y$ et $y\mathcal{R}z$, on a alors $x\mathcal{R}z$.

Exemple 1.3 Soient E, F deux ensembles et φ une application de E dans F . La relation $\mathcal{R}$ définie sur E par $x\mathcal{R}y$ si, et seulement si, on a $\varphi(x) = \varphi(y)$ est une relation d'équivalence. Pour $F = E$ et $\varphi = \text{Id}$, on a la relation d'égalité.

Si $\mathcal{R}$ est une relation d'équivalence sur E , on note pour tout $x \in E$:

$$\bar{x} = \{y \in E \mid x\mathcal{R}y\}$$

un tel ensemble est non vide car il contient x (propriété de symétrie) et on dit que c'est la classe d'équivalence de x . L'ensemble de toutes ces classes d'équivalence est appelé l'ensemble quotient de E par $\mathcal{R}$ et noté $E/\mathcal{R}$.

Il est facile de vérifier que l'application :

$$\begin{aligned} \pi : E &\rightarrow E/\mathcal{R} \\ x &\mapsto \bar{x} \end{aligned}$$

est surjective. On dit que c'est la surjection canonique de E sur $E/\mathcal{R}$.

Théorème 1.8.

Si $\mathcal{R}$ est une relation d'équivalence sur E , ses classes d'équivalence forment alors une partition de E , c'est-à-dire que $E = \bigcup_{C \in E/\mathcal{R}} C$.

Preuve. Toute classe d'équivalence est dans E , donc $\bigcup_{C \in E/\mathcal{R}} C \subset E$ et tout $x \in E$ appartient à $C = \bar{x}$, donc $E \subset \bigcup_{C \in E/\mathcal{R}} C$ et on a l'égalité. Soient C et C' deux classes

d'équivalence distinctes. Si $C \cap C' \neq \emptyset$, il existe alors $x \in C \cap C'$ avec $C = \bar{y}$ et $C' = \bar{z}$ pour deux éléments y et z de E , ce qui implique que $x\mathcal{R}y$ et $x\mathcal{R}z$ et donc que $y\mathcal{R}z$ par symétrie et transitivité, mais cela implique que $C = C'$, ce qui n'est pas. En conclusion, on a la partition $E = \bigcup_{C \in E/\mathcal{R}} C$. $\square$

1.10 Exercices

Exercice 1.1. Montrer que les assertions $P \rightarrow Q$ et $\bar{P} \vee Q$ sont équivalentes, ainsi que les assertions $\bar{P} \rightarrow \bar{Q}$ et $P \wedge \bar{Q}$.

Solution. Ces deux assertions ont même table de vérité.

P	Q	$\bar{P}$	$\bar{P} \vee Q$	$P \rightarrow Q$	P	Q	$P \wedge \bar{Q}$	$\bar{P} \rightarrow \bar{Q}$
V	V	F	V	V	V	V	F	F
V	F	F	F	F	V	F	V	V
F	V	V	V	V	F	V	F	F
F	F	V	V	V	F	F	F	F

Exercice 1.2. Montrer que les assertions $P \leftrightarrow P$, $(P \wedge Q) \rightarrow P$, $P \rightarrow (P \vee Q)$, $P \vee (P \rightarrow Q)$, $P \rightarrow (Q \rightarrow P)$ et $((P \rightarrow Q) \rightarrow P) \rightarrow P$ sont des tautologies (i. e. toujours vraies).

Solution. Pour $P \leftrightarrow P$, $(P \wedge Q) \rightarrow P$, $P \rightarrow (P \vee Q)$, c'est évident et pour les autres, on utilise la table de vérité :

P	Q	$P \rightarrow Q$	$Q \rightarrow P$	$P \vee (P \rightarrow Q)$	$P \rightarrow (Q \rightarrow P)$	$((P \rightarrow Q) \rightarrow P) \rightarrow P$
V	V	V	V	V	V	V
V	F	F	V	V	V	V
F	V	V	F	V	V	V
F	F	V	V	V	V	V

Exercice 1.3. Simplifier l'expression $R = (\bar{P} \wedge Q) \vee (\bar{P} \wedge \bar{Q}) \vee (P \wedge Q)$.

Solution. En utilisant les tables de vérité, on a :

P	Q	$\bar{P} \wedge Q$	$\bar{P} \wedge \bar{Q}$	$(\bar{P} \wedge Q) \vee (\bar{P} \wedge \bar{Q})$	$P \wedge Q$	R
V	V	F	F	F	V	V
V	F	F	F	F	F	F
F	V	V	F	V	F	V
F	F	F	V	V	F	V

Donc R a la même table de vérité que $P \rightarrow Q$, ce qui signifie que R est équivalent à $P \rightarrow Q$.

Exercice 1.4. On dit qu'une théorie est non contradictoire si $P \wedge \overline{P}$ est faux pour toute proposition P . Montrer que si dans une théorie une propriété P est contradictoire, c'est-à-dire si $P \wedge \overline{P}$ est vraie, alors $Q \wedge \overline{Q}$ est vraie pour toute propriété Q .

Solution. Nous allons montrer que s'il existe un énoncé contradictoire P , alors tout énoncé Q est vrai, donc $\overline{Q}$ aussi et $Q \wedge \overline{Q}$ est vraie. On vérifie tout d'abord que $R = \overline{P} \rightarrow (P \rightarrow Q)$ est une tautologie avec la table de vérité :

P	Q	$\overline{P}$	$P \rightarrow Q$	$\overline{P} \rightarrow (P \rightarrow Q)$
V	V	F	V	V
V	F	F	F	V
F	V	V	V	V
F	F	V	V	V

Comme R et $\overline{P}$ sont vraies, $P \rightarrow Q$ est vraie et Q est vraie puisque P est vraie.

Exercice 1.5. En raisonnant par l'absurde, montrer que $\sqrt{2}$ et $\frac{\ln(2)}{\ln(3)}$ sont irrationnels.

Solution. Supposons que $\sqrt{2} = \frac{p}{q}$ avec p, q entiers naturels non nuls premiers entre eux. On a alors $p^2 = 2q^2$ qui entraîne que p est pair, soit $p = 2p'$ et $q^2 = 2p'^2$ entraîne q pair, ce qui contredit p et q premiers entre eux. Supposons que l'on ait $\frac{\ln(2)}{\ln(3)} = \frac{p}{q}$ avec p, q entiers naturels non nuls premiers entre eux. On a alors $\ln(2^q) = \ln(3^p)$ et $2^p = 3^q$, ce qui est impossible puisque 2^p est un entier pair et 3^q est un entier impair.

Exercice 1.6. Soit n un entier naturel non carré, c'est-à-dire ne s'écrivant pas sous la forme $n = p^2$ avec p entier. En raisonnant par l'absurde et en utilisant le théorème de Bézout, montrer que $\sqrt{n}$ est irrationnel.

Solution. Si n est non carré, on a alors $n \geq 2$. Supposons que $\sqrt{n} = \frac{p}{q}$ avec p, q premiers entre eux dans $\mathbb{N}^*$. Le théorème de Bézout nous dit qu'il existe un couple (u, v) d'entiers relatifs tels que $up + vq = 1$. On a alors :

$$1 = (up + vq)^2 = u^2p^2 + 2uvpq + v^2q^2$$

avec $u^2p^2 = u^2nq^2$. L'égalité précédente s'écrit alors $qr = 1$ avec $r = u^2nq + 2uvp + v^2q$ dans $\mathbb{Z}$, ce qui implique que $q = 1$ et $\sqrt{n} = p$, en contradiction avec n non carré.

Exercice 1.7. Montrer que $P_n = \prod_{k=1}^n \left(1 + \frac{1}{k}\right)^k = \frac{(n+1)^n}{n!}$ pour tout $n \in \mathbb{N}^*$.

Solution. Il revient au même de calculer $S_n = \ln(P_n)$. On a :

$$\begin{aligned} S_n &= \ln \left(\prod_{k=1}^n \left(\frac{k+1}{k} \right)^k \right) = \sum_{k=1}^n (k \ln(k+1) - k \ln(k)) \\ &= \sum_{k=1}^n k \ln(k+1) - \sum_{k=1}^n k \ln(k) \end{aligned}$$

et le changement d'indice $j = k + 1$ dans la première somme donne :

$$\begin{aligned} S_n &= \sum_{j=2}^{n+1} (j-1) \ln(j) - \sum_{k=1}^n k \ln(k) = \sum_{j=2}^{n+1} j \ln(j) - \sum_{j=2}^{n+1} \ln(j) - \sum_{k=1}^n k \ln(k) \\ &= \sum_{k=2}^{n+1} k \ln(k) - \sum_{k=2}^{n+1} \ln(k) - \sum_{k=1}^n k \ln(k) = (n+1) \ln(n+1) - \sum_{k=2}^{n+1} \ln(k) \end{aligned}$$

(on a utilisé le fait que l'indice est muet dans une somme). On a donc en définitive :

$$\begin{aligned} S_n &= \ln(P_n) = \ln \left((n+1)^{n+1} \right) - \sum_{k=2}^{n+1} \ln(k) \\ &= \ln \left((n+1)^{n+1} \right) - \ln \left(\prod_{k=2}^n k \right) = \ln \left((n+1)^{n+1} \right) - \ln(n!) = \ln \left(\frac{(n+1)^{n+1}}{n!} \right) \end{aligned}$$

et $P_n = \frac{(n+1)^n}{n!}$.

Exercice 1.8. Montrer que si φ est une fonction strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$, on a alors $\varphi(n) \geq n$ pour tout n .

Solution. Comme φ est une fonction de $\mathbb{N}$ dans $\mathbb{N}$, $\varphi(0)$ est un entier naturel et donc $\varphi(0) \geq 0$. Supposant le résultat acquis pour $n \geq 0$, sachant que φ est strictement croissante, on a $\varphi(n+1) > \varphi(n) \geq n$, donc $\varphi(n+1) > n$, ce qui équivaut à $\varphi(n+1) \geq n+1$ puisque $\varphi(n+1)$ est un entier.

Exercice 1.9. Montrer par récurrence que pour tout entier naturel non nul n , on a :

$$U_n = \sum_{k=1}^n k = \frac{n(n+1)}{2}, \quad V_n = \sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6},$$

$$W_n = \sum_{k=1}^n k^3 = \left(\frac{n(n+1)}{2} \right)^2 = U_n^2$$

Solution. Pour $n = 1$ c'est clair. En supposant les résultats acquis pour $n \geq 1$, on a :

$$U_{n+1} = \frac{n(n+1)}{2} + n + 1 = \frac{n^2 + 3n + 2}{2} = \frac{(n+1)(n+2)}{2}$$

$$V_{n+1} = \frac{n(n+1)(2n+1)}{6} + (n+1)^2 = \frac{(n+1)(2n^2 + 7n + 6)}{6}$$

$$= \frac{(n+1)(n+2)(2n+3)}{6}$$

$$W_{n+1} = \left(\frac{n(n+1)}{2} \right)^2 + (n+1)^3 = \left(\frac{(n+1)(n+2)}{2} \right)^2$$

On a aussi $U_n = 1 + 2 + \dots + (n-1) + n = n + (n-1) + \dots + 2 + 1$ et en additionnant terme à terme on obtient $2U_n = n(n+1)$. Le calcul de U_n peut aussi se faire en passant par V_{n+1} et en utilisant l'identité $(k+1)^2 = k^2 + 2k + 1$. Précisément, en effectuant le changement d'indice $k = j + 1$, on a :

$$V_{n+1} = \sum_{k=1}^{n+1} k^2 = \sum_{j=0}^n (j+1)^2 = \sum_{j=0}^n j^2 + 2 \sum_{j=0}^n j + \sum_{j=0}^n 1$$

soit $V_{n+1} = V_n + 2U_n + n + 1$ et :

$$2U_n = V_{n+1} - V_n - (n+1) = (n+1)^2 - (n+1) = n(n+1)$$

ce qui donne bien $U_n = \frac{n(n+1)}{2}$. De même, le calcul de V_n peut aussi se faire en passant par W_{n+1} et en utilisant l'identité $(k+1)^3 = k^3 + 3k^2 + 3k + 1$. Précisément, en effectuant le changement d'indice $k = j + 1$, on a :

$$W_{n+1} = \sum_{k=1}^{n+1} k^3 = \sum_{j=0}^n (j+1)^3 = \sum_{j=0}^n j^3 + 3 \sum_{j=0}^n j^2 + 3 \sum_{j=0}^n j + \sum_{j=0}^n 1$$

soit $W_{n+1} = W_n + 3V_n + 3U_n + n + 1$ et :

$$3V_n = W_{n+1} - W_n - 3U_n - (n+1) = (n+1)^3 - 3 \frac{n(n+1)}{2} - (n+1)$$

$$= \frac{n(n+1)(2n+1)}{2}$$

ce qui donne bien $V_n = \frac{n(n+1)(2n+1)}{6}$.

Exercice 1.10. Montrer que pour tout entier naturel n et tous nombres complexes a et b on a $b^{n+1} - a^{n+1} = (b - a) \sum_{k=0}^n a^k b^{n-k}$. En particulier, pour $a \neq 1$ et $b = 1$, on a $\sum_{k=0}^n a^k = \frac{a^{n+1} - 1}{a - 1}$.

Solution. Pour $n = 0$, c'est évident. En supposant le résultat acquis au rang $n \geq 0$, on a :

$$\begin{aligned} b^{n+2} - a^{n+2} &= (b^{n+1} - a^{n+1})b + ba^{n+1} - a^{n+2} \\ &= (b - a) \sum_{k=0}^n a^k b^{n+1-k} + (b - a)a^{n+1} \\ &= (b - a)(b^{n+1} + ab^n + \dots + a^{n-1}b^2 + a^n b) + (b - a)a^{n+1} \\ &= (b - a) \sum_{k=0}^{n+1} a^k b^{n+1-k} \end{aligned}$$

Le résultat est donc vrai pour tout $n \geq 0$.

Exercice 1.11. Montrer que pour tout entier naturel n et tous nombres complexes a et b on a $(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k$, où $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ pour k compris entre 0 et n avec la convention $0! = 1$ (formule du binôme de Newton).

Solution. Pour $n = 0$ et $n = 1$, c'est évident. En supposant le résultat acquis au rang $n \geq 1$, on a :

$$\begin{aligned} (a + b)^{n+1} &= (a + b)^n (a + b) = \left(\sum_{k=0}^n \binom{n}{k} a^{n-k} b^k \right) (a + b) \\ &= \sum_{k=0}^n \binom{n}{k} a^{n-(k-1)} b^k + \sum_{k=0}^n \binom{n}{k} a^{n-k} b^{k+1} \\ &= \sum_{k=0}^n \binom{n}{k} a^{n-(k-1)} b^k + \sum_{k=1}^{n+1} \binom{n}{k-1} a^{n-(k-1)} b^k \\ &= a^{n+1} + \sum_{k=1}^n \left(\binom{n}{k} + \binom{n}{k-1} \right) a^{n+1-k} b^k + b^{n+1} \end{aligned}$$

et tenant compte de $\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}$ (triangle de Pascal), cela s'écrit :

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

Le résultat est donc vrai pour tout $n \geq 0$.

Exercice 1.12. Montrer par récurrence, que pour tout entier naturel non nul n et tout nombre complexe λ différent de 1, on a :

$$\sum_{k=1}^n k\lambda^k = n \frac{\lambda^{n+1}}{\lambda-1} + \lambda \frac{1-\lambda^n}{(\lambda-1)^2}$$

Solution. Pour $n = 1$, c'est clair. Si c'est vrai pour $n \geq 1$, alors :

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

Exercice 1.13. Soit $x_1, \dots, x_n$ des réels dans $[0, 1]$. Montrer par récurrence que $\prod_{k=1}^n (1 - x_k) \geq 1 - \sum_{k=1}^n x_k$.

Solution. Notons $u_n = \prod_{k=1}^n (1 - x_k)$ et $v_n = 1 - \sum_{k=1}^n x_k$. Pour $n = 1$, on a $u_1 = v_1$. Supposant le résultat acquis au rang $n \geq 1$ et tenant compte de $1 - x_{n+1} \geq 0$, on a :

$$\begin{aligned} u_{n+1} &= u_n (1 - x_{n+1}) \geq \left(1 - \sum_{k=1}^n x_k\right) (1 - x_{n+1}) \\ &\geq 1 - \sum_{k=1}^n x_k - x_{n+1} + x_{n+1} \sum_{k=1}^n x_k \geq 1 - \sum_{k=1}^{n+1} x_k = v_{n+1} \end{aligned}$$

puisque tous les x_k sont positifs.

Exercice 1.14. Pour tout entier naturel n supérieur ou égal à 2, on note

$$H_n = \sum_{k=1}^n \frac{1}{k}.$$

1. Soit p un entier naturel non nul. Montrer que $H_{2p} = \frac{1}{2}H_p + \frac{a}{2b+1}$ où a, b sont des entiers naturels avec a non nul.

2. Montrer par récurrence que pour tout entier naturel non nul H_n est le quotient d'un entier impair par un entier pair et qu'en conséquence ce n'est pas un entier.

Solution.

1. On a $H_{2p} = \sum_{k=1}^p \frac{1}{2k} + \sum_{k=0}^{p-1} \frac{1}{2k+1} = \frac{1}{2}H_p + \frac{N}{D}$ avec $D = \text{ppcm}(1, 3, \dots, 2p-1)$ qui est impair et N entier naturel non nul.

2. On a $H_2 = \frac{3}{2} \notin \mathbb{N}$. Supposons le résultat acquis au rang $n \geq 2$. Si $n = 2p$, on a alors :

$$H_{n+1} = H_n + \frac{1}{2p+1} = \frac{2a+1}{2b} + \frac{1}{2p+1} = \frac{(2a+1)(2p+1)+2b}{2b(2p+1)} = \frac{2a'+1}{2b'}$$

avec $a' = a + b + p + 2ap$ et $b' = b(2p+1)$. Si $n = 2p+1$, on a alors :

$$\begin{aligned} H_{n+1} &= H_{2(p+1)} = \frac{c}{2d+1} + \frac{1}{2}H_{p+1} \\ &= \frac{c}{2d+1} + \frac{1}{2} \frac{2a+1}{2b} = \frac{4bc + (2d+1)(2a+1)}{4b(2d+1)} = \frac{2a'+1}{2b'} \end{aligned}$$

avec $a' = a + d + 2ad + 2bc$ et $b' = 2b(2d+1)$. Dans tous les cas, H_n est le quotient d'un entier impair par un entier pair et en conséquence, ce n'est pas un entier.

Exercice 1.15. Simplifier les expressions suivantes, où A et B sont des sous-ensembles d'un ensemble E :

1. $C = (\overline{A} \cap B) \cup (\overline{A} \cap \overline{B}) \cup (A \cap B), \overline{C}.$

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

Solution.

1. Avec la distributivité de $\cap$ sur $\cup$, on a :

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

(on a mis $\overline{A}$ en facteur) et avec la distributivité de $\cup$ sur $\cap$, on a :

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

$$\overline{C} = A \cap \overline{B}.$$

2. En posant :

$$X = A \cap \overline{B}, Y = \overline{X} \cap (\overline{A} \cap B), Z = \overline{Y} \cup (A \cap B), T = \overline{Z} \cap (A \cap B)$$

on a :

$$D = \overline{T} = Z \cup (\overline{A} \cup \overline{B}) = \overline{Y} \cup (A \cap B) \cup (\overline{A} \cup \overline{B}) = X \cup (A \cup \overline{B}) \cup (A \cap B) \cup (\overline{A} \cup \overline{B})$$

avec $(A \cup \overline{B}) \cup (\overline{A} \cup \overline{B}) = E$, donc $D = E$.

Exercice 1.16. Soient $A_1, \dots, A_p$ des ensembles deux à deux distincts. Montrer que l'un de ces ensembles ne contient aucun des autres.

Solution. On raisonne par l'absurde, c'est-à-dire qu'on suppose que chacun des ensembles A_k contient un ensemble A_j différent de A_k . Donc A_1 contient un ensemble $A_{j_1} \neq A_1$, soit $A_{j_1} \subsetneq A_1$, A_{j_1} contient un ensemble $A_{j_2} \neq A_{j_1}$, soit $A_{j_2} \subsetneq A_{j_1}$, et on peut continuer indéfiniment, ce qui est impossible puisque la famille d'ensembles est finie.

Exercice 1.17. Que dire de 2 ensembles A, B tels que $A \cap B = A \cup B$?

Solution. On a toujours $A \cap B \subset A \cup B$. Si de plus $A \cup B \subset A \cap B$, on a alors :

$$A \subset A \cup B \subset A \cap B \subset B \text{ et } B \subset A \cup B \subset A \cap B \subset A$$

ce qui donne $A = B$.

Exercice 1.18. Soient A, B, C trois ensembles. Montrer que :

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

Solution. On a $(A \cup B) \cap (B \cup C) = B \cup (A \cap C)$ et en notant $D = (A \cup B) \cap (B \cup C) \cap (C \cup A)$, on a :

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

Ou alors on part de $x \in D$ et on montre que $x \in E = (A \cap B) \cup (B \cap C) \cup (C \cap A)$, puis partant de $x \in E$, on montre que $x \in D$.

Exercice 1.19. Soient E un ensemble et une application $f : \mathcal{P}(E) \rightarrow \mathbb{R}$ telle que pour toutes parties disjointes de E on ait :

$$f(A \cup B) = f(A) + f(B)$$

Montrer que $f(\emptyset) = 0$, puis que pour toutes parties A, B de E , on a :

$$f(A \cup B) + f(A \cap B) = f(A) + f(B)$$

Pour E fini, la fonction f qui associe à une partie A de E son cardinal (c'est-à-dire le nombre de ses éléments) vérifie l'équation fonctionnelle de cet exercice.

Solution. On a $f(\emptyset) = f(\emptyset \cup \emptyset) = f(\emptyset) + f(\emptyset)$ dans $\mathbb{R}$, donc $f(\emptyset) = 0$. Avec les partitions $A \cup B = A \cup (B \setminus A)$ et $B = (A \cap B) \cup (B \setminus A)$, on a :

$$\begin{cases} f(A \cup B) = f(A) + f(B \setminus A) \\ f(B) = f(A \cap B) + f(B \setminus A) \end{cases}$$

et par soustraction $f(A \cup B) - f(B) = f(A) - f(A \cap B)$, ce qui donne le résultat.

Exercice 1.20. Montrer qu'une application f strictement monotone de $\mathbb{R}$ dans $\mathbb{R}$ est injective.

Solution. Supposons que f soit strictement croissante (au besoin on remplace f par $-f$). Si $x \neq y$, on a nécessairement $x > y$ ou $y > x$ et donc $f(x) > f(y)$ ou $f(x) < f(y)$, soit $f(x) \neq f(y)$ dans tous les cas.

Exercice 1.21. Soit m un entier naturel. Montrer que s'il existe un entier naturel n et une injection φ de $E_n = \{1, \dots, n\}$ dans $E_m = \{1, \dots, m\}$, on a alors nécessairement $n \leq m$.

Solution. On procède par récurrence sur $m \geq 0$. Si $m = 0$, on a alors $E_m = \emptyset$ et $E_n = \emptyset$ (en effet, si $E_n \neq \emptyset$, l'ensemble $f(E_n)$ est alors non vide et contenu dans l'ensemble vide, ce qui est impossible), donc $n = 0$. Supposons le résultat acquis pour $m \geq 0$. Soit φ une injection de E_n dans E_{m+1} . Si $n = 0$, on a bien $n \leq m + 1$. Si $n \geq 1$, on distingue alors deux cas de figure :

- soit $\varphi(n) = m + 1$ et dans ce cas φ induit une bijection de E_{n-1} dans E_m (la restriction de φ à E_{n-1}) et $n - 1 \leq m$, soit $n \leq m + 1$;
- soit $\varphi(n) \neq m + 1$ et dans ce cas, en désignant par ψ l'application de E_{m+1} dans lui même définie par $\psi(\varphi(n)) = m + 1$, $\psi(m + 1) = \varphi(n)$ et $\psi(k) = k$ pour $k \in E_{m+1} \setminus \{\varphi(n), m + 1\}$, l'application $\psi \circ \varphi$ est injective de E_n dans E_{m+1} (composée de deux injections puisque φ est injective et ψ bijective) avec $\psi \circ \varphi(n) = m + 1$, ce qui nous ramène au cas précédent.

Exercice 1.22. Soient n, m deux entiers naturels. Montrer que s'il existe une bijection φ de $E_n = \{1, \dots, n\}$ sur $E_m = \{1, \dots, m\}$, on a alors nécessairement $n = m$.

Solution. On a $n \leq m$ puisque φ est une injection de E_n dans E_m et $m \leq n$ puisque φ^{-1} est une injection de E_m dans E_n , ce qui donne $n = m$.

Exercice 1.23. Soient E, F deux ensembles et f une bijection de E sur F . Montrer que si g [resp. h] est une application de F sur E telle que $g \circ f = \text{Id}_E$ [resp. $f \circ h = \text{Id}_F$], alors g [resp. h] est bijective et $g = f^{-1}$ [resp. $h = f^{-1}$].

Solution. Résulte de $g = (g \circ f) \circ f^{-1} = Id_E \circ f^{-1} = f^{-1}$ et $h = f^{-1} \circ (f \circ h) = f^{-1} \circ Id_F = f^{-1}$.

Exercice 1.24. Soient E, F, G des ensembles, f une application de E dans F et g une application de F dans G . Montrer que :

1. si $g \circ f$ est injective, alors f est injective ;
2. si $g \circ f$ est surjective, alors g est surjective ;
3. si $g \circ f$ est surjective et g injective, alors f est surjective ;
4. Si $g \circ f$ est injective et f surjective, alors g est injective.

Solution.

1. Si x, x' dans E sont tels que $f(x) = f(x')$, alors $g \circ f(x) = g \circ f(x')$ et $x = x'$ puisque $g \circ f$ est injective. L'application f est donc injective.
2. Pour tout z dans G , il existe x dans E tel que $z = g \circ f(x)$ puisque $g \circ f$ est surjective et en notant $y = f(x)$, on a $y \in F$ et $z = g(y)$, ce qui prouve que g est surjective.
3. Soit $y \in F$. Comme $g \circ f$ est surjective, il existe $x \in E$ tel que $z = g(y) = (g \circ f)(x) = g(f(x))$ et $y = f(x)$ si on suppose de plus que g est injective. En conséquence, f est surjective.
4. Soient y, y' dans F tels que $g(y) = g(y')$. Comme f est surjective, il existe x, x' dans E tels que $y = f(x)$ et $y' = f(x')$, ce qui donne $g \circ f(x) = g \circ f(x')$ et $x = x'$ puisque $g \circ f$ est injective, donc $y = y'$.

Exercice 1.25. Soient m un entier naturel non nul et E un ensemble non vide. Montrer que s'il existe une surjection φ de $E_m = \{1, \dots, m\}$ sur E , on peut alors construire une injection de E dans E_m . Pour $E = E_n = \{1, \dots, n\}$, on a alors nécessairement $n \leq m$.

Solution. Comme φ est surjective de E_m sur E , on a $\varphi^{-1}\{x\} \neq \emptyset$ pour tout $x \in E$ et chacun de ces sous-ensembles de E_m a un plus petit élément $j_x = \min \varphi^{-1}\{x\} \in E_m$, ce qui permet de définir l'application ψ de E dans E_m par $\psi(x) = j_x$. On a alors $\varphi \circ \psi(x) = \varphi(j_x) = x$ pour tout $x \in E$, c'est-à-dire que $\varphi \circ \psi = Id_E$ et l'application ψ est injective.

Exercice 1.26. Soient E un ensemble et f une application de E dans E . Montrer que f est injective si, et seulement si, $f(A \cap B) = f(A) \cap f(B)$ pour toutes parties A et B de E .

Solution. On a toujours $f(A \cap B) \subset f(A) \cap f(B)$ pour toutes parties A et B de E , que f soit injective ou pas. En effet un élément y de $f(A \cap B)$ s'écrit $y = f(x)$ avec $x \in A \cap B$ et donc $y \in f(A) \cap f(B)$. Réciproquement si $y \in f(A) \cap f(B)$,

il existe $x \in A$ et $x' \in B$ tels que $y = f(x) = f(x')$ et dans le cas où f est injective, on a nécessairement $x = x' \in A \cap B$, donc $y \in f(A \cap B)$. On a donc $f(A \cap B) = f(A) \cap f(B)$ pour toutes parties A et B de E , si f est injective. Réciproquement supposons que $f(A \cap B) = f(A) \cap f(B)$ pour toutes parties A et B de E . Si f n'est pas injective, il existe $x \neq x'$ dans E tels que $f(x) = f(x')$ et :

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

ce qui est impossible. Donc f est injective.

Exercice 1.27. Soient E un ensemble et f une application de E dans E . Montrer que f est bijective si, et seulement si, $f(\overline{A}) = \overline{f(A)}$ pour toute partie A de E .

Solution. Supposons f bijective. Un élément y de E est dans $f(\overline{A})$ si, et seulement si, il s'écrit $y = f(x)$ où x est uniquement déterminé dans $\overline{A}$, ce qui implique $y \notin f(A)$ (sinon $y = f(x') = f(x)$ avec $x' \in A$ et $x = x' \in A$, ce qui contredit $x \in \overline{A}$). On a donc $f(\overline{A}) \subset \overline{f(A)}$. Si $y \notin f(A)$, il s'écrit $y = f(x)$ (f est bijective) et $x \notin A$, donc $y \in \overline{f(A)}$. On a donc $f(\overline{A}) \subset \overline{f(A)}$ et $f(\overline{A}) = \overline{f(A)}$. Supposons que $f(\overline{A}) = \overline{f(A)}$ pour toute partie A de E . En particulier, on a $f(E) = f(\overline{\emptyset}) = \overline{f(\emptyset)} = \overline{\emptyset} = E$ et f est surjective. Si $x \neq x'$ dans E , en remarquant que $x' \in \overline{\{x\}}$, on a $f(x') \in f(\overline{\{x\}}) = \overline{f(\{x\})}$ et $f(x) \neq f(x')$. Donc f est injective.

Exercice 1.28. Soient E, F, G, H des ensembles, f une application de E dans F , g une application de F dans G et h une application de G dans H . Montrer que si $g \circ f$ et $h \circ g$ sont bijectives, alors f, g et h sont bijectives.

Solution. Si $g \circ f$ est bijective, elle est alors surjective et il en est de même de g (exercice 1.24). Si $h \circ g$ est bijective, elle est alors injective et il en est de même de g (exercice 1.24). Donc g est bijective. Il en résulte que $f = g^{-1} \circ (g \circ f)$ et $h = (h \circ g) \circ g^{-1}$ sont bijectives comme composées.

Exercice 1.29. On désigne par f l'application définie sur $\mathbb{N}^2 = \mathbb{N} \times \mathbb{N}$ par $f(n, m) = 2^n 3^m$. Montrer que f est injective. Il résulte que $\mathbb{N}^2$ est en bijection avec le sous ensemble $f(\mathbb{N}^2)$ de $\mathbb{N}$. Ce résultat se traduit en disant que $\mathbb{N}^2$ est dénombrable.

Solution. L'égalité $f(n, m) = f(n', m')$ avec (n, m) et (n', m') dans $\mathbb{N}^2$ équivaut à $2^n 3^m = 2^{n'} 3^{m'}$ et l'unicité de la décomposition en facteurs premiers d'un entier naturel non nul nous dit que $(n, m) = (n', m')$. L'application f est donc injective de $\mathbb{N}^2$ dans $\mathbb{N}$ et bijective de $\mathbb{N}^2$ dans $f(\mathbb{N}^2) \subset \mathbb{N}$.

Exercice 1.30. Soient E un ensemble et A une partie de E . On définit une relation $\mathcal{R}$ sur $\mathcal{P}(E)$ par :

$$(X\mathcal{R}Y) \Leftrightarrow (X \cup A = Y \cup A)$$

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

Solution.

1. La relation est clairement réflexive, symétrique, transitive.
2. Soit $X \in \mathcal{P}(E)$. on a :

$$(Y \in \overline{X}) \Leftrightarrow (X \cup A = Y \cup A)$$

Pour tout $x \in Y \setminus A$, on a $x \in Y \cup A = X \cup A$ avec $x \notin A$, donc $x \in X \setminus A$. De même on montre que si $x \in X \setminus A$ alors $x \in Y \setminus A$. Ainsi $Y \setminus A = X \setminus A$. De plus $Y = (Y \setminus A) \cup (Y \cap A)$ donc $Y = (X \setminus A) \cup B$ avec $B \in \mathcal{P}(A)$. Inversement soit $Y = (X \setminus A) \cup B$ avec $B \in \mathcal{P}(A)$. On a :

$$Y \cup A = (X \setminus A) \cup A = X \cup A$$

Ainsi $\overline{X} = \{(X \setminus A) \cup B \mid B \in \mathcal{P}(A)\}$.

Exercice 1.31. Soient $(X, \leq)$ et $(Y, \leq)$ deux ensembles ordonnés (on note abusivement les deux ordres de la même manière). On définit sur $X \times Y$ la relation $\mathcal{R}$ par :

$$(x, y) \mathcal{R} (x', y') \Leftrightarrow (x < x') \text{ ou } (x = x' \text{ et } y \leq y')$$

Montrer que $\mathcal{R}$ est une relation d'ordre et qu'il est total si, et seulement si, X et Y sont totalement ordonnés.

Solution. $\mathcal{R}$ est clairement réflexive, antisymétrique et transitive. Pour l'ordre total, considérons deux couples (x, y) et (x', y') . Comme X et Y sont d'ordre total, il y a plusieurs cas :

- si $x < x'$, alors $(x, y) \mathcal{R} (x', y')$;
- si $x > x'$, alors $(x', y') \mathcal{R} (x, y)$;
- si $x = x'$ et $y < y'$, alors $(x, y) \mathcal{R} (x', y')$;
- si $x = x'$ et $y > y'$, alors $(x', y') \mathcal{R} (x, y)$;
- si $x = x'$ et $y = y'$, alors $(x', y') = (x, y)$ (et donc $(x, y) \mathcal{R} (x', y')$ et $(x', y') \mathcal{R} (x, y)$).

Tous les couples étant comparables on a une relation d'ordre total. Réciproquement, on vérifie que si $\mathcal{R}$ est total, il en est alors de même des ordres sur X et Y .

Chapitre 2

Structure de groupe

2.1 Loi de composition interne

Définition 2.1. On appelle loi de composition interne sur un ensemble non vide G toute application φ définie sur $G \times G$ et à valeurs dans G .

Si φ est loi de composition interne sur G , on notera souvent $a \star b = \varphi(a, b)$ pour tous a, b dans G . Il sera parfois commode de noter une telle loi sous la forme additive $(a, b) \mapsto a + b$ ou sous la forme multiplicative $(a, b) \mapsto a \cdot b$ ou plus simplement $(a, b) \mapsto ab$. On note $(G, \star)$ l'ensemble G muni de la loi de composition interne $\star$.

Exemples 2.1

1. L'addition et la multiplication usuelles sont des lois de composition interne sur $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$.
2. Si E est un ensemble non vide et $\mathcal{P}(E)$ l'ensemble de toutes les parties de E , les applications :

$$(A, B) \mapsto A \cap B, (A, B) \mapsto A \cup B, (A, B) \mapsto A \triangle B = (A \cup B) \setminus (A \cap B)$$

sont des lois de composition interne sur $\mathcal{P}(E)$.

3. Si E est un ensemble non vide et $\mathcal{F}(E)$ l'ensemble de toutes les applications de E dans E , alors l'application de composition $(f, g) \mapsto f \circ g$ est une loi de composition interne sur $\mathcal{F}(E)$.

Définition 2.2. Soit G un ensemble non vide muni d'une loi de composition interne $(a, b) \mapsto a \star b$. On dit que :

1. cette loi est associative si $(a \star b) \star c = a \star (b \star c)$ pour tous a, b, c dans G ;
2. cette loi est commutative si $a \star b = b \star a$ pour tous a, b dans G ;
3. e est un élément neutre pour cette loi si $a \star e = e \star a = a$ pour tout $a \in G$;

4. un élément a de G est dit régulier (ou simplifiable) si :

$$\forall (b, c) \in G^2, \begin{cases} a \star b = a \star c \Rightarrow b = c \\ b \star a = c \star a \Rightarrow b = c \end{cases}$$

Dire qu'un élément $a \in G$ est régulier à gauche [resp. à droite] signifie que l'application $g \mapsto a \star g$ [resp. $g \mapsto g \star a$] est injective.

Si $\star$ est une loi de composition interne associative sur G , on écrira $a \star b \star c$ pour $(a \star b) \star c$ ou $a \star (b \star c)$. De manière plus générale, toujours dans le cas d'une loi associative, on peut effectuer les opérations $a_1 \star a_2 \star \cdots \star a_n$ où les a_j sont des

éléments de G , ce que l'on notera $\prod_{j=1}^n a_j$ dans le cas d'une loi multiplicative ou

$\sum_{j=1}^n a_j$ dans le cas d'une loi additive. Ce produit (ou cette somme) est donc défini

par $a_1 \in G$ et supposant $\prod_{j=1}^{n-1} a_j$ construit pour $n \geq 2$, on a $\prod_{j=1}^n a_j = \prod_{j=1}^{n-1} a_j \star a_n$, le parenthésage étant sans importance du fait de l'associativité.

Pour $n = 0$, il sera commode de noter $\prod_{j=1}^n a_j = 1$ (ou $\sum_{j=1}^n a_j = 0$ dans le cas d'une loi additive), où 1 [resp. 0 pour une loi additive] est l'élément neutre.

Dans le cas où tous les a_j sont égaux à un même élément a , ce produit est noté a^n et on dit que c'est la puissance n -ième de a . On retiendra que ces éléments de G sont donc définis par la relation de récurrence $a^0 = 1$ et $a^{n+1} = a^n \star a$ pour tout $n \in \mathbb{N}$. Dans le cas où la loi est notée additivement, on note plutôt na au lieu de a^n . On vérifie facilement que $a^n \star a^m = a^m \star a^n = a^{n+m}$ [resp. $(na) + (ma) = (ma) + (na) = (n+m)a$ pour une loi additive] pour tous n, m dans $\mathbb{N}^*$ (voir le théorème 2.9).

Exemples 2.2

1. Les opérations usuelles d'addition et de multiplication sont commutatives et associatives sur $G = \mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$. 0 est un élément neutre pour l'addition et 1 est un élément neutre pour la multiplication pour chacun de ces ensembles. Tous les éléments de G sont simplifiables pour l'addition et tous les éléments de $G^* = G \setminus \{0\}$ sont simplifiables pour la multiplication.
2. Si E est un ensemble non vide, les opérations $\cap$ et $\cup$ sont commutatives et associatives sur $\mathcal{P}(E)$. L'ensemble vide $\emptyset$ est un élément neutre pour $\cup$ et E est un élément neutre pour l'intersection $\cap$.
3. Si E est un ensemble non vide la composition des applications est associative et non commutative dans $\mathcal{F}(E)$. L'identité est un élément neutre pour cette loi.

Théorème 2.1.

Soit $(G, \star)$ un ensemble non vide muni d'une loi de composition interne. Si G admet un élément neutre, alors ce dernier est unique.

Preuve. Soient e, e' deux éléments neutres. On a alors $e = e \star e'$ puisque e' est neutre et $e' = e \star e'$ puisque e est neutre, ce qui implique $e = e'$. $\square$

Définition 2.3. Soit $(G, \star)$ un ensemble non vide muni d'une loi de composition interne et admettant un élément neutre e . On dit qu'un élément a de G est inversible s'il existe un élément a' dans G tel que $a \star a' = a' \star a = e$. On dit alors que a' est un inverse (ou un symétrique) de a dans G .

Théorème 2.2.

Soit $(G, \star)$ un ensemble non vide muni d'une loi de composition interne associative et admettant un élément neutre e . Si $a \in G$ admet un inverse dans G , alors ce dernier est unique.

Preuve. Supposons que $a \in G$ admette deux inverses a' et a'' . On a alors :

$$a' \star a \star a'' = (a' \star a) \star a'' = e \star a'' = a''$$

puisque la loi est associative et a' est inverse de a et :

$$a' \star a \star a'' = a' \star (a \star a'') = a' \star e = a'$$

puisque a'' est inverse de a , ce qui implique $a' = a''$. $\square$

Pour une loi non associative, l'unicité du symétrique n'est pas assurée. Par exemple dans l'ensemble $G = \{0, -1, 1\}$ muni de la loi définie par la table :

$\star$	0	-1	1
0	0	-1	1
-1	-1	0	0
1	1	0	0

0 est neutre et $1 \star 1 = 1 \star (-1) = 0$.

En cas d'existence, on notera a^{-1} un inverse de a dans $(G, \star)$, la loi $\star$ étant associative. Dans le cas d'une loi de composition interne notée de façon additive, on notera plutôt $-a$ un inverse de a et on l'appellera opposé.

Exemples 2.3

1. Dans $(\mathbb{N}, +)$ seul 0 a un opposé et dans $(\mathbb{N}, \cdot)$ seul 1 a un inverse.
2. Dans $(\mathbb{Z}, +)$ tout élément admet un opposé et dans $(\mathbb{Z}, \cdot)$ les seuls éléments inversibles sont 1 et -1.

2.2 Groupes

Définition 2.4. Un groupe est un ensemble non vide G muni d'une loi de composition interne $\star$ possédant les propriétés suivantes :

- la loi $\star$ est associative ;

- il existe un élément neutre e pour la loi $\star$;
- tout élément de G admet un symétrique.

Si de plus la loi $\star$ est commutative, on dit que le groupe G est commutatif ou abélien.

En général, s'il n'y pas de confusion possible, on dira tout simplement que G est un groupe et on notera ab ou $a + b$ le résultat de l'opération $a \star b$. Pour un groupe multiplicatif $(G, \cdot)$, on notera 1 l'élément neutre, a^{-1} le symétrique d'un élément a de G et pour un groupe additif $(G, +)$, on notera 0 l'élément neutre, $-a$ le symétrique qu'on appelle opposé.

Exemples 2.4

1. Les ensembles $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$ munis de l'addition usuelle sont des groupes abéliens.
2. L'ensemble $\mathbb{N}$ muni de l'addition usuelle n'est pas un groupe du fait qu'un élément non nul de $\mathbb{N}$ n'a pas d'opposé dans $\mathbb{N}$ (l'équation $a + x = 0$ avec $a \neq 0$ dans $\mathbb{N}$ n'a pas de solution dans $\mathbb{N}$).
3. Les ensembles $\mathbb{Q}^*$, $\mathbb{R}^*$ et $\mathbb{C}^*$ munis de la multiplication usuelle sont des groupes abéliens.
4. L'ensemble $\mathbb{Z}^*$ muni de la multiplication usuelle n'est pas un groupe du fait qu'un élément de $\mathbb{Z} \setminus \{-1, 0, 1\}$ n'a pas d'inverse dans $\mathbb{Z}$ (l'équation $ax = 1$ avec $a \in \mathbb{Z} \setminus \{-1, 0, 1\}$ n'a pas de solution dans $\mathbb{Z}$).
5. Si E est un ensemble non vide, l'ensemble $\mathcal{P}(E)$ est alors un groupe pour l'opération de différence symétrique : $(A, B) \mapsto A \triangle B = (A \cup B) \setminus (A \cap B)$.
6. Si E est un ensemble non vide, l'ensemble des bijections de E dans lui même muni de la composition des applications est un groupe (en général non abélien). Ce groupe est le groupe des permutations de E , il est noté $S(E)$ ou $\mathfrak{S}(E)$.
7. Si $H_1, \dots, H_n$ sont des groupes multiplicatifs, alors le produit direct $\prod_{k=1}^n H_k$ muni de la loi $((a_1, \dots, a_n), (b_1, \dots, b_n)) \mapsto (a_1 b_1, \dots, a_n b_n)$ est un groupe et ce groupe est commutatif si, et seulement si, tous les H_i le sont.

Théorème 2.3.

Dans un groupe $(G, \star)$ tout élément est simplifiable.

Preuve. Soient a, b, c dans G . Si $a \star b = a \star c$, on a alors $a^{-1} \star a \star b = a^{-1} \star a \star c$, soit $b = c$. De même si $b \star a = c \star a$, alors $b \star a \star a^{-1} = c \star a \star a^{-1}$, soit $b = c$. $\square$

Si $(G, \star)$ est un groupe tel que G ait un nombre fini $n \geq 1$ d'éléments, on dira alors que G est un groupe fini d'ordre (ou de cardinal) n . Pour un tel groupe fini d'ordre petit, on peut dresser sa table de composition. Cette table est appelée table de Pythagore.

Théorème 2.4.

Soit $(G, \star)$ un groupe. Pour tous a, b dans G , on a $(a^{-1})^{-1} = a$ et $(a \star b)^{-1} = b^{-1} \star a^{-1}$.

Preuve. La première égalité se déduit immédiatement de la définition de a^{-1} et la deuxième résulte de $b^{-1} \star a^{-1} \star a \star b = b^{-1} \star e \star b = b^{-1} \star b = e$. $\square$

Plus généralement, on vérifie facilement par récurrence sur $p \geq 2$ que si $a_1, \dots, a_p$ sont des éléments d'un groupe G , on a alors $(a_1 \star \dots \star a_p)^{-1} = a_p^{-1} \star \dots \star a_1^{-1}$.

2.3 Sous-groupes

Définition 2.5. *Soit $(G, \star)$ un groupe. Un sous-groupe de G est un sous-ensemble H de G tel que :*

- H est non vide ;
- pour tous a, b dans H , $a \star b^{-1}$ est dans H .

Le résultat qui suit nous donne une définition équivalente de la notion de sous-groupe.

Théorème 2.5.

Soit $(G, \star)$ un groupe. Un sous-ensemble H de G est sous-groupe si, et seulement si :

- H contient l'élément neutre e de G ;
- H est stable pour la loi $\star$, c'est-à-dire que $a \star b \in H$ pour tous a, b dans H ;
- H est stable par passage à l'inverse, c'est-à-dire que $a^{-1} \in H$ pour tout $a \in H$.

Preuve. Soit H un sous-groupe de G . Pour $a \in H$, on a $e = a \star a^{-1} \in H$, $a^{-1} = e \star a^{-1} \in H$ et pour $b \in H$, $a \star b = a \star (b^{-1})^{-1} \in H$. Réciproquement si H contient e , il est non vide et s'il est stable par multiplication et passage à l'inverse, on a pour a, b dans H , $b^{-1} \in H$ et $a \star b^{-1} \in H$. $\square$

On vérifie facilement qu'un sous-groupe H d'un groupe G est lui même un groupe et H est commutatif si G l'est.

La notion de sous-groupe est commode pour montrer qu'un ensemble est un groupe : on peut essayer de le voir comme sous-groupe d'un groupe connu, ce qui évite de prouver l'associativité.

Exemples 2.5

1. Si $(G, \star)$ est un groupe d'élément neutre e , alors $H = \{e\}$ et G sont des sous-groupes de G . On dit que ce sont les sous-groupes triviaux de G .

2. L'ensemble Γ des nombres complexes de module égal à 1 (le cercle unité) est un sous-groupe du groupe multiplicatif $\mathbb{C}^*$.
3. Pour tout entier $n \geq 1$, l'ensemble $\Gamma_n = \{z \in \mathbb{C} \mid z^n = 1\}$ des racines n -èmes de l'unité est un sous-groupe de Γ .

En admettant le théorème de division euclidienne dans $\mathbb{Z}$ (théorème 4.1), on a le résultat suivant.

Théorème 2.6.

Si G est un sous-groupe de $(\mathbb{Z}, +)$, il existe alors un unique entier naturel n tel que $G = n\mathbb{Z} = \{qn \mid q \in \mathbb{Z}\}$. Cet entier n est le plus petit élément de $G \cap \mathbb{N}^*$.

Preuve. Si $G = \{0\}$, on a $G = 0\mathbb{Z}$. Si $G \neq \{0\}$, il existe dans G un entier a non nul et comme G est un sous-groupe de $(\mathbb{Z}, +)$ $-a$ est aussi dans G et l'un des entiers a ou $-a$ est dans $G^+ = G \cap \mathbb{N}^*$. L'ensemble G^+ est donc une partie non vide de $\mathbb{N}^*$ et en conséquence admet un plus petit élément $n \geq 1$. Comme $n \in G$ et G est un groupe additif, on a $n\mathbb{Z} \subset G$. D'autre part, pour tout $m \in G$, la division euclidienne par n donne $m = qn + r$ avec $r = m - nq \in G^+$ et $r \leq n - 1$, ce qui impose $r = 0$ par définition de n . On a donc $G \subset n\mathbb{Z}$ et $G = n\mathbb{Z}$. L'unicité provient du fait que $n\mathbb{Z} = m\mathbb{Z}$ si, et seulement si, $n = \pm m$ et pour n, m positifs, on a nécessairement $n = m$. $\square$

Le résultat précédent se traduit en disant que $\mathbb{Z}$ est un anneau principal et il a de nombreuses applications.

2.4 Sous-groupe engendré par une partie

Théorème 2.7.

Soit $(G, \star)$ un groupe. L'intersection d'une famille quelconque $(H_i)_{i \in I}$ de sous-groupes de G est un sous-groupe de G .

Preuve. Soit $H = \bigcap_{i \in I} H_i$. Comme l'élément neutre e est dans tous les H_i , il est aussi dans H et $H \neq \emptyset$. Si a, b sont dans H , ils sont alors dans tous les H_i , donc $a \star b^{-1} \in H_i$ pour tout $i \in I$, ce qui signifie que $a \star b^{-1} \in H$. On a donc montré que H un sous-groupe de G . $\square$

La réunion d'une famille de sous-groupes de G n'est pas nécessairement un sous-groupe. Par exemple $2\mathbb{Z}$ et $3\mathbb{Z}$ sont des sous-groupes de $(\mathbb{Z}, +)$, mais la réunion H ne l'est pas puisque 2 et 3 sont dans H alors que $2 + 3 = 5 \notin H$.

Corollaire 2.1. *Si X est une partie d'un groupe $(G, \star)$, l'intersection de tous les sous-groupes de G qui contiennent X est un sous-groupe de G .*

Preuve. L'ensemble des sous-groupes de G qui contiennent X est non vide puisque G en fait partie et le théorème précédent nous dit que l'intersection de tous ces sous-groupes est un sous-groupe de G . $\square$

Définition 2.6. Si X est une partie d'un groupe $(G, \star)$, le sous-groupe de G engendré par X est l'intersection de tous les sous-groupes de G qui contiennent X .

On note $\langle X \rangle$ le sous-groupe de G engendré par X et ce sous-groupe $\langle X \rangle$ est le plus petit (pour l'ordre de l'inclusion) des sous-groupes de G qui contiennent X .

Dans le cas où X est l'ensemble vide, on a $\langle X \rangle = \{e\}$.

Définition 2.7. Si X est une partie d'un groupe $(G, \star)$, on dit que X engendre G si $G = \langle X \rangle$.

Théorème 2.8.

Soient $(G, \star)$ un groupe et X, Y deux parties de G .

1. On a $X \subset \langle X \rangle$ et l'égalité est réalisée si, et seulement si X est un sous-groupe de G .
2. Si $X \subset Y$, on a alors $\langle X \rangle \subset \langle Y \rangle$.
3. En notant, pour X non vide, X^{-1} l'ensemble formé des symétriques des éléments de X , soit $X^{-1} = \{x^{-1} \mid x \in X\}$, les éléments de $\langle X \rangle$ sont de la forme $x_1 \star \cdots \star x_n$ où $n \in \mathbb{N}^*$ et les x_k sont dans $X \cup X^{-1}$ pour tout k compris entre 1 et n .

Preuve. Les points 1. et 2. se déduisent immédiatement des définitions. Pour le point 3. on montre tout d'abord que l'ensemble :

$$H = \{x_1 \star \cdots \star x_n \mid n \in \mathbb{N} \text{ et } x_k \in X \cup X^{-1} \text{ pour } 1 \leq k \leq n\}$$

est un sous-groupe de $(G, \star)$. Pour $x_1 \in X$, on a $e = x_1 \star x_1^{-1} \in H$ et pour $x = x_1 \star \cdots \star x_n$, $y = y_1 \star \cdots \star y_m$ dans H , on a :

$$x \star y^{-1} = x_1 \star \cdots \star x_n \star y_m^{-1} \star \cdots \star y_1^{-1} \in H$$

Donc H est bien un sous-groupe de G . Comme $X \subset H$, il nous suffit de montrer que H est contenu dans tout sous-groupe de G qui contient X . Si K est un tel sous-groupe, tout produit $x_1 \star \cdots \star x_n$ de H est un produit d'éléments de $X \cup X^{-1} \subset K$, donc dans K , ce qui prouve que $H \subset K$. On a donc bien $\langle X \rangle = H$. $\square$

Le point 3. du théorème précédent nous dit aussi que $\langle X \rangle = \langle X^{-1} \rangle = \langle X \cup X^{-1} \rangle$.

2.5 Groupes monogènes

On se donne un groupe multiplicatif $(G, \cdot)$. Si X est une partie de G formée d'un nombre fini d'éléments, $x_1, \dots, x_n$, on note alors $\langle X \rangle = \langle x_1, \dots, x_n \rangle$.

Définition 2.8. On dit que G est un groupe monogène s'il existe $x_1 \in G$ tel que $G = \langle x_1 \rangle$. Si de plus, G est fini, on dit alors qu'il est cyclique (ce

terme sera justifié après avoir défini la notion d'ordre d'un élément d'un groupe).

Pour tout $a \in G$ nous avons déjà défini les puissances entières positives de a (paragraphe 2.1). Dans un groupe, on définit les puissances entières, positives ou négatives, de $a \in G$ par $a^0 = 1$, $a^{n+1} = a^n a$ et $a^{-n} = (a^n)^{-1}$ pour tout $n \in \mathbb{N}$. On peut remarquer que pour $n \in \mathbb{N}^*$, on a aussi $a^{-n} = (a^{-1})^n$, ce qui résulte de :

$$(a^{-1})^n a^n = a^{-1} \cdots a^{-1} a \cdots a = 1$$

En notation additive, a^n est noté na pour $n \in \mathbb{Z}$.

Théorème 2.9.

Pour a dans G et n, m dans $\mathbb{Z}$, on a $a^n a^m = a^{n+m}$ et pour $b \in G$ qui commute avec a , on a $(ab)^n = a^n b^n = b^n a^n$.

Preuve. On montre tout d'abord le résultat pour n, m dans $\mathbb{N}$ par récurrence sur $m \geq 0$ à n fixé. Le résultat est évident pour $m = 0$ et le supposant acquis pour $m \geq 0$, on a $a^n a^{m+1} = a^n a^m a = a^{n+m} a = a^{n+m+1}$. On en déduit que pour n', m' dans $\mathbb{N}$, on a :

$$\begin{aligned} a^{-n'} a^{-m'} &= (a^{n'})^{-1} (a^{m'})^{-1} = (a^{m'} a^{n'})^{-1} \\ &= (a^{m'+n'})^{-1} = (a^{n'+m'})^{-1} = a^{-n'-m'} \end{aligned}$$

c'est-à-dire que le résultat est valable pour $n \leq 0$ et $m \leq 0$. Pour n, m' dans $\mathbb{N}$ tels que $n \geq m'$ on a $a^{n-m'} a^{m'} = a^n \Rightarrow a^n (a^{m'})^{-1} = a^n a^{-m'} = a^{n-m'}$ et pour $n \leq m'$, on a $a^{n-m'} = (a^{m'-n})^{-1} = (a^{m'} a^{-n})^{-1} = a^n a^{-m'}$, donc le résultat est valable pour $n \geq 0$ et $m \leq 0$. On procède de manière analogue pour $n \leq 0$ et $m \geq 0$. En définitive, c'est valable pour tous n, m dans $\mathbb{Z}$. En supposant que a et b commutent, on montre par récurrence sur $n \geq 0$ que $(ab)^n = a^n b^n$ et $ab^{n+1} = b^{n+1}a$. C'est clair pour $n = 0$ et supposant le résultat acquis pour $n \geq 0$, on a :

$$(ab)^{n+1} = (ab)^n ab = a^n b^n ab = a^n b^n ba = a^n b^{n+1} a = a^n ab^{n+1} = a^{n+1} b^{n+1}$$

Et avec $ab = ba$, on déduit que $(ab)^n = (ba)^n = b^n a^n$. Ensuite, pour $n' \geq 0$, on a :

$$(ab)^{-n'} = ((ab)^{n'})^{-1} = (a^{n'} b^{n'})^{-1} = (b^{n'} a^{n'})^{-1} = a^{-n'} b^{-n'} = b^{-n'} a^{-n'}$$

et le résultat est valable pour $n \leq 0$. □

Théorème 2.10.

Pour tout $a \in G$, le sous-groupe engendré par a est $\langle a \rangle = \{a^n \mid n \in \mathbb{Z}\}$.

Preuve. En notant $H = \{a^n \mid n \in \mathbb{Z}\}$, on a $\{a\} \subset H$, $1 = a^0 \in H$ et pour n, m dans $\mathbb{Z}$, $a^n (a^m)^{-1} = a^{n-m} \in H$, donc H est un sous-groupe de G qui contient $\{a\}$ et c'est le plus petit du fait que pour tout sous-groupe K de G qui contient $\{a\}$, on a $a^n \in K$ pour tout $n \in \mathbb{Z}$, ce qui implique $H \subset K$. On a donc bien $H = \langle a \rangle$. $\square$

Pour une loi de groupe notée additivement, on a dans le cas où G est commutatif :

$$\langle x_1, \dots, x_n \rangle = \left\{ \sum_{k=1}^n \alpha_k x_k \mid (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}^n \right\}$$

Par exemple pour le groupe additif $G = \mathbb{Z}$, on a $\langle x_1, \dots, x_n \rangle = \sum_{k=1}^n x_k \mathbb{Z} = \delta \mathbb{Z}$, où $\delta \in \mathbb{N}$ est pgcd de $x_1, \dots, x_n$. Cette notion est étudiée au paragraphe 4.4.

2.6 Groupes finis. Théorème de Lagrange

Pour ce paragraphe, on se donne un groupe multiplicatif $(G, \cdot)$.

Théorème 2.11.

Pour tout sous-groupe H de G , la relation $\sim$ définie sur G par :

$$(x \sim y) \Leftrightarrow (x^{-1}y \in H)$$

est une relation d'équivalence.

Preuve. Pour tout $x \in G$, on a $x^{-1}x = 1 \in H$, donc $\sim$ est réflexive. Si x, y dans G sont tels que $x^{-1}y \in H$, on a alors $(x^{-1}y)^{-1} = y^{-1}x \in H$, ce qui signifie que $y \sim x$. Cette relation est donc symétrique. Si x, y, z dans G sont tels que $x^{-1}y \in H$ et $y^{-1}z \in H$, on a alors $x^{-1}z = (x^{-1}y)(y^{-1}z) \in H$, ce qui signifie que $x \sim z$. Cette relation est donc transitive. $\square$

Avec les notations du théorème précédent, on note, pour tout $g \in G$, $\bar{g}$ la classe d'équivalence de g et on dit que $\bar{g}$ est la classe à gauche modulo H de g .

On a, pour $g \in G$:

$$(h \in \bar{g}) \Leftrightarrow (g \sim h) \Leftrightarrow (k = g^{-1}h \in H) \Leftrightarrow (\exists k \in H \mid h = gk) \Leftrightarrow (h \in gH)$$

soit $\bar{g} = gH$. L'ensemble de ces classes d'équivalence est noté G/H et on l'appelle l'ensemble des classes à gauche modulo H . On a donc :

$$G/H = \{\bar{g} \mid g \in G\} = \{gH \mid g \in G\}$$

L'application :

$$\begin{array}{ccc} \pi : G & \rightarrow & G/H \\ g & \mapsto & \bar{g} = gH \end{array}$$

est surjective. On dit que c'est la surjection canonique de G sur G/H .

Dans le cas où G est le groupe additif $\mathbb{Z}$ tout sous-groupe de G est de la forme $n\mathbb{Z}$ où n est un entier naturel et cette construction aboutit au groupe $\frac{\mathbb{Z}}{n\mathbb{Z}}$ des

classes résiduelles modulo n (ces groupes seront étudiés plus en détails au chapitre 4.5).

Définition 2.9. Si G est un groupe ayant un nombre fini d'éléments son cardinal est appelé l'ordre de G .

Théorème 2.12. Lagrange

Soient G un groupe fini d'ordre $n \geq 2$ et H un sous-groupe de G .

1. Les classes à gauche modulo H forment une partition de G .
2. Pour tout $g \in G$ on a $\text{card}(\bar{g}) = \text{card}(H)$.
3. L'ordre du sous-groupe H divise l'ordre du groupe G .

Preuve.

1. Comme G est fini, il en est de même de G/H . Notons $G/H = \{\bar{g}_1, \dots, \bar{g}_p\}$, où $1 \leq p \leq n$ et $\bar{g}_j \neq \bar{g}_k$, pour $1 \leq j \neq k \leq p$. Pour tout $g \in G$, il existe un unique indice j tel que $\bar{g} = \bar{g}_j$ et $g \in \bar{g}_j$. On a donc $G = \bigcup_{j=1}^p \bar{g}_j$. Dire que g est dans $\bar{g}_j \cap \bar{g}_k$ signifie que g est équivalent modulo H à g_j et g_k et donc par transitivité g_j et g_k sont équivalents, ce qui revient à dire que $\bar{g}_j = \bar{g}_k$. Les classes à gauche modulo H forment donc bien une partition de G .
2. Pour $g \in G$, l'application $h \mapsto gh$ est injective (dans un groupe tout élément est simplifiable) et en restriction à H elle réalise une bijection de H sur $gH = \bar{g}$. Il en résulte que $\bar{g}$ est de même cardinal que H .
3. Avec la partition $G = \bigcup_{j=1}^p \bar{g}_j$ et $\text{card}(\bar{g}_j) = \text{card}(H)$ pour tout j , on déduit que $\text{card}(G) = p \text{card}(H)$ et $\text{card}(H)$ divise $\text{card}(G)$.

□

Le cardinal p de l'ensemble G/H est noté $[G : H]$ et on l'appelle l'indice de H dans G . Le théorème de Lagrange peut aussi se traduire par :

$$[G : H] = \text{card}(G/H) = \frac{\text{card}(G)}{\text{card}(H)}$$

2.7 Morphismes de groupes

On désigne par $(G, \star)$ et $(H, \cdot)$ deux groupes et on note respectivement e et 1 les éléments neutres de $(G, \star)$ et $(H, \cdot)$.

Définition 2.10. On dit que φ est un morphisme de groupes de G dans H si φ est une application de G dans H telle que :

$$\forall (a, b) \in G^2, \varphi(a \star b) = \varphi(a) \cdot \varphi(b)$$

Dans le cas où φ est de plus bijective, on dit que φ est un isomorphisme du groupe G sur le groupe H . Dans le cas où $H = G$, on dit que φ est un endomorphisme du groupe $(G, \star)$ et que c'est un automorphisme du groupe $(G, \star)$ si φ est de plus bijective.

Si G et H sont deux groupes isomorphes, on notera $G \simeq H$.

Théorème 2.13.

Soient G, H, K trois groupes, φ un morphisme de groupes de G dans H et ψ un morphisme de groupes de H dans K . L'application $\psi \circ \varphi$ est un morphisme de groupes de G dans K . Si φ est un automorphisme de G , alors φ^{-1} est également un automorphisme de G .

Preuve. En notant les lois de chacun des groupes sous forme multiplicative, on a pour tout $(a, b) \in G^2$:

$$\begin{aligned} (\psi \circ \varphi)(a \cdot b) &= \psi(\varphi(a \cdot b)) = \psi(\varphi(a) \cdot \varphi(b)) \\ &= \psi(\varphi(a)) \psi(\varphi(b)) = \psi \circ \varphi(a) \psi \circ \varphi(b) \end{aligned}$$

Si φ est un automorphisme de G , on a alors pour tous a', b' dans G , en notant $a = \varphi^{-1}(a')$, $b = \varphi^{-1}(b')$:

$$\varphi^{-1}(a' \star b') = \varphi^{-1}(\varphi(a) \star \varphi(b)) = \varphi^{-1}(\varphi(a \star b)) = a \star b = \varphi^{-1}(a') \star \varphi^{-1}(b')$$

ce qui signifie que φ^{-1} est un morphisme de groupe. Et on sait déjà qu'il est bijectif, c'est donc un automorphisme de G $\square$

On déduit du théorème précédent que l'ensemble $(\text{Aut}(G), \circ)$ des automorphismes de G dans lui même est un sous-groupe du groupe symétrique $(S(G), \circ)$ formé des bijections (ou permutations) de G .

Exemples 2.6

1. La fonction exponentielle est un isomorphisme de groupes de $(\mathbb{R}, +)$ sur $(\mathbb{R}^{+,*}, \cdot)$.
2. La fonction logarithme népérien est un isomorphisme de groupes de $(\mathbb{R}^{+,*}, \cdot)$ sur $(\mathbb{R}, +)$.

Théorème 2.14.

Si φ est un morphisme de groupes de G dans H , on a alors $\varphi(e) = 1$ et $\varphi(a)^{-1} = \varphi(a^{-1})$ pour tout $a \in G$.

Preuve. Pour tout $a \in G$, on a $\varphi(a) = \varphi(a \star e) = \varphi(a) \cdot \varphi(e)$ et multipliant par $\varphi(a)^{-1}$, on obtient $1 = \varphi(e)$. Pour tout $a \in G$, on a :

$$1 = \varphi(e) = \varphi(a \star a^{-1}) = \varphi(a) \cdot \varphi(a^{-1})$$

et multipliant par $\varphi(a)^{-1}$, on obtient $\varphi(a)^{-1} = \varphi(a^{-1})$. $\square$

Définition 2.11. Soit φ un morphisme de groupes de G dans H .

1. Le noyau de φ est l'ensemble $\ker(\varphi) = \{x \in G \mid \varphi(x) = 1\}$.
2. L'image de φ est l'ensemble $\text{Im}(\varphi) = \{\varphi(x) \mid x \in G\}$.

Théorème 2.15.

Si φ est un morphisme de groupes de G dans H , alors :

1. $\ker(\varphi)$ est un sous-groupe de G ;
2. φ est injectif si, et seulement si, $\ker(\varphi) = \{e\}$;
3. $\text{Im}(\varphi)$ est un sous-groupe de H ;
4. φ est surjectif si, et seulement si, $\text{Im}(\varphi) = H$;
5. pour tout sous-groupe G' de G , $\varphi(G')$ est un sous-groupe de H ;
6. pour tout sous-groupe H' de H , $\varphi^{-1}(H')$ est un sous-groupe de G .

Preuve.

1. On a $\ker(\varphi) \neq \emptyset$ puisque $e \in \ker(\varphi)$ ($\varphi(e) = 1$) et pour x, y dans $\ker(\varphi)$:

$$\varphi(x \star y^{-1}) = \varphi(x) \cdot \varphi(y^{-1}) = \varphi(x) \cdot \varphi(y)^{-1} = 1$$

c'est-à-dire que $x \star y^{-1} \in \ker(\varphi)$ et $\ker(\varphi)$ est un sous-groupe de G .

2. Si φ est injectif, on a alors :

$$(\forall x \in \ker(\varphi), \varphi(x) = 1 = \varphi(e)) \Rightarrow (x = e)$$

et donc $\ker(\varphi) = \{e\}$. Réciproquement si $\ker(\varphi) = \{e\}$, pour x, y dans G tels que $\varphi(x) = \varphi(y)$, on a :

$$1 = \varphi(x)^{-1} \cdot \varphi(x) = \varphi(x^{-1}) \cdot \varphi(y) = \varphi(x^{-1} \star y)$$

donc $x^{-1} \star y \in \ker(\varphi)$ et $x^{-1} \star y = e$, ce qui équivaut à $x = y$.

3. On a $\text{Im}(\varphi) \neq \emptyset$ puisque $\varphi(e) \in \text{Im}(\varphi)$ et pour $\varphi(x), \varphi(y)$ dans $\text{Im}(\varphi)$ avec x, y dans G :

$$\varphi(x) \cdot \varphi(y)^{-1} = \varphi(x) \cdot \varphi(y^{-1}) = \varphi(x \star y^{-1}) \in \text{Im}(\varphi)$$

et $\text{Im}(\varphi)$ est un sous-groupe de H .

4. C'est la définition de la surjectivité.

5. On a $e \in G'$, donc $1 = \varphi(e) \in \varphi(G')$ et pour $a' = \varphi(a), b' = \varphi(b)$ dans $\varphi(G')$ avec a, b dans G' , on a :

$$a' \star (b')^{-1} = \varphi(a) \star (\varphi(b))^{-1} = \varphi(a) \star \varphi(b^{-1}) = \varphi(a \star b^{-1}) \in \varphi(G')$$

Prenant $G' = G$, on retrouve le fait que $\text{Im}(\varphi)$ est un sous-groupe de H .

6. On a $1 = \varphi(e) \in H'$, donc $e \in \varphi^{-1}(H')$ et pour a, b dans $\varphi^{-1}(H')$, on a :

$$\varphi(a \star b^{-1}) = \varphi(a) \star \varphi(b^{-1}) = \varphi(a) \star (\varphi(b))^{-1} \in H'$$

donc $a \star b^{-1} \in \varphi^{-1}(H')$. Prenant $H' = \{1\}$, on retrouve le fait que $\ker(\varphi)$ est un sous-groupe de G .

□

2.8 Ordre d'un élément dans un groupe

Pour ce paragraphe, on se donne un groupe multiplicatif $(G, \cdot)$ et pour tout $a \in G$, $\langle a \rangle = \{a^n \mid n \in \mathbb{Z}\}$ est le sous-groupe de G engendré par a .

Définition 2.12. L'ordre de $a \in G$ est l'élément $\theta(a) \in \mathbb{N}^* \cup \{+\infty\}$ défini par $\theta(a) = \text{card}(\langle a \rangle)$. Si $\theta(a)$ est dans $\mathbb{N}^*$, on dit alors que a est d'ordre fini, sinon on dit qu'il est d'ordre infini.

Seul l'unité $1 \in G$ est d'ordre 1 dans G . En effet, si $a = 1$, alors $\langle a \rangle = \{1\}$ et si $a \neq 1$, alors $a^0 \neq a^1$ et $\langle a \rangle$ a au moins deux éléments.

Pour tout $a \in G$, on a $\theta(a) = \theta(a^{-1})$ puisque :

$$\langle a^{-1} \rangle = \{(a^{-1})^n \mid n \in \mathbb{Z}\} = \{a^{-n} \mid n \in \mathbb{Z}\} = \{a^n \mid n \in \mathbb{Z}\} = \langle a \rangle$$

Dans le cas où le groupe G est fini, le théorème de Lagrange (théorème 2.12) nous dit que, pour tout $a \in G$, l'ordre de a divise l'ordre de G .

Un groupe fini G d'ordre n est cyclique si, et seulement si, il existe dans G un élément d'ordre n .

Pour $a \in G$, le sous-groupe de G engendré par a peut être vu comme l'image du morphisme de groupes :

$$\begin{array}{ccc} \varphi_a : \mathbb{Z} & \rightarrow & G \\ k & \mapsto & a^k \end{array}$$

(pour j, k dans $\mathbb{Z}$, on a $\varphi_a(j+k) = a^{j+k} = a^j a^k = \varphi_a(j) \varphi_a(k)$ et φ_a est bien un morphisme de groupes).

Connaissant les sous-groupes additifs de $\mathbb{Z}$, on a le résultat suivant.

Théorème 2.16.

Pour $a \in G$, on a $\theta(a) = +\infty$ si, et seulement si, φ_a est injective et pour a d'ordre fini, on a $\ker(\varphi_a) = \theta(a)\mathbb{Z}$.

Preuve. Le noyau de φ_a étant un sous-groupe de $\mathbb{Z}$, il existe un unique entier $n \geq 0$ tel que $\ker(\varphi_a) = n\mathbb{Z}$. On aura $n = 0$ si, et seulement si, φ_a est injective, ce qui revient à dire que $\varphi_a(k) = a^k \neq 1$ pour tout $k \in \mathbb{Z}^*$ ou encore que $\varphi_a(k) = a^k \neq \varphi_a(j) = a^j$ pour tous $j \neq k$ dans $\mathbb{Z}$ et le sous-groupe $\langle a \rangle = \text{Im}(\varphi_a)$ est infini. Si $n \geq 1$, en effectuant, pour $k \in \mathbb{Z}$, la division euclidienne de k par n , on a $k = qn + r$ avec $0 \leq r \leq n-1$ et $a^k = (a^n)^q a^r = a^r$, ce qui nous donne :

$$\langle a \rangle = \text{Im}(\varphi_a) = \{a^r \mid 0 \leq r \leq n-1\}$$

De plus pour $1 \leq r \leq n-1$, on a $a^r \neq 1$ puisque $n = \inf(\ker(\varphi_a) \cap \mathbb{N}^*)$, ce qui entraîne $a^r \neq a^s$ pour $0 \leq r \neq s \leq n-1$ (pour $s \geq r$, l'égalité $a^r = a^s$ équivaut à $a^{s-r} = 1$ avec $s-r$ compris entre 0 et $n-1$, ce qui équivaut à $r = s$). Le groupe $\langle a \rangle$ a donc exactement n éléments. $\square$

Une autre définition de l'ordre d'un élément d'un groupe est donnée par le résultat suivant.

Corollaire 2.2. Dire que $a \in G$ est d'ordre fini $n \geq 1$ équivaut à dire que $a^n = 1$ et $a^k \neq 1$ pour tout k est compris entre 1 et $n - 1$ ($\theta(a)$ est le plus petit entier naturel non nul tel que $a^n = 1$).

Preuve. Si a est d'ordre $n \geq 1$, on a vu avec la démonstration du théorème précédent que $a^n = 1$ et $a^k \neq 1$ pour tout k est compris entre 1 et $n - 1$. Réciproquement s'il existe un entier $n \geq 1$ tel que $a^n = 1$ et $a^k \neq 1$ pour k est compris entre 1 et $n - 1$, le morphisme de groupes φ_a est non injectif, donc a est d'ordre fini et $\ker(\varphi_a) = \theta(a)\mathbb{Z}$ avec $\theta(a) = \inf(\ker(\varphi_a) \cap \mathbb{N}^*) = n$. $\square$

Corollaire 2.3. Dire que $a \in G$ est d'ordre fini $n \geq 1$ équivaut à dire que, pour $k \in \mathbb{Z}$, on a $a^k = 1$ si, et seulement si, k est multiple de n .

Preuve. Si a est d'ordre n , on a alors $a^n = 1$ et pour $k = qn + r \in \mathbb{Z}$ avec $q \in \mathbb{Z}$ et $0 \leq r \leq n - 1$ (division euclidienne), on a $a^k = a^r = 1$ si, et seulement si $r = 0$. Réciproquement supposons que $a^k = 1$ si, et seulement si, k est multiple de n . On a alors $a^n = 1$ et $a^k \neq 1$ si k est compris entre 1 et $n - 1$, ce qui signifie que a est d'ordre n . $\square$

En résumé, on retiendra que :

$$\begin{aligned} (\theta(a) = +\infty) &\Leftrightarrow (\varphi_a \text{ injective}) \Leftrightarrow (\forall k \in \mathbb{Z}^*, a^k \neq 1) \\ &\Leftrightarrow (\langle a \rangle \text{ est infini isomorphe à } \mathbb{Z}) \end{aligned}$$

$$\begin{aligned} (\theta(a) = n \in \mathbb{N}^*) &\Leftrightarrow (\ker(\varphi_a) = n\mathbb{Z}) \Leftrightarrow (\langle a \rangle = \{a^r \mid 0 \leq r \leq n - 1\}) \\ &\Leftrightarrow (k \in \mathbb{Z} \text{ et } a^k = 1 \text{ équivaut à } k \equiv 0 \pmod{n}) \\ &\Leftrightarrow (n \text{ est le plus petit entier naturel non nul tel que } a^n = 1) \end{aligned}$$

Pour a d'ordre fini, le groupe $\langle a \rangle$ est dit cyclique, ce qui est justifié par le fait que $a^{qn+r} = a^r$ pour $q \in \mathbb{Z}$ et $0 \leq r \leq n - 1$.

Exemple 2.1 Le groupe multiplicatif Γ_n des racines n -èmes de l'unité, qui est cyclique d'ordre n , est isomorphe à $\frac{\mathbb{Z}}{n\mathbb{Z}}$ par l'application $\bar{k} \mapsto e^{\frac{2ik\pi}{n}}$.

Dans le cas où le groupe G est additif, l'ordre de $a \in G$ est défini comme le plus petit entier $n \geq 1$ tel que $na = 0$, quand cet ordre est fini. L'égalité $ma = 0$ équivaut alors à dire que m est multiple de n . Le groupe engendré par a est alors :

$$\langle a \rangle = \{ka \mid k \in \mathbb{Z}\} = \{ra \mid 0 \leq r \leq n - 1\}$$

Corollaire 2.4. Si G est fini d'ordre m , on a alors $a^m = 1$ pour tout $a \in G$.

Preuve. $\theta(a)$ est fini et divise m . $\square$

En se reportant au paragraphe 4.4 pour la définition du pgcd et du ppcm de deux entiers, on a le résultat suivant.

Théorème 2.17.

Soient $a, b \in G$ d'ordre fini et $k \in \mathbb{Z}^*$.

1. On a $\theta(a^k) = \frac{\theta(a)}{\theta(a) \wedge k}$ (en particulier $\theta(a^{-1}) = \theta(a)$).
2. Si k divise $\theta(a)$, on a alors $\theta(a^k) = \frac{\theta(a)}{|k|}$.
3. Si k est premier avec $\theta(a)$, on a alors $\theta(a^k) = \theta(a)$.
4. Si $ab = ba$, alors ab est d'ordre fini divisant $\theta(a) \vee \theta(b)$. Dans le cas où $\langle a \rangle \cap \langle b \rangle = \{1\}$, on a $\theta(ab) = \theta(a) \vee \theta(b)$. Si $\theta(a)$ et $\theta(b)$ sont premiers entre eux, on a alors $\langle a \rangle \cap \langle b \rangle = \{1\}$ et $\theta(ab) = \theta(a) \vee \theta(b) = \theta(a) \theta(b)$.

Preuve.

1. Soit $\delta = \theta(a) \wedge k$ et n', k' premiers entre eux tels que $\theta(a) = \delta n', k = \delta k'$. Pour tout entier relatif j , on a :

$$\begin{aligned} \left((a^k)^j = a^{kj} = 1 \right) &\Leftrightarrow (\exists q \in \mathbb{Z} \mid kj = q\theta(a)) \\ &\Leftrightarrow (\exists q \in \mathbb{Z} \mid k'j = qn') \Leftrightarrow (n' \text{ divise } j \text{ (Gauss)}) \end{aligned}$$

et en conséquence $\theta(a^k) = n' = \frac{\theta(a)}{\theta(a) \wedge k}$.

2. Si k divise $\theta(a)$, on a alors $\theta(a) \wedge k = |k|$ et $\theta(a^k) = \frac{\theta(a)}{|k|}$.
3. Si k est premier avec $\theta(a)$, on a alors $\theta(a) \wedge k = 1$ et $\theta(a^k) = \theta(a)$.
4. Soit $\mu = \theta(a) \vee \theta(b)$. Dans le cas où a et b commutent, on a $(ab)^\mu = a^\mu b^\mu = 1$ avec $\mu \geq 1$ et ab est d'ordre fini et cet ordre divise μ . En désignant par $n = \theta(ab)$ l'ordre de ab , on a $a^n b^n = (ab)^n = 1$ et $a^n = b^{-n} \in \langle a \rangle \cap \langle b \rangle$. Si $\langle a \rangle \cap \langle b \rangle = \{1\}$, on a alors $a^n = b^n = 1$ et n est multiple de $\theta(a)$ et $\theta(b)$, donc de $\theta(a) \vee \theta(b)$ et $n = \theta(a) \vee \theta(b)$. Si $\theta(a) \wedge \vee \theta(b) = 1$, on a alors $\theta(a) \vee \theta(b) = \theta(a) \theta(b)$. De plus avec $\langle a \rangle \cap \langle b \rangle \subset \langle a \rangle$ et $\langle a \rangle \cap \langle b \rangle \subset \langle b \rangle$, on déduit que $\text{card}(\langle a \rangle \cap \langle b \rangle)$ divise $\theta(a) = \text{card}(\langle a \rangle)$ et $\theta(b) = \text{card}(\langle b \rangle)$, donc $\text{card}(\langle a \rangle \cap \langle b \rangle) = 1$ et $\langle a \rangle \cap \langle b \rangle = \{1\}$, ce qui implique que $\theta(ab) = \theta(a) \vee \theta(b) = \theta(a) \theta(b)$.

□

Si $\theta(a)$ et $\theta(b)$ ne sont pas premiers entre eux, avec a, b commutant et d'ordre fini, l'ordre de ab n'est pas nécessairement le ppcm de $\theta(a)$ et $\theta(b)$. En prenant par exemple a d'ordre $n \geq 2$ dans G et $b = a^{-1}$ qui est également d'ordre n , on a $ab = ba = 1$ d'ordre $1 \neq \text{ppcm}(n, n) = n$.

Pour a et b ne commutant pas, le produit ab peut être d'ordre infini, même si a et b sont d'ordre fini.

Pour la fin de ce paragraphe, $G = \langle a \rangle$ un groupe cyclique d'ordre $n \geq 2$.

Si H est un sous-groupe de G , le théorème de Lagrange nous dit que l'ordre de H est un diviseur de n . Le théorème qui suit nous dit que les sous-groupes d'un groupe cyclique sont cycliques et que pour tout diviseur d de n , il existe un sous-groupe de G d'ordre d .

Théorème 2.18.

Pour tout diviseur d de n , il existe un unique sous-groupe d'ordre d du groupe cyclique $G = \langle a \rangle$, c'est le groupe cyclique $H = \langle a^{\frac{n}{d}} \rangle$.

Preuve. Pour tout diviseur d de n , $H = \langle a^{\frac{n}{d}} \rangle$ est un sous-groupe cyclique de G et le théorème 2.17 nous dit qu'il est d'ordre $\theta(a^{\frac{n}{d}}) = \frac{n}{n \wedge \frac{n}{d}} = d$. Réciproquement soit H un sous-groupe de G d'ordre d , un diviseur de n . Si $d = 1$, on a alors $H = \{1\} = \langle a^n \rangle$. Si $d \geq 2$, H n'est pas réduit à $\{1\}$, donc il existe un entier k compris entre 1 et $n - 1$ tel que $a^k \in H$ et on peut poser :

$$p = \min \{k \in \{1, \dots, n-1\} \mid a^k \in H\}$$

En écrivant, pour tout $h = a^k \in H$, $k = pq + r$ avec $0 \leq r \leq p - 1$ (division euclidienne par p), on a $a^r = a^k (a^{pq})^{-1} \in H$ et nécessairement $r = 0$. On a donc $H \subset \langle a^p \rangle \subset H$, soit $H = \langle a^p \rangle$. Avec $a^n = 1 \in H$, on déduit que n est multiple de p et l'ordre de H est $d = \frac{n}{n \wedge p} = \frac{n}{p}$, c'est-à-dire que $H = \langle a^{\frac{n}{d}} \rangle$. Un tel sous-groupe d'ordre d est donc unique. $\square$

Exemples 2.7

1. Les sous-groupes de $\frac{\mathbb{Z}}{n\mathbb{Z}}$ (voir le paragraphe 4.12) sont les $\langle \frac{n}{d} \bar{1} \rangle = \langle \frac{\bar{n}}{d} \rangle$ où d est un diviseur de n . Un tel sous-groupe est isomorphe à $\frac{\mathbb{Z}}{d\mathbb{Z}}$ et il y en a autant que de diviseurs de n .
2. Les sous-groupes de $\Gamma_n = \{z \in C \mid z^n = 1\} = \langle e^{\frac{2i\pi}{n}} \rangle$ sont les $\langle (e^{\frac{2i\pi}{n}})^{\frac{n}{d}} \rangle = \langle e^{\frac{2i\pi}{d}} \rangle = \Gamma_d$ où d est un diviseur de n et il y en a autant que de diviseurs de n .

2.9 Exercices

Exercice 2.1. Montrer que si $(G, \star)$ est un groupe, alors pour tout $a \in G$, la translation à gauche $g_a : x \mapsto a \star x$ [resp. à droite $d_a : x \mapsto x \star a$] est une bijection de G d'inverse $g_{a^{-1}}$ [resp. $d_{a^{-1}}$].

Solution. L'égalité $g_a(x) = g_a(y)$ équivaut à $a \star x = a \star y$ et multipliant à gauche par a^{-1} , on en déduit que $x = y$. L'application g_a est donc injective. Pour $y \in G$, l'équation $g_a(x) = y$ équivaut à $a \star x = y$, ce qui entraîne $x = a^{-1} \star y$. L'application g_a est donc surjective. En fait comme, pour tout $y \in G$, l'équation $g_a(x) = y$ a pour unique solution $x = a^{-1} \star y$, on déduit immédiatement que g_a est bijective d'inverse $g_{a^{-1}}$.

Exercice 2.2. Montrer que si $(G, \star)$ est un groupe et E un ensemble non vide, alors l'ensemble G^E des applications de E dans G muni de la loi $\cdot$ définie par :

$$\forall (f, g) \in G^E \times G^E, \forall x \in E, (f \cdot g)(x) = f(x) \star g(x)$$

est un groupe et que ce groupe est commutatif si G l'est.

Solution. Pour f, g dans G^E , $f \cdot g$ est bien une application de E dans G , donc un élément de G^E . L'application $1 : x \mapsto e$ est le neutre pour cette loi. Si $f \in G^E$, l'application $f' : x \mapsto (f(x))^{-1}$ est l'inverse de f . Pour f, g, h dans G^E et $x \in E$, on a :

$$\begin{aligned} (f \cdot (g \cdot h))(x) &= f(x) \star (g \cdot h)(x) = f(x) \star (g(x) \star h(x)) \\ &= (f(x) \star g(x)) \star h(x) = (f \cdot g)(x) \star h(x) = ((f \cdot g) \cdot h)(x) \end{aligned}$$

et donc $f \cdot (g \cdot h) = (f \cdot g) \cdot h$. L'ensemble G^E muni de la loi $\cdot$ est donc un groupe. Si G est commutatif, on a alors pour f, g dans G^E et tout $x \in E$, $(f \cdot g)(x) = f(x) \star g(x) = g(x) \star f(x) = (g \cdot f)(x)$, ce qui revient à dire que $f \cdot g = g \cdot f$. Le groupe $(G^E, \cdot)$ est donc commutatif si G l'est.

Exercice 2.3. Soit $(G, \star)$ un ensemble non vide muni d'une loi de composition interne associative, admettant un élément neutre e à gauche, c'est-à-dire que :

$$\forall a \in G, e \star a = a$$

et telle que tout élément de G admette un symétrique à gauche, c'est-à-dire que :

$$\forall a \in G, \exists a' \in G \mid a' \star a = e.$$

Montrer alors que $(G, \star)$ est un groupe (e est alors le neutre de $(G, \star)$ et a' le symétrique de a).

Solution. Soient $a \in G$ et $a' \in G$ tel que $a' \star a = e$. En désignant par a'' un inverse à gauche de a' , on a $a \star a' = a'' \star a' \star a \star a' = a'' \star (a' \star a) \star a' = a'' \star a' = e$, ce qui signifie que a' est aussi un inverse à droite de a . Et avec :

$$a \star e = a \star (a' \star a) = (a \star a') \star a = e \star a = a$$

on déduit que e est aussi un neutre à droite. En définitive, e est un élément neutre dans $(G, \star)$ et a' est le symétrique de a . Avec l'associativité de $\star$, il en résulte que $(G, \star)$ est un groupe.

Exercice 2.4. Montrer que l'ensemble $G =]-1, 1[$ muni de la loi $\star$ définie par $x \star y = \frac{x+y}{1+xy}$ est un groupe commutatif.

Solution. Pour x, y dans $] -1, 1[$, on a $|xy| < 1$, donc $1 + xy > 0$ et $x \star y$ est bien défini. De plus avec :

$$(x + y)^2 - (1 + xy)^2 = x^2 + y^2 - 1 - x^2y^2 = (x^2 - 1)(1 - y^2) < 0$$

on déduit que $|x \star y| < 1$ et $\star$ défini bien une loi interne sur G . De la commutativité de la somme et du produit sur $\mathbb{R}$ on déduit que $\star$ est commutative. Pour tout $x \in G$, on a $x \star 0 = x$ et $x \star (-x) = 0$, donc 0 est neutre et tout élément de G est inversible. Enfin pour x, y, z dans G , on a :

$$x \star (y \star z) = \frac{x + y \star z}{1 + x \cdot y \star z} = \frac{x + \frac{y+z}{1+yz}}{1 + x \frac{y+z}{1+yz}} = \frac{x + y + z + xyz}{xy + xz + yz + 1}$$

et :

$$(x \star y) \star z = \frac{x \star y + z}{1 + x \star y \cdot z} = \frac{\frac{x+y}{1+xy} + z}{1 + \frac{x+y}{1+xy} z} = \frac{x + y + z + xyz}{xy + xz + yz + 1}$$

donc $\star$ est associative.

Exercice 2.5. Montrer que l'ensemble $G =]-\frac{\pi}{2}, \frac{\pi}{2}[$ muni de la loi $\star$ définie par $x \star y = \arctan(\tan(x) + \tan(y))$ est un groupe commutatif.

Solution. La fonction $\arctan$ étant bijective de $\mathbb{R}$ sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$ l'application $\star$ défini bien une loi interne sur G . De la commutativité de la somme sur $\mathbb{R}$ on déduit que $\star$ est commutative. Pour tout $x \in G$, on a $x \star 0 = x$ et $x \star (-x) = 0$ (la fonction $\tan$ est impaire), donc 0 est neutre et tout élément de G est inversible. Enfin pour x, y, z dans G , on a :

$$x \star (y \star z) = \arctan(\tan(x) + \tan(y \star z)) = \arctan(\tan(x) + (\tan(y) + \tan(z)))$$

et :

$$(x \star y) \star z = \arctan(\tan(x \star y) + \tan(z)) = \arctan(\tan(x) + (\tan(y) + \tan(z)))$$

ce qui montre que $\star$ est associative.

Exercice 2.6. Soit G un groupe multiplicatif d'élément neutre 1. Montrer que si on a $a^2 = 1$ pour tout a dans G , alors G est commutatif.

Solution. Pour a, b dans G , de $abab = (ab)^2 = 1$, on déduit que $a(abab)b = ab$, soit $a^2bab^2 = ab$ ou encore $ba = ab$.

Exercice 2.7. Soit G un groupe multiplicatif d'élément neutre 1.

1. Montrer que G est commutatif si, et seulement si, on a $(ab)^2 = a^2b^2$ pour tous a, b dans G (ce qui revient à dire que l'application $a \mapsto a^2$ est un morphisme de groupes).

2. Montrer que G est commutatif si, et seulement si, on a $(ab)^{-1} = a^{-1}b^{-1}$ pour tous a, b dans G (ce qui revient à dire que l'application $a \mapsto a^{-1}$ est un morphisme de groupes).

Solution.

1. Dans le cas où G est commutatif, on a $(ab)^2 = abab = aabb = a^2b^2$ pour tous a, b dans G . Réciproquement, si $(ab)^2 = a^2b^2$ pour tous a, b dans G , de $abab = (ab)^2 = a^2b^2 = aabb$, on déduit par simplification à gauche par a et à droite par b que $ba = ab$. On peut retrouver le résultat de l'exercice précédent avec ce résultat. Si $a^2 = 1$ pour tout a dans G , on a alors pour tous a, b dans G , $(ab)^2 = 1 = a^2b^2$ et G est commutatif.
2. Dans le cas où G est commutatif, on a $a^{-1}b^{-1}ab = a^{-1}b^{-1}ba = 1$ pour tous a, b dans G , donc $a^{-1}b^{-1} = (ab)^{-1}$. Réciproquement, si $(ab)^{-1} = a^{-1}b^{-1}$ pour tous a, b dans G , on a alors $ab = \left((ab)^{-1}\right)^{-1} = (a^{-1}b^{-1})^{-1} = ba$ et G est commutatif.

Exercice 2.8. Montrer que pour tout groupe $(G, \star)$ et tout élément a de G , le centralisateur de a formé des éléments Z_a de G qui commutent avec a , soit $Z_a = \{b \in G \mid a \star b = b \star a\}$, est un sous-groupe de G .

Solution. On a $Z_a \neq \emptyset$ puisque $e \in Z_a$. Pour b, c dans Z_a , on a :

$$(b \star c) \star a = b \star (c \star a) = b \star (a \star c) = (b \star a) \star c = (a \star b) \star c = a \star (b \star c)$$

c'est-à-dire que $b \star c \in Z_a$. Pour b dans Z_a , de $a \star b = b \star a$, on déduit que

$$b^{-1} \star a = b^{-1} \star a \star b \star b^{-1} = b^{-1} \star b \star a \star b^{-1} = a \star b^{-1}$$

c'est-à-dire que $b^{-1} \in Z_a$. En définitive, Z_a est un sous-groupe de G . Dans le cas où G est commutatif, on a $Z_a = G$ pour tout $a \in G$.

Exercice 2.9. Montrer que pour tout groupe $(G, \star)$, le centre (ou commutateur) $Z(G)$ de G formé des éléments de G qui commutent à tous les autres éléments de G , soit $Z(G) = \{a \in G \mid \forall b \in G, a \star b = b \star a\}$, est un sous-groupe de G .

Solution. On a $Z(G) \neq \emptyset$ puisque $e \in Z(G)$. Pour a, b dans $Z(G)$, on a pour tout $c \in G$:

$$(a \star b^{-1}) \star c = a \star (c^{-1} \star b)^{-1} = a \star (b \star c^{-1})^{-1} = (a \star c) \star b^{-1} = c \star (a \star b^{-1})$$

c'est-à-dire que $a \star b^{-1} \in Z(G)$. En définitive, $Z(G)$ est un sous-groupe de G . On peut remarquer que $Z(G) = G$ si, et seulement si, G est commutatif.

Exercice 2.10. Soit H une partie finie non vide d'un groupe $(G, \star)$. Montrer que H est un sous-groupe de G si, et seulement si, il est stable pour la multiplication.

Solution. Il est clair que la condition est nécessaire. Supposons que H soit fini et stable pour la multiplication. Il s'agit alors de montrer que pour tout $a \in H$, l'inverse a^{-1} est aussi dans H . La translation à gauche $g_a : x \mapsto a \star x$ est une bijection de G et comme H est stable pour la multiplication, cette translation est injective de H dans H et donc bijective puisque H est fini. Il existe donc $x \in H$ tel que $a \star x = a$, ce qui entraîne $x = e$ (en multipliant à gauche par a^{-1}). On a donc $e \in H$ et il existe $x \in H$ tel que $a \star x = e$, ce qui entraîne $x = a^{-1}$ et $a^{-1} \in H$.

Exercice 2.11. Soient H, K deux sous-groupes d'un groupe multiplicatif G . On définit les sous-ensembles HK et KH de G par :

$$HK = \{hk \mid (h, k) \in H \times K\}, \quad KH = \{kh \mid (k, h) \in K \times H\}.$$

Montrer que :

$$(HK \text{ est un sous-groupe de } G) \Leftrightarrow (HK = KH)$$

Solution. Supposons que HK soit un sous-groupe de G . Si $g \in HK$, alors g^{-1} est aussi dans HK puisque HK est un groupe, il existe donc (h, k) dans $H \times K$ tel que $g^{-1} = hk$ et $g = (g^{-1})^{-1} = k^{-1}h^{-1} \in KH$ (H et K sont des groupes). On a donc $HK \subset KH$. Si $g \in KH$, il existe alors (h, k) dans $H \times K$ tel que $g = kh$ et $g^{-1} = h^{-1}k^{-1} \in HK$ et comme HK est un groupe, il en résulte que $g = (g^{-1})^{-1} \in HK$. On a donc $KH \subset HK$ et l'égalité $HK = KH$. Réciproquement supposons que $HK = KH$. On a $1 = 1 \cdot 1 \in HK$. Si $g_1 = h_1k_1$ et $g_2 = h_2k_2$ sont dans HK avec h_1, h_2 dans H et k_1, k_2 dans K , alors $g_1g_2^{-1} = h_1k_1k_2^{-1}h_2^{-1}$ avec $h_1 \in H, k_1k_2^{-1} \in K$ (K est un groupe), donc $h_1(k_1k_2^{-1}) \in HK = KH$ et il existe $(k_3, h_3) \in K \times H$ tel que $h_1(k_1k_2^{-1}) = k_3h_3$, ce qui donne $g_1g_2^{-1} = k_3(h_3h_2^{-1}) \in KH = HK$. On a donc prouvé que HK est un sous-groupe de G . Dans le cas où G est commutatif, on a toujours $HK = KH$ et HK est un sous-groupe de G si H et K le sont.

Exercice 2.12. Soient H, K deux sous-groupes d'un groupe G . Montrer que $H \cup K$ est un sous-groupe de G si, et seulement si, $H \subset K$ ou $K \subset H$.

Solution. Si $H \subset K$ ou $K \subset H$, on a alors $H \cup K = K$ ou $H \cup K = H$ et c'est un sous-groupe de G . Réciproquement, supposons que $H \cup K$ soit un sous-groupe de G . Si $H \subset K$ c'est terminé, sinon il existe $g \in H \setminus K$. Pour tout $k \in K \subset H \cup K$, gk est dans $H \cup K$ (c'est un groupe) et gk ne peut être dans K (sinon $g = (gk)k^{-1} \in K$, ce qui n'est pas), donc $gk \in H$ et $k = g^{-1}(gk) \in H$. On a donc $K \subset H$.

Exercice 2.13. Soient G un groupe fini, H, K deux sous-groupes de G et φ l'application :

$$\begin{aligned}\varphi: H \times K &\rightarrow HK \\ (h, k) &\mapsto hk\end{aligned}$$

1. Montrer que pour tout $g \in HK$, on a $\text{card}(\varphi^{-1}(g)) = \text{card}(H \cap K)$.
2. En déduire que $\text{card}(H) \text{card}(K) = \text{card}(HK) \text{card}(H \cap K)$, puis que :

$$(HK \text{ est un sous-groupe de } G) \Leftrightarrow (HK \subset KH) \Leftrightarrow (HK = KH)$$

Solution.

1. Soit $g = h_1 k_1 \in HK$. L'égalité $g = hk$ avec $(h, k) \in H \times K$ équivaut à $h_1 k_1 = hk$, ce qui entraîne $h = h_1 k_1 k^{-1} = h_1 g$ avec $g = k_1 k^{-1} = h_1^{-1} h \in H \cap K$ et $k = h^{-1} h_1 k_1 = g^{-1} k_1$. On a donc $\varphi^{-1}(g) \subset \{(h_1 g, g^{-1} k_1) \mid g \in H \cap K\}$. Réciproquement si $(h, k) = (h_1 g, g^{-1} k_1)$ avec $g \in H \cap K$, on a alors $(h, k) \in H \times K$ et $hk = h_1 g g^{-1} k_1 = g$. On a donc $\varphi^{-1}(g) = \{(h_1 g, g^{-1} k_1) \mid g \in H \cap K\}$ et $\text{card}(\varphi^{-1}(g)) = \text{card}(H \cap K)$.
2. En écrivant qu'on a la partition $H \times K = \bigcup_{g \in HK} \varphi^{-1}(g)$, on déduit que :

$$\begin{aligned}\text{card}(H) \text{card}(K) &= \text{card}(H \times K) = \sum_{g \in HK} \text{card}(\varphi^{-1}(g)) \\ &= \sum_{g \in HK} \text{card}(H \cap K) = \text{card}(HK) \text{card}(H \cap K)\end{aligned}$$

Il en résulte que $\text{card}(HK) = \text{card}(KH) = \frac{\text{card}(H) \text{card}(K)}{\text{card}(H \cap K)}$.

3. On en déduit que $(HK \subset KH) \Leftrightarrow (HK = KH)$ et l'exercice précédent permet de conclure.

Exercice 2.14. Soit G un groupe. Montrer que pour tout n -uplet $(x_1, \dots, x_n)$ d'éléments de G qui commutent deux à deux (avec $n \geq 1$),

$$\text{on a } \langle x_1, \dots, x_n \rangle = \left\{ \prod_{k=1}^n x_k^{\alpha_k} \mid (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}^n \right\}.$$

Solution. En notant $X = \{x_1, \dots, x_n\}$, on a $X^{-1} = \{x_1^{-1}, \dots, x_n^{-1}\}$ et comme les x_k commutent, on déduit que :

$$\begin{aligned}\langle x_1, \dots, x_n \rangle &= \left\{ \prod_{k=1}^m y_k \mid m \in \mathbb{N} \text{ et } y_k \in X \cup X^{-1} \text{ pour } 1 \leq k \leq m \right\} \\ &= \left\{ \prod_{k=1}^n x_k^{\alpha_k} \mid (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}^n \right\}\end{aligned}$$

($x_k x_j = x_j x_k$ entraîne $x_j^{-1} x_k x_j x_j^{-1} = x_j^{-1} x_j x_k x_j^{-1}$, soit $x_j^{-1} x_k = x_k x_j^{-1}$ et les éléments de $X \cup X^{-1}$ commutent).

Exercice 2.15. Montrer qu'un groupe G engendré par deux éléments a et b qui commutent est commutatif.

Solution. Comme $ab = ba$, on a $G = \langle a, b \rangle = \{a^\alpha b^\beta \mid (\alpha, \beta) \in \mathbb{Z}^2\}$ et ce groupe est commutatif.

Exercice 2.16. Soit $X = \{r_1, \dots, r_n\}$ une partie finie de $\mathbb{Q}$ et $G = \langle X \rangle$ le sous-groupe de $(\mathbb{Q}, +)$ engendré par X . Montrer que G est monogène.

Solution. En désignant par μ le ppcm des dénominateurs de $r_1, \dots, r_n$, il existe des entiers relatifs $a_1, \dots, a_n$ tels que $r_k = \frac{a_k}{\mu}$ pour tout k compris entre 1 et n et en désignant par δ le pgcd de $a_1, \dots, a_n$, on a :

$$G = \left\{ \sum_{k=1}^n \alpha_k \frac{a_k}{\mu} \mid (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}^n \right\} = \left\{ \frac{\delta}{\mu} \sum_{k=1}^n \alpha_k b_k \mid (\alpha_1, \dots, \alpha_n) \in \mathbb{Z}^n \right\}$$

où $b_1, \dots, b_n$ sont des entiers relatifs premiers entre eux dans leur ensemble. On a donc $G = \frac{\delta}{\mu} \mathbb{Z}$, ce qui signifie que G est monogène engendré par $\frac{\delta}{\mu}$.

Exercice 2.17. Soient G un groupe et H, K deux sous-groupes distincts de G d'ordre un même nombre premier $p \geq 2$. Montrer que $H \cap K = \{1\}$.

Solution. $H \cap K$ est un sous-groupe de H , il est donc d'ordre 1 ou p . S'il est d'ordre p , il est égal à H et $H = H \cap K \subset K$ entraîne $H = K$, puisque ces deux ensembles ont le même nombre d'éléments. On a donc, pour $H \neq K$, $p = 1$ et $H \cap K = \{1\}$.

Exercice 2.18. Soient G un groupe, H un sous-groupe de G et K un sous-groupe de H . Montrer que si l'indice de K dans G est fini, alors l'indice de H dans G et celui de K dans H sont aussi finis et on a $[G : K] = [G : H][H : K]$.

Solution. On note respectivement $(g_i H)_{i \in I}$ et $(h_j K)_{j \in J}$ les classes à gauches modulo H dans G et modulo K dans H deux à deux distinctes. Nous allons alors montrer que la famille des classes à gauches modulo K dans G deux à deux distinctes est $(g_i h_j K)_{(i,j) \in I \times J}$. Dans le cas où $[G : K]$ est fini, il n'y a qu'un nombre fini de telles classes, ce qui impose que I et J sont finis et on a :

$$[G : K] = \text{card}(I \times J) = \text{card}(I) \text{card}(J) = [G : H][H : K]$$

Montrons le résultat annoncé. Si g est un élément de G , il existe un unique indice $i \in I$ tel que $gH = g_iH$ et il existe $h \in H$ tel que $g = g_ih$. De même il existe un unique indice $j \in J$ tel que $hK = h_jK$ et h s'écrit $h = h_jk$ avec $k \in K$, ce qui donne $g = g_ih_jk \in g_ih_jK$ et $gK = g_ih_jK$. Les classes à gauche dans G modulo K sont donc les g_ih_jK pour $(i, j) \in I \times J$. Il reste à montrer que ces classes sont deux à deux distinctes. Si (i, j) et (i', j') dans $I \times J$ sont tels que $g_ih_jK = g_{i'}h_{j'}K$, il existe $k \in K$ tel que $g_ih_j = g_{i'}h_{j'}k$ et $g_i = g_{i'}(h_{j'}kh_j^{-1})$ avec $h_{j'}kh_j^{-1} \in H$, ce qui impose $g_iH = g_{i'}H$ et $i = i'$. Il en résulte que $h_j = h_{j'}k$ et $h_jK = h_{j'}K$, qui équivaut à $j = j'$.

Exercice 2.19.

1. Soient $(G, \cdot)$ un groupe, E un ensemble non vide et $f : G \rightarrow E$ une application bijective. Montrer que l'ensemble E muni de la loi $\star$ définie par $x \star y = f(f^{-1}(x) \cdot f^{-1}(y))$ est un groupe isomorphe à $(G, \cdot)$ (on dit qu'on a transporté la structure de groupe de G sur E).
2. Retrouver les résultats des exercices 2.4 et 2.5.
3. Montrer que pour tout entier impair $n \geq 1$ impair l'application $(x, y) \mapsto x \star y = \sqrt[n]{x^n + y^n}$ définit une structure de groupe commutatif sur $\mathbb{R}$.

Solution.

1. La fonction f étant bijective de G sur E l'application $\star$ définit bien une loi interne sur E . Pour tout $x \in E$, on a $x \star f(1) = f(1) \star x = x$ et $x \star f((f^{-1}(x))^{-1}) = f((f^{-1}(x))^{-1}) \star x = f(1)$ donc $f(1)$ est neutre et tout élément de E est inversible. Enfin pour x, y, z dans E , on a :

$$x \star (y \star z) = f(f^{-1}(x) \cdot f^{-1}(y \star z)) = f(f^{-1}(x) \cdot f^{-1}(y) \cdot f^{-1}(z))$$

et :

$$(x \star y) \star z = f(f^{-1}(x \star y) \cdot f^{-1}(z)) = f(f^{-1}(x) \cdot f^{-1}(y) \star f^{-1}(z))$$

ce qui montre que $\star$ est associative. Avec $f^{-1}(x \star y) = f^{-1}(x) \cdot f^{-1}(y)$, on déduit que f^{-1} est un morphisme de groupes de $(E, \star)$ sur $(G, \cdot)$ et f est un morphisme de groupes de $(G, \cdot)$ sur $(E, \star)$.

on dit qu'on a transporté la structure de groupe de $(G, \cdot)$ sur E par la bijection f .

2. Les exercices 2.4 et 2.5 sont des exemples de telle situation avec le groupe $(\mathbb{R}, +)$, $f(x) = \operatorname{th}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ pour $x \in \mathbb{R}$ qui réalise une bijection de $\mathbb{R}$ sur $] -1, 1[$ avec pour bijection réciproque argth et $f(x) = \arctan(x)$ pour $x \in \mathbb{R}$ qui réalise une bijection de $\mathbb{R}$ sur $] -\frac{\pi}{2}, \frac{\pi}{2}[$ avec pour bijection réciproque $\tan$. Dans le premier cas, on a :

$$\begin{aligned} x \star y &= f(f^{-1}(x) + f^{-1}(y)) = \operatorname{th}(\operatorname{argth}(x) + \operatorname{argth}(y)) \\ &= \frac{\operatorname{th}(\operatorname{argth}(x)) + \operatorname{th}(\operatorname{argth}(y))}{1 + \operatorname{th}(\operatorname{argth}(x)) \operatorname{th}(\operatorname{argth}(y))} = \frac{x + y}{1 + xy} \end{aligned}$$

et dans le second, on a :

$$x \star y = f(f^{-1}(x) + f^{-1}(y)) = \arctan(\tan(x) + \tan(y))$$

3. L'application $f : x \mapsto \sqrt[n]{x}$ est bijective de $\mathbb{R}$ sur $\mathbb{R}$ pour n impair, son inverse étant l'application $x \mapsto x^n$ et on a $x \star y = \sqrt[n]{x^n + y^n} = f(f^{-1}(x) + f^{-1}(y))$.

Exercice 2.20.

1. Montrer que si f est un endomorphisme du groupe additif $\mathbb{R}$, on a alors :

$$\forall a \in \mathbb{R}, \forall r \in \mathbb{Q}, f(ra) = rf(a)$$

2. Montrer que les seuls endomorphismes du groupe additif $\mathbb{R}$ qui sont monotones sont les homothéties (i. e. les applications $x \mapsto \lambda x$, où λ est une constante réelle).

Solution. Un endomorphisme du groupe additif $\mathbb{R}$ est une application $f : \mathbb{R} \rightarrow \mathbb{R}$ qui vérifie l'équation fonctionnelle de Cauchy :

$$\forall (x, y) \in \mathbb{R}^2, f(x + y) = f(x) + f(y) \quad (2.1)$$

1. En prenant $(x, y) = (0, 0)$ dans (2.1), on obtient $f(0) = 2f(0)$, ce qui équivaut à $f(0) = 0$ (un morphisme de groupes transforme le neutre en neutre).

En prenant $(x, y) = (x, -x)$ dans (2.1), on obtient $f(x) + f(-x) = 0$. On a donc $f(-x) = -f(x)$ pour tout $x \in \mathbb{R}$, c'est-à-dire que la fonction f est impaire (un morphisme de groupes transforme l'opposé en opposé). De (2.1) on déduit par récurrence que pour tout $a \in \mathbb{R}$ on a $f(na) = nf(a)$ pour tout $n \in \mathbb{N}$. En effet, le résultat est vrai pour $n = 0$ et le supposant vrai pour $n \geq 0$, on a :

$$f((n+1)a) = f(na) + f(a) = nf(a) + f(a) = (n+1)f(a),$$

il est donc vrai pour tout $n \in \mathbb{N}$. En écrivant, pour tout $n \in \mathbb{N} \setminus \{0\}$, que $f(a) = f\left(n \frac{a}{n}\right) = nf\left(\frac{a}{n}\right)$, on déduit que $f\left(\frac{a}{n}\right) = \frac{1}{n}f(a)$ pour tout $a \in \mathbb{R}$ et tout $n \in \mathbb{N} \setminus \{0\}$. Il en résulte que pour tout rationnel positif $r = \frac{p}{q}$, avec $p \in \mathbb{N}$

et $q \in \mathbb{N} \setminus \{0\}$, on a $f(ra) = f\left(p \frac{a}{q}\right) = pf\left(\frac{a}{q}\right) = \frac{p}{q}f(a) = rf(a)$. Enfin avec l'imparité de f , on déduit que ce dernier résultat est encore vrai pour les rationnels négatifs. On a donc $f(ra) = rf(a)$ pour tout $a \in \mathbb{R}$ et tout $r \in \mathbb{Q}$.

2. Soit f un endomorphisme croissant du groupe additif $\mathbb{R}$. En particulier, on a $\lambda = f(1) \geq f(0) = 0$. En désignant, pour $x \in \mathbb{R}$, par $(r_n)_{n \in \mathbb{N}}$ et $(s_n)_{n \in \mathbb{N}}$ des suites d'approximations décimales par défaut et par excès de ce réel, on a pour tout $n \in \mathbb{N}$:

$$\lambda r_n = f(r_n) \leq f(x) \leq f(s_n) = \lambda s_n$$

et faisant tendre n vers l'infini, on en déduit que $f(x) = \lambda x$. On procède de manière analogue pour f décroissante.

Exercice 2.21. Déterminer l'ordre d'un élément du groupe $\mathbb{C}^*$.

Solution. Tout nombre complexe non nul s'écrit $z = \rho e^{i\alpha}$ où $\rho \in \mathbb{R}^{+,*}$ et $\alpha \in [0, 2\pi[$ (avec un tel choix de α , cette écriture est unique). Si $\rho \neq 1$, on a $|z^k| = \rho^k \neq 1$ pour tout entier relatif k , donc $z^k \neq z^j$ pour $k \neq j$ dans $\mathbb{Z}$ et $\langle z \rangle$ est infini. Si $\rho = 1$, on a alors, pour k entier relatif non nul, $z^k = e^{ik\alpha} = 1$ si, et seulement si, il existe un entier relatif q tel que $k\alpha = 2q\pi$, ce qui signifie que $\frac{\alpha}{2\pi}$ est rationnel. On en déduit donc que :

- pour $\frac{\alpha}{2\pi}$ irrationnel, $z^k \neq 1$ pour tout entier relatif k et $\langle z \rangle$ est infini ;
- pour $\frac{\alpha}{2\pi} = \frac{p}{q}$ rationnel avec $(p, q) \in \mathbb{Z} \times \mathbb{N}^*$ et $p \wedge q = 1$, en effectuant la division euclidienne d'un entier relatif k par q , on a $k = mq + r$ avec $0 \leq r \leq q - 1$ et :

$$z^k = e^{ik\alpha} = (e^{iq\alpha})^m e^{ir\alpha} = (e^{2ip\pi})^m e^{ir\alpha} = e^{ir\alpha}$$

et $\langle z \rangle = \{e^{ir\alpha} \mid 0 \leq r \leq q - 1\}$ a au plus q éléments. Pour $0 \leq r \neq s \leq q - 1$ l'égalité $e^{ir\alpha} = e^{is\alpha}$ équivaut à $e^{i(s-r)\alpha} = 1$, ce qui revient à dire $(s - r)\alpha = 2m\pi$ avec $m \in \mathbb{Z}$, qui tenant compte de $\alpha = 2\pi\frac{p}{q}$, donne $(s - r)\frac{p}{q} = m$, soit q divise $p(s - r)$ sachant que q est premier avec p , donc q divise $s - r$ (théorème de Gauss) et nécessairement $r = s$ puisque $|r - s| \leq q - 1$. On a donc exactement q éléments dans $\langle z \rangle$ et z est d'ordre q . En fait $\langle z \rangle$ est le groupe Γ_q des racines q -èmes de l'unité. En définitive, on a :

$$\theta(\rho e^{i\alpha}) = \begin{cases} +\infty & \text{si } \rho \neq 1 \text{ ou } \rho = 1 \text{ et } \frac{\alpha}{2\pi} \text{ irrationnel} \\ q & \text{si } \frac{\alpha}{2\pi} = \frac{p}{q} \in \mathbb{Q} \text{ avec } p \wedge q = 1 \end{cases}$$

Exercice 2.22. Déterminer les sous-groupes finis du groupe $\mathbb{C}^*$.

Solution. Si $G \subset \mathbb{C}^*$ est un groupe d'ordre $n \geq 1$, on a alors $z^n = 1$ pour $z \in G$ et G est contenu dans l'ensemble $\Gamma_n = \left\{ e^{2i\frac{k\pi}{n}} \mid 0 \leq k \leq n - 1 \right\}$ des racines n -èmes de l'unité qui est lui même un groupe d'ordre n . On a donc $G = \Gamma_n$.

Exercice 2.23. Soit G un groupe fini d'ordre m . Montrer que pour tout entier relatif n premier avec m , l'application $g \mapsto g^n$ est une bijection de G sur lui même (c'est donc une permutation de G).

Solution. Comme $m \wedge n = 1$, le théorème de Bézout nous dit qu'il existe deux entiers relatifs u et v tels que $un + vm = 1$ et pour tout $g \in G$, on a $g = g^{un+vm} = (g^u)^n (g^m)^v = (g^u)^n$, ce qui signifie que l'application $g \mapsto g^n$ est surjective. Comme G est fini, cette application est bijective.

Exercice 2.24. Montrer qu'un groupe G est fini si, et seulement si, l'ensemble de ses sous-groupes est fini.

Solution. Si G est un groupe fini alors l'ensemble $\mathcal{P}(G)$ des parties de G est fini (de cardinal $2^{\text{card}(G)}$) et il en est de même de l'ensemble des sous-groupes de G .

Réciproquement soit $(G, \cdot)$ un groupe tel que l'ensemble de ses sous-groupes soit fini. On peut écrire $G = \bigcup_{g \in G} \langle g \rangle$ et cette réunion est finie, soit $G = \bigcup_{k=1}^r \langle g_k \rangle$. Si l'un

de ces sous-groupes $\langle g_k \rangle$ est infini, alors les $\langle g_k^n \rangle$ où n décrit $\mathbb{N}$ forment une famille infinie de sous-groupes de G : en effet l'égalité $\langle g_k^n \rangle = \langle g_k^m \rangle$ entraîne $g_k^n = g_k^{jm}$, soit $g_k^{n-jm} = 1$ et $n - jm = 0$ (g_k est d'ordre infini), c'est-à-dire que m divise n . Comme n et m jouent des rôles symétriques, on a aussi n qui divise m et en définitive $n = m$ (on peut aussi dire que $\langle g_k \rangle$ est isomorphe à $\mathbb{Z}$ et de ce fait a une infinité de sous-groupes). On a donc une contradiction si l'un des $\langle g_k \rangle$ est infini. Donc tous les $\langle g_k \rangle$ sont finis et aussi G .

Chapitre 3

Structures d'anneau et de corps

3.1 Anneaux

Définition 3.1. Soit A un ensemble non vide muni de deux lois de composition interne notées $+$ (une addition) et $\cdot$ (une multiplication). On dit que $(A, +, \cdot)$ est un anneau si :

- $(A, +)$ est un groupe commutatif ;
- la loi $\cdot$ est associative ;
- la loi $\cdot$ est distributive par rapport à la loi $+$, ce qui signifie que :

$$\forall (a, b, c) \in A^3, \begin{cases} a \cdot (b + c) = a \cdot b + a \cdot c \\ (b + c) \cdot a = b \cdot a + c \cdot a \end{cases}$$

Si la loi $\cdot$ est commutative, on dit alors que l'anneau A est commutatif. S'il existe un élément neutre pour la loi $\cdot$, on dira alors que A est un anneau unitaire.

Si $(A, +, \cdot)$ est un anneau, on notera 0 le neutre pour l'addition et s'il existe on notera 1 le neutre pour la multiplication. L'opposé d'un élément a (i. e. le symétrique pour $+$) sera noté $-a$ et on notera $a - b$ pour $a + (-b)$. On écrira souvent ab pour $a \cdot b$ dans un anneau. Dans un anneau unitaire, on supposera que $0 \neq 1$ (sans quoi l'anneau est réduit à $\{0\}$). Un anneau unitaire a donc au moins deux éléments.

Exemples 3.1

1. Les ensembles $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ muni des opérations usuelles sont des anneaux commutatifs et unitaires.
2. Soient E un ensemble non vide et A un anneau. On vérifie facilement que l'ensemble A^E des applications de E dans A muni des opérations d'addition et

de multiplication définies par :

$$\forall (f, g) \in A^E \times A^E, \forall x \in E, \begin{cases} (f + g)(x) = f(x) + g(x) \\ (f \cdot g)(x) = f(x) \cdot g(x) \end{cases}$$

est un anneau. Cet anneau est commutatif si A l'est et il est unitaire si A l'est avec comme élément neutre pour le produit l'application constante égale à 1. En particulier l'ensemble $\mathbb{R}^{\mathbb{N}}$ des suites réelles est un anneau commutatif unitaire et pour toute partie non vide I de R , l'ensemble R^I des fonctions définies sur I et à valeurs réelles est un anneau commutatif unitaire.

Si $A_1, \dots, A_p$ sont des anneaux, on peut alors munir le produit direct $\prod_{k=1}^p A_k$ d'une structure d'anneau comme dans le cas où $p = 2$. Si $A_k = A$ pour tout k compris entre 1 et p , on note alors A^p cet anneau produit.

Avec le théorème qui suit, on donne un résumé des règles de calculs utilisables dans un anneau.

Théorème 3.1.

Dans un anneau $(A, +, \cdot)$, on a les règles de calcul suivantes, où $a, b, c, a_1, \dots, a_p$ sont des éléments de A , p un entier naturel non nul et n un entier relatif :

$$\begin{aligned} & - a \cdot 0 = 0 \cdot a = 0; (-a) \cdot b = a \cdot (-b) = -a \cdot b; \\ & - (-a) \cdot (-b) = a \cdot b; (a - b) \cdot c = a \cdot c - b \cdot c; \\ & - a \cdot (b - c) = a \cdot b - a \cdot c; n(a \cdot b) = (na) \cdot b = a \cdot (nb); \\ & - a \cdot \left(\sum_{k=1}^p b_k \right) = \sum_{k=1}^p a \cdot b_k; \left(\sum_{k=1}^p b_k \right) \cdot a = \sum_{k=1}^p b_k \cdot a; \end{aligned}$$

Preuve.

- On a $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$ et simplifiant par $a \cdot 0$ dans le groupe $(A, +)$, on en déduit que $a \cdot 0 = 0$.
- De $0 = 0 \cdot b = (a + (-a)) \cdot b = a \cdot b + (-a) \cdot b$, on déduit que $(-a) \cdot b = -a \cdot b$.
- On en déduit que $(-a) \cdot (-b) = -a \cdot (-b) = -(-a \cdot b) = a \cdot b$ (dans un groupe, l'opposé de l'opposé est l'élément).
- On en déduit aussi que $(a - b) \cdot c = a \cdot c + (-b) \cdot c = a \cdot c - b \cdot c$.
- On montre d'abord par récurrence sur $n \geq 0$ que $n(a \cdot b) = (na) \cdot b$. C'est vrai pour $n = 0$ et supposant que c'est vrai pour $n \geq 0$, on a :

$$(n + 1)(a \cdot b) = n(a \cdot b) + a \cdot b = (na) \cdot b + a \cdot b = (na + a) \cdot b = ((n + 1)a) \cdot b$$

Ensuite avec $(-n)(a \cdot b) = -n(a \cdot b) = -(na) \cdot b = (-na) \cdot b$, on déduit que le résultat est valable pour les entiers relatifs.

- Les deux derniers résultats se montrent facilement par récurrence sur $p \geq 1$.

□

Sur $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ ou $\mathbb{C}$, une formule intéressante est celle du binôme de Newton. Elle est en fait valable sur anneau unitaire quand les éléments commutent.

Théorème 3.2.

Soit $(A, +, \cdot)$ un anneau unitaire. Si a et b commutent dans A , on a alors $(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$ pour tout entier naturel n (formule du binôme de Newton).

Preuve. Pour $n = 0$ et $n = 1$, c'est évident. En supposant le résultat acquis au rang $n \geq 1$, on a :

$$\begin{aligned}
 (a + b)^{n+1} &= (a + b)^n (a + b) = \left(\sum_{k=0}^n \binom{n}{k} a^k b^{n-k} \right) (a + b) \\
 &= \sum_{k=0}^n \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^n \binom{n}{k} a^k b^{n-(k-1)} \\
 &= \sum_{k=1}^{n+1} \binom{n}{k-1} a^k b^{n-(k-1)} + \sum_{k=0}^n \binom{n}{k} a^k b^{n-(k-1)} \\
 &= b^{n+1} + \sum_{k=1}^n \left(\binom{n}{k-1} + \binom{n}{k} \right) a^k b^{n+1-k} + a^{n+1}
 \end{aligned}$$

et tenant compte de $\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$ (triangle de Pascal), cela s'écrit :

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

Le résultat est donc vrai pour tout $n \geq 0$. $\square$

Si a et b ne commutent pas, la formule du binôme n'est plus nécessairement vraie (voir le calcul matriciel au paragraphe 6.6).

La formule du binôme peut aussi se montrer en utilisant un argument de dénombrement. Comme a et b commutent, $(a + b)^n = (a + b) \cdots (a + b)$, le produit étant effectué n fois, est une somme de monômes $a^k b^{n-k}$ et, pour k fixé entre 0 et n , il y a autant de monômes $a^k b^{n-k}$ que de produits $aabaa \cdots$ où a intervient k fois et b intervient $n - k$ fois. Dans une telle liste, il y a $\binom{n}{k}$ façons de choisir la position des k éléments a (les a étant placés, les b le sont automatiquement), ce qui donne la formule.

L'identité remarquable qui suit, pour a et b qui commutent, est aussi intéressante.

Théorème 3.3.

Soit $(A, +, \cdot)$ un anneau unitaire. Si a et b commutent dans A , on a alors $b^{n+1} - a^{n+1} = (b - a) \sum_{k=0}^n a^k b^{n-k}$ pour tout entier naturel n .

Preuve. On procède par récurrence sur $n \geq 0$. Pour $n = 0$, c'est évident. En supposant le résultat acquis au rang $n \geq 0$, on a :

$$\begin{aligned}
 b^{n+2} - a^{n+2} &= (b^{n+1} - a^{n+1})b + ba^{n+1} - a^{n+2} \\
 &= (b - a) \sum_{k=0}^n a^k b^{n+1-k} + (b - a) a^{n+1} \\
 &= (b - a) (b^{n+1} + ab^n + \dots + a^{n-1}b^2 + a^n b) + (b - a) a^{n+1} \\
 &= (b - a) \sum_{k=0}^{n+1} a^k b^{n+1-k}
 \end{aligned}$$

Le résultat est donc vrai pour tout $n \geq 0$. $\square$

Là encore, si a et b ne commutent pas, ce résultat n'est plus nécessairement vraie (voir le calcul matriciel au paragraphe 6.6).

3.2 Éléments inversibles dans un anneau unitaire

Pour ce paragraphe $(A, +, \cdot)$ est un anneau unitaire. On note $A^\times$ l'ensemble des éléments de A inversibles pour la multiplication, c'est-à-dire l'ensemble des éléments $a \in A$ pour lesquels il existe un élément $a' \in A$ tel que $a \cdot a' = a' \cdot a = 1$. Quand il existe un tel inverse est unique et on le note a^{-1} . Comme $1 \neq 0$ dans A , on a $A^\times \subset A \setminus \{0\}$ et cette inclusion peut être stricte.

L'exercice 2.3 nous dit que pour vérifier qu'un élément a de l'anneau unitaire A est inversible, il suffit de vérifier qu'il a un inverse à gauche (ou à droite) puisque la loi multiplicative est associative.

Théorème 3.4.

Soit $(A, +, \cdot)$ un anneau unitaire. L'ensemble $A^\times$ des éléments inversibles de A est un groupe pour le produit.

Preuve. $A^\times$ est non vide puisqu'il contient 1. Si a, b sont dans $A^\times$, on a alors :

$$b^{-1}a^{-1}ab = b^{-1}b = 1, \quad abb^{-1}a^{-1} = aa^{-1} = 1$$

c'est-à-dire que ab est inversible d'inverse $b^{-1}a^{-1}$. La multiplication définit donc une loi interne sur $A^\times$. On sait déjà que cette loi est associative, que 1 en est le neutre et tout $a \in A^\times$ est inversible par construction d'inverse $a^{-1} \in A^\times$ (on $(a^{-1})^{-1} = a$). $(A^\times, \cdot)$ est donc un groupe. $\square$

Définition 3.2. On dit que $A^\times$ est le groupe des unités de A .

Définition 3.3. On dit que $a \in A$ est un diviseur de 0 si $a \neq 0$ et s'il existe $b \neq 0$ dans A tel que $a \cdot b = 0$.

Un diviseur de 0 dans un anneau unitaire n'est jamais inversible (pour la multiplication) et, par contraposée, un élément inversible ne peut être un diviseur de 0. Si a est un diviseur de 0, une égalité de la forme $a \cdot b = a \cdot c$ ne peut être simplifiée a priori. Un élément simplifiable pour le produit ne peut donc être un diviseur de 0.

Définition 3.4. Un anneau est dit *intègre* s'il est commutatif et n'admet pas de diviseur de 0.

Dans un anneau intègre, on a :

$$(a \cdot b = 0) \Leftrightarrow (a = 0 \text{ ou } b = 0)$$

Exemples 3.2

1. Dans un anneau de Boole $(\mathcal{P}(E), \Delta, \cap)$, on a $A \cap (E \setminus A) = \emptyset$ pour toute partie A de E , donc tout $A \neq \emptyset$ est un diviseur de 0 (le 0 pour la loi Δ). Donc cet anneau n'est pas intègre.
2. Les anneaux $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ sont intègres.
3. Soient A_1, A_2 deux anneaux. Dans l'anneau produit $A_1 \times A_2$, on a $(a_1, 0) \cdot (0, a_2) = (0, 0)$, c'est-à-dire que pour $a_1 \neq 0$ et $a_2 \neq 0$, $(a_1, 0)$ et $(0, a_2)$ sont des diviseurs de 0. Donc, pour A_1 et A_2 non réduits à $\{0\}$, $A_1 \times A_2$ n'est jamais intègre.

3.3 Sous-anneaux

Définition 3.5. Soit $(A, +, \cdot)$ un anneau. Un sous-anneau de A est une partie non vide B de A telle que $(B, +)$ est un sous-groupe de A et B est stable pour la multiplication, c'est-à-dire que pour tous a, b dans B , $a \cdot b$ est aussi dans B . Si l'anneau A est unitaire, B doit contenir 1.

Il est facile de vérifier qu'un sous-anneau d'un anneau est lui-même un anneau.

Théorème 3.5.

Soit $(A, +, \cdot)$ un anneau et B une partie non vide de A . B est un sous-anneau de A si, et seulement si :

$$\forall (a, b) \in B^2, \begin{cases} a - b \in B \\ a \cdot b \in B \end{cases}$$

(pour A unitaire, il faut ajouter $1 \in B$).

Preuve. Laisée au lecteur. □

Exemple 3.1 Les ensembles $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ munis des opérations usuelles sont des sous-anneaux de $\mathbb{C}$.

3.4 Morphismes d'anneaux

Les anneaux considérés sont supposés unitaires.

On désigne par $(A, +, \cdot)$ et $(B, +, \cdot)$ deux anneaux unitaires. On note respectivement 0 et 1 les éléments neutres de ces anneaux pour l'addition et la multiplication (en cas d'ambiguïté, on les notera 0_A , 0_B , 1_A et 1_B).

Définition 3.6. On dit que φ est un morphisme d'anneaux de A dans B si φ est une application de A dans B telle que :

- $\varphi(1) = 1$;
- $\forall (a, b) \in A^2, \varphi(a + b) = \varphi(a) + \varphi(b)$;
- $\forall (a, b) \in A^2, \varphi(a \cdot b) = \varphi(a) \cdot \varphi(b)$

Dans le cas où φ est de plus bijective, on dit que φ est un isomorphisme d'anneaux A sur B . Dans le cas où $A = B$, on dit que φ est un endomorphisme de l'anneau A et que c'est un automorphisme de l'anneau A si φ est de plus bijective.

On peut remarquer qu'un morphisme d'anneaux de A dans B est en particulier un morphisme de groupes de $(A, +)$ dans $(B, +)$. On a donc $\varphi(0) = 0$ et $\varphi(-a) = -\varphi(a)$ pour tout $a \in A$.

Définition 3.7. Soit φ un morphisme d'anneaux de A dans B . Le noyau de φ est l'ensemble $\ker(\varphi) = \{x \in A \mid \varphi(x) = 0\}$ et l'image de φ est l'ensemble $\text{Im}(\varphi) = \{\varphi(x) \mid x \in A\}$.

Il est facile de vérifier que $\ker(\varphi)$ est un sous-anneau de A et $\text{Im}(\varphi)$ un sous-anneau de B . En fait pour tout $x \in \ker(\varphi)$ et tout $y \in A$, on a $\varphi(xy) = \varphi(x)\varphi(y) = 0 \cdot \varphi(y) = 0$, c'est-à-dire que $xy \in \ker(\varphi)$. Un tel morphisme est injectif [resp. surjectif] si, et seulement si, $\ker(\varphi) = \{0\}$ [resp. $\text{Im}(\varphi) = B$].

3.5 Corps

Définition 3.8. Soit $\mathbb{K}$ un ensemble non vide muni de deux lois de composition interne notées $+$ (une addition) et $\cdot$ (une multiplication). On dit que $(\mathbb{K}, +, \cdot)$ est un corps si :

- $(\mathbb{K}, +, \cdot)$ est un anneau unitaire (avec $1 \neq 0$) ;
- tous les éléments de $\mathbb{K} \setminus \{0\}$ sont inversibles pour la multiplication (ce qui revient à dire que $\mathbb{K}^\times = \mathbb{K} \setminus \{0\}$).

Si de plus l'anneau $(\mathbb{K}, +, \cdot)$ est commutatif, on dit que le corps $(\mathbb{K}, +, \cdot)$ est commutatif.

Pour un corps on note aussi $\mathbb{K}^*$ l'ensemble $\mathbb{K}^\times = \mathbb{K} \setminus \{0\}$.

Dire que $(\mathbb{K}, +, \cdot)$ est un corps équivaut aussi à dire que :

- $(\mathbb{K}, +, \cdot)$ est un anneau unitaire ;
- $(\mathbb{K}^*, \cdot)$ est un groupe.

Dans un corps, on notera $-a$ l'opposé d'un élément a (i. e. le symétrique pour la loi $+$) et a^{-1} ou $\frac{1}{a}$ l'inverse d'un élément non nul a (i. e. le symétrique pour la loi $\cdot$).

Dans un corps tout élément non nul est simplifiable et il n'y a pas de diviseurs de 0. Un corps est donc en particulier un anneau intègre.

Les règles de calcul valables dans un anneau (exercice 3.1) le sont aussi dans un corps avec de plus l'équivalence :

$$(ab = 0) \Leftrightarrow (a = 0 \text{ ou } b = 0)$$

Dans un corps commutatif, pour $(a, b) \in \mathbb{K}^* \times \mathbb{K}$, on écrira $a^{-1} \cdot b = \frac{b}{a}$ (si le corps n'est pas commutatif on a, a priori, $a^{-1} \cdot b \neq b \cdot a^{-1}$ et l'écriture $\frac{b}{a}$ est ambiguë).

Exemple 3.2 Les ensembles $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ muni des opérations usuelles sont des corps commutatifs. Mais $\mathbb{Z}$ n'est pas un corps.

Définition 3.9. Soit $(\mathbb{K}, +, \cdot)$ un corps. On dit qu'une partie $\mathbb{L}$ de $\mathbb{K}$ est un sous-corps de $\mathbb{K}$ si :

- $\mathbb{L}$ est un sous-anneau de $\mathbb{K}$;
- $\mathbb{L}^* = \mathbb{L} \setminus \{0\}$ est stable par passage à l'inverse, c'est-à-dire que pour tout $x \in \mathbb{L}^*$, x^{-1} est dans $\mathbb{L}^*$.

On vérifie facilement qu'un sous-corps d'un corps est lui même un corps.

Théorème 3.6.

Soit $(\mathbb{K}, +, \cdot)$ un corps et $\mathbb{L}$ une partie non vide de $\mathbb{K}$. $\mathbb{L}$ est un sous-corps de $\mathbb{K}$ si, et seulement si :

- $1 \in \mathbb{L}$;
- $\forall (x, y) \in \mathbb{L}^2, x - y \in \mathbb{L}$;
- $\forall (x, y) \in \mathbb{L} \times \mathbb{L}^*, xy^{-1} \in \mathbb{L}$.

Preuve. Laissée au lecteur. □

Si $\mathbb{L}$ est un sous-corps d'un corps $\mathbb{K}$, on dit alors que $\mathbb{K}$ est une extension de $\mathbb{L}$.

Exemple 3.3 Les ensembles $\mathbb{Q}, \mathbb{R}$ muni des opérations usuelles sont des sous-corps de $\mathbb{C}$.

3.6 Morphismes de corps

On désigne par $(\mathbb{K}, +, \cdot)$ et $(\mathbb{L}, +, \cdot)$ deux corps. On note respectivement 0 et 1 les éléments neutres de ces corps pour l'addition et la multiplication (en cas d'ambiguïté, on les notera $0_{\mathbb{K}}$, $0_{\mathbb{L}}$, $1_{\mathbb{K}}$ et $1_{\mathbb{L}}$).

Définition 3.10. On dit que φ est un morphisme de corps de $\mathbb{K}$ dans $\mathbb{L}$ si φ est une application de $\mathbb{K}$ dans $\mathbb{L}$ telle que :

- $\varphi(1_{\mathbb{K}}) = 1_{\mathbb{L}}$;
- $\forall (a, b) \in \mathbb{K}^2, \varphi(a + b) = \varphi(a) + \varphi(b)$;
- $\forall (a, b) \in \mathbb{K}^2, \varphi(a \cdot b) = \varphi(a) \cdot \varphi(b)$

Dans le cas où φ est de plus bijective, on dit que φ est un isomorphisme de corps de $\mathbb{K}$ sur $\mathbb{L}$. Dans le cas où $\mathbb{K} = \mathbb{L}$, on dit que φ est un endomorphisme du corps $\mathbb{K}$ et que c'est un automorphisme du corps $\mathbb{K}$ si φ est de plus bijective.

On peut remarquer qu'un morphisme de corps est en fait un morphisme d'anneaux unitaires. On a, pour un tel morphisme, $\varphi(0) = 0$, $\varphi(1) = 1$, $\varphi(-a) = -\varphi(a)$ pour tout $a \in \mathbb{K}$ et $\varphi(a^{-1}) = \varphi(a)^{-1}$ pour tout $a \in \mathbb{K}^*$.

3.7 Exercices

Exercice 3.1. Soit E un ensemble non vide. Montrer que l'ensemble $\mathcal{P}(E)$ des parties de E muni des opérations Δ de différence symétrique et $\cap$ d'intersection est un anneau commutatif et unitaire (c'est l'anneau de Boole).

Solution. On rappelle que pour A, B dans $\mathcal{P}(E)$, on a $A \Delta B = (A \cup B) \setminus (A \cap B) = (A \setminus B) \cup (B \setminus A)$, la réunion étant disjointe. De la commutativité des opérateurs $\cup$ et $\cap$, on déduit que Δ est commutative. Pour A, B, C dans $\mathcal{P}(E)$, on a :

$$\begin{aligned}
 (x \in (A \Delta B) \Delta C) &\Leftrightarrow (x \in A \Delta B \text{ et } x \notin C) \text{ ou } (x \in C \text{ et } x \notin A \Delta B) \\
 &\Leftrightarrow (x \in A \text{ et } x \notin B \text{ et } x \notin C) \text{ ou } (x \in B \text{ et } x \notin A \text{ et } x \notin C) \\
 &\text{ou } (x \in C \text{ et } x \notin A \text{ et } x \notin B) \text{ ou } (x \in C \text{ et } x \in A \cap B) \\
 &\Leftrightarrow (x \in A \text{ et } x \notin B \text{ et } x \notin C) \text{ ou } (x \in B \text{ et } x \notin A \text{ et } x \notin C) \\
 &\text{ou } (x \in C \text{ et } x \notin A \text{ et } x \notin B) \text{ ou } (x \in A \cap B \cap C)
 \end{aligned}$$

et :

$$\begin{aligned}
 (x \in A \Delta (B \Delta C)) &\Leftrightarrow (x \in A \text{ et } x \notin B \Delta C) \text{ ou } (x \in B \Delta C \text{ et } x \notin A) \\
 &\Leftrightarrow (x \in A \text{ et } x \notin B \text{ et } x \notin C) \text{ ou } (x \in A \text{ et } x \in B \cap C) \\
 &\text{ou } (x \in B \text{ et } x \notin C \text{ et } x \notin A) \text{ ou } (x \in C \text{ et } x \notin B \text{ et } x \notin A) \\
 &\Leftrightarrow (x \in A \text{ et } x \notin B \text{ et } x \notin C) \text{ ou } (x \in A \cap B \cap C) \\
 &\text{ou } (x \in B \text{ et } x \notin C \text{ et } x \notin A) \text{ ou } (x \in C \text{ et } x \notin B \text{ et } x \notin A)
 \end{aligned}$$

D'où l'égalité $(A \Delta B) \Delta C = A \Delta (B \Delta C)$. L'ensemble vide est le neutre pour Δ et pour tout $A \in \mathcal{P}(E)$, on a $A \Delta A = (A \cup A) \setminus (A \cap A) = A \setminus A = \emptyset$, c'est-à-dire que A est l'opposé de A pour la loi Δ (tous les éléments de $\mathcal{P}(E) \setminus \{\emptyset\}$ sont d'ordre 2, ce qui permet de retrouver la commutativité de $(\mathcal{P}(E), \Delta)$ et le fait que $\mathcal{P}(E)$ est de cardinal une puissance de 2 si E est fini). En définitive, $(\mathcal{P}(E), \Delta)$ est un groupe commutatif. On vérifie facilement que $\cap$ est commutative et associative. L'ensemble E est le neutre pour $\cap$. Pour A, B, C dans $\mathcal{P}(E)$, on a :

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

c'est-à-dire que $\cap$ est distributive par rapport à Δ . En définitive, $(\mathcal{P}(E), \Delta, \cap)$ est un anneau commutatif et unitaire.

Exercice 3.2. Soit $(A, +, \cdot)$ un anneau unitaire. Montrer que si $a \in A$ est tel que $a^n = 0$ pour un entier $n \geq 1$ (on dit alors que a est nilpotent), alors $1 - a$ est inversible d'inverse $\sum_{k=0}^{n-1} a^k$.

Solution. On a $1 - a^n = (1 - a) \sum_{k=0}^{n-1} a^k$ pour tout $n \geq 1$ et $a^n = 0$ nous donne

$$(1 - a) \sum_{k=0}^{n-1} a^k = 1, \text{ ce qui signifie que } 1 - a \text{ est inversible d'inverse } \sum_{k=0}^{n-1} a^k.$$

Exercice 3.3. On appelle nombre décimal tout nombre rationnel de la forme $\frac{a}{10^m}$ où a est un entier relatif et m un entier naturel.

1. Montrer que l'ensemble D des nombres décimaux est un anneau unitaire, commutatif et intègre.
2. Montrer que l'ensemble des nombres décimaux inversibles est $D^\times = \{r = \pm 2^\alpha 5^\beta \mid (\alpha, \beta) \in \mathbb{Z}^2\}$.

Solution.

1. Facile.
2. Un rationnel $r = \frac{a}{10^m}$ est inversible dans $\mathbb{D}$ si, et seulement si, il existe un entier relatif b et un entier naturel n tels que $\frac{a}{10^m} \frac{b}{10^n} = 1$, ce qui revient à dire que $ab = 10^{n+m}$ ou encore que 2 et 5 sont les seuls diviseurs premiers possibles de a et b .

Exercice 3.4. Soit $p \geq 2$ un entier sans facteurs carrés dans sa décomposition en produit de nombres premiers (c'est-à-dire que $p = \prod_{k=1}^r p_k$ où les p_k sont premiers deux à deux distincts).

1. Montrer que l'ensemble $\mathbb{Z}[\sqrt{p}] = \{n + m\sqrt{p} \mid (n, m) \in \mathbb{Z}^2\}$ est un sous-anneau de $\mathbb{R}$.
2. Montrer que $n + m\sqrt{p} = 0$ dans $\mathbb{Z}[\sqrt{p}]$ si, et seulement si, $n = m = 0$ (ce qui signifie que l'écriture $a = n + m\sqrt{p}$ d'un élément de $\mathbb{Z}[\sqrt{p}]$ est unique).
3. Quels sont les éléments de $\mathbb{Z}$ (qui est contenu dans $\mathbb{Z}[\sqrt{p}]$) qui sont inversibles.
4. Montrer que si $n + m\sqrt{p}$ est inversible dans $\mathbb{Z}[\sqrt{p}]$, il en est alors de même de $n - m\sqrt{p}$.
5. Montrer que le groupe des éléments inversibles de $\mathbb{Z}[\sqrt{p}]$ est :

$$(\mathbb{Z}[\sqrt{p}])^\times = \{n + m\sqrt{p} \in \mathbb{Z}[\sqrt{p}] \mid n^2 - pm^2 = \pm 1\}$$

Solution.

1. On a $1 = 1 + 0\sqrt{p} \in \mathbb{Z}[\sqrt{p}]$. Pour $a = n + m\sqrt{p}$ et $a' = n' + m'\sqrt{p}$ dans $\mathbb{Z}[\sqrt{p}]$, on a :

$$\begin{cases} a - a' = (n - n') + (m - m')\sqrt{p} \in \mathbb{Z}[\sqrt{p}] \\ aa' = (nn' + pmm') + (nm' + mn')\sqrt{p} \in \mathbb{Z}[\sqrt{p}] \end{cases}$$

Donc $\mathbb{Z}[\sqrt{p}]$ est un sous-anneau de $\mathbb{R}$.

2. Si $a = n + m\sqrt{p} = 0$ avec $m \neq 0$, on a alors $\sqrt{p} = -\frac{n}{m} \in \mathbb{Q}$, ce qui n'est pas possible si p est sans facteurs carrés. En effet $\sqrt{p} = \frac{a}{b}$ avec a, b premiers entre eux dans $\mathbb{N}^*$, donne $a^2 = pb^2$, donc p_1 divise a , soit $a = p_1 a_1$ et $p_1^2 a_1^2 = pb^2$, soit $p_1 a_1^2 = \prod_{k=2}^r p_k b^2$ et p_1 va diviser b (il est premier avec $\prod_{k=2}^r p_k$ dans le cas où $r \geq 2$), ce qui contredit $a \wedge b = 1$. L'égalité $n + m\sqrt{p} = 0$ entraîne donc $m = 0$ et $n = 0$.
3. Si $n \in \mathbb{Z} \subset \mathbb{Z}[\sqrt{p}]$ est inversible, il existe alors $n' + m'\sqrt{p} \in \mathbb{Z}[\sqrt{p}]$ tel que $n(n' + m'\sqrt{p}) = nn' + nm'\sqrt{p} = 1$, ce qui entraîne $nn' = 1$ et $nm' = 0$, soit

$m' = 0$ et $nn' = 1$ dans $\mathbb{Z}$, ce qui donne $n = n' = \pm 1$. Donc :

$$\mathbb{Z} \cap (\mathbb{Z}[\sqrt{p}])^\times = \mathbb{Z}^\times = \{-1, 1\}$$

4. Si $a = n + m\sqrt{p}$ est inversible dans $\mathbb{Z}[\sqrt{p}]$, il existe alors $a' = n' + m'\sqrt{p}$ dans $\mathbb{Z}[\sqrt{p}]$ tel que $aa' = 1$, soit $(nn' + pmm') + (nm' + mn')\sqrt{p} = 1$, ce qui entraîne $nn' + pmm' = 1$ et $nm' + mn' = 0$ (unicité de l'écriture $n + m\sqrt{p}$ dans $\mathbb{Z}[\sqrt{p}]$). Il en résulte que $(n - m\sqrt{p})(n' - m'\sqrt{p}) = (nn' + pmm') - (nm' + mn')\sqrt{p} = 1$ et $n - m\sqrt{p}$ est inversible dans $\mathbb{Z}[\sqrt{p}]$.
5. Si $a = n + m\sqrt{p}$ est inversible dans $\mathbb{Z}[\sqrt{p}]$, il en est alors de même de $n - m\sqrt{p}$ et du produit $(n + m\sqrt{p})(n - m\sqrt{p}) = n^2 - pm^2$ ($\mathbb{Z}[\sqrt{p}]$ est un groupe multiplicatif), ce qui entraîne $n^2 - pm^2 = \pm 1$. Réciproquement, si n et m sont tels que $n^2 - pm^2 = \pm 1$, on a alors $(n + m\sqrt{p})(n - m\sqrt{p}) = \pm 1$ et $n + m\sqrt{p}$ est inversible d'inverse $\pm(n - m\sqrt{p})$. On peut montrer que les éléments inversibles de $\mathbb{Z}[\sqrt{2}]$ sont les éléments de la forme $\pm(1 + \sqrt{2})^n$ où n est un entier relatif, l'inverse de $\pm(1 + \sqrt{2})^n$ étant $\pm(-1 + \sqrt{2})^n$.

Exercice 3.5. On désigne par p un entier naturel non nul et par $\mathbb{Z}[i\sqrt{p}]$ l'ensemble des nombres complexes défini par :

$$\mathbb{Z}[i\sqrt{p}] = \{a + ib\sqrt{p} \mid (a, b) \in \mathbb{Z}^2\}.$$

1. Montrer que $\mathbb{Z}[i\sqrt{p}]$ est un anneau unitaire commutatif et intègre (pour $p = 1$, $\mathbb{Z}[i]$ est l'anneau des entiers de Gauss).
2. Montrer que $\mathbb{Z}[i\sqrt{p}]$ est contenu dans tout sous-anneau unitaire de $\mathbb{C}$ qui contient $i\sqrt{p}$. L'anneau $\mathbb{Z}[i\sqrt{p}]$ est donc le plus petit sous-anneau de $\mathbb{C}$ (pour l'ordre de l'inclusion) qui contient $i\sqrt{p}$, on dit que c'est le sous-anneau de $\mathbb{C}$ engendré par $i\sqrt{p}$.
3. Montrer que $\mathbb{Z}[i\sqrt{p}]$ est égal à l'intersection de tous les sous-anneaux de $\mathbb{C}$ qui contiennent i .
4. Déterminer le groupe $\mathbb{Z}[i\sqrt{p}]^\times$ des éléments inversibles de $\mathbb{Z}[i\sqrt{p}]$.

Solution.

1. Il suffit de montrer que $\mathbb{Z}[i\sqrt{p}]$ est un sous-anneau de $\mathbb{C}$. On a $1 = 1 + i \cdot 0 \cdot \sqrt{p} \in \mathbb{Z}[i\sqrt{p}]$. Pour $z = a + ib\sqrt{p}$ et $z' = a' + ib'\sqrt{p}$, où a, a', b, b' sont des entiers relatifs, on a :

$$\begin{cases} z - z' = (a - a') + (b - b')i\sqrt{p} \in \mathbb{Z}[i\sqrt{p}] \\ zz' = (aa' - pbb') + (ab' + ba')i\sqrt{p} \in \mathbb{Z}[i\sqrt{p}] \end{cases}$$

Donc $\mathbb{Z}[i\sqrt{p}]$ est un sous-anneau de $\mathbb{C}$ et comme $\mathbb{C}$, il est unitaire commutatif et intègre.

2. Si un anneau A contient $i\sqrt{p}$, il contient également 1 (il s'agit d'anneaux unitaires) et en conséquence il contient tout élément de la forme $a + ib\sqrt{p}$ avec $(a, b) \in \mathbb{Z}^2$. On a donc $\mathbb{Z}[i\sqrt{p}] \subset A$.

3. En désignant par $(A_i)_{i \in I}$ la famille de tous les sous-anneaux de $\mathbb{C}$ qui contiennent $i\sqrt{p}$, on a $A = \bigcap_{i \in I} A_i \subset \mathbb{Z}[i\sqrt{p}]$ puisque $\mathbb{Z}[i\sqrt{p}]$ est l'un de ces sous-anneaux et $\mathbb{Z}[i\sqrt{p}] \subset A$ puisque A est un anneau. On a donc bien $\mathbb{Z}[i\sqrt{p}] = A$.
4. Si $z = a + ib\sqrt{p}$ est inversible dans $\mathbb{Z}[i\sqrt{p}]$, il existe alors $z' \in \mathbb{Z}[i\sqrt{p}]$ tel que $zz' = 1$ et $|z|^2 |z'|^2 = 1$ avec $|z|^2 = a^2 + b^2 p^2 \in \mathbb{N}$ et $|z'|^2 \in \mathbb{N}$, ce qui impose $|z|^2 = |z'|^2 = 1$. On a donc $a^2 + b^2 p^2 = 1$ avec $(a^2, b^2 p^2) \in \mathbb{N}^2$, ce qui équivaut à $(a^2, b^2 p^2) = (1, 0)$ ou $(a^2, b^2 p^2) = (0, 1)$ ou encore à $a = \pm 1$ et $b = 0$ ou $a = 0$ et $b^2 p^2 = 1$. Pour $p = 1$, la condition $b^2 p^2 = 1$ équivaut à $b = \pm 1$ et pour $p \geq 2$, elle n'est jamais réalisée puisque, pour tout $b \in \mathbb{Z}$, on a $b^2 p^2 = 0$ ou $b^2 p^2 \geq p^2 \geq 4$. On a donc $\mathbb{Z}[i]^\times \subset \{-1, 1, -i, i\}$ et $\mathbb{Z}[i\sqrt{p}]^\times \subset \{-1, 1\}$ pour $p \geq 2$. Les inclusions réciproques se vérifiant facilement. En définitive, on a :

$$\mathbb{Z}[i\sqrt{p}]^\times = \begin{cases} \{-1, 1, -i, i\} & \text{si } p = 1 \\ \{-1, 1\} & \text{si } p \geq 2 \end{cases}$$

Exercice 3.6. Soit $(\mathbb{K}, +, \cdot)$ un corps commutatif.

1. Montrer que pour tout $(a, b) \in \mathbb{K}^* \times \mathbb{K}$ l'équation $ax + b = 0$ a une unique solution.
2. Soit $\lambda \in \mathbb{K}$. Montrer que s'il existe $\alpha \in \mathbb{K}$ tel que $\alpha^2 = \lambda$, alors l'équation $x^2 = \lambda$ a deux solutions exactement dans $\mathbb{K}$, à savoir α et $-\alpha$.
3. Soit $(a, b, c) \in \mathbb{K}^* \times \mathbb{K}^2$. Montrer que si l'équation $ax^2 + bx + c = 0$ a une solution x_1 dans $\mathbb{K}$, elle en a alors une seconde x_2 . Dans ce cas, on a $x_1 + x_2 = -\frac{b}{a}$, $x_1 x_2 = \frac{c}{a}$ et pour tout $x \in \mathbb{K}$, $ax^2 + bx + c = a(x - x_1)(x - x_2)$ (forme factorisée de $ax^2 + bx + c$). Dans un corps commutatif, une équation de degré 2 a donc 0 ou 2 solutions.

Solution.

1. Dans le groupe $(\mathbb{K}, +)$, l'équation $ax + b = 0$ équivaut à $ax = -b$ (unicité de l'opposé) et comme $a \in \mathbb{K}^*$ est inversible, l'équation $ax = -b$ équivaut à $a^{-1}ax = a^{-1}(-b)$, encore équivalent à $x = -a^{-1}b$. D'où l'existence et l'unicité dans $\mathbb{K}$ de la solution de l'équation $ax + b = 0$.
2. L'équation $x^2 = \lambda = \alpha^2$ équivaut à $x^2 - \alpha^2 = (x - \alpha)(x + \alpha) = 0$ encore équivalente à $x = \alpha$ ou $x = -\alpha$.
3. De $ax_1^2 + bx_1 + c = 0$, on déduit que pour tout $x \in \mathbb{K}$, on a :

$$\begin{aligned} ax^2 + bx + c &= ax^2 + bx + c - (ax_1^2 + bx_1 + c) \\ &= a(x^2 - x_1^2) + b(x - x_1) = (x - x_1)(a(x + x_1) + b) \end{aligned}$$

donc l'équation $ax^2 + bx + c = 0$ est équivalente à $(x - x_1)(a(x + x_1) + b) = 0$ encore équivalent à $x - x_1 = 0$ ou $a(x + x_1) + b = 0$, la dernière équation ayant

pour unique solution $x_2 = -a^{-1}b - x_1$. Notre équation a donc exactement deux solutions, à savoir x_1 et $x_2 = -\frac{b}{a} - x_1$. On a donc $x_1 + x_2 = -\frac{b}{a}$ et :

$$x_1 x_2 = -\frac{1}{a} (bx_1 + ax_1^2) = -\frac{1}{a} (-c) = \frac{c}{a}.$$

Pour tout $x \in \mathbb{K}$, on a :

$$\begin{aligned} ax^2 + bx + c &= (x - x_1) (a(x + x_1) + b) \\ &= a(x - x_1) \left(x + x_1 + \frac{b}{a} \right) = a(x - x_1)(x - x_2). \end{aligned}$$

On a donc montré que $ax^2 + bx + c$ est factorisable dans K , si, et seulement si, l'équation $ax^2 + bx + c = 0$ a des solutions dans $\mathbb{K}$. Par exemple sur $\mathbb{R}$, l'équation $x^2 + 1$ n'est pas factorisable.

Exercice 3.7. Montrer qu'un anneau unitaire intègre et fini est un corps.

Solution. Soit A un anneau unitaire intègre. Pour tout $a \neq 0$ dans A , l'application $x \mapsto ax$ est injective. En effet si $ax = ay$, alors $a(x - y) = 0$ et $x - y = 0$ puisque A est intègre et $a \neq 0$. Si de plus A est fini, alors cette application est bijective et en particulier il existe $b \in A$ tel que $ab = 1$, ce qui prouve que a est inversible à droite. On montre de même que a est inversible à gauche. On a donc montré que tout élément non nul de a est inversible, ce qui revient à dire que A est un corps.

Exercice 3.8. Soit p un entier sans facteurs carrés dans sa décomposition en produit de nombres premiers. Montrer que l'ensemble :

$$\mathbb{Q}[\sqrt{p}] = \{r + s\sqrt{p} \mid (r, s) \in \mathbb{Q}^2\}$$

est un sous-corps de $\mathbb{R}$.

Solution. On vérifie facilement que $\mathbb{Q}[\sqrt{p}]$ est un sous-anneau de $\mathbb{R}$ (même démonstration que pour $\mathbb{Z}[\sqrt{p}]$ déjà rencontré). Comme $\sqrt{p}$ est irrationnel, on a $a = r + s\sqrt{p} = 0$ si, et seulement si, $r = s = 0$. Pour $a \neq 0$ dans $\mathbb{Q}[\sqrt{p}]$, on a :

$$a^{-1} = \frac{1}{r + s\sqrt{p}} = \frac{r - r\sqrt{p}}{r^2 - ps^2} \in \mathbb{Q}[\sqrt{p}].$$

En conclusion, $\mathbb{Q}[\sqrt{p}]$ est un sous-corps de $\mathbb{R}$.

Exercice 3.9. Montrer que l'ensemble $\mathbb{Q}[i] = \{r + si \mid (r, s) \in \mathbb{Q}^2\}$ est un sous-corps de $\mathbb{C}$.

Solution. On vérifie facilement que $\mathbb{Q}[i]$ est un sous-anneau de $\mathbb{C}$ (même démonstration que pour $\mathbb{Z}[i]$ déjà rencontré). Pour $z \neq 0$ dans $\mathbb{Q}[i]$, on a $a^{-1} = \frac{1}{r+si} = \frac{r-si}{r^2+s^2} \in \mathbb{Q}[i]$. En conclusion, $\mathbb{Q}[i]$ est un sous-corps de $\mathbb{C}$.

Exercice 3.10. Montrer que l'identité est le seul endomorphisme de corps non identiquement nul de $\mathbb{R}$.

Solution. Si f est endomorphisme du corps $\mathbb{R}$, on a alors $f(x+y) = f(x) + f(y)$ et $f(xy) = f(x)f(y)$ pour tous x, y dans $\mathbb{R}$. Avec $f(1) = (f(1))^2$, on déduit que $f(1) = 0$ ou $f(1) = 1$. Si $f(1) = 0$, alors pour tout $x \in \mathbb{R}$ on a $f(x) = f(x)f(1) = 0$ et f est identiquement nulle. C'est une homothétie de rapport 0. On suppose donc que f n'est pas identiquement nulle et on a alors $f(1) = 1$. Avec $f(x^2) = (f(x))^2 \geq 0$, on déduit que $f(x) \geq 0$ pour tout $x \geq 0$ et pour $x \geq y$ dans $\mathbb{R}$, on a $f(x) - f(y) = f(x-y) \geq 0$, ce qui signifie que f est croissante. On déduit alors de l'exercice 2.20 que $f(x) = x$ pour tout $x \in \mathbb{R}$ ($\lambda = f(1) = 1$). L'identité est donc le seul morphisme de corps non identiquement nul de $\mathbb{R}$ dans lui même.

Chapitre 4

Division euclidienne dans $\mathbb{Z}$

On rappelle que l'ensemble $(\mathbb{Z}, +, \cdot)$ des entiers relatifs est un anneau unitaire, commutatif et intègre. On dit qu'une partie A non vide de $\mathbb{Z}$ est minorée s'il existe un entier m tel que $n \geq m$ pour tout $n \in A$. Si de plus $m \in A$ on dit alors que c'est un plus petit élément. Dans ce cas il est uniquement déterminé. On dit qu'une partie A non vide de $\mathbb{Z}$ est majorée s'il existe un entier M tel que $n \leq M$ pour tout $n \in A$. Si de plus $M \in A$ on dit alors que c'est un plus grand élément. Dans ce cas il est uniquement déterminé.

L'ensemble $\mathbb{Z}$ est bien ordonné, c'est-à-dire que toute partie non vide et minorée de $\mathbb{Z}$ admet un plus petit élément et toute partie non vide et majorée de $\mathbb{Z}$ admet un plus grand élément.

4.1 Divisibilité et congruences

Définition 4.1. On dit que l'entier relatif a est divisible par l'entier relatif d , ou que a est un multiple de d , s'il existe un entier relatif q tel que $a = qd$. On note d/a .

Pour $d \neq 0$ l'entier q est uniquement déterminé (une égalité $a = dq = dq'$ entraîne $d(q - q') = 0$ et $q - q' = 0$ puisque $\mathbb{Z}$ est intègre). On se limitera donc au cas où $d \in \mathbb{Z}^*$.

La relation de divisibilité est une relation d'ordre non totale sur $\mathbb{N}$, c'est-à-dire qu'elle est :

- réflexive : pour tout $a \in \mathbb{N}$, a/a ;
- anti-symétrique : si a/b et b/a dans $\mathbb{N}$ alors $a = b$;
- transitive : si a/b et b/c dans $\mathbb{N}$ alors a/c .

Deux éléments quelconques de $\mathbb{N}$ ne sont pas toujours comparables. Par exemple on n'a aucune relation de divisibilité entre 3 et 5 dans $\mathbb{N}$.

Sur $\mathbb{Z}$ on a les propriétés suivantes :

- les seuls diviseurs de 1 sont 1 et -1 ;
- si d/a et $a \neq 0$, alors $|d| \leq |a|$ (si $a = 0$, on a $0 = 0 \cdot d$ pour tout $d \in \mathbb{Z}$)

- si a/b et b/a dans $\mathbb{Z}$ alors $|a| = |b|$, (si $a = 0$, alors $b = 0$), soit $a = \pm b$ (la relation de divisibilité n'est donc pas anti-symétrique sur $\mathbb{Z}$, elle est seulement réflexive et transitive et ce n'est pas une relation d'ordre);
- si d/a et d/b dans $\mathbb{Z}$ alors $d/(\lambda a + \mu b)$ pour tous λ, μ dans $\mathbb{Z}$.

Pour tout entier relatif n , on note $n\mathbb{Z} = \{n \cdot q \mid q \in \mathbb{Z}\}$ l'ensemble de tous les multiples de n et $\mathcal{D}_n = \{q \in \mathbb{Z} \mid q \text{ divise } n\}$ l'ensemble de tous les diviseurs de n . Nous verrons plus loin que les $n\mathbb{Z}$ sont les seuls sous-groupes de $(\mathbb{Z}, +)$.

Définition 4.2. Soient $n \in \mathbb{N}$ et a, b deux entiers relatifs. On dit que a est congru à b modulo n si n divise $a - b$. On note $a \equiv b \pmod{n}$.

Dire que a est congru à b modulo n équivaut aussi à dire que $a - b \in n\mathbb{Z}$.

Pour $n = 0$, on a $0\mathbb{Z} = \{0\}$ et $a \equiv b \pmod{0}$ revient à dire que $a = b$.

Pour $n = 1$, on a $1\mathbb{Z} = \mathbb{Z}$ et la relation $a \equiv b \pmod{1}$ est toujours vérifiée. On suppose donc, dans ce qui suit que $n \geq 2$.

4.2 Le théorème de division euclidienne dans $\mathbb{Z}$

Théorème 4.1.

Soit $(a, b) \in \mathbb{Z} \times \mathbb{Z}^*$. Il existe un unique couple $(q, r) \in \mathbb{Z} \times \mathbb{N}$ tel que :

$$\begin{cases} a = bq + r \\ 0 \leq r < |b| \end{cases} \quad (4.1)$$

Preuve. On suppose que $b > 0$ et on pose $A = \{k \in \mathbb{Z} \mid bk \leq a\}$. Cet ensemble est non vide (pour $a \geq 0$, 0 est dans A et pour $a < 0$, a est dans A) et majoré (pour $a \geq 0$, a majore A et pour $a < 0$, 0 majore A). Il admet donc un plus grand élément q qui vérifie $qb \leq a < (q+1)b$. Il suffit alors de poser $r = a - bq$. Pour $b < 0$ on travaille avec $-b$ et on a l'existence de (q', r') vérifiant $a = -bq' + r'$, et $0 \leq r' < -b$. Et il suffit de poser $q = -q'$, $r = r'$. Supposons qu'il existe deux couples d'entiers (q, r) et (q', r') vérifiant (4.1) avec $q \neq q'$. On a alors $|r - r'| = |b(q - q')| \geq |b|$ avec r et r' dans $] -|b|, |b|$ ce qui est impossible. On a donc $q = q'$ et $r = r'$. Le couple (q, r) vérifiant (4.1) est donc unique. $\square$

Définition 4.3. Avec les notations du théorème 4.1 on dit que a est le dividende, b le diviseur, q le quotient et r le reste dans la division euclidienne de a par b .

Dire que le reste dans la division euclidienne de a par b est nul revient aussi à dire que b divise a .

En utilisant la division euclidienne par un entier naturel non nul n , $a = qn + r$ avec $q \in \mathbb{Z}$ et $r \in \{0, 1, \dots, n-1\}$, on voit que a est congru modulo n au reste r . Réciproquement, si a est congru à un entier $r \in \{0, 1, \dots, n-1\}$, alors r est

le reste dans la division euclidienne de a par n . C'est-à-dire que le reste dans la division euclidienne de a par n est l'unique entier r vérifiant $a \equiv r \pmod{n}$ et $0 \leq r < n$.

La démonstration précédente du théorème de division euclidienne n'est pas constructive. Un algorithme de détermination du quotient et du reste est donné par la méthode de descente infinie de Fermat qui revient à faire une démonstration par récurrence du théorème 4.1. Le principe est le suivant en supposant a et b strictement positifs. Si $a < b$, on prend alors $(q, r) = (0, a)$. Si $b \leq a$, il existe alors un entier $q_1 \geq 1$ tel que $q_1 b \leq a$ et on pose $r_1 = a - bq_1$. Si $r_1 < b$ on prend alors $(q, r) = (q_1, r_1)$. Si $b \leq r_1$, il existe alors un entier $q_2 \geq 1$ tel que $q_2 b \leq r_1$ et on pose $r_2 = r_1 - bq_2$. En continuant ainsi de suite on construit deux suites d'entiers (q_n) et (r_n) par la relation de récurrence :

si $r_n < b$ alors $(q_{n+1}, r_{n+1}) = (q_n, r_n)$;

si $b \leq r_n$ alors q_{n+1} est choisi tel que $q_{n+1}b \leq r_n$ et on pose $r_{n+1} = r_n - bq_{n+1}$.

La suite (r_n) est une suite strictement décroissante d'entiers naturels. Le procédé s'arrêtera donc au bout d'un nombre fini d'étapes, c'est-à-dire qu'il existe un entier p tel que $r_p < b$ et dans ce cas le couple $(q, r) = (q_1 + \dots + q_p, r_p)$ est la solution cherchée. En effet on a $0 \leq r_p = r_{p-1} - q_p b = a - (q_1 + \dots + q_p) b < b$.

4.3 Les systèmes de numération

Une première application importante du théorème de division euclidienne est le théorème de numération dans une base.

Théorème 4.2.

Soit b un entier supérieure ou égal à 2. Pour tout $n \in \mathbb{N}^*$, il existe un unique entier p et un unique $(p+1)$ -uplet $(n_0, \dots, n_p) \in \mathbb{N}^{p+1}$ tels que $n_p \neq 0$, $0 \leq n_k \leq b-1$ pour tout $k \in \{0, 1, \dots, p\}$ et :

$$n = \sum_{k=0}^p n_k b^k \quad (4.2)$$

Preuve. En remarquant que $\mathbb{N}^* = \bigcup_{j=0}^{+\infty} [b^j, b^{j+1}[$, il suffit de montrer le résultat

pour tout entier n dans $[b^j, b^{j+1}[$ où j décrit $\mathbb{N}$. Pour $j = 0$ tout $n \in [1, b[$ s'écrit sous la forme (4.2) avec $p = 0$ et $n_0 = n$. Supposons le résultat acquis pour $j \geq 0$ et soit $n \in [b^{j+1}, b^{j+2}[$. En utilisant le théorème de division euclidienne on peut écrire $n = bq + n_0$ avec $0 \leq n_0 \leq b-1$. On a alors $bq = n - n_0 > b^{j+1} - b = b(b^j - 1)$, donc $q > b^j - 1$, soit $q \geq b^j$. On a également $q = \frac{n - n_0}{b} < b^{j+1} - \frac{n_0}{b} \leq b^{j+1}$. En

définitive $q \in [b^j, b^{j+1}[$ et avec l'hypothèse de récurrence il s'écrit $q = \sum_{k=0}^p q_k b^k$ avec

$q_p \neq 0$. D'où $n = bq + n_0 = \sum_{k=0}^{p+1} n_k b^k$ avec $n_k = q_{k-1}$ pour tout $k \in \{1, \dots, p+1\}$.

En particulier $n_{p+1} = q_p \neq 0$. Supposons que l'on ait deux écritures $n = \sum_{k=0}^p n_k b^k =$

$\sum_{k=0}^{p'} n'_k b^k$ avec $p' \geq p$, $0 \leq n_k \leq b-1$, $0 \leq n'_k \leq b-1$, $n_p \neq 0$ et $n'_{p'} \neq 0$. On a alors :

$$b^p \leq n \leq \sum_{k=0}^p (b-1) b^k = b^{p+1} - 1 < b^{p+1}$$

De même $b^{p'} \leq n < b^{p'+1}$. Donc $b^{p'} < b^{p+1}$ soit $b^{p'-p} < b$ et nécessairement $p = p'$. En remarquant que n_0 est le reste dans la division euclidienne de n par b , on déduit que $n_0 = n'_0$ puis par récurrence que $n_k = n'_k$ pour tout $k \in \{1, \dots, p\}$. D'où l'unicité de la décomposition. $\square$

Dans la décomposition (4.2) on a $b^p \leq n < b^{p+1}$, c'est-à-dire que p est le plus grand entier vérifiant $b^p \leq n$.

Définition 4.4. Avec les notations du théorème 4.2 on dit que (4.2) est la représentation en base b de l'entier n . On note $n = \overline{n_p \cdots n_1 n_0}^b$ et on dit que les n_k sont les chiffres dans l'écriture en base b de n .

Si $n = \overline{n_p \cdots n_1 n_0}^b$, alors n_0 est le reste dans la division euclidienne de n par b et $\overline{n_p \cdots n_1}^b$ est le quotient. Cette remarque nous permet de donner un algorithme de calcul des chiffres dans l'écriture en base b de n : on divise n par b , puis le quotient par b et ainsi de suite, un quotient nul indique la fin du processus et les restes successifs donnent, de droite à gauche, l'écriture en base b de n .

On peut remarquer que l'écriture en base b de l'entier b est $b = \overline{10}^b$ et que plus généralement, pour tout entier $p \geq 1$, l'écriture de l'entier b^p en base b est $b = \overline{10 \cdots 0}^b$ (1 suivi de p zéros).

On peut également remarquer que si $n = \overline{n_p \cdots n_1 n_0}^b$, alors pour tout entier k compris entre 1 et p , $\overline{n_{k-1} \cdots n_1 n_0}^b$ est le reste dans la division euclidienne de n par b^k et $\overline{n_p \cdots n_k}^b$ est le quotient.

L'écriture en base $b = 10$ permet d'obtenir les critères classiques de divisibilité résumés avec l'exercice 4.5.

4.4 Plus grand commun diviseur et plus petit commun multiple

La caractérisation des sous-groupes de $(\mathbb{Z}, +)$ (théorème 2.6) peut être utilisée pour définir le pgcd de deux ou plusieurs entiers relatifs, non tous nuls.

Théorème 4.3.

Soient a, b deux entiers relatifs non tous deux nuls. Il existe un unique entier naturel δ tel que $a\mathbb{Z} + b\mathbb{Z} = \delta\mathbb{Z}$. Cet entier s'écrit $\delta = au + bv$ avec $(u, v) \in \mathbb{Z}^2$ et c'est le plus grand entier naturel qui divise a et b .

Preuve. $a\mathbb{Z} + b\mathbb{Z}$ étant un sous-groupe de $(\mathbb{Z}, +)$, le théorème 2.6 nous dit qu'il existe un unique entier naturel δ tel que $a\mathbb{Z} + b\mathbb{Z} = \delta\mathbb{Z}$. Comme $\delta \in \delta\mathbb{Z} = a\mathbb{Z} + b\mathbb{Z}$, il existe $(u, v) \in \mathbb{Z}^2$ tel que $\delta = au + bv$. De $a\mathbb{Z} \subset \delta\mathbb{Z}$ et $b\mathbb{Z} \subset \delta\mathbb{Z}$, on déduit que δ est un diviseur commun à a et b . Si $d \in \mathbb{N}$ est un diviseur commun à a et b , il divise aussi $\delta = au + bv$ et $d \leq \delta$ (a et b n'étant pas tous les deux nuls, on a $\delta \neq 0$). Donc δ est bien le plus grand entier naturel qui divise a et b . $\square$

On peut donc donner la définition suivante.

Définition 4.5. Soient a, b deux entiers relatifs non tous deux nuls. On appelle plus grand commun diviseur de a et b le plus grand entier naturel qui divise a et b . On le note $\text{pgcd}(a, b)$ ou $a \wedge b$.

La relation $a \wedge b = au + bv$ avec $(u, v) \in \mathbb{Z}^2$ est l'identité de Bézout.

On définit de manière analogue le pgcd d'une famille $a_1, \dots, a_p$ formée de p entiers non tous nuls comme le plus grand des diviseurs communs à $a_1, \dots, a_p$. On le note $\text{pgcd}(a_1, \dots, a_p)$ ou $a_1 \wedge \dots \wedge a_p$ et c'est un entier supérieur ou égal à 1. Cette définition est justifiée par le théorème suivant.

Théorème 4.4.

Soient $a_1, \dots, a_p$ des entiers relatifs non tous nuls. Il existe un unique entier naturel δ tel que $a_1\mathbb{Z} + \dots + a_p\mathbb{Z} = \delta\mathbb{Z}$. Cet entier s'écrit $\delta = \sum_{k=1}^p u_k a_k$ avec $(u_1, \dots, u_p) \in \mathbb{Z}^p$ et c'est le plus grand entier naturel qui divise $a_1, \dots, a_p$.

Preuve. Analogue au cas où $p = 2$. $\square$

Comme dans le cas où $p = 2$, on vérifie que $a_1 \wedge \dots \wedge a_p$ est aussi le plus grand élément pour l'ordre de la division dans $\mathbb{N}$ de l'ensemble des diviseurs positifs communs à $a_1, \dots, a_p$.

L'égalité $\delta = \sum_{k=1}^p u_k a_k$ est l'identité de Bézout.

La notation $a_1 \wedge \dots \wedge a_p$ ne pose pas de problème du fait de la commutativité et l'associativité du pgcd (elle est indépendante de l'ordre des a_k).

Le théorème qui suit, où l'on résume les propriétés du pgcd , se démontre facilement.

Théorème 4.5.

Pour a, b, c entiers relatifs, on a :

- pour $(a, b) \neq (0, 0)$, $a \wedge b \in \mathbb{N}^*$;
- pour $a \neq 0$, $a \wedge 0 = a \wedge a = |a|$;
- $a \wedge 1 = 1$;
- pour $(a, b) \neq (0, 0)$, $a \wedge b = |a| \wedge |b| = |a| \wedge b = a \wedge |b|$;
- pour $(a, b) \neq (0, 0)$, $a \wedge b = b \wedge a$ (commutativité du pgcd) ;

- pour $b \neq 0$, on a $a \wedge b = |b|$ si, et seulement si, b divise a ;
- pour $(a, b) \neq (0, 0)$ et $c \neq 0$, $(ac) \wedge (bc) = |c| (b \wedge a)$;
- si $d \neq 0$ est un diviseur commun de a et b non tous deux nuls, alors $\frac{a}{d} \wedge \frac{b}{d} = \frac{a \wedge b}{|d|}$;
- pour $(a, b, c) \neq (0, 0, 0)$, on a $a \wedge (b \wedge c) = (a \wedge b) \wedge c$ (associativité du pgcd).

Par définition du pgcd, on a $a \wedge b = au + bv$ avec u, v entiers relatifs, mais en général la réciproque est fausse, c'est-à-dire que si δ est entier naturel tel que $\delta = au + bv$ avec u, v entiers relatifs, il n'y a aucune raison pour que δ soit le pgcd de a et b . Par exemple $2 = 3 \cdot 2 + 2 \cdot (-2)$ et $3 \wedge 2 = 1$. Mais pour $\delta = 1$, cette réciproque est vraie et ce résultat est très souvent utilisé.

Définition 4.6. Soient $(a, b) \in \mathbb{Z}^2 \setminus \{(0, 0)\}$. On dit que a et b sont premiers entre eux si leur pgcd est égal à 1.

De manière équivalente, on peut dire que a et b sont premiers entre eux si, et seulement si, -1 et 1 sont leurs seuls diviseurs communs.

De manière plus générale, on peut donner la définition suivante.

Définition 4.7. Soient $a_1, \dots, a_p$ des entiers relatifs non tous nuls. On dit que $a_1, \dots, a_p$ sont premiers entre eux dans leur ensemble si leur pgcd est égal à 1.

Théorème 4.6.

Soient $(a, b) \in \mathbb{Z}^2 \setminus \{(0, 0)\}$ et $\delta = a \wedge b$. Il existe deux entiers p et q premiers entre eux tels que $a = \delta p$ et $b = \delta q$.

Preuve. Comme δ divise a et b il existe deux entiers p et q tels que $a = \delta p$ et $b = \delta q$. Le pgcd $\delta' = p \wedge q$ est un diviseur de p et q , donc l'entier naturel $\delta\delta'$ divise $a = \delta p$ et $b = \delta q$ et nécessairement $\delta\delta' \leq \delta$, soit $\delta(1 - \delta') \geq 0$ avec $\delta > 0$. On a donc $\delta' \leq 1$. Mais δ' est supérieur ou égal à 1 comme tout pgcd qui se respecte. En définitive, on a $\delta' = 1$, c'est-à-dire que p et q sont premiers entre eux. $\square$

De manière plus générale, on a le résultat suivant.

Théorème 4.7.

Soient $a_1, \dots, a_p$ des entiers relatifs non tous nuls. et $\delta = a_1 \wedge \dots \wedge a_p$. Il existe des entiers $a'_1, \dots, a'_p$ premiers entre eux dans leur ensemble tels que $a_k = \delta a'_k$ pour tout k compris entre 1 et p .

Preuve. Analogie au cas où $p = 2$. $\square$

Théorème 4.8. Bézout

Deux entiers relatifs a et b non tous deux nuls sont premiers entre eux si, et seulement si, il existe deux entiers relatifs u et v tels que $au + bv = 1$.

Preuve. On sait déjà, par définition, que la condition est nécessaire. Réciproquement s'il existe deux entiers relatifs u et v tels que $au + bv = 1$ alors $\delta = a \wedge b$ est un entier naturel qui divise a et b , il divise donc $1 = au + bv$ et $\delta = 1$, c'est-à-dire que a et b sont premiers entre eux. $\square$

La relation de Bézout $au + bv = 1$ implique que u et v sont aussi premiers entre eux. On a aussi $|a| \wedge |b| = |a| \wedge b = a \wedge |b| = a \wedge b = 1$.

Ce théorème peut se généraliser comme suit.

Théorème 4.9. Bézout

Des entiers relatifs $a_1, \dots, a_p$ non tous nuls sont premiers entre eux dans leur ensemble si, et seulement si, il existe deux entiers relatifs $u_1, \dots, u_p$ tels que $\sum_{k=1}^p u_k a_k = 1$.

Preuve. On sait déjà, par définition, que la condition est nécessaire. Réciproquement s'il existe deux entiers relatifs $u_1, \dots, u_p$ tels que $\sum_{k=1}^p u_k a_k = 1$ alors $\delta = a_1 \wedge \dots \wedge a_p$ est un entier naturel qui divise tous les a_k , il divise donc $1 = \sum_{k=1}^p u_k a_k$ et $\delta = 1$, c'est-à-dire que $a_1, \dots, a_p$ sont premiers entre eux dans leur ensemble. $\square$

Corollaire 4.1. *Soient a, b, c des entiers relatifs non nuls. Si c est premier avec a alors $a \wedge b = a \wedge (bc)$ (le pgcd de deux entiers est inchangé si on multiplie l'un d'eux par un nombre premier avec l'autre).*

Preuve. Soient $\delta = a \wedge b$ et $\delta' = a \wedge (bc)$. Comme δ divise a et b , il divise a et bc ainsi que leur pgcd δ' . De $au + cv = 1$, on déduit que $abu + bcv = b$ et δ' qui divise a et bc va diviser a et b ainsi que leur pgcd δ . On a donc $\delta = \delta'$. $\square$

Corollaire 4.2. *Soient $a_1, \dots, a_p$ et c des entiers relatifs non nuls. Si c est premier avec chacun des a_k , pour k compris entre 1 et p , il est alors premier avec leur produit $\prod_{k=1}^p a_k$.*

Preuve. Avec le corollaire précédent, on a $c \wedge \prod_{k=1}^p a_k = c \wedge \left(a_1 \prod_{k=2}^p a_k \right) = c \wedge \prod_{k=2}^p a_k$ puisque a_1 est premier avec c et par récurrence finie, on déduit que :

$$c \wedge \prod_{k=1}^p a_k = c \wedge \prod_{k=2}^p a_k = c \wedge \prod_{k=3}^p a_k = \dots = c \wedge a_p = 1$$

puisque chaque a_k , pour k compris entre 1 et p , est premier avec c . $\square$

Une conséquence importante du théorème de Bézout est le résultat suivant.

Théorème 4.10. Gauss

Soient a, b, c des entiers relatifs non nuls. Si a divise bc et a est premier avec b alors a divise c .

Preuve. Comme a et b sont premiers entre eux, il existe deux entiers u, v tels que $au + bv = 1$ et pour tout entier c , on a $acu + bcv = c$, de sorte que si a divise bc , il va diviser $c = acu + bcv$. $\square$

Ce résultat peut être utilisé pour donner une unique représentation des nombres rationnels non nuls.

Corollaire 4.3. *Tout nombre rationnel non nul r s'écrit de manière unique $r = \frac{p}{q}$ avec $p \in \mathbb{Z}^*$ et $q \in \mathbb{N}^*$ premiers entre eux.*

Preuve. Un nombre rationnel non nul r s'écrit $r = \frac{a}{b}$ avec (a, b) dans $\mathbb{Z}^* \times \mathbb{N}^*$. En notant $\delta = a \wedge b$ on a $a = \delta p$, $b = \delta q$ et $r = \frac{p}{q}$ avec p et q premiers entre eux.

Si $r = \frac{p}{q} = \frac{p'}{q'}$ avec $(p, q), (p', q')$ dans $\mathbb{Z}^* \times \mathbb{N}^*$ tels que $p \wedge q = p' \wedge q' = 1$, on a alors $pq' = p'q$ avec q premier avec p et q qui divise pq' , donc q divise q' d'après le théorème de Gauss. De manière analogue, on voit que q' divise q . On a donc $q = q'$ (q, q' sont des entiers naturels non nuls) et $p = p'$. L'écriture est donc unique. $\square$

Corollaire 4.4. *Si un entier relatif non nul n est divisible par des entiers $a_1, \dots, a_p$ deux à deux premiers entre eux, il est alors divisible par leur produit.*

Preuve. Supposons que n soit divisible par a_1 et a_2 premiers entre eux. On a alors $n = a_1 q_1$ et a_2 divise n en étant premier avec a_1 , il va donc diviser q_1 (théorème de Gauss), c'est-à-dire que $q_1 = a_2 q_2$ et $n = a_1 a_2 q_2$ est divisible par $a_1 a_2$. En supposons le résultat acquis au rang $p - 1 \geq 2$, soient $a_1, a_2, \dots, a_p$ deux à deux premiers entre eux qui divisent n . L'hypothèse de récurrence nous dit que

n est divisible par $a = \prod_{k=1}^{p-1} a_k$. Comme a_p est premier avec chacun des a_k , pour k compris entre 1 et $p - 1$, il est premier avec leur produit a (corollaire 4.2) et n qui est divisible par a et a_p est aussi divisible par leur produit $\prod_{k=1}^p a_k$. $\square$

La caractérisation des sous-groupes de $(\mathbb{Z}, +)$ peut aussi être utilisée pour définir le ppcm de deux ou plusieurs entiers relatifs, non tous nuls.

Théorème 4.11.

Soient a, b deux entiers relatifs. Il existe un unique entier naturel μ tel que $a\mathbb{Z} \cap b\mathbb{Z} = \mu\mathbb{Z}$. Si $a = 0$ ou $b = 0$, alors $\mu = 0$. Si $a \neq 0$ et $b \neq 0$, alors μ est le plus petit entier naturel non nul multiple de a et de b .

Preuve. $a\mathbb{Z} \cap b\mathbb{Z}$ étant un sous-groupe de $(\mathbb{Z}, +)$, l'existence et l'unicité de μ se déduit du théorème 2.6. Si $a = 0$ ou $b = 0$, on a $\mu\mathbb{Z} \subset 0\mathbb{Z} = \{0\}$ et $\mu = 0$. Si $a \neq 0$ et $b \neq 0$, de $\mu\mathbb{Z} \subset a\mathbb{Z}$ et $\mu\mathbb{Z} \subset b\mathbb{Z}$, on déduit que μ est multiple de a et b . Si $m \in \mathbb{N}$ est un multiple commun à a et b , il est dans $a\mathbb{Z} \cap b\mathbb{Z} = \mu\mathbb{Z}$ et c'est donc un multiple de μ , ce qui implique $m \geq \mu$. Donc, μ est bien le plus petit entier naturel non nul multiple de a et de b . $\square$

On peut donc donner la définition qui suit.

Définition 4.8. *Soient a, b deux entiers relatifs. On appelle plus petit commun multiple de a et b le plus petit entier naturel multiple de a et b . On le note $\text{ppcm}(a, b)$ ou $a \vee b$.*

La définition de $\text{ppcm}(a, b)$ peut aussi se justifier directement sans référence directe aux sous-groupes de $\mathbb{Z}$. Pour ce faire, on désigne par $\mathcal{M}_{a,b}$ l'ensemble des multiples communs à a et b . Si $a \neq 0$ et $b \neq 0$, alors l'ensemble $\mathcal{M}_{a,b} \cap \mathbb{N}^*$ des multiples communs à a et b qui sont strictement positifs est non vide car il contient $|ab|$, il admet donc un plus petit élément μ qui est bien plus petit commun multiple de a et b . Pour $a = 0$ ou $b = 0$, on a $\mathcal{M}_{a,b} = \{0\}$ et $\mu = 0$.

Le ppcm de a et b est aussi le plus petit élément pour l'ordre de la division dans $\mathbb{Z}$ de l'ensemble $\mathcal{M}_{a,b}$ des multiples communs à a et b . En effet, $a \vee b$ est un multiple de a et b et tout multiple commun m à a et b qui est dans $a\mathbb{Z} \cap b\mathbb{Z} = \mu\mathbb{Z}$ est un multiple de $a \vee b$.

Le théorème qui suit, où l'on résume les propriétés du ppcm , se démontre facilement.

Théorème 4.12.

Pour a, b, c entiers relatifs, on a :

- pour a, b non nuls, $a \vee b \in \mathbb{N}^*$;
- pour a non nul, $a \vee 1 = a \vee a = |a|$;
- $a \vee b = |a| \vee |b| = |a| \vee b = a \vee |b|$;
- $a \vee b = b \vee a$ (commutativité du ppcm) ;
- $a \vee b = |b|$ si, et seulement si, b est multiple de a ;
- $(ac) \vee (bc) = |c| (b \vee a)$;
- si $d \in \mathbb{Z}^*$ est un diviseur commun de a et b , alors $\frac{a}{d} \vee \frac{b}{d} = \frac{a \vee b}{|d|}$;
- pour $(a, b, c) \neq (0, 0, 0)$, on a $a \vee (b \vee c) = (a \vee b) \vee c$ (associativité du ppcm).

Lemme 4.1 *Soient a, b deux entiers relatifs premiers entre eux. On a $a \vee b = |ab|$.*

Preuve. Du fait que $|ab|$ est un multiple de a et b on déduit que $\mu = a \vee b$ divise ab . D'autre part il existe deux entiers k, k' tels que $\mu = ka = k'b$ et comme a est premier avec b et divise $k'b$, il divise k' (théorème de Gauss). Ce qui donne $\mu = k''ab$ et ab divise μ . D'où l'égalité $\mu = ab$. $\square$

Nous verrons que la réciproque du résultat précédent est vraie.

Théorème 4.13.

Soient a, b deux entiers relatifs. On a alors $(a \wedge b)(a \vee b) = |ab|$.

Preuve. On note $\delta = a \wedge b$ et on a $|a| = \delta a'$, $|b| = \delta b'$ avec a' et b' premiers entre eux. Ce qui donne $\mu = a \vee b = (\delta a') \vee (\delta b') = \delta(a' \vee b') = \delta a'b'$ et $\delta\mu = \delta a'\delta b' = |ab|$. $\square$

Du lemme et du théorème précédent, on déduit que $a \wedge b = 1 \Leftrightarrow a \vee b = |ab|$. Pour a, b dans $\mathbb{Z}^*$, on a $a \vee b = \frac{|ab|}{a \wedge b}$, donc on peut définir de manière naturelle le ppcm de deux entiers relatifs non tous deux nuls par $a \vee b = \frac{|ab|}{a \wedge b}$.

On définit de manière analogue le ppcm d'une famille $a_1, \dots, a_p$ formée de p entiers non tous nuls comme le plus petit des multiples communs à $a_1, \dots, a_p$. On le note $\text{ppcm}(a_1, \dots, a_p)$ ou $a_1 \vee \dots \vee a_p$ et c'est un entier supérieur ou égal à 1. Cette définition est justifiée par le théorème suivant.

Théorème 4.14.

Soient $a_1, \dots, a_p$ des entiers relatifs non tous nuls. Il existe un unique entier naturel μ tel que $a_1\mathbb{Z} \cap \dots \cap a_p\mathbb{Z} = \mu\mathbb{Z}$. μ est le plus petit entier naturel divisible par $a_1, a_2, \dots$ et a_p .

Preuve. Analogie au cas où $p = 2$. $\square$

La notation $a_1 \vee \dots \vee a_p$ ne pose pas de problème du fait de la commutativité et l'associativité du ppcm (elle est indépendante de l'ordre des a_k).

Comme dans le cas où $p = 2$, on vérifie que $a_1 \vee \dots \vee a_p$ est aussi le plus petit élément pour l'ordre de la division dans $\mathbb{N}$ de l'ensemble des multiples positifs communs à $a_1, \dots, a_p$.

4.5 L'algorithme d'Euclide

Le lemme qui suit permet de déduire du théorème de division euclidienne un algorithme de calcul du pgcd de deux entiers positifs. C'est l'algorithme d'Euclide. Il permet également de déterminer des entiers u et v tels que $au + bv = a \wedge b$.

Théorème 4.15.

Soient a, b deux entiers naturels non nuls et r le reste dans la division euclidienne de a par b . On a alors $a \wedge b = b \wedge r$.

Preuve. Par division euclidienne, on a $a = bq + r$ avec $0 \leq r < b$. L'entier naturel $\delta = a \wedge b$ qui est un diviseur commun à a et b va diviser $r = a - bq$, c'est donc un

diviseur commun à b et r et $\delta \leq \delta' = b \wedge r$. L'entier naturel $\delta' = b \wedge r$ qui est un diviseur commun à b et r va diviser $a = bq + r$, c'est donc un diviseur commun à a et b et $\delta' \leq \delta$. On a donc bien $\delta = \delta'$. $\square$

Le principe de l'algorithme d'Euclide est le suivant pour $a > b$ dans $\mathbb{N}^*$ (par symétrie, on peut supposer que $a \geq b$ et pour $a = b$, on a $a \wedge a = a$).

On note $r_0 = b$ et on désigne par r_1 le reste dans la division euclidienne de a par b . On a alors $0 \leq r_1 < r_0$ et d'après le lemme précédent $a \wedge b = r_0 \wedge r_1$.

Si $r_1 = 0$ alors $r_0 \wedge r_1 = r_0 = b$ et c'est terminé.

Si $r_1 \neq 0$ on désigne alors par r_2 le reste dans la division euclidienne de r_0 par r_1 et on a $0 \leq r_2 < r_1$ et $a \wedge b = r_0 \wedge r_1 = r_1 \wedge r_2$.

Si $r_2 = 0$ alors $r_1 \wedge r_2 = r_1$ et c'est terminé. Sinon on continue.

On définit donc ainsi une suite d'entiers $(r_n)_{n \geq 0}$ par :

- $r_0 = b$;
- r_1 est le reste dans la division euclidienne de a par b ; on a donc $0 \leq r_1 < b$;
- pour $n \geq 2$, si $r_{n-1} = 0$ alors $r_n = 0$ et sinon r_n est le reste dans la division euclidienne de r_{n-2} par r_{n-1} et on a $0 \leq r_n < r_{n-1}$. Dans tous les cas on $r_n \leq r_{n-1}$ l'égalité étant réalisée si, et seulement si, les deux termes sont nuls.

La suite $(r_n)_{n \geq -1}$ ainsi construite est donc une suite décroissante d'entiers positifs, elle est donc stationnaire à partir d'un certain rang. Précisément il existe un entier $p \geq 1$ tel que $r_p = 0 < r_{p-1} < \dots < r_1 < r_0$ et :

$$a \wedge b = r_0 \wedge r_1 = \dots = r_{p-1} \wedge r_p = r_{p-1}$$

c'est-à-dire que $a \wedge b$ est le dernier reste non nul dans cette suite de divisions euclidiennes.

On a donc construit, avec l'algorithme d'Euclide, deux suites d'entiers $(r_n)_{0 \leq n \leq p}$ et $(q_n)_{1 \leq n \leq p}$ de la manière suivante :

$$\begin{cases} a = q_1 r_0 + r_1 & (0 < r_1 < r_0 = b) \\ r_0 = q_2 r_1 + r_2 & (0 < r_2 < r_1) \\ r_1 = q_3 r_2 + r_3 & (0 < r_3 < r_2) \\ \vdots \\ r_{p-3} = q_{p-1} r_{p-2} + r_{p-1} & (0 < r_{p-1} < r_{p-2}) \\ r_{p-2} = q_p r_{p-1} + r_p & (r_p = 0) \end{cases}$$

On vérifie alors, par récurrence finie sur $n \in \{0, 1, \dots, p-1\}$, qu'il existe des entiers u_n et v_n tels que $r_n = au_n + bv_n$.

Pour $n = 0$ et $n = 1$ on a $r_0 = b = a \cdot 0 + b \cdot 1$ et $r_1 = a \cdot 1 + b(-q_1)$. En supposant le résultat acquis jusqu'à l'ordre $n-1$ pour $0 \leq n-1 \leq p-2$ on a :

$$\begin{aligned} r_n &= -q_n r_{n-1} + r_{n-2} = -q_n (au_{n-1} + bv_{n-1}) + au_{n-2} + bv_{n-2} \\ &= a(u_{n-2} - q_n u_{n-1}) + b(v_{n-2} - q_n v_{n-1}) = au_n + bv_n \end{aligned}$$

En particulier pour $n = p-1$ on a $a \wedge b = r_{p-1} = au_{p-1} + bv_{p-1} = au + bv$.

Un tel couple d'entiers (u, v) n'est pas unique puisque si (u, v) est une solution, pour tout entier λ , le couple $(u', v') = (u + \lambda v, v - \lambda a)$ est aussi solution. On a en effet $a(u + \lambda b) + b(v - \lambda a) = au + bv = a \wedge b$.

Les suites $(r_n)_{0 \leq n \leq p-1}$, $(u_n)_{0 \leq n \leq p-1}$ et $(v_n)_{0 \leq n \leq p-1}$ vérifient la même relation de récurrence :

$$x_n = x_{n-2} - q_n x_{n-1} \quad (2 \leq n \leq p-1)$$

avec les conditions initiales $(r_0, r_1) = (b, 1)$, $(u_0, u_1) = (0, 1)$, $(v_0, v_1) = (1, -q_1)$, où q_1, r_1 sont, respectivement, le quotient et le reste dans la division euclidienne de a par b .

4.6 Equations diophantiennes $ax + by = c$

Soient a, b, c trois entiers relatifs, avec a et b non nuls. On s'intéresse ici à l'équation diophantienne $ax + by = c$ dans $\mathbb{Z}^2$. En notant δ le pgcd de a et b , on a $a = \delta a'$, $b = \delta b'$ avec a' et b' premiers entre eux.

Lemme 4.2 *L'équation diophantienne $ax + by = c$ a des solutions entières si, et seulement si, δ divise c .*

Preuve. Si c n'est pas un multiple de δ , comme δ divise $ax + by$ pour tous entiers x, y , l'équation $ax + by = c$ n'a pas de solutions. Si $c = \delta c'$ est un multiple de δ , en écrivant que $\delta = au_0 + bv_0$ avec u_0, v_0 dans $\mathbb{Z}$ (théorème de Bézout) on déduit que $(x_0, y_0) = (u_0 c', v_0 c')$ est une solution de $ax + by = c$. $\square$

Théorème 4.16.

Si c est multiple de δ , alors l'ensemble des solutions de $ax + by = c$ est $S = \{(x_0 - kb', y_0 + ka') \mid k \in \mathbb{Z}\}$, où (x_0, y_0) est une solution particulière.

Preuve. Si (x, y) est une solution de $ax + by = c$, on a alors $ax_0 + by_0 = c$ et $ax + by = c$, ce qui donne par soustraction $a(x_0 - x) = b(y - y_0)$ et divisant par δ , on obtient $a'(x_0 - x) = b'(y - y_0)$. Avec le théorème de Gauss on en déduit alors que a' divise $y - y_0$. On a donc $y - y_0 = ka'$ avec $k \in \mathbb{Z}$, ce qui entraîne $a'(x_0 - x) = b'ka'$ et $x_0 - x = kb'$. En définitive on a $(x, y) = (x_0 - kb', y_0 + ka')$ avec $k \in \mathbb{Z}$. Réciproquement on vérifie que pour tout $k \in \mathbb{Z}$, $(x_0 - kb', y_0 + ka')$ est bien solution de $ax + by = c$. En effet on a :

$$ax + by = ax_0 + by_0 + k(a'b - ab') = c + k\delta(a'b' - a'b') = c$$

$\square$

L'algorithme d'Euclide vu au paragraphe précédent nous permet d'obtenir une solution particulière $(x_0, y_0) = \left(u_0 \frac{c}{\delta}, v_0 \frac{c}{\delta}\right)$.

Exemple 4.1 *Soit à résoudre l'équation $370x + 45y = 15$. Le pgcd de 370 et 45 est égal à 5 qui divise 15. On cherche tout d'abord une solution particulière de $74x + 9y = 1$. En utilisant l'algorithme d'Euclide, on a $74 = 8 \cdot 9 + 2$ et $9 = 4 \cdot 2 + 1$, donc $1 = 9 - 4 \cdot 2 = 9 - 4 \cdot (74 - 8 \cdot 9) = 74 \cdot (-4) + 9 \cdot 33$. Le couple $(-12, 99)$ est solution de $74x + 9y = 3$ et de $370x + 45y = 15$.*

4.7 Equations $ax \equiv b \pmod{n}$

Soient n un entier supérieur ou égal à 2, a un entier supérieur ou égal à 1 et b un entier relatif. On veut résoudre dans $\mathbb{Z}$ l'équation diophantienne $ax \equiv b \pmod{n}$. On s'intéresse tout d'abord au cas où $b = 1$.

Lemme 4.3 *Soient n un entier supérieur ou égal à 2, a un entier supérieur ou égal à 1. L'équation $ax \equiv 1 \pmod{n}$ a des solutions dans $\mathbb{Z}$ si, et seulement si, a est premier avec n .*

Preuve. Le théorème de Bézout nous dit que a est premier avec n si, et seulement si, il existe des entiers relatifs x et k tels que $ax - kn = 1$, ce qui équivaut à dire que $x \in \mathbb{Z}$ est solution de $ax \equiv 1 \pmod{n}$. $\square$

Pour a et n premiers entre eux, l'algorithme d'Euclide nous permet de trouver une solution $x_0 \in \mathbb{Z}$ de $ax \equiv 1 \pmod{n}$ et pour tout autre solution $x \in \mathbb{Z}$ l'entier $a(x - x_0)$ est divisible par n . Comme n est premier avec a , le théorème de Gauss nous dit que nécessairement n divise $x - x_0$. Il est clair que réciproquement pour tout $k \in \mathbb{Z}$, $x_0 + kn$ est solution de $ax \equiv 1 \pmod{n}$. En définitive, l'ensemble des solutions de $ax \equiv 1 \pmod{n}$ est $S = \{x_0 + kn \mid k \in \mathbb{Z}\}$, où x_0 est une solution particulière de $ax \equiv 1 \pmod{n}$.

Si a et n sont premiers entre eux, alors il existe une unique solution de $ax \equiv 1 \pmod{n}$ dans $\{1, \dots, n-1\}$. En effet, si S est l'ensemble des solutions de $ax \equiv 1 \pmod{n}$ alors $S \cap \mathbb{N}^*$ est non vide et donc admet un plus petit élément $x > 0$. Si $x \geq n$ on a alors $x = qn + r$ avec $q \geq 1$ et $0 \leq r < n$. Comme $ar \equiv ax \equiv 1 \pmod{n}$, on a $r \in S \cap \mathbb{N}$ et nécessairement $r = 0$. Mais $x = qn$ entraîne $ax \equiv 0 \pmod{n}$ en contradiction avec $x \in S$. On a donc $x < n$.

On s'intéresse maintenant au cas où les entiers a et n sont premiers entre eux et b est un entier relatif. Dans ce cas on peut trouver une solution x_0 de l'équation $ax \equiv 1 \pmod{n}$ et pour tout entier relatif k , $x = bx_0 + kn$ est solution de $ax \equiv b \pmod{n}$. Réciproquement si x est solution de $ax \equiv b \pmod{n}$ alors $a(x - bx_0)$ est divisible par n avec n premier avec a . Le théorème de Gauss nous dit alors que $x - bx_0$ est divisible par n . En définitive, pour a et n premiers entre eux, l'ensemble des solutions de $ax \equiv b \pmod{n}$ est $S = \{bx_0 + kn \mid k \in \mathbb{Z}\}$, où x_0 est une solution particulière de $ax \equiv 1 \pmod{n}$.

Considérons maintenant le cas où $\delta = a \wedge n$ n'est pas nécessairement égal à 1 et b est un entier relatif. On a alors $a = \delta a'$, $n = \delta n'$ et si l'équation $ax \equiv b \pmod{n}$ admet une solution $x \in \mathbb{Z}$ alors $\delta n'$ divise $\delta a'x - b$, donc δ divise $\delta a'x - b$ et δ divise b . En conclusion si b n'est pas un multiple de δ alors l'équation $ax \equiv b \pmod{n}$ n'a pas de solution dans $\mathbb{Z}$. On suppose donc que b est un multiple de δ , soit $b = \delta b'$. Si $x \in \mathbb{Z}$ est solution de $ax \equiv b \pmod{n}$ alors x est solution de $a'x \equiv b' \pmod{n'}$ avec a' et n' premiers entre eux. On sait alors que x est de la forme $x = b'x'_0 + kn'$ où x'_0 est une solution de $a'x \equiv 1 \pmod{n'}$ et k est un entier relatif. Réciproquement on peut vérifier que pour tout entier $k \in \mathbb{Z}$, $x = b'x'_0 + kn'$ est solution de $ax \equiv b \pmod{n}$. En effet on a $ax = a'x'_0\delta b' + a'k\delta n' = (1 + k'n')\delta b' + a'kn = b + n(k'b' + ka') \equiv b \pmod{n}$. En définitive, si $b = \delta b'$ où $\delta = a \wedge n$, alors l'ensemble des solutions de $ax \equiv b \pmod{n}$ est $S = \{b'x'_0 + kn' \mid k \in \mathbb{Z}\}$, où x'_0 est une solution particulière de $a'x \equiv 1 \pmod{n'}$, où $a = \delta a'$, $n = \delta n'$.

4.8 Le théorème Chinois

On s'intéresse ici aux système d'équations diophantiennes :

$$\begin{cases} x \equiv a \pmod{n} \\ x \equiv b \pmod{m} \end{cases} \quad (4.3)$$

où n, m sont deux entiers naturels supérieur ou égal à 2.

Théorème 4.17. chinois

Soient n, m deux entiers supérieur ou égal à 2 premiers entre eux. Quels que soient les entiers relatifs a et b le système (4.3) a une infinité de solutions dans $\mathbb{Z}$.

Preuve. Comme n et m sont premiers entre eux on peut trouver une infinité de couples d'entiers relatifs (u, v) tels que $nu + mv = 1$. En posant $x = bnu + amv$ on obtient une infinité de solutions de (4.3). $\square$

Dans le cas où n et m sont premiers entre eux on vient de voir que si (u_0, v_0) est solution de $nu + mv = 1$ (un tel couple peut être obtenu par l'algorithme d'Euclide) alors $x_0 = bnu_0 + amv_0$ est une solution particulière de (4.3).

Si $x \in \mathbb{Z}$ est solution de (4.3) alors x est congru à x_0 modulo n et modulo m , soit $x - x_0 = pn = qm$. Mais m est premier avec n , le théorème de Gauss nous dit alors que m divise p . On a donc $x = x_0 + knm$ avec $k \in \mathbb{Z}$. Et réciproquement on vérifie que pour tout entier relatif k , $x_0 + knm$ est solution de (4.3). En définitive, si n et m sont premiers entre eux, alors l'ensemble des solutions de (4.3) est $S = \{x_0 + knm \mid k \in \mathbb{Z}\}$, où x_0 est une solution particulière de (4.3).

Dans le cas général où m et n ne sont pas nécessairement premiers entre eux on note $\delta = m \wedge n$, $n = \delta n'$, $m = \delta m'$ avec n', m' premiers entre eux et $\mu = m \vee n$. Si $x \in \mathbb{Z}$ est une solution de (4.3) alors δ qui divise n et m va diviser $x - a$ et $x - b$, il divise donc $a - b$. Donc si $a - b$ n'est pas un multiple de $\delta = m \wedge n$ le système d'équations (4.3) n'a pas de solutions. On suppose donc que $a - b$ est multiple de δ , c'est-à-dire que $b - a = \delta c'$. Les entiers n' et m' étant premiers entre eux, le théorème de Bézout nous dit qu'il existe des entiers u_0 et v_0 tels que $n'u_0 + m'v_0 = 1$. En posant $x_0 = bn'u_0 + am'v_0$, on a :

$$x_0 = b(1 - m'v_0) + am'v_0 = b - m'v_0(b - a) = b - m'v_0\delta c' = b - mv_0c' \equiv b \pmod{m}$$

Et de manière analogue on voit que x_0 est congru à a modulo n . L'entier x_0 est donc solution de (4.3).

Si $x \in \mathbb{Z}$ est solution de (4.3) alors x est congru à x_0 modulo n et modulo m , soit $x - x_0 = pn = qm = p\delta n' = q\delta m'$.

On déduit donc que $\frac{x - x_0}{\delta}$ est un entier et $\frac{x - x_0}{\delta} = pn' = qm'$. Mais m' est premier avec n' , le théorème de Gauss nous dit alors que m' divise p . On a donc $\frac{x - x_0}{\delta} = kn'm'$ avec $k \in \mathbb{Z}$. Ce qui peut aussi s'écrire $x - x_0 = knm' = k \frac{nm}{\delta} = k\mu$ avec $k \in \mathbb{Z}$. Réciproquement on vérifie facilement que pour tout entier relatif k , $x_0 + k\mu$ est solution de (4.3). En définitive, l'ensemble des solutions de (4.3) est $S = \{x_0 + k(m \vee n) \mid k \in \mathbb{Z}\}$, où x_0 est une solution particulière de (4.3).

4.9 L'ensemble $\mathcal{P}$ des nombres premiers

L'ensemble $\mathcal{D}_n$ des diviseurs dans $\mathbb{N}^*$ d'un entier $n \geq 2$ contient toujours 1 et n , il est donc de cardinal supérieur ou égal à 2. On s'intéresse ici aux entiers $p \geq 2$ tels que $\mathcal{D}_p$ soit de cardinal minimal, à savoir 2.

Définition 4.9. On dit qu'un entier naturel p est premier s'il est supérieur ou égal à 2 et si les seuls diviseurs positifs de p sont 1 et p .

On note $\mathcal{P}$ l'ensemble de tous les nombres premiers.

Le résultat qui suit se déduit du fait que toute partie non vide de $\mathbb{N}$ admet un plus petit élément.

Théorème 4.18. Euclide

Tout entier n supérieur ou égal à 2 a au moins un diviseur premier.

Preuve. Pour tout entier $n \geq 2$ l'ensemble $\mathcal{D}_n$ des diviseurs strictement positifs de n a au moins deux éléments, 1 et n , donc $\mathcal{D}_n \setminus \{1\}$ est non vide dans $\mathbb{N} \setminus \{0, 1\}$ et il admet un plus petit élément p qui est nécessairement premier. En effet si p n'est pas premier il admet un diviseur q tel que $2 \leq q < p$ avec q qui divise n , ce qui contredit le caractère minimal de p . $\square$

Corollaire 4.5. *Tout entier relatif $n \in \mathbb{Z} \setminus \{-1, 0, 1\}$ a au moins un diviseur premier.*

Preuve. Tout diviseur premier de $|n| \geq 2$ convient. $\square$

Un entier naturel non premier s'écrit donc $n = pq$ avec $p \geq 2$ premier et $q \geq 2$. On dit alors qu'il est composé.

Un diviseur premier de $n \geq 2$ est nécessairement inférieur ou égal à n . En fait, pour n non premier, on peut toujours en trouver un qui est inférieur ou égal à $\sqrt{n}$, c'est $p = \min(\mathcal{D}_n \setminus \{1\})$.

Théorème 4.19.

Tout entier n supérieur ou égal à 2 qui est composé a au moins un diviseur premier p tel que $2 \leq p \leq \sqrt{n}$.

Preuve. En supposant n composé et en gardant les notations de la démonstration du théorème précédent, on a vu que $p = \min(\mathcal{D}_n \setminus \{1\})$ est un diviseur premier de n . On a donc $n = pq$ avec $2 \leq q \leq n$ (on a $q \neq 1$ puisque n n'est pas premier) et $q \in \mathcal{D}_n \setminus \{1\}$, ce qui implique que $p \leq q$ et $p^2 \leq pq = n$, soit $p \leq \sqrt{n}$. $\square$

Le théorème précédent nous donne un premier algorithme, relativement simple, permettant de savoir si un entier $n \geq 2$ est premier ou non : on effectue successivement la division euclidienne de n par tous les entiers $p \leq \sqrt{n}$: si l'une de ces divisions donne un reste nul, alors n n'est pas premier, sinon, n est premier. Une petite amélioration peut être apportée à cet algorithme en remarquant que si 2 ne divise pas n , il est inutile de tester les divisibilités par les entiers $p \leq \sqrt{n}$

pairs. Pour tester la divisibilité de n par les nombres premiers $p \leq \sqrt{n}$, on doit disposer de la liste de tous ces nombres premiers. Le crible d'Eratosthène nous permet d'obtenir une telle liste. Le principe est le suivant :

- on se donne la liste de tous les entiers compris entre 2 et m (m est la partie entière de $\sqrt{n}$ pour l'algorithme précédent) ;
- on garde 2 et on supprime tous les autres multiples de 2 de cette liste ;
- le premier entier strictement supérieur à 2 est 3 et comme il ne possède pas de diviseur strict, il est premier, on le garde et on supprime tous les autres multiples de 3 de la liste ;
- on continue ainsi de suite et on s'arrête dès que l'on tombe sur un nombre premier strictement plus grand que $\sqrt{m}$. La liste finale contient alors tous les nombres premiers inférieurs à m .

Les résultats qui suivent sont élémentaires, mais souvent utiles.

Lemme 4.4 *Soit p un entier naturel premier. Pour tout entier naturel non nul n , on a soit p qui divise n , soit p qui est premier avec n .*

Preuve. Comme $\delta = p \wedge n$ divise p , on a soit $\delta = p$ et p divise n , soit $\delta = 1$ et p est premier avec n . $\square$

Lemme 4.5 *Deux nombres premiers distincts sont premiers entre eux.*

Preuve. Soient p, q deux nombres premiers. Si $\delta = p \wedge q \neq 1$, le lemme précédent nous dit que p divise q et q divise p , donc $p = q$. $\square$

Lemme 4.6 *Un entier $p \geq 2$ est premier si, et seulement si, il est premier avec tout entier compris entre 1 et $p - 1$.*

Preuve. Si p est premier, comme il ne divise pas $k \in \{1, \dots, p - 1\}$, il est premier avec k . Réciproquement si p n'est pas premier, il s'écrit alors $p = ab$ avec $a \geq 2$, $b \geq 2$ et p n'est pas premier avec $a \in \{2, \dots, p - 1\}$. $\square$

Le théorème de Gauss nous donne le résultat suivant qui nous sera utile pour prouver l'unicité (à l'ordre près) de la décomposition en facteurs premiers d'un entier $n \geq 2$.

Lemme 4.7 *Soit p un nombre premier et r un entier naturel supérieur ou égal à 2. Si p divise le produit $n_1 \cdots n_r$ de r entiers naturels non nuls, alors p divise l'un des n_k .*

Preuve. Si p divise $n_1 n_2$, on a soit p qui divise n_1 , soit p qui est premier avec n_1 et il va alors diviser n_2 (théorème de Gauss).

Supposons le résultat acquis au rang $n - 1 \geq 2$. Si p divise $n_1 n_2 \cdots n_r$, on a soit p qui divise n_1 , soit p qui est premier avec n_1 et il va alors diviser $n_2 \cdots n_r$ et l'un des n_k où k est compris entre 2 et n . $\square$

Dans le cas où tous les n_k sont égaux à un même entier n , on a :

$$(p \text{ premier divise } n^r) \Rightarrow (p \text{ divise } n)$$

On peut montrer de nombreuses manières que l'ensemble $\mathcal{P}$ des nombres premiers est infini. La démonstration élémentaire qui suit, conséquence de l'existence de diviseurs premiers pour $n \geq 2$, est due à Euclide.

Théorème 4.20. Euclide

L'ensemble $\mathcal{P}$ des nombres premiers est infini.

Preuve. On sait déjà que $\mathcal{P}$ est non vide (il contient 2). Supposons que $\mathcal{P}$ soit fini avec $\mathcal{P} = \{p_1, \dots, p_r\}$. L'entier $n = p_1 \cdots p_r + 1$ qui est supérieur 2 admet un diviseur premier $p_k \in \mathcal{P}$. L'entier p_k divise alors $n = p_1 \cdots p_r + 1$ et $p_1 \cdots p_r$, il divise donc la différence qui est égale à 1, ce qui est impossible. En conclusion $\mathcal{P}$ est infini. $\square$

En rangeant les nombres premiers dans l'ordre croissant, on constate que les entiers $n_r = p_1 \cdots p_r + 1$ sont premiers pour r compris entre 1 et 5 ($n_1 = 3$, $n_2 = 7$, $n_3 = 31$, $n_4 = 211$, $n_5 = 2311$). Pour $r = 6$, $n_6 = 30031 = 59 \times 509$ n'est pas premier. On ne sait pas si la suite $(n_r)_{r \geq 1}$ contient une infinité de nombres premiers.

Pour tout réel $x \geq 2$, on désigne par $\pi(x) = \text{card}(\mathcal{P} \cap [0, x])$ le cardinal de l'ensemble des nombres premiers contenus dans l'intervalle $[0, x]$. Du théorème d'Euclide on déduit que $\lim_{x \rightarrow +\infty} \pi(x) = +\infty$. Le théorème des nombres premiers (conjecturé par Gauss, puis montré par Hadamard et de la Vallée-Poussin) nous dit plus précisément que $\pi(x) \underset{x \rightarrow +\infty}{\sim} \frac{x}{\ln(x)}$.

4.10 Décomposition en facteurs premiers

Le théorème qui suit est parfois appelé « théorème fondamental de l'arithmétique ».

Théorème 4.21.

Tout entier naturel $n \geq 2$ se décompose de manière unique sous la forme $n = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$, où $2 \leq p_1 < p_2 < \cdots < p_r$ sont des nombres premiers et les α_k sont des entiers naturels non nuls.

Preuve. On démontre tout d'abord l'existence d'une telle décomposition par récurrence sur $n \geq 2$. Pour $n = 2$, on a déjà la décomposition. Supposons le résultat acquis pour tout entier k compris entre 2 et $n \geq 2$. Si $n + 1$ est premier, on a déjà la décomposition, sinon on écrit $n + 1 = ab$ avec a et b compris entre 2 et n et il suffit d'utiliser l'hypothèse de récurrence pour a et b . L'unicité d'une telle décomposition peut aussi se montrer par récurrence sur $n \geq 2$. Le résultat est évident pour $n = 2$. Supposons le acquis pour tout entier k compris entre 2 et $n \geq 2$. Si $n + 1$ a deux décompositions :

$$n + 1 = p_1^{\alpha_1} \cdots p_r^{\alpha_r} = q_1^{\beta_1} \cdots q_s^{\beta_s}$$

où les p_j [resp. q_i] sont premiers deux à deux distincts et les α_j [resp. β_i] entiers naturels non nuls. L'entier p_1 est premier et divise le produit $q_1^{\beta_1} \cdots q_s^{\beta_s}$, il divise donc nécessairement l'un des q_k . L'entier q_k étant également premier la seule possibilité est $p_1 = q_k$. En simplifiant par p_1 on se ramène à la décomposition d'un entier inférieur ou égal à n et il suffit alors d'utiliser l'hypothèse de récurrence pour conclure. $\square$

L'écriture $n = p_1^{\alpha_1} \cdots p_r^{\alpha_r}$ est la décomposition en facteurs premiers de l'entier n .

Le théorème précédent se traduit en disant que l'anneau $\mathbb{Z}$ des entiers relatifs est factoriel.

Si $n = \prod_{k=1}^r p_k^{\alpha_k}$ est un entier décomposé en produit de facteurs premiers, alors les diviseurs de n sont de la forme $d = \prod_{k=1}^r p_k^{\gamma_k}$ où les γ_k sont des entiers naturels tels

que $\gamma_k \leq \alpha_k$ pour tout k compris entre 1 et r . Il y a donc $\prod_{k=1}^r (\alpha_k + 1)$ diviseurs positifs possibles de n .

La décomposition en facteurs premiers peut être utilisée pour calculer le pgcd et le ppcm de deux entiers naturels supérieur ou égal à 2. Tout est basé sur le lemme qui suit.

Lemme 4.8 Soient n, m deux entiers naturels supérieur ou égal à 2 de décompositions en facteurs premiers $n = \prod_{k=1}^r p_k^{\alpha_k}$, $m = \prod_{k=1}^r p_k^{\beta_k}$ avec les p_k premiers deux à deux distincts et les α_k, β_k entiers naturels (certains de ces entiers pouvant être nuls). On a alors :

$$(a \text{ divise } b) \Leftrightarrow (\forall k \in \{1, 2, \dots, r\}, \alpha_k \leq \beta_k)$$

Preuve. Dire que n divise m équivaut à dire qu'il existe un entier $q \geq 1$ tel que $m = qn$ et en écrivant $q = p_1^{\gamma_1} \cdots p_r^{\gamma_r}$ où les γ_k sont des entiers naturels, on a :

$$m = p_1^{\beta_1} \cdots p_r^{\beta_r} = p_1^{\alpha_1 + \gamma_1} \cdots p_r^{\alpha_r + \gamma_r}$$

L'unicité de la décomposition en facteurs premiers de m nous dit alors que :

$$\forall k \in \{1, 2, \dots, r\}, \beta_k = \alpha_k + \gamma_k \geq \alpha_k$$

□

Théorème 4.22.

Avec les hypothèses du lemme précédent, on a :

$$n \wedge m = \prod_{k=1}^r p_k^{\min(\alpha_k, \beta_k)}, \quad n \vee m = \prod_{k=1}^r p_k^{\max(\alpha_k, \beta_k)}$$

Preuve. L'entier $\delta = \prod_{k=1}^r p_k^{\min(\alpha_k, \beta_k)}$ divise n et m d'après le lemme précédent. Si

d est un diviseur de n et m , il s'écrit sous la forme $d = \prod_{k=1}^r p_k^{\gamma_k}$ où les γ_k sont des entiers naturels tels que $\gamma_k \leq \alpha_k$ et $\gamma_k \leq \beta_k$ pour tout k compris entre 1 et r , on a

donc $\gamma_k \leq \min(\alpha_k, \beta_k)$ pour tout k compris entre 1 et r et d divise δ . Donc δ est bien le pgcd de n et m .

Pour ce qui est du ppcm, on a :

$$n \vee m = \frac{nm}{n \wedge m} = \prod_{k=1}^r p_k^{\alpha_k + \beta_k - \min(\alpha_k, \beta_k)}$$

avec $\alpha_k + \beta_k - \min(\alpha_k, \beta_k) = \max(\alpha_k, \beta_k)$ pour tout k compris entre 1 et r . $\square$

Le résultat précédent se généralise au calcul du pgcd et du ppcm de $p \geq 2$ entiers naturels.

4.11 Les théorèmes de Fermat et de Wilson

Le lemme qui suit nous donne une démonstration relativement simple du théorème de Fermat.

Lemme 4.9 *Un entier naturel $p \geq 2$ est premier si, et seulement si, pour tout entier k compris entre 1 et $p-1$, p divise $\binom{p}{k} = \frac{p!}{k!(p-k)!}$.*

Preuve. Si $p \geq 2$ est premier, comme il divise $p! = k!(p-k)!\binom{p}{k}$ et est premier avec $k!(p-k)!$ (sinon il diviserait ce produit et donc l'un des entiers j compris entre 1 et $p-1$, ce qui est impossible), il divise $\binom{p}{k}$ (théorème de Gauss). Réciproquement, supposons que p divise $\binom{p}{k}$ pour tout entier k compris entre 1 et $p-1$.

Pour tout k compris entre 1 et $p-1$, on a $\binom{p-1}{k-1} + \binom{p-1}{k} = \binom{p}{k}$ (triangle de Pascal) et avec $\binom{p}{k} \equiv 0$ modulo p , on déduit que $\binom{p-1}{k} \equiv -\binom{p-1}{k-1}$ modulo p et par récurrence finie sur k compris entre 1 et $p-1$, on déduit que $\binom{p-1}{k}$ est congru à $(-1)^k$ modulo p . En effet, pour $k=1$, on a $\binom{p-1}{1} \equiv -\binom{p-1}{0} = -1$ modulo p et en supposant le résultat acquis pour $k-1$ compris entre 1 et $p-2$, on a $\binom{p-1}{k} \equiv -\binom{p-1}{k-1} \equiv -(-1)^{k-1} = (-1)^k$. Si p n'est pas premier, il admet un diviseur d compris entre 2 et $p-1$ et on a $\binom{p}{d} = \frac{p}{d} \binom{p-1}{d-1} = q \binom{p-1}{d-1}$, où q est un entier compris entre 2 et $p-1$, avec $\binom{p}{d} \equiv 0$ modulo p et $\binom{p-1}{d-1} \equiv (-1)^{d-1}$ modulo p , ce qui donne $q(-1)^{d-1} \equiv 0$ modulo p , encore équivalent à dire que p divise $q(-1)^{d-1}$ avec $2 \leq |q(-1)^{d-1}| \leq p-1$, ce qui est impossible. Donc p est premier. $\square$

Théorème 4.23. Fermat

Soit p un entier naturel premier. Pour tout entier relatif n on a $n^p \equiv n \pmod{p}$.

Preuve. On démontre tout d'abord ce résultat sur les entiers naturels par récurrence sur $n \geq 0$. Pour $n = 0$ le résultat est évident. On le supposant acquis pour $n \geq 0$, on a :

$$(n+1)^p = n^p + \sum_{k=1}^{p-1} \binom{p}{k} n^k + 1 \equiv n^p + 1 \equiv n+1 \pmod{p}$$

Pour $n < 0$, on a $(-n)^p \equiv -n$ modulo p . Si $p = 2$ alors $-n \equiv n$ modulo 2 et $(-n)^2 = n^2$. Pour $p \geq 3$, p est impair et $(-n)^p = -n^p$ est congru à $-n$ modulo p , donc n^p est congru à n modulo p . $\square$

On peut aussi déduire du lemme 4.9 que si p est premier, alors pour tout couple (a, b) d'entiers relatifs, on a :

$$(a+b)^p = a^p + \sum_{k=1}^{p-1} \binom{p}{k} a^{p-k} b^k + b^p \equiv a^p + b^p \pmod{p}$$

Par récurrence sur l'entier $n \geq 1$, on déduit alors que pour tout n -uplet $(a_1, \dots, a_n)$ d'entiers relatifs, on a :

$$(a_1 + \dots + a_n)^p \equiv a_1^p + \dots + a_n^p \pmod{p}$$

Prenant tous les a_k égaux à 1, on en déduit que n^p est congru à n modulo p . Ce résultat est encore valable pour $n = 0$ et $n < 0$.

Théorème 4.24. Fermat

Soit p un entier naturel premier. Pour tout entier relatif n non multiple de p , on a $n^{p-1} \equiv 1 \pmod{p}$.

Preuve. L'entier premier p divise $n^p - n = n(n^{p-1} - 1)$ et est premier avec n si n n'est pas un multiple de p , il divise donc $n^{p-1} - 1$. $\square$

Si p est un entier pour lequel il existe un entier n compris entre 1 et $p-1$ tel que n^{p-1} ne soit pas congru à 1 modulo p , alors p n'est pas premier puisque n^{p-1} est congru à 1 modulo p pour p premier et $1 \leq n \leq p-1$ d'après le théorème de Fermat.

La réciproque du théorème de Fermat est fausse. On peut en fait montrer que pour $p \geq 2$, la condition $n^{p-1} \equiv 1 \pmod{p}$ pour tout n premier avec p est équivalente à p premier ou $p = \prod_{k=1}^r p_k$ avec $r \geq 3$, $3 \leq p_1 < \dots < p_r$ premiers

tels que pour tout k compris entre 1 et r , $p_k - 1$ divise $p - 1$ (un tel entier est appelé nombre de Carmichael ou nombre pseudo-premier). Par exemple 561, 1105, 1729, sont des nombres de Carmichael. En 1999, Alford, Granville et Pomerance ont montré qu'il y a une infinité de nombres de Carmichael.

Le théorème de Fermat peut être utilisé pour calculer des congruences avec des grands nombres. Si p est un nombre premier impair, n, m deux entiers naturels, l'entier n n'étant pas multiple de p , en effectuant les divisions euclidiennes par p et par $p-1$, on $n = qp + r$, $m = q'(p-1) + s$ avec $1 \leq r \leq p-1$, $0 \leq s \leq p-2$ et $n^m \equiv r^s \pmod{p}$. Par exemple on a $2003^{2003} \equiv 4 \pmod{5}$. En effet $2003 = 5 \cdot 400 + 3$ et $2003 = 4 \cdot 500 + 3$.

Lemme 4.10 *Soit $p \geq 5$ un nombre premier. Pour tout entier k compris entre 1 et $p-1$, il existe un unique entier r compris entre 1 et $p-1$ tel que kr soit congru à 1 modulo p .*

Preuve. Pour $k = 1$, on peut prendre $r = 1$. Tout entier k compris entre 2 et $p-1$ étant premier avec p , il existe deux entiers relatifs u et v tels que $ku + pv = 1$ et ku est congru à 1 modulo p (on peut aussi utiliser le théorème de Fermat : comme k est premier avec p , k^{p-1} est congru à 1 modulo p et $u = k^{p-2}$ convient). En effectuant la division euclidienne de u par p , on a $u = qp + r$ avec r compris entre 0 et $p-1$ et kr est congru à ku , donc à 1, modulo p , ce qui exclut la valeur $r = 0$. Supposons que l'on ait trouvé un autre entier $s \neq r$ vérifiant la même condition que r . On peut supposer que $s > r$. On a alors $kr \equiv ks \equiv 1 \pmod{p}$ avec r, s compris entre 1 et $p-1$, ce qui implique que $k(s-r) \equiv 0 \pmod{p}$, donc p divise $k(s-r)$ en étant premier avec k et en conséquence il doit diviser $s-r$ avec $1 \leq s-r \leq p-1$, ce qui est impossible. L'entier r est donc unique. $\square$

Dans le lemme précédent, on aura $k = r$, si, et seulement si, $k^2 - 1 = (k-1)(k+1)$ est divisible par p , donc p doit diviser $k-1$ ou $k+1$, ce qui n'est possible que si $k = 1$ ou $k = p-1$.

Théorème 4.25. Wilson

Un entier p supérieur ou égal à 2 est premier si, et seulement si, $(p-1)!$ est congru à -1 modulo p .

Preuve. Si p n'est pas premier il s'écrit $p = ab$ avec a et b entiers compris entre 2 et $p-1$. L'entier a est alors un diviseur de $(p-1)!$ et de p qui divise $(p-1)! + 1$, donc a divise $(p-1)! + 1$ et a divise 1, ce qui est impossible. Soit $p \geq 2$ un nombre premier. Pour $p = 2$, $(p-1)! + 1 = 2$ est congru à 0 modulo 2 et pour $p = 3$, $(p-1)! + 1 = 3$ est congru à 0 modulo 3. Pour $p \geq 5$, en utilisant le lemme précédent, on partitionne l'ensemble $E = \{2, 3, \dots, p-2\}$ en deux sous-ensembles E_1 et E_2 à $\frac{p-3}{2}$ éléments de sorte que :

$$\forall k \in E_1, \exists r \in E_2 \mid kr \equiv 1 \pmod{p}$$

et on a alors :

$$(p-1)! = 1 \cdot \left(\prod_{k \in E_1} k \right) \left(\prod_{r \in E_2} r \right) (p-1) \equiv p-1 \equiv -1 \pmod{p}$$

$\square$

4.12 Les anneaux $\frac{\mathbb{Z}}{n\mathbb{Z}}$ et la fonction indicatrice d'Euler

On peut facilement vérifier que la relation de congruence est une relation d'équivalence. Cette relation est compatible avec l'addition et la multiplication sur $\mathbb{Z}$. C'est-à-dire que :

$$(a \equiv b \pmod{n}, c \equiv d \pmod{n}) \Rightarrow (a + c \equiv b + d \pmod{n}, ac \equiv bd \pmod{n})$$

En effet $a - b \in n\mathbb{Z}$ et $c - d \in n\mathbb{Z}$ entraîne $a + c - (b + d) = (a - b) + (c - d) \in n\mathbb{Z}$ puisque $n\mathbb{Z}$ est un groupe et $ac - bd = a(c - d) + d(a - b) \in n\mathbb{Z}$ puisque $n\mathbb{Z}$ est un idéal de $\mathbb{Z}$.

Cette compatibilité permet de munir l'ensemble $\frac{\mathbb{Z}}{n\mathbb{Z}}$ des classes d'équivalence modulo n d'une structure d'anneau.

Pour $n = 0$, la congruence modulo 0 est tout simplement la relation d'égalité et $\frac{\mathbb{Z}}{0\mathbb{Z}} = \{\{a\} \mid a \in \mathbb{Z}\}$ est en bijection avec $\mathbb{Z}$. On identifie donc $\frac{\mathbb{Z}}{0\mathbb{Z}}$ à $\mathbb{Z}$.

Pour $n = 1$, deux entiers relatifs quelconques sont toujours congrus modulo 1, donc $\frac{\mathbb{Z}}{1\mathbb{Z}} = \{\mathbb{Z}\} = \{\overline{0}\}$ est identifié à $\{0\}$.

Théorème 4.26.

Pour tout $n \in \mathbb{N}^*$, on a $\frac{\mathbb{Z}}{n\mathbb{Z}} = \{\overline{0}, \overline{1}, \dots, \overline{n-1}\}$. Cet ensemble est donc de cardinal égal à n et il est en bijection avec l'ensemble de tous les restes modulo n .

Preuve. Le théorème de division euclidienne nous permet d'écrire tout entier relatif a sous la forme $a = qn + r$ avec $0 \leq r \leq n - 1$, ce qui entraîne que $\overline{a} = \overline{r}$. On a donc $\frac{\mathbb{Z}}{n\mathbb{Z}} = \{\overline{0}, \overline{1}, \dots, \overline{n-1}\}$. Pour montrer que cet ensemble est de cardinal égal à n , il nous reste à montrer que tous ses éléments sont distincts. Si $\overline{r} = \overline{s}$ avec r et s compris entre 0 et $n - 1$, on a alors $s - r = qn$ avec $q \in \mathbb{Z}$ et l'encadrement $0 \leq |s - r| = |q|n \leq n - 1$ dans $\mathbb{N}$ impose $q = 0$, ce qui équivaut à $r = s$. $\square$

Considérant qu'un anneau a au moins deux éléments et que $\frac{\mathbb{Z}}{1\mathbb{Z}} = \{\overline{0}\}$, on suppose dans ce qui suit que $n \geq 2$.

La compatibilité de la relation de congruence modulo n avec l'addition et la multiplication sur $\mathbb{Z}$ va nous permettre de transporter la structure d'anneau de $\mathbb{Z}$ à $\frac{\mathbb{Z}}{n\mathbb{Z}}$, un tel prolongement étant unique.

On désigne par π_n la surjection canonique de $\mathbb{Z}$ sur $\frac{\mathbb{Z}}{n\mathbb{Z}}$, c'est l'application qui associe à tout entier relatif sa classe modulo n . Tout antécédent par π_n d'un élément x de $\frac{\mathbb{Z}}{n\mathbb{Z}}$ est appelé un représentant de x .

Théorème 4.27.

Il existe une unique structure d'anneau commutatif unitaire sur $\frac{\mathbb{Z}}{n\mathbb{Z}}$ telle que la surjection canonique π_n soit un morphisme d'anneaux.

Preuve. On vérifie tout d'abord qu'on définit deux opérations internes sur $\frac{\mathbb{Z}}{n\mathbb{Z}}$ avec :

$$\forall (x, y) \in \left(\frac{\mathbb{Z}}{n\mathbb{Z}}\right)^2, \begin{cases} x + y = \overline{a + b} \\ xy = \overline{ab} \end{cases}$$

où $a \in \mathbb{Z}$ est un représentant de x et $b \in \mathbb{Z}$ un représentant de y . En effet, si a' est un autre représentant de x et b' un représentant de y , on a alors $a \equiv a'$ et $b \equiv b'$ modulo n , ce qui entraîne $a + b \equiv a' + b'$ et $ab \equiv a'b'$ modulo n , soit $\overline{a + b} = \overline{a' + b'}$ et $\overline{ab} = \overline{a'b'}$, ce qui prouve que ces définitions de $x + y$ et xy ne dépendent pas des choix des représentants de x et y . On vérifie ensuite facilement que ces deux lois confèrent à $\frac{\mathbb{Z}}{n\mathbb{Z}}$ une structure d'anneau commutatif unitaire et que π_n est bien un morphisme d'anneaux. Réciproquement s'il existe une structure d'anneau commutatif unitaire sur $\frac{\mathbb{Z}}{n\mathbb{Z}}$ qui fait de π_n un morphisme d'anneaux, on a alors pour tous $x = \pi_n(a)$, $y = \pi_n(b)$ dans $\frac{\mathbb{Z}}{n\mathbb{Z}}$:

$$\begin{cases} x + y = \pi_n(a) + \pi_n(b) = \pi_n(a + b) = \overline{a + b} \\ xy = \pi_n(a) \pi_n(b) = \pi_n(ab) = \overline{ab} \end{cases}$$

ce qui prouve l'unicité. □

Le résultat qui suit donne une généralisation du petit théorème de Fermat.

Définition 4.10. *On dit que deux polynômes P et Q à coefficients entiers sont congrus modulo un nombre premier p s'ils sont de mêmes degré et tous leurs coefficients sont égaux modulo p .*

Théorème 4.28.

p est premier si, et seulement si, il existe un entier relatif n premier avec p tel que $(X + n)^p$ soit congru à $X^p + n$ modulo p .

Preuve. On sait que si p est premier, alors pour tout entier k compris entre 1 et $p - 1$, p divise $\binom{p}{k} = \frac{p!}{k!(p-k)!}$ (lemme 4.9), ce qui implique en utilisant la formule du binôme de Newton que, pour tout $n \in \mathbb{Z}$, $(X + n)^p$ est congru à $X^p + n^p$ modulo p et le théorème de Fermat nous dit que n^p est congru à n modulo p . Si $(X + n)^p$ est congru à $X^p + n$ modulo p , on a alors $\binom{p}{k} n^k \equiv 0$ modulo p pour tout k compris entre 1 et $p - 1$ et $n^p \equiv n$ modulo p . Comme p est premier avec n et divise $\binom{p}{k} n^k$, pour k compris entre 1 et $p - 1$, il va diviser $\binom{p}{k}$ (théorème de Gauss). On déduit alors du lemme 4.9 que p est premier. □

Définition 4.11. On dit qu'un élément $\bar{a}$ de $\frac{\mathbb{Z}}{n\mathbb{Z}}$ est inversible s'il existe $\bar{b}$ dans $\frac{\mathbb{Z}}{n\mathbb{Z}}$ tel que $\bar{a}\bar{b} = \bar{1}$.

On note $\left(\frac{\mathbb{Z}}{n\mathbb{Z}}\right)^\times$ l'ensemble des éléments inversibles de $\frac{\mathbb{Z}}{n\mathbb{Z}}$. C'est un groupe pour la loi multiplicative.

Théorème 4.29.

Soit a un entier relatif. Les propriétés suivantes sont équivalentes :

$$\begin{aligned} \left(\bar{a} \text{ est inversible dans } \frac{\mathbb{Z}}{n\mathbb{Z}}\right) &\Leftrightarrow (a \text{ est premier avec } n) \\ &\Leftrightarrow \left(\bar{a} \text{ est un générateur de } \left(\frac{\mathbb{Z}}{n\mathbb{Z}}, +\right)\right) \end{aligned}$$

Preuve. Dire que $\bar{a}$ est inversible dans $\frac{\mathbb{Z}}{n\mathbb{Z}}$ équivaut à dire qu'il existe $\bar{b}$ dans $\frac{\mathbb{Z}}{n\mathbb{Z}}$ tel que $\bar{a}\bar{b} = \bar{1}$, encore équivalent à dire qu'il existe b, q dans $\mathbb{Z}$ tels que $ab + qn = 1$, ce qui équivaut à dire que a et n sont premiers entre eux (théorème de Bézout). En traduisant le fait que $\bar{a}$ est inversible dans $\frac{\mathbb{Z}}{n\mathbb{Z}}$ par l'existence d'un entier relatif b tel que $\bar{a}\bar{b} = \bar{b}\bar{a} = \bar{1}$, on déduit que cela équivaut à dire que $\bar{1}$ est dans le groupe engendré par $\bar{a}$ et donc que ce groupe est $\frac{\mathbb{Z}}{n\mathbb{Z}}$. $\square$

Définition 4.12. On appelle fonction indicatrice d'Euler la fonction qui associe à tout entier naturel non nul n , le nombre, noté $\varphi(n)$, d'entiers compris entre 1 et n qui sont premiers avec n .

Le théorème précédent nous dit que pour tout entier $n \geq 2$, $\varphi(n)$ est le nombre de générateurs du groupe $\left(\frac{\mathbb{Z}}{n\mathbb{Z}}, +\right)$, ou encore que c'est le nombre d'éléments inversibles de $\frac{\mathbb{Z}}{n\mathbb{Z}}$.

Du théorème de Lagrange, on déduit immédiatement le résultat suivant.

Théorème 4.30. Euler

Pour tout entier relatif a premier avec n , on a $a^{\varphi(n)} \equiv 1 \pmod{n}$.

Preuve. Si a est premier avec n , alors $\bar{a}$ appartient à $\mathbb{Z}_n^*$ qui est un groupe d'ordre $\varphi(n)$ et en conséquence son ordre divise $\varphi(n)$ (théorème de Lagrange), ce qui entraîne $\bar{a}^{\varphi(n)} = \bar{1}$, ou encore $a^{\varphi(n)} \equiv 1 \pmod{n}$. $\square$

Si n est premier, alors tout entier compris entre 1 et $n - 1$ est premier avec n , ce qui implique que $\varphi(n) = n - 1$ et le théorème d'Euler devient le petit théorème de Fermat.

Théorème 4.31. Fermat

Soit p un entier naturel premier. Pour tout entier relatif a on a $a^p \equiv a \pmod{p}$.

Preuve. Le théorème d'Euler nous dit que $a^{p-1} \equiv 1 \pmod{p}$ si a est premier avec p , c'est-à-dire si a n'est pas multiple de p , ce qui entraîne $a^p \equiv a \pmod{p}$. Pour a multiple de p , on a $a^p \equiv a \equiv 0 \pmod{p}$. $\square$

Dans le cas où n est premier tous les éléments de $\frac{\mathbb{Z}}{n\mathbb{Z}} \setminus \{\bar{0}\}$ sont inversibles et en conséquence $\frac{\mathbb{Z}}{n\mathbb{Z}}$ est un corps. En fait on a le résultat plus précis suivant.

Théorème 4.32.

Pour $n \geq 2$ il y a équivalence entre :

$$(n \text{ est premier}) \Leftrightarrow \left(\frac{\mathbb{Z}}{n\mathbb{Z}} \text{ est un corps} \right) \Leftrightarrow \left(\frac{\mathbb{Z}}{n\mathbb{Z}} \text{ est intègre} \right)$$

Preuve. On vient de voir que pour n premier $\frac{\mathbb{Z}}{n\mathbb{Z}}$ est un corps. De manière générale, tout corps est intègre. Supposons $\frac{\mathbb{Z}}{n\mathbb{Z}}$ intègre et soit d un diviseur de n différent de n dans $\mathbb{N}$. Il existe donc un entier q compris entre 2 et n tel que $n = qd$ et dans $\frac{\mathbb{Z}}{n\mathbb{Z}}$ on a $\bar{q}\bar{d} = \bar{0}$ avec $\bar{d} \neq \bar{0}$, ce qui impose $\bar{q} = \bar{0}$, donc $q = n$ et $d = 1$. L'entier n est donc premier. $\square$

Ce résultat nous permet de retrouver le petit théorème de Fermat. On peut également en déduire le théorème de Wilson.

Théorème 4.33. Wilson

Un entier n est premier si, et seulement si, $(n-1)! \equiv -1 \pmod{n}$.

Preuve. Si n est premier alors $\frac{\mathbb{Z}}{n\mathbb{Z}}$ est un corps commutatif et tout élément $\bar{k}$ de $\left(\frac{\mathbb{Z}}{n\mathbb{Z}} \right)^\times$ est racine du polynôme $X^{n-1} - \bar{1}$, on a donc $X^{n-1} - \bar{1} = \prod_{k=1}^{n-1} (X - \bar{k})$ dans $\frac{\mathbb{Z}}{n\mathbb{Z}}[X]$ et en évaluant ce polynôme en $\bar{0}$, il vient $-\bar{1} = \prod_{k=1}^{n-1} (-\bar{k}) = (-1)^{n-1} \overline{(n-1)!}$.

Pour $n = 2$, on a $-\bar{1} = \bar{1}$ et pour $n \geq 2$ premier on a n impair et $-\bar{1} = \overline{(n-1)!}$ dans $\frac{\mathbb{Z}}{n\mathbb{Z}}$. Réciproquement si $n \geq 2$ est tel que $\overline{(n-1)!} = -\bar{1}$ dans $\frac{\mathbb{Z}}{n\mathbb{Z}}$, alors tout diviseur d de n compris entre 1 et $n-1$ divisant $(n-1)! = -1 + kn$ va diviser -1 , ce qui donne $d = 1$ et l'entier n est premier. $\square$

Le calcul de $\varphi(n)$ pour $n \geq 2$ peut se faire en utilisant la décomposition de n en facteurs premiers grâce à la version suivante du théorème chinois.

Théorème 4.34. Théorème chinois

Les entiers n et m sont premiers entre eux si, et seulement si, les anneaux $\frac{\mathbb{Z}}{nm\mathbb{Z}}$ et $\frac{\mathbb{Z}}{n\mathbb{Z}} \times \frac{\mathbb{Z}}{m\mathbb{Z}}$ sont isomorphes.

Preuve. Pour tout entier relatif k , on note $\bar{k}$ sa classe modulo nm , $\dot{k}$ sa classe modulo n et $\ddot{k}$ sa classe modulo m . Le produit cartésien $\frac{\mathbb{Z}}{n\mathbb{Z}} \times \frac{\mathbb{Z}}{m\mathbb{Z}}$ est naturellement muni d'une structure d'anneau commutatif unitaire avec les lois $+$ et $\cdot$ définies par :

$$\left(\dot{j}, \ddot{k}\right) + \left(\dot{j}', \ddot{k}'\right) = \left(\dot{j} + \dot{j}', \ddot{k} + \ddot{k}'\right) \text{ et } \left(\dot{j}, \ddot{k}\right) \cdot \left(\dot{j}', \ddot{k}'\right) = \left(\dot{j} \cdot \dot{j}', \ddot{k} \cdot \ddot{k}'\right)$$

Supposons n et m premiers entre eux. L'application $\varphi : k \mapsto \left(\dot{k}, \ddot{k}\right)$ est un morphisme d'anneaux de $\mathbb{Z}$ dans $\frac{\mathbb{Z}}{n\mathbb{Z}} \times \frac{\mathbb{Z}}{m\mathbb{Z}}$ et son noyau est formé des entiers divisibles par n et m donc par nm puisque ces entiers sont premiers entre eux, il se factorise donc en un morphisme injectif d'anneaux de $\frac{\mathbb{Z}}{nm\mathbb{Z}}$ dans $\frac{\mathbb{Z}}{n\mathbb{Z}} \times \frac{\mathbb{Z}}{m\mathbb{Z}}$ par $\bar{\varphi} : \bar{k} \mapsto \left(\dot{k}, \ddot{k}\right)$. Ces deux anneaux ayant même cardinal, l'application $\bar{\varphi}$ réalise en fait un isomorphisme d'anneaux de $\frac{\mathbb{Z}}{nm\mathbb{Z}}$ dans $\frac{\mathbb{Z}}{n\mathbb{Z}} \times \frac{\mathbb{Z}}{m\mathbb{Z}}$. Si n et m ne sont pas premiers entre eux les groupes additifs $\frac{\mathbb{Z}}{nm\mathbb{Z}}$ et $\frac{\mathbb{Z}}{n\mathbb{Z}} \times \frac{\mathbb{Z}}{m\mathbb{Z}}$ ne peuvent être isomorphes puisque 1 est d'ordre nm dans $\frac{\mathbb{Z}}{nm\mathbb{Z}}$ et tous les éléments de $\frac{\mathbb{Z}}{n\mathbb{Z}} \times \frac{\mathbb{Z}}{m\mathbb{Z}}$ ont un ordre qui divise le ppcm de n et m qui est strictement inférieur à nm . $\square$

Corollaire 4.6. *Si n et m sont deux entiers naturels non nuls premiers entre eux, on a alors $\varphi(nm) = \varphi(n)\varphi(m)$.*

Preuve. On utilise les notations de la démonstration précédente. La restriction de l'isomorphisme $\bar{\varphi}$ à $\left(\frac{\mathbb{Z}}{nm\mathbb{Z}}\right)^\times$ réalise un isomorphisme de groupes multiplicatifs de $\left(\frac{\mathbb{Z}}{nm\mathbb{Z}}\right)^\times$ sur $\left(\frac{\mathbb{Z}}{n\mathbb{Z}}\right)^\times \times \left(\frac{\mathbb{Z}}{m\mathbb{Z}}\right)^\times$, ce qui entraîne que $\varphi(nm) = \text{card}\left(\left(\frac{\mathbb{Z}}{nm\mathbb{Z}}\right)^\times\right) = \text{card}\left(\left(\frac{\mathbb{Z}}{n\mathbb{Z}}\right)^\times\right) \text{card}\left(\left(\frac{\mathbb{Z}}{m\mathbb{Z}}\right)^\times\right) = \varphi(n)\varphi(m)$. $\square$

Le calcul de $\varphi(n)$ est alors ramené à celui de $\varphi(p^\alpha)$ où p est un nombre premier et α un entier naturel non nul.

Lemme 4.11 *Soient p un nombre premier et α un entier naturel non nul. On a $\varphi(p^\alpha) = (p-1)p^{\alpha-1}$.*

Preuve. Si p est premier, alors un entier k compris entre 1 et p^α n'est pas premier avec p^α si, et seulement si, il est divisible par p , ce qui équivaut à $k = mp$ avec

$1 \leq m \leq p^{\alpha-1}$, il y a donc $p^{\alpha-1}$ possibilités. On en déduit alors que $\varphi(p^\alpha) = p^\alpha - p^{\alpha-1} = (p-1)p^{\alpha-1}$. $\square$

Théorème 4.35.

Si $n \geq 1$ a pour décomposition en facteurs premiers $n = \prod_{i=1}^r p_i^{\alpha_i}$ avec $2 \leq p_1 < \dots < p_r$ premiers et les α_i entiers naturels non nuls, alors :

$$\varphi(n) = \prod_{i=1}^r p_i^{\alpha_i-1} (p_i - 1) = n \prod_{i=1}^r \left(1 - \frac{1}{p_i}\right)$$

Preuve. En utilisant les résultats précédents, on a :

$$\varphi(n) = \prod_{i=1}^r \varphi(p_i^{\alpha_i}) = \prod_{i=1}^r (p_i - 1) p_i^{\alpha_i-1} = n \prod_{i=1}^r \left(1 - \frac{1}{p_i}\right)$$

$\square$

De ce résultat on déduit que pour tout $n \geq 3$, $\varphi(n)$ est un entier pair. En effet, pour $n = 2^\alpha$ avec $\alpha \geq 2$, on a $\varphi(n) = 2^{\alpha-1}$ qui est pair et pour $n = 2^\alpha \prod_{i=1}^r p_i^{\alpha_i} = p_1^{\alpha_1} m$ avec $\alpha \geq 0$, $r \geq 1$, tous les p_i étant premiers impairs, on a $\varphi(n) = (p_1 - 1) p_1^{\alpha_1-1} \varphi(m)$ qui est pair.

4.13 Exercices

Exercice 4.1. Montrer que pour tous a, b dans $\mathbb{Z}$, on a :

$$(a\mathbb{Z} \subset b\mathbb{Z}) \Leftrightarrow (b/a) \Leftrightarrow (\mathcal{D}_b \subset \mathcal{D}_a) \text{ et } (a\mathbb{Z} = b\mathbb{Z}) \Leftrightarrow (a = \pm b) \Leftrightarrow (\mathcal{D}_a = \mathcal{D}_b)$$

Solution. Si $a\mathbb{Z} \subset b\mathbb{Z}$, on a alors $a \in b\mathbb{Z}$, c'est-à-dire qu'il existe un entier q tel que $a = bq$ et b/a . Si b/a , on a alors $a = qb$ avec $q \in \mathbb{Z}$ et tout diviseur δ de b va diviser a , ce qui signifie que $\mathcal{D}_b \subset \mathcal{D}_a$. Si $\mathcal{D}_b \subset \mathcal{D}_a$, on a alors $b \in \mathcal{D}_a$, c'est-à-dire qu'il existe un entier q tel que $a = bq$ et pour tout pa dans $a\mathbb{Z}$, on a $pa = pqb \in b\mathbb{Z}$, c'est-à-dire que $a\mathbb{Z} \subset b\mathbb{Z}$. On a donc ainsi montré la première série d'équivalence. Si $a\mathbb{Z} = b\mathbb{Z}$, on a alors $a\mathbb{Z} \subset b\mathbb{Z}$ et $b\mathbb{Z} \subset a\mathbb{Z}$, donc b/a et a/b et $a = \pm b$. Si $a = \pm b$, les entiers a et b ont les mêmes diviseurs, ce qui signifie que $\mathcal{D}_a = \mathcal{D}_b$. Si $\mathcal{D}_a = \mathcal{D}_b$, on a alors $\mathcal{D}_a \subset \mathcal{D}_b$ et $\mathcal{D}_b \subset \mathcal{D}_a$, donc b/a et a/b et $a = \pm b$ qui équivaut à $a\mathbb{Z} = b\mathbb{Z}$.

Exercice 4.2. Soient a et b dans $\mathbb{Z}$. Montrer que si $p = a^2 + b^2$ est impair supérieur ou égal à 3 alors $p - 1$ est multiple de 4.

Solution. Tout entier k est congru à 0, 1, 2 ou 3 modulo 4, donc k^2 est congru à 0 ou 1 modulo 4 et $a^2 + b^2$ est congru à 0, 1 ou 2 modulo 4. Si p est impair et $p = a^2 + b^2$ alors p est congru à 1 modulo 4 et $p - 1$ est multiple de 4.

Exercice 4.3. Calculer, pour tout entier naturel n , le reste dans la division euclidienne par 13 de l'entier $x_n = 4^{2n+1} + 3^{n+2}$.

Solution. On a :

$$\begin{aligned} x_n &= 4 \cdot 4^{2n} + 9 \cdot 3^n = 4 \cdot 16^n + 9 \cdot 3^n = 4(16^n - 3^n) + (4 + 9)3^n \\ &= 4(16 - 3) \sum_{k=1}^n 16^{n-k} 3^{k-1} + 13 \cdot 3^n = 13y_n \end{aligned}$$

Le reste dans la division euclidienne par 13 de x_n est donc nul, le quotient étant donné par $q_n = 4 \sum_{k=1}^n 16^{n-k} 3^{k-1} + 3^n$

Exercice 4.4. Soient a et b deux entiers naturels non nuls tels que $a > b$. Donner une condition suffisante sur les entiers a et b pour que tous les entiers $x_n = a^{2n+1} + b^{n+2}$, où n est un entier naturel, soient divisibles par $a + b^2$.

Solution. On a :

$$\begin{aligned} x_n &= a \cdot a^{2n} + b^2 \cdot b^n = a \cdot (a^2)^n + b^2 \cdot b^n = a \left((a^2)^n - b^n \right) + (a + b^2) b^n \\ &= a (a^2 - b) \sum_{k=1}^n (a^2)^{n-k} b^{k-1} + (a + b^2) b^n \end{aligned}$$

Si $a = b + 1$ (i. e. b et a sont deux entiers consécutifs), alors $a^2 - b = b^2 + b + 1 = a + b^2$ et x_n est divisible par $a + b^2$ pour tout n (la condition $a^2 = b$ n'est pas possible puisqu'on suppose que $b < a$). Pour $(a, b) = (4, 3)$ on retrouve l'exercice précédent).

Exercice 4.5. Soit n un entier naturel et $n = n_p \cdots n_1 n_0$ son écriture décimale. Montrer que :

- n est divisible par 2 si, et seulement si, n_0 est pair ;
- n est divisible par 5 si, et seulement si, n_0 est égal à 0 ou 5 ;
- n est divisible par 3 si, et seulement si, $\sum_{k=0}^p n_k$ est divisible par 3 ;
- n est divisible par 9 si, et seulement si, $\sum_{k=0}^p n_k$ est divisible par 9 ;

— n est divisible par 11 si, et seulement si, $\sum_{k=0}^p (-1)^k n_k$ est divisible par 11.

Solution. Ces critères de divisibilité se déduisent de la connaissance du reste dans la division euclidienne de 10 par 2, 3, 5, 9 et 11 respectivement. Comme 10 est congru à 0 modulo 2 et modulo 5, on déduit que n est congru à n_0 modulo 2 et modulo 5 et donc n est divisible par 2 (resp. par 5) si, et seulement si, son chiffre des unités n_0 est pair, c'est-à-dire égal à 0, 2, 4, 6 ou 8 (resp. multiple de 5, c'est-à-dire égal à 0 ou 5). Du fait que 10 est congru à 1 modulo 3 et modulo 9, on déduit que 10^k est congru à 1 modulo 3 et modulo 9 pour tout entier k et n est congru à $\sum_{k=0}^p n_k$ modulo 3 et modulo 9. Donc n est divisible par 3 (resp. par 9) si, et seulement si, la somme de ses chiffres est divisible par 3 (resp. par 9). Enfin du fait que 10 est congru à -1 modulo 11 on déduit que 10^k est congru à $(-1)^k$ modulo 11 pour tout entier k et n est congru à $\sum_{k=0}^p (-1)^k n_k$ modulo 11. Donc n est divisible par 11 si, et seulement si, la somme alternée de ses chiffres est divisible par 11.

Exercice 4.6. Soient a, b deux entiers relatifs non tous deux nuls et $\mathcal{D}_{a,b}$ l'ensemble des diviseurs communs à a et b dans $\mathbb{N}^*$, c'est-à-dire :

$$\mathcal{D}_{a,b} = \{d \in \mathbb{N}^* \mid d/a \text{ et } d/b\}.$$

Montrer que $a \wedge b$ est le plus grand élément pour l'ordre de la division dans $\mathbb{N}$ de $\mathcal{D}_{a,b}$ et que $\mathcal{D}_{a,b} = \mathcal{D}_{a \wedge b}$ (ensemble des diviseurs strictement positifs de $a \wedge b$). Sans référence directe aux sous-groupes de $\mathbb{Z}$, montrer que l'ensemble $\mathcal{D}_{a,b}$ admet donc un plus grand élément δ (δ est alors le plus grand diviseur communs à a et b).

Solution. On sait déjà que $\delta = a \wedge b$ divise a et b , donc $\delta \in \mathcal{D}_{a,b}$ et tout entier $d \in \mathcal{D}_{a,b}$ divisant a et b va diviser $\delta = au + bv$. Comme tout $d \in \mathcal{D}_{a,b}$ divise δ , on a $\mathcal{D}_{a,b} \subset \mathcal{D}_\delta$ et comme tout $d \in \mathcal{D}_\delta$ divise δ qui divise lui-même a et b , d va diviser a et b , soit $d \in \mathcal{D}_{a,b}$. On a donc $\mathcal{D}_{a,b} = \mathcal{D}_{a \wedge b}$. L'ensemble $\mathcal{D}_{a,b}$ est non vide car il contient 1. Comme a et b ne sont pas tous deux nuls, cet ensemble est fini puisqu'un entier relatif non nul n'a qu'un nombre fini de diviseurs dans $\mathbb{N}^*$. L'ensemble $\mathcal{D}_{a,b}$ est donc non vide et majoré dans $\mathbb{N}^*$, il admet donc un plus grand élément $\delta \in \mathbb{N}^*$ qui est bien le plus grand diviseur communs à a et b .

Exercice 4.7. Soient a et b deux entiers relatifs non tous deux nuls. Montrer que $a \wedge b = a \wedge (a + b) = b \wedge (a + b)$.

Solution. Si $(a, b) \neq (0, 0)$, on a alors $(a, a + b) \neq (0, 0)$ et $(b, a + b) \neq (0, 0)$. En notant $\delta = a \wedge b$ et $\delta' = a \wedge (a + b)$, on a :

$$\delta = au + bv = a(u - v) + (a + b)v \in a\mathbb{Z} + (a + b)\mathbb{Z} = \delta'\mathbb{Z}$$

donc $\delta\mathbb{Z} \subset \delta'\mathbb{Z}$ et :

$$\delta' = au' + (a + b)v' = a(u' + v') + bv' \in a\mathbb{Z} + b\mathbb{Z} = \delta\mathbb{Z}$$

donc $\delta'\mathbb{Z} \subset \delta\mathbb{Z}$ et donc $\delta\mathbb{Z} = \delta'\mathbb{Z}$, ce qui équivaut à $\delta = \delta'$. On peut aussi dire que comme δ divise a et b , il divise a et $a + b$, donc il divise leur pgcd δ' . De même δ' qui divise a et $a + b$ va diviser a et $b = (a + b) - a$, il divise donc δ et $\delta = \delta'$. En permutant les rôles de a et b , on a $a \wedge b = a \wedge (a + b) = b \wedge (a + b)$.

Exercice 4.8. Montrer que si $a_1, \dots, a_p$ sont des entiers relatifs non nuls deux à deux premiers entre eux alors $a_1 \vee \dots \vee a_p = |a_1 \dots a_p|$. Ce résultat est-il encore valable si on suppose que $a_1, \dots, a_p$ sont premiers entre eux dans leur ensemble. A-t-on $(a_1 \wedge \dots \wedge a_p)(a_1 \vee \dots \vee a_p) = |a_1 \dots a_p|$ dans $\mathbb{Z}^*$?

Solution. On sait déjà que si a_1 et a_2 sont premiers entre eux alors $a_1 \vee a_2 = |a_1 a_2|$. Supposons le résultat acquis pour $p - 1 \geq 2$ et soient $a_1, \dots, a_p$ deux à deux premiers entre eux. Les entiers $a_1 \dots a_{p-1}$ et a_p sont alors premiers entre eux (corollaire 4.2) et en utilisant l'associativité du ppcm, on a :

$$a_1 \vee \dots \vee a_p = (a_1 \vee \dots \vee a_{p-1}) \vee a_p = |a_1 \dots a_{p-1}| \vee a_p = |a_1 \dots a_p|.$$

Ce résultat n'est plus valable si on suppose seulement que les a_k sont premiers entre eux dans leur ensemble comme le montre l'exemple suivant :

$$2 \vee 3 \vee 4 = 12 \neq 2 \cdot 3 \cdot 4 = 24$$

La réponse à la deuxième question est non pour $n \geq 3$ comme le montre ce qui précède.

Exercice 4.9. Déterminer tous les couples (a, b) d'entiers naturels non nuls tels que $a \wedge b = 3$ et $a \vee b = 12$.

Solution. De $a \vee b = 12$ on déduit que a, b sont des diviseurs de 12 donc dans $\{1, 2, 3, 4, 6, 12\}$. de $a \wedge b = 3$, on déduit que a et b sont multiples de 3, donc dans $\{3, 6, 12\}$. De $ab = (a \wedge b)(a \vee b) = 36$, on déduit que :

- $a = 3$ [resp. $b = 3$] donne $b = 12$ [resp. $a = 12$] et $(3, 12)$, $(12, 3)$ sont deux solutions possibles ;
- $a = 6$ [resp. $b = 6$] donne $b = 6$ [resp. $a = 6$] et $a \wedge b = 6 \neq 3$.

En définitive, on a $(a, b) \in \{(3, 12), (12, 3)\}$.

Exercice 4.10. On se propose de montrer que pour tout entier naturel $n \geq 2$, on a $\mu_n = \text{ppcm}(1, 2, \dots, n) \geq 2^{n-2}$. Pour tout entier naturel n , on, note $I_n = \int_0^1 x^n (1-x)^n dx$.

1. Montrer que pour tout $n \in \mathbb{N}$, on a $0 < I_n \leq \frac{1}{4^n}$.
2. Montrer que, pour tout $n \in \mathbb{N}^*$, il existe un entier naturel non nul a_n tel que $I_n = \frac{a_n}{\mu_{2n+1}}$.
3. En déduire que pour tout $n \in \mathbb{N}^*$, on a $\mu_{2n+1} \geq 2^{2n}$.
4. En déduire le résultat annoncé.

Solution. Pour $n = 2$ et $n = 3$, on a $\mu_2 = 2 \geq 1$ et $\mu_3 = 6 \geq 2$.

1. Pour $0 < x < 1$, on a $0 < x(x-1) \leq \sup_{[0,1]} x(1-x) = \frac{1}{4}$, ce qui donne le résultat.
2. On a $I_n = \sum_{k=0}^n \binom{n}{k} (-1)^k \int_0^1 x^{n+k} dx = \sum_{k=0}^n \binom{n}{k} \frac{(-1)^k}{n+k+1}$ et en réduisant au même dénominateur, cela donne $I_n = \frac{a_n}{\mu_{2n+1}}$, où $a_n \in \mathbb{N}^*$.
3. On a alors $\mu_{2n+1} I_n = a_n \geq 1$ et $\mu_{2n+1} \geq \frac{1}{I_n} \geq 4^n = 2^{2n}$.
4. Pour $n \in \mathbb{N}^*$, on a $\mu_{2n+2} = \mu_{2n+1} \vee (2n+2) \geq 2^{2n}$. On a donc montré que $\mu_n \geq 2^{n-2}$ pour tout $n \geq 4$. On peut en fait montrer que $\mu_n \geq 2^n$ pour tout $n \geq 7$.

Exercice 4.11. Résoudre l'équation diophantienne $128x + 28y = 8$ dans $\mathbb{Z}^2$.

Solution. En notant $(a, b) = (128, 28)$, $c = 8$ et $\delta = a \wedge b$, on a vu au paragraphe précédent que $\delta = 4 = 2 \cdot 128 - 9 \cdot 28 = au_0 + bv_0$ et $(x_0, y_0) = \left(u_0 \frac{c}{\delta}, v_0 \frac{c}{\delta}\right) = (4, -18)$ est une solution particulière de notre équation. Toutes les solutions étant données par $(x, y) = (x_0 - kb', y_0 + ka') = (4 - 7k, -18 + 32k)$, où k décrit $\mathbb{Z}$.

Exercice 4.12. Soient $(a_k)_{1 \leq k \leq p}$ et $(b_k)_{1 \leq k \leq q}$ deux suites finies d'entiers relatifs non nuls. Montrer que si $n = \prod_{k=1}^p a_k$ et $m = \prod_{k=1}^q b_k$ sont premiers entre eux, alors chaque a_k , pour k compris entre 1 et p , est premier avec chacun des b_j , pour j compris entre 1 et q .

Solution. Soit $\delta = a_k \wedge b_j$ où $1 \leq k \leq p$ et $1 \leq j \leq q$. Comme δ est un entier naturel non nul qui divise a_k et b_j , il divise n et m et vaut nécessairement 1.

Exercice 4.13. Est-il équivalent de dire $a_1, \dots, a_p$ sont premiers entre eux dans leur ensemble et $a_1, \dots, a_p$ sont deux à deux premiers entre eux ?

Solution. On vérifie immédiatement que la réponse est négative en considérant le triplet $(2, 3, 8)$.

Exercice 4.14. Soient a, n deux entiers naturels non nuls. Montrer que $\frac{(a+1)^n - 1}{a} \wedge a = a \wedge n$.

Solution. On remarque d'abord que $(a+1)^n - 1 = a \sum_{k=0}^{n-1} (a+1)^k$ est divisible par a , donc $\frac{(a+1)^n - 1}{a}$ est un entier. Soit $\delta = \frac{(a+1)^n - 1}{a} \wedge a$. Pour tout $k \geq 0$, on a $(a+1)^k \equiv 1 \pmod{a}$ (pour $k = 0$, c'est clair et pour $k \geq 1$, on utilise la formule du binôme), donc $b = \frac{(a+1)^n - 1}{a} = \sum_{k=0}^{n-1} (a+1)^k \equiv n \pmod{a}$, de sorte que δ qui divise a et b divise aussi $n = b - pa$ ($p \in \mathbb{Z}$). Il en résulte que δ divise $\delta' = a \wedge n$. Comme δ' divise a et n , il divise aussi $b = n + pa$ et en conséquence δ' divise δ . On a donc bien $\delta = \delta'$.

Exercice 4.15. On se donne un entier naturel $a \geq 2$ et on définit la suite $(u_n)_{n \in \mathbb{N}}$ par $u_n = a^{2^n} + 1$.

1. Montrer que $u_{n+1} = (u_n - 1)^2 + 1$ pour tout $n \in \mathbb{N}$.
2. Montrer que $u_{n+1} = (a - 1) \prod_{k=0}^n u_k + 2$ pour tout $n \in \mathbb{N}$.
3. Montrer que, pour $n \neq m$ dans $\mathbb{N}$, on a :

$$u_n \wedge u_m = \begin{cases} 1 & \text{si } a \text{ est pair} \\ 2 & \text{si } a \text{ est impair} \end{cases}$$

4. Calculer $u_n^p \wedge u_m^p$ pour $n \neq m$ dans $\mathbb{N}$ et p dans $\mathbb{N}^*$.

Solution.

1. On a $u_{n+1} = a^{2^{n+1}} + 1 = (a^{2^n})^2 + 1 = (u_n - 1)^2 + 1$.
2. Pour $n = 0$, on a $u_1 = a^2 + 1 = (a^2 - 1) + 2 = (a - 1)u_0 + 2$. En supposant le résultat acquis pour $n - 1 \geq 0$, on a :

$$u_{n+1} = u_n(u_n - 2) + 2 = u_n(a - 1) \prod_{k=0}^{n-1} u_k + 2 = (a - 1) \prod_{k=0}^n u_k + 2$$

3. Supposons que $m > n$. On a :

$$u_m = (a-1) \prod_{k=0}^{m-1} u_k + 2 = (a-1) u_n \prod_{\substack{k=0 \\ k \neq n}}^{m-1} u_k + 2 = qu_n + 2$$

Le pgcd de u_n et u_m divise alors 2 et il vaut 2 ou 1. Si a est pair, alors u_n est impair et $\delta = 1$ puisqu'il divise u_n , ce qui signifie que u_n et u_m sont premiers entre eux (pour $a = 2$, c'est le cas des nombres de Fermat). Si a est impair, alors tous les u_n sont pairs et δ vaut 2.

4. On a :

$$u_n^p \wedge u_m^p = (u_n \wedge u_m)^p = \begin{cases} 1 & \text{si } a \text{ est pair} \\ 2^p & \text{si } a \text{ est impair} \end{cases}$$

Exercice 4.16.

1. Montrer que pour tout entier naturel n , il existe deux entiers p_n et q_n premiers entre eux tels que $(\sqrt{2} + 1)^n = p_n + q_n\sqrt{2}$.
2. En utilisant l'application φ définie sur l'anneau $\mathbb{Z}[\sqrt{2}]$ par $\varphi(a + b\sqrt{2}) = a - b\sqrt{2}$ pour tout $(a, b) \in \mathbb{Z}^2$, montrer que, pour tout $n \in \mathbb{N}$, on a $(\sqrt{2} - 1)^n = (-1)^n (p_n - q_n\sqrt{2})$.
3. En déduire que, pour tout $n \in \mathbb{N}$, il existe un entier naturel r_n tel que :

$$\begin{cases} (\sqrt{2} + 1)^n = \sqrt{r_n} + \sqrt{r_n - 1} \\ (\sqrt{2} - 1)^n = \sqrt{r_n} - \sqrt{r_n - 1} \end{cases}$$

Solution.

1. Le réel $\theta = \sqrt{2} + 1$ est dans l'anneau $\mathbb{Z}[\sqrt{2}]$, il en est donc de même de θ^n pour tout $n \in \mathbb{N}$, ce qui prouve l'existence des suites d'entiers $(p_n)_{n \in \mathbb{N}}$ et $(q_n)_{n \in \mathbb{N}}$. On peut aussi retrouver ce résultat par récurrence en écrivant que pour tout $n \in \mathbb{N}$, on a $\theta^{n+1} = (\sqrt{2} + 1)(p_n + q_n\sqrt{2}) = (p_n + 2q_n) + (p_n + q_n)\sqrt{2}$, ce qui donne $p_{n+1} = p_n + 2q_n$, $q_{n+1} = p_n + q_n$ et montre en outre que les p_n et q_n sont des entiers naturels non nuls sauf $q_0 = 0$. On a alors $p_{n+1} \wedge q_{n+1} = (p_n + 2q_n) \wedge (p_n + q_n)$. En utilisant la relation $a \wedge b = a \wedge (a + b) = b \wedge (a + b)$ (exercice 4.7), on a :

$$\begin{aligned} p_n \wedge q_n &= q_n \wedge (p_n + q_n) = (q_n + p_n) \wedge q_n \\ &= (q_n + p_n) \wedge (p_n + q_n + q_n) = p_{n+1} \wedge q_{n+1} \end{aligned}$$

On a donc $p_n \wedge q_n = p_0 \wedge q_0 = 1 \wedge 0 = 1$ pour tout $n \in \mathbb{N}$.

2. L'application φ réalise un automorphisme de l'anneau $A = \mathbb{Z}[\sqrt{2}]$. En effet, pour tous $x = a + b\sqrt{2}$ et $x' = a' + b'\sqrt{2}$ dans A , on a :

$$\begin{cases} \varphi(x + x') = (a + a') - (b + b')\sqrt{2} = \varphi(x) + \varphi(x') \\ \varphi(xx') = (aa' + 2bb') - (ab' + a'b)\sqrt{2} = \varphi(x)\varphi(x') \end{cases}$$

et $x = a + b\sqrt{2}$ a pour unique antécédent $a - b\sqrt{2}$ par φ . On a donc, pour tout $n \in \mathbb{N}$:

$$\begin{aligned} (\sqrt{2} - 1)^n &= (-1)^n (1 - \sqrt{2})^n = (-1)^n \left(\varphi(\sqrt{2} + 1) \right)^n \\ &= (-1)^n \varphi \left((\sqrt{2} + 1)^n \right) = (-1)^n \varphi(p_n + q_n \sqrt{2}) = (-1)^n (p_n - q_n \sqrt{2}) \end{aligned}$$

3. On a $(\sqrt{2} + 1)^n (\sqrt{2} - 1)^n = 1$, soit $(-1)^n (p_n + q_n \sqrt{2}) (p_n - q_n \sqrt{2}) = 1$, ou encore $p_n^2 - 2q_n^2 = (-1)^n$ (c'est une relation de Bézout pour p_n et q_n qui sont premiers entre eux) et :

$$\begin{cases} (\sqrt{2} + 1)^n = p_n + q_n \sqrt{2} = p_n + \sqrt{p_n^2 - (-1)^n} \\ (\sqrt{2} - 1)^n = (-1)^n (p_n - q_n \sqrt{2}) = (-1)^n (p_n - \sqrt{p_n^2 - (-1)^n}) \end{cases}$$

En posant $s_n = p_n^2$, on a :

$$\begin{cases} (\sqrt{2} + 1)^n = \sqrt{s_n} + \sqrt{s_n - (-1)^n} \\ (\sqrt{2} - 1)^n = (-1)^n (\sqrt{s_n} - \sqrt{s_n - (-1)^n}) \end{cases}$$

Pour n pair, cela s'écrit :

$$\begin{cases} (\sqrt{2} + 1)^n = \sqrt{s_n} + \sqrt{s_n - 1} = \sqrt{r_n} + \sqrt{r_n - 1} \\ (\sqrt{2} - 1)^n = \sqrt{s_n} - \sqrt{s_n - 1} = \sqrt{r_n} - \sqrt{r_n - 1} \end{cases}$$

avec $r_n = s_n$ et pour n impair :

$$\begin{cases} (\sqrt{2} + 1)^n = \sqrt{s_n + 1} + \sqrt{s_n} = \sqrt{r_n} + \sqrt{r_n - 1} \\ (\sqrt{2} - 1)^n = \sqrt{s_n + 1} - \sqrt{s_n} = \sqrt{r_n} - \sqrt{r_n - 1} \end{cases}$$

avec $r_n = s_n - 1$.

Exercice 4.17. Soient $a \geq 2$ et $m \geq 2$ deux entiers et $p = a^m - 1$. Montrer que si p est premier alors $a = 2$ et m est premier. La réciproque est-elle vraie ? On appelle nombre de Mersenne tout entier de la forme $2^m - 1$.

Solution. Supposons que p soit premier. On a $p = a^m - 1 = (a - 1) \sum_{k=0}^{m-1} a^k = (a - 1)q$. Si $a > 2$, on a $a - 1 \geq 2$ et $q \geq 2$ puisque $m \geq 2$ et p ne peut être premier. On a donc nécessairement $a = 2$ et $p = 2^m - 1$. Si m n'est pas premier, il s'écrit $m = ab$ avec $a \geq 2$, $b \geq 2$ et :

$$p = 2^{ab} - 1 = (2^a)^b - 1 = (2^a - 1) \sum_{k=0}^{b-1} (2^a)^k = (2^a - 1)q$$

avec $2^a - 1 \geq 2$ (puisque $2^a \geq 4$) et $q \geq 2$ (puisque $b \geq 2$) et p ne peut être premier. L'entier m est donc nécessairement premier. Pour $m = 2, 3, 5, 7$, on a $p = 3, 7, 31, 127$ qui sont premiers et pour $m = 11$, on a $p = 2^{11} - 1 = 2047 = 23 \times 89$. La réciproque est donc fausse.

Exercice 4.18. Soient $p \geq 2$ un nombre premier et n, m des entiers naturels non nuls. Montrer que p divise n ou p^m est premier avec n .

Solution. Si p divise n c'est fini. Sinon p est premier avec n et le théorème de Bézout nous dit qu'il existe deux entiers relatifs u et v tels que $up + vn = 1$. On a alors $1 = (up + vn)^m = u^m p^m + v_m n$, ce qui signifie que p^m et n sont premiers entre eux.

Exercice 4.19. Soient a et b deux entiers relatifs non nuls et n un entier naturel non nul. Montrer que $a^n \wedge b^n = (a \wedge b)^n$ et $a^n \vee b^n = (a \vee b)^n$. On peut en déduire que deux entiers relatifs non nuls a et b sont premiers entre eux si, et seulement si, a^n et b^n sont premiers entre eux, quel que soit l'entier $n \geq 1$.

Solution. Soient $\delta = a \wedge b$ et p, q premiers entre eux tels que $a = \delta p$ et $b = \delta q$. On a $a^n \wedge b^n = (\delta^n p^n) \wedge (\delta^n q^n) = \delta^n (p^n \wedge q^n)$ et $(a \wedge b)^n = \delta^n$. Il s'agit alors de montrer que $\delta' = p^n \wedge q^n = 1$ si $p \wedge q = 1$ (on s'est ramené en fait au cas où a et b sont premiers entre eux). Supposons que $\delta' \geq 2$, il admet alors un diviseur premier $d \geq 2$ et d qui divise p^n et q^n va diviser p et q ce qui est impossible. On a donc nécessairement $\delta' = 1$. Pour ce qui est du ppcm, on a :

$$a^n \vee b^n = \frac{|a|^n |b|^n}{a^n \wedge b^n} = \frac{(|a| |b|)^n}{(a \wedge b)^n} = (a \vee b)^n$$

Exercice 4.20. Soient a, b, c, d des entiers relatifs non nuls. Montrer que si $a \wedge b = c \wedge d = 1$, alors $(ac) \wedge (bd) = (a \wedge d) (b \wedge c)$.

Solution. En notant $\delta_1 = a \wedge d$ et $\delta_2 = b \wedge c$, on a $a = \delta_1 p_1$, $d = \delta_1 q_1$, $b = \delta_2 p_2$ et $c = \delta_2 q_2$ avec $p_1 \wedge q_1 = p_2 \wedge q_2 = 1$, ce qui donne $(ac) \wedge (bd) = \delta_1 \delta_2 ((p_1 q_2) \wedge (p_2 q_1))$. Si $\delta' = ((p_1 q_2) \wedge (p_2 q_1)) \geq 2$, il admet alors un diviseur premier p qui divise $p_1 q_2$ et $p_2 q_1$ et on a quatre possibilités :

- soit p divise p_1 et q_1 , ce qui est impossible puisque $p_1 \wedge q_1 = 1$;
- soit p divise q_2 et p_2 , ce qui est impossible puisque $p_2 \wedge q_2 = 1$;
- soit p divise p_1 et p_2 et il divise alors a et b ce qui est impossible puisque $a \wedge b = 1$;
- soit p divise q_2 et q_1 et il divise alors c et d ce qui est impossible puisque $c \wedge d = 1$.

La seule possibilité est donc $\delta' = 1$.

Exercice 4.21. Montrer que pour tout entier naturel n , on peut trouver un nombre premier p plus grand que n . Conclure.

Solution. Pour tout $n \in \mathbb{N}$, l'entier $m = n! + 1 \geq 2$ admet un diviseur premier p_n . Si $p_n < n$ alors p_n est un diviseur de $n!$, donc de $1 = m - n!$, ce qui est impossible. On peut donc construire une suite strictement croissante $(p_n)_{n \in \mathbb{N}}$ de nombres premiers, ce qui implique que $\mathcal{P}$ est infini.

Exercice 4.22. On note $\mathcal{P}_1 = \{p \in \mathcal{P} \mid \exists n \in \mathbb{N} ; p = 4n + 3\}$ et $\mathcal{P}_2 = \{p \in \mathcal{P} \mid \exists n \in \mathbb{N} ; p = 6n + 5\}$. Montrer, en s'inspirant de la démonstration du théorème d'Euclide, que $\mathcal{P}_1$ [resp. $\mathcal{P}_2$] est infini et conclure. De manière plus générale on peut montrer que si a et b sont deux entiers premiers entre eux alors il existe une infinité de nombres premiers de la forme $an + b$ (théorème de Dirichlet).

Solution. On remarque qu'un nombre premier différent de 2 est nécessairement impair et son reste dans la division euclidienne par 4 [resp. par 6] ne peut être que 1 ou 3 [resp. 1, 3 ou 5]. Supposons que $\mathcal{P}_1$ [resp. $\mathcal{P}_2$] soit fini et notons $p_1 = 3$ [resp. $p_1 = 5$] $< p_2 < \dots < p_r$ tous ses éléments. L'entier :

$$m = 4p_1 \cdots p_r - 1 = 4(p_1 \cdots p_r - 1) + 3$$

$$\text{resp. } m = 6p_1 \cdots p_r - 1 = 6(p_1 \cdots p_r - 1) + 5$$

qui est de la forme $4n + 3$ [resp. $6n + 5$] avec $n \geq 2$ n'est pas premier puisque strictement supérieur à tous les p_k pour k compris entre 1 et r ($m > 4p_k - 1 > p_k$ puisque $p_k \geq 3$) [resp. $m > 6p_k - 1 > p_k$ puisque $p_k \geq 5$]. Comme m est impair [resp. impair non multiple de 3 puisque congru à 5 modulo 3] ses diviseurs premiers sont de la forme $4k + 1$ avec $k \in \mathbb{N}^*$ ou $4k + 3$ avec $k \in \mathbb{N}$ [resp. $6k + 1$ avec $k \in \mathbb{N}^*$ ou $6k + 5$ avec $k \in \mathbb{N}$] et ils ne peuvent pas être tous de la forme $4k + 1$ [resp. $6k + 1$] sans quoi m serait aussi de cette forme, donc congru à 1 modulo 4 [resp. modulo 6] ce qui contredit le fait qu'il est congru à 3 [resp. à 5] (ou à -1) modulo 4 [resp. modulo 6]. L'entier m a donc un diviseur p_k dans $\mathcal{P}_1$ [resp. dans $\mathcal{P}_2$] et comme p_k divise $p_1 \cdots p_r$, il va aussi diviser -1 , ce qui est impossible avec p_k premier. L'ensemble $\mathcal{P}_1$ [resp. $\mathcal{P}_2$] est donc infini. De $\mathcal{P}_1 \subset \mathcal{P}$ [resp. $\mathcal{P}_2 \subset \mathcal{P}$] on déduit que $\mathcal{P}$ est infini.

Exercice 4.23. L'unicité dans la décomposition en facteurs premiers peut être utilisée pour montrer que $\mathbb{Q}$ (ou $\mathbb{N}^2$) est dénombrable. On désigne par f l'application définie sur $\mathbb{N}^2 = \mathbb{N} \times \mathbb{N}$ par $f(n, m) = 2^n 3^m$. Montrer que f est injective. Il résulte que $\mathbb{N}^2$ est dénombrable.

Solution. L'égalité $f(n, m) = f(n', m')$ avec (n, m) et (n', m') dans $\mathbb{N}^2$ équivaut à $2^n 3^m = 2^{n'} 3^{m'}$ et l'unicité de la décomposition en facteurs premiers d'un entier naturel non nul nous dit que $(n, m) = (n', m')$. L'application f est donc injective de $\mathbb{N}^2$ dans $\mathbb{N}$ et bijective de $\mathbb{N}^2$ dans $f(\mathbb{N}^2) \subset \mathbb{N}$.

Exercice 4.24. Soit $n \geq 2$ un entier sans facteur carré, c'est-à-dire que n a une décomposition en facteurs premiers de la forme $n = \prod_{k=1}^r p_k$ où les p_k sont premiers deux à deux distincts. Montrer que $\sqrt{n}$ est irrationnel.

Solution. Si $\sqrt{n}$ est rationnel, il s'écrit alors $\sqrt{n} = \frac{a}{b}$, où a, b sont deux entiers naturels non nuls premiers entre eux. On a $a^2 = nb^2$ et si p est un diviseur premier de a (on a bien $a \geq 2$ puisque $\sqrt{n} > 1$), p^2 divise nb^2 en étant premier avec b^2 (a et b sont premiers entre eux), il divise n (théorème de Gauss), ce qui contredit le fait que n est sans facteur carré. Donc $\sqrt{n}$ est irrationnel.

Exercice 4.25. Soit n un entier de la forme $n = 2^m + 1$ avec $m \geq 0$. Montrer que si n est premier alors $m = 0$ ou m est une puissance de 2 (ce qui revient à dire que $n = 2^{2^p} + 1$ est un nombre de Fermat).

Solution. Si $m = 0$, alors $n = 2$ est premier. Si $m = 1 = 2^0$, alors $n = 3$ est premier. On suppose que $m \geq 2$. La décomposition en facteurs premiers de m permet d'écrire que $m = 2^p(2q + 1)$ où p et q sont des entiers naturels. Si q est non nul, on a alors :

$$n = \left(2^{2^p}\right)^{2q+1} + 1 = a^{2q+1} + 1 = (a+1) \sum_{k=0}^{2q} (-1)^k a^{2q-k} = (a+1)b$$

avec $a+1 = 2^{2^p} + 1 \geq 3$ et $b = \frac{n}{a+1} = \frac{a \cdot a^{2q} + 1}{a+1} > 1$ ($a \geq 2$ et $q \geq 1$), donc n n'est pas premier.

Exercice 4.26. Soient $p \neq q$ deux nombres premiers et $n = pq$. Montrer que si a et b sont deux entiers naturels tels que $ab \equiv 1 \pmod{\varphi(n)}$, alors pour tout entier relatif c , on a $c^{ab} \equiv c \pmod{n}$. Ce résultat est à la base du système cryptographique R.S.A.

Solution. Si $ab \equiv 1 \pmod{\varphi(n)}$, il existe alors $k \in \mathbb{Z}$ tel que $ab = 1 + k\varphi(n) = 1 + k(p-1)(q-1)$. Si c est un entier relatif premier avec p , on a alors $c^{p-1} \equiv 1 \pmod{p}$ (théorème de Fermat) et $c^{ab} = c^{k(p-1)(q-1)+1} \equiv c \pmod{p}$. Si l'entier relatif c n'est pas premier avec p , c'est nécessairement un multiple de p (qui est premier) et $c^{ab} \equiv 0 \equiv c \pmod{p}$. De manière analogue, on a $c^{ab} \equiv c \pmod{q}$ et avec p et q premiers entre eux il en résulte que $c^{ab} \equiv c \pmod{pq}$.

Chapitre 5

Le corps $\mathbb{C}$ des nombres complexes

Les ensembles $\mathbb{Z}$ d'entiers relatifs et $\mathbb{Q}$ de nombres rationnels peuvent être construits à partir de problèmes analogues. Pour l'ensemble $\mathbb{Z}$ il s'agit des équations $x + a = 0$ qui n'ont pas de solution dans $\mathbb{N}$ pour a entier naturel non nul et pour l'ensemble $\mathbb{Q}$ il s'agit des équations $ax = 1$ qui n'ont pas de solution dans $\mathbb{Z}$ pour a entier relatif différent de $-1, 0$ et 1 . Le passage de l'ensemble $\mathbb{Q}$ de nombres rationnels à l'ensemble $\mathbb{R}$ de nombres réels est plus délicat. Les problèmes sont de nature algébrique (par exemple l'équation $x^2 = 2$ n'a pas de solution dans $\mathbb{Q}$) mais aussi de nature topologique : l'existence de borne supérieure pour les ensembles non vides et majorés n'est pas assurée dans $\mathbb{Q}$ alors qu'elle l'est dans $\mathbb{R}$ (par exemple l'ensemble $A = \{r \in \mathbb{Q} \mid r^2 \leq 2\}$ n'a pas de borne supérieure dans $\mathbb{Q}$). On consultera le cours d'analyse pour de plus amples détails sur la construction de l'ensemble $\mathbb{R}$ des nombres réels.

La construction de l'ensemble des nombres complexes est motivée par le fait que certaines équations polynomiales telles que l'équation $x^2 + 1 = 0$ n'ont pas de solutions réelles.

Le but de ce chapitre est de construire un ensemble que nous noterons $\mathbb{C}$ qui contient $\mathbb{R}$ et qui est muni d'opérations d'addition et de multiplication ayant les mêmes propriétés que leurs analogues sur $\mathbb{R}$, ce qui se traduira en disant que $\mathbb{C}$ est un corps commutatif. De plus dans cet ensemble $\mathbb{C}$ toute équation algébrique $P(x) = 0$, où P est un polynôme non constant, a des solutions, ce qui se traduira en disant que $\mathbb{C}$ est algébriquement clos. Dans un premier temps, on se contentera de décrire les solutions des équations de degré 2, $x^2 + bx + c = 0$. Pour les équations de degré supérieur, on dispose du théorème de d'Alembert-Gauss dit théorème fondamental de l'algèbre dont la démonstration classique nécessite des outils d'analyse réelle tels que le fait qu'une fonction continue sur un compact de $\mathbb{C}$ est bornée et atteint ses bornes. Nous verrons aussi que contrairement à $\mathbb{R}$, l'ensemble $\mathbb{C}$ que nous aurons construit ne peut pas être muni d'une relation d'ordre compatible avec la multiplication, c'est-à-dire telle que si $x \leq y$ et $0 \leq z$, alors $x \cdot z \leq y \cdot z$.

5.1 Conditions nécessaires à la construction de $\mathbb{C}$

Supposons que nous ayons construit un ensemble $\mathbb{C}$ contenant $\mathbb{R}$ muni d'opérations d'addition et multiplication qui prolongent celles que nous connaissons sur les réels avec les mêmes propriétés (mises à part celle relative à la relation d'ordre $\leq$ sur $\mathbb{R}$) et tel que l'équation $x^2 + 1 = 0$ admette au moins une solution i dans $\mathbb{C}$.

Pour tous réels x, y , le nombre $z = x + iy$ sera alors dans $\mathbb{C}$ et l'égalité $z = 0$ est réalisée si, et seulement si, $x = y = 0$. En effet, si $y = 0$, alors $x = 0$ et si $y \neq 0$, alors $i = -\frac{x}{y}$ est réel, ce qui n'est pas possible puisque l'équation $x^2 + 1 = 0$ n'a pas de solution réelle. Il en résulte que pour x, x', y, y' réels l'égalité $x + iy = x' + iy'$ est réalisée si, et seulement si, $x = x'$ et $y = y'$.

De plus pour l'addition et la multiplication de deux éléments $z = x + iy$ et $z' = x' + iy'$ de $\mathbb{C}$, on doit avoir :

$$\begin{cases} z + z' = (x + x') + i(y + y') \\ zz' = (xx' - yy') + i(xy' + yx') \end{cases}$$

5.2 Construction de $\mathbb{C}$

Les considérations précédentes nous conduisent à définir sur l'ensemble $\mathbb{R}^2$ des couples de réels les opérations d'addition et de multiplication suivantes, en désignant par $z = (x, y)$ et $z' = (x', y')$ deux éléments quelconques de $\mathbb{R}^2$:

$$\begin{cases} z + z' = (x + x', y + y') \\ z \cdot z' = (xx' - yy', xy' + yx') \end{cases}$$

On notera $\mathbb{C}$ l'ensemble $\mathbb{R}^2$ muni de ces deux opérations (ou lois de composition interne) et on l'appelle ensemble des nombres complexes.

La multiplication de deux nombres complexes z et z' sera notée $z \cdot z'$ ou plus simplement zz' .

L'égalité de deux nombres complexes $z = (x, y)$ et $z' = (x', y')$ est réalisée si, et seulement si, on a les égalités $x = x'$ et $y = y'$ (c'est ce qui se passe dans tout produit cartésien $E \times F$).

En particulier, pour $y = y' = 0$, on a $(x, 0) = (x', 0)$ si, et seulement si, $x = x'$, ce qui signifie que l'application $x \in \mathbb{R} \mapsto (x, 0) \in \mathbb{C}$ est injective, ce qui permet de réaliser une bijection de $\mathbb{R}$ sur le sous ensemble $\mathbb{R}'$ de $\mathbb{C}$ formé des couples $(x, 0)$. Cette bijection permet d'identifier $\mathbb{R}$ à $\mathbb{R}'$, ce qui signifie qu'un nombre réel x est identifié à son image $(x, 0)$ dans $\mathbb{C}$. Cette identification $x = (x, 0)$ est bien compatible avec les opérations d'addition et de multiplication des réels dans le sens où :

$$\begin{cases} x + x' = (x, 0) + (x', 0) = (x + x', 0) = x + x' \\ xx' = (x, 0)(x', 0) = (xx', 0) = xx' \end{cases}$$

L'opération d'addition vérifie les propriétés suivantes, déduites des propriétés analogues sur $\mathbb{R}$:

- elle est commutative, c'est-à-dire que pour tous nombres complexes $z = (x, y)$ et $z' = (x', y')$, on a $z + z' = (x + x', y + y') = (x' + x, y' + y) = z' + z$;

- elle est associative, c'est-à-dire que pour tous nombres complexes $z = (x, y)$, $z' = (x', y')$ et $z'' = (x'', y'')$, on a :

$$\begin{aligned} z + (z' + z'') &= (x, y) + (x' + x'', y' + y'') = (x + x' + x'', y + y' + y'') \\ &= ((x + x') + x'', (y + y') + y'') = (x + x', y + y') + (x'', y'') \\ &= (z + z') + z'' \end{aligned}$$

- le réel $0 = (0, 0)$ est un élément neutre, c'est-à-dire que pour tout nombre complexe $z = (x, y)$, on a $z + 0 = 0 + z = z$;
- tout nombre complexe $z = (x, y)$ admet un opposé donné par $z' = (-x, -y)$, ce qui signifie que $z + z' = z' + z = 0$. On note $-z$ cet opposé.

Tout cela se traduit en disant que $(\mathbb{C}, +)$ est un groupe commutatif comme $(\mathbb{R}, +)$, $(\mathbb{Q}, +)$ et $(\mathbb{Z}, +)$ et comme pour n'importe quel groupe, on peut vérifier que :

- l'élément neutre est unique ;
- pour tout $z \in \mathbb{C}$, l'opposé est unique ;
- tout élément de $\mathbb{C}$ est simplifiable (ou régulier) pour l'addition, c'est-à-dire que si $z + z' = z + z''$, alors $z' = z''$.

Pour ce qui est de l'autre opération de multiplication, on a les propriétés suivantes, encore déduites des propriétés analogues sur $\mathbb{R}$:

- elle est commutative, c'est-à-dire que pour tous nombres complexes $z = (x, y)$ et $z' = (x', y')$, on a $zz' = (xx' - yy', xy' + yx') = (x'x - y'y, y'x + x'y) = z'z$;
- elle est associative, c'est-à-dire que pour tous nombres complexes $z = (x, y)$, $z' = (x', y')$ et $z'' = (x'', y'')$, on a $z(z'z'') = (zz')z''$. En effet, on a :

$$\begin{aligned} z(z'z'') &= (x, y)(x'x'' - y'y'', x'y'' + y'x'') \\ &= (x(x'x'' - y'y'') - y(x'y'' + y'x''), x(x'y'' + y'x'') + y(x'x'' - y'y'')) \\ &= (xx'x'' - xy'y'' - yx'y'' - yy'x'', xx'y'' + xy'x'' + yx'x'' - yy'y'') \end{aligned}$$

et :

$$\begin{aligned} (zz')z'' &= (xx' - yy', xy' + yx')(x'', y'') \\ &= ((xx' - yy')x'' - (xy' + yx')y'', (xx' - yy')y'' + (xy' + yx')x'') \\ &= (xx'x'' - xy'y'' - yx'y'' - yy'x'', xx'y'' + xy'x'' + yx'x'' - yy'y'') \\ &= z(z'z'') \end{aligned}$$

on peut remarquer que seule la commutativité de l'addition des réels a été utilisée ici ;

- elle est distributive par rapport à l'addition, c'est-à-dire que pour tous nombres complexes z , z' et z'' , on a $z(z' + z'') = zz' + zz''$, ce qui se vérifie encore sans problème ;
- le réel $1 = (1, 0)$ est un élément neutre, c'est-à-dire que pour tout nombre complexe $z = (x, y)$, on a $z \cdot 1 = 1 \cdot z = z$;

- tout nombre complexe $z = (x, y)$ différent de 0 admet un inverse donné par $z' = \left(\frac{x}{x^2 + y^2}, -\frac{y}{x^2 + y^2} \right)$, ce qui signifie que $zz' = z'z = 1$. En effet l'égalité $zz' = 1$ équivaut à $xx' - yy' = 1$ et $xy' + yx' = 0$, ce qui entraîne :

$$\begin{cases} x^2x' - xyy' = x \\ yxy' + y^2x' = 0 \end{cases} \quad \text{et} \quad \begin{cases} yxx' - y^2y' = y \\ x^2y' + xyx' = 0 \end{cases}$$

et en additionnant les deux premières égalités [resp. en soustrayant les deux dernières], on obtient $(x^2 + y^2)x' = x$, $(x^2 + y^2)y' = -y$, ce qui donne compte tenu de $x^2 + y^2 \neq 0$ pour $z \neq 0$, $x' = \frac{x}{x^2 + y^2}$ et $y' = -\frac{y}{x^2 + y^2}$. Réciproquement, on vérifie facilement que cette solution convient. On note z^{-1} ou $\frac{1}{z}$ cet inverse. On peut remarquer qu'on a utilisé ici la commutativité du produit des réels.

Tout cela se traduit en disant que $(\mathbb{C}, +, \cdot)$ est un corps commutatif comme $(\mathbb{R}, +, \cdot)$, $(\mathbb{Q}, +, \cdot)$. Là encore le neutre et l'inverse sont uniques et tout nombre complexe non nul est simplifiable pour le produit.

On note $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$ et ce qui précède nous dit que $(\mathbb{C}^*, \cdot)$ est un groupe commutatif.

On peut remarquer que pour x réel non nul on a bien $\frac{1}{x} = \left(\frac{1}{x}, 0 \right) = \frac{1}{(x, 0)}$.

Ces opérations d'addition et multiplication prolongent bien celles de $\mathbb{R}$.

L'associativité de la multiplication permet de définir les puissances n -ièmes d'un nombre complexe z par $z^0 = 1$ et $z^{n+1} = z^n \cdot z$ pour tout $n \in \mathbb{N}^*$.

Pour $z \neq 0$ et $n \in \mathbb{N}^*$, on a $z^n \neq 0$ et $(z^n)^{-1} = \frac{1}{z^n} = (z^{-1})^n$. On note alors $z^{-n} = \frac{1}{z^n}$.

Comme sur $\mathbb{R}$, on a $z^{p+q} = z^p z^q$ et $(z^p)^q = z^{pq}$ pour tous entiers relatifs p et q .

Comme sur $\mathbb{R}$, une égalité $zz' = 0$ équivaut à $z = 0$ ou $z' = 0$. En effet si $z \neq 0$, il admet un inverse et $0 = z^{-1} \cdot 0 = z^{-1}zz' = z'$.

En posant $i = (0, 1)$, on a $i^2 = (0, 1)(0, 1) = (-1, 0) = -1$. De cette égalité, on déduit que $\frac{1}{i} = -i$.

Le nombre complexe $-i$ est aussi solution de l'équation $x^2 + 1 = 0$ et $i, -i$ sont les seules solutions de cette équation. En effet si $\alpha^2 + 1 = 0$, on a $\alpha^2 = -1 = i^2$ et $\alpha^2 - i^2 = (\alpha - i)(\alpha + i) = 0$ de sorte que $\alpha = i$ ou $\alpha = -i$.

Théorème 5.1.

Tout nombre complexe s'écrit de manière unique $z = x + iy$, où x et y sont deux réels.

Preuve. Un nombre complexe s'écrit de manière unique :

$$z = (x, y) = (x, 0) + (0, y) = (x, 0)(1, 0) + (y, 0)(0, 1) = x \cdot 1 + y \cdot i = x + iy$$

□

Définition 5.1. Avec les notations du théorème qui précède, on dit que x est la partie réelle de z et y sa partie imaginaire, ce qui se note $x = \operatorname{Re}(z)$ et $y = \operatorname{Im}(z)$.

Définition 5.2. On dit qu'un nombre complexe est un imaginaire pur si sa partie réelle est nulle.

En résumé un nombre complexe s'écrit $z = x + iy$ où $i^2 = -1$ et pour $z = x + iy$, $z' = x' + iy'$ dans $\mathbb{C}$, on a :

$$\begin{cases} (z = z') \Leftrightarrow (x = x' \text{ et } y = y') \\ (z \in \mathbb{R}) \Leftrightarrow (y = \operatorname{Im}(z) = 0) \\ z + z' = (x + x') + i(y + y') \\ zz' = (xx' - yy') + i(xy' + x'y) \\ \frac{1}{z} = \frac{x}{x^2 + y^2} - \frac{y}{x^2 + y^2}i \text{ si } z \neq 0 \end{cases}$$

5.3 Conjugué et module d'un nombre complexe

Définition 5.3. Le conjugué du nombre complexe $z = x + iy$ est le nombre complexe $\bar{z} = x - iy$.

On déduit immédiatement de cette définition les propriétés suivantes du conjugué.

Théorème 5.2.

Pour tous nombres complexes z et z' , on a :

1. $\bar{\bar{z}} = z$, $\overline{z + z'} = \bar{z} + \bar{z'}$, $\overline{zz'} = \bar{z}\bar{z'}$.
2. $z + \bar{z} = 2\operatorname{Re}(z)$, $z - \bar{z} = 2i\operatorname{Im}(z)$.
3. $z \in \mathbb{R} \Leftrightarrow z = \bar{z}$.
4. z est imaginaire pur si, et seulement si, $\bar{z} = -z$.
5. $z\bar{z} = (\operatorname{Re}(z))^2 + (\operatorname{Im}(z))^2 \in \mathbb{R}^+$.

Le dernier point du théorème précédent nous permet de donner la définition suivante.

Définition 5.4. Le module du nombre complexe $z = x + iy$ est le réel $|z| = \sqrt{z\bar{z}} = \sqrt{x^2 + y^2}$.

Dans le cas où z est réel, $|z|$ est la valeur absolue de z .

On vérifie facilement les propriétés suivantes liées au module.

Théorème 5.3.

Pour tous nombres complexes z et z' , on a :

1. $|z| \geq 0$ et $|z| = 0$ si, et seulement si, $z = 0$;
2. $|z| = |\bar{z}|$;
3. $|zz'| = |z| |z'|$;
4. pour tout nombres complexes $z_1, \dots, z_n$, on a $\left| \prod_{k=1}^n z_k \right| = \prod_{k=1}^n |z_k|$.
5. si $z \neq 0$, alors $\frac{1}{z} = \frac{\bar{z}}{|z|^2} = \frac{x-iy}{x^2+y^2}$;
6. si $z' \neq 0$, alors $\left| \frac{z}{z'} \right| = \frac{|z|}{|z'|}$;
7. $|z| = 1$ si, et seulement si, $\frac{1}{z} = \bar{z}$;
8. $|\operatorname{Re}(z)| \leq |z|$ et l'égalité est réalisée si, et seulement si, z est réel ;
9. $|\operatorname{Im}(z)| \leq |z|$ et l'égalité est réalisée si, et seulement si, z est imaginaire pur.

Du point 4 on déduit que pour tout nombre complexe z et tout entier naturel n , on a $|z^n| = |z|^n$. Cette égalité étant encore valable pour n entier relatif et z non nul.

Du point 8 on déduira l'inégalité de Cauchy-Schwarz avec son cas d'égalité.

Comme pour la valeur absolue réelle, on dispose de l'inégalité triangulaire. Cette inégalité est conséquence des résultats suivants.

Théorème 5.4.

Pour tous nombres complexes z et z' , on a :

$$|z + z'|^2 = |z|^2 + 2 \operatorname{Re}(z\bar{z}') + |z'|^2$$

Preuve. Il suffit d'écrire, pour $z = x + iy$, $z' = x' + iy'$:

$$\begin{aligned} |z + z'|^2 &= (x + x')^2 + (y + y')^2 = x^2 + y^2 + 2(xx' + yy') + x'^2 + y'^2 \\ &= |z|^2 + 2 \operatorname{Re}(z\bar{z}') + |z'|^2 \end{aligned}$$

□

Théorème 5.5. Inégalité de Cauchy-Schwarz

Pour tous nombres complexes z et z' , on a $|\operatorname{Re}(z\bar{z}')| \leq |z| |z'|$, l'égalité étant réalisée si, et seulement si, z et z' sont liés sur $\mathbb{R}$ (c'est-à-dire si $z' = 0$ ou $z \neq 0$ et $\frac{z}{z'} \in \mathbb{R}$), ce qui est encore équivalent à dire que $z\bar{z}'$ est réel.

Preuve. On a $|\operatorname{Re}(zz')| \leq |zz'| = |z||z'|$ et l'égalité est réalisée si, et seulement si, zz' est réel. Pour $z' = 0$, c'est le cas et pour $z' \neq 0$, il existe un réel $\lambda \in \mathbb{R}$ tel que $z = \lambda \frac{1}{z'} = \frac{\lambda}{|z'|^2} z'$. La réciproque est évidente. $\square$

Théorème 5.6.

Pour tous nombres complexes z et z' , on a $|z + z'| \leq |z| + |z'|$, l'égalité étant réalisée si, et seulement si, z et z' sont positivement liées sur $\mathbb{R}$ (i.e. $z = 0$ ou $z \neq 0$ et $\frac{z}{z'} \in \mathbb{R}^+$), ce qui est encore équivalent à dire que $z\overline{z'} \in \mathbb{R}^+$.

Preuve. On a :

$$|z + z'|^2 = |z|^2 + 2\operatorname{Re}(z\overline{z'}) + |z'|^2 \leq |z|^2 + 2|z||z'| + |z'|^2 = (|z| + |z'|)^2$$

ce qui équivaut à $|z + z'| \leq |z| + |z'|$. L'égalité est réalisée si, et seulement si, $\operatorname{Re}(z\overline{z'}) = |z||z'|$, ce qui implique $|\operatorname{Re}(z\overline{z'})| = |z||z'|$ et z, z' sont liés sur $\mathbb{R}$. Pour $z' \neq 0$, on a $z = \lambda z'$ avec $\lambda \in \mathbb{R}$ et $\operatorname{Re}(z\overline{z'}) = \lambda |z'|^2 = |z||z'|$ impose $\lambda \geq 0$. La réciproque est évidente. $\square$

Corollaire 5.1. *Pour tous nombres complexes z et z' , on a :*

$$||z| - |z'||| \leq |z - z'| \leq |z| + |z'|$$

Preuve. Il suffit d'écrire que $|z| \leq |z - z'| + |z'|$ et $|z'| \leq |z - z'| + |z|$. $\square$

Pour ce qui est du module d'une somme de nombres complexes, on a de manière plus générale le résultat suivant qui se montre facilement par récurrence.

Théorème 5.7.

Pour toute suite finie $(z_k)_{1 \leq k \leq n}$ de nombres complexes non nuls avec $n \geq 2$, on a $\left| \sum_{k=1}^n z_k \right| \leq \sum_{k=1}^n |z_k|$, l'égalité étant réalisée si, et seulement si, il existe des réels $\lambda_2, \dots, \lambda_n$ tels que $z_k = \lambda_k z_1$ pour $k = 2, \dots, n$.

Preuve. On procède par récurrence sur $n \geq 2$. Pour $n = 2$, c'est connu. Supposons le résultat acquis au rang $n - 1 \geq 2$. Pour $z_1, \dots, z_n$ dans $\mathbb{C}$ avec $n \geq 3$, en utilisant les résultats pour $n = 2$ et l'hypothèse de récurrence, on a :

$$\left| \sum_{k=1}^n z_k \right| = \left| z_1 + \sum_{k=2}^n z_k \right| \leq |z_1| + \left| \sum_{k=2}^n z_k \right| \leq \sum_{k=1}^n |z_k|$$

Si l'égalité $\left| \sum_{k=1}^n z_k \right| = \sum_{k=1}^n |z_k|$ est réalisée, en posant $Z_2 = \sum_{k=2}^n z_k$, on a :

$$\left| \sum_{k=1}^n z_k \right| = |z_1 + Z_2| \leq |z_1| + |Z_2| \leq \sum_{k=1}^n |z_k|$$

et l'égalité $\left| \sum_{k=1}^n z_k \right| = \sum_{k=1}^n |z_k|$ nous dit que toutes les inégalités précédentes sont des égalités. On a donc $|z_1 + Z_2| = |z_1| + |Z_2|$ et $Z_2 = \lambda_1 z_1$ avec $\lambda_1 \in \mathbb{R}^{+,*}$ ($Z_2 = 0$ entraîne $|z_1| = \sum_{k=1}^n |z_k|$, donc $\sum_{k=2}^n |z_k| = 0$ et tous les z_k sont nuls, ce qui est contraire à l'hypothèse de départ), puis de $|z_1| + |Z_2| = \sum_{k=1}^n |z_k|$, on déduit que $|Z_2| = \sum_{k=2}^n |z_k|$ et avec l'hypothèse de récurrence qu'il existe des réels $\lambda_k > 0$ tels que $z_k = \lambda_k z_2$ pour $k = 3, \dots, n$. On a alors $Z_2 = \sum_{k=2}^n z_k = \left(1 + \sum_{k=3}^n \lambda_k \right) z_2 = \lambda_1 z_1$ et $z_2 = \mu_2 z_1$, $z_k = \lambda_k z_2 = \mu_k z_1$ pour $k = 3, \dots, n$, tous les μ_k étant strictement positifs. $\square$

5.4 Les équations de degré 2

On s'intéresse ici aux équations algébriques de degré 2 à coefficients complexes, c'est-à-dire aux équations de la forme $ax^2 + bx + c = 0$ où a, b, c sont des nombres complexes avec $a \neq 0$.

Si on se place dans le cadre réel (*i. e.* les coefficients a, b, c sont réels et on cherche des solutions réelles), on sait qu'une telle équation n'a pas nécessairement de solution (c'est l'exemple de $x^2 + 1 = 0$ qui nous a conduit aux nombres complexes).

En divisant par a , on se ramène au cas où $a = 1$.

On remarque tout d'abord qu'une telle équation a au plus deux solutions complexes. En effet, si $z_1 \in \mathbb{C}$ est solution de $x^2 + bx + c = 0$, en désignant par z une autre solution, on a le système d'équations :

$$\begin{cases} z^2 + bz + c = 0 \\ z_1^2 + bz_1 + c = 0 \end{cases}$$

et par soustraction on aboutit à $(z - z_1)(z + z_1 + b) = 0$ qui donne $z = x_1$ ou $z = -z_1 - b$. Il suffit donc de trouver une solution (s'il en existe) de cette équation pour avoir les deux.

Nous allons voir que sur $\mathbb{C}$ une équation de degré 2 a toujours deux solutions, distinctes ou confondues.

Connaissant $i \in \mathbb{C}$ solution de $x^2 + 1 = 0$, on déduit que pour tout réel a non nul l'équation $x^2 + a = 0$ a exactement deux solutions distinctes. En effet si a est négatif, cette équation équivaut à $x^2 = -a$ avec $-a > 0$, et dans l'ensemble des réels, on sait que cela équivaut à dire que $x = \pm\sqrt{-a}$, ce qui fournit deux solutions réelles distinctes. Si a est positif, cette équation équivaut à $x^2 = -a = i^2(\sqrt{a})^2$, soit à $x^2 - (i\sqrt{a})^2 = 0$ ce qui équivaut à $x = \pm i\sqrt{a}$. On a donc le résultat suivant.

Théorème 5.8.

Pour tout nombre réel non nul a l'équation $x^2 + a = 0$ a exactement deux solutions données par :

$$\begin{cases} x_1 = -\sqrt{-a} \text{ et } x_2 = \sqrt{-a} \text{ pour } a < 0 \\ x_1 = -i\sqrt{a} \text{ et } x_2 = i\sqrt{a} \text{ pour } a > 0 \end{cases}$$

Pour $a = 0$, $x = 0$ est la seule solution de cette équation.

Définition 5.5. Si α est un nombre complexe, on dit que le nombre complexe u est une racine carrée de α si $u^2 = \alpha$.

Le théorème précédent nous dit que tout nombre réel non nul a a exactement deux racines carrées complexes, ce sont les réels $\pm\sqrt{a}$ pour $a > 0$ et les complexes $\pm i\sqrt{-a}$ pour $a < 0$. Ce résultat est en fait valable pour tout nombre complexe α .

Théorème 5.9.

Tout nombre complexe non nul $\alpha = a + ib$ a exactement deux racines carrées.

Preuve. Il s'agit de résoudre l'équation $z^2 = \alpha$ et pour ce faire il nous suffit de trouver une solution. Si $b = 0$, alors $\alpha = a$ est réel et le problème a été résolu. On suppose donc que $b \neq 0$. En notant $z = x + iy$, l'équation $z^2 = \alpha = a + ib$ équivaut au système de deux équations aux inconnues x, y :

$$\begin{cases} x^2 - y^2 = a \\ 2xy = b \end{cases}$$

En utilisant le module de z , le système :

$$\begin{cases} x^2 - y^2 = a \\ |z^2| = |z|^2 = x^2 + y^2 = |\alpha| \end{cases}$$

nous donne immédiatement $x^2 = \frac{a + |\alpha|}{2}$ avec

$$a + |\alpha| = a + \sqrt{a^2 + b^2} > a + \sqrt{a^2} = a + |a| \geq 0$$

et en conséquence, $x = \pm \frac{\sqrt{a + |\alpha|}}{\sqrt{2}}$, la partie imaginaire y étant déterminée par l'équation $2xy = b$ avec $b \neq 0$ (donc $x \neq 0$ et $y \neq 0$). En définitive l'équation $z^2 = \alpha$ a deux solutions complexes données par :

$$\begin{aligned} z_1 &= \sqrt{\frac{a + |\alpha|}{2}} + i \frac{b}{2\sqrt{\frac{a + |\alpha|}{2}}} = \frac{1}{\sqrt{2}} \frac{a + ib + |\alpha|}{\sqrt{a + |\alpha|}} \\ &= \frac{1}{\sqrt{2}} \frac{\alpha + |\alpha|}{\sqrt{\operatorname{Re}(\alpha) + |\alpha|}} = \frac{1}{\sqrt{2}} \frac{a + ib + \sqrt{a^2 + b^2}}{\sqrt{a + \sqrt{a^2 + b^2}}} \end{aligned}$$

et $z_2 = -z_1$. □

On en déduit alors le résultat suivant.

Théorème 5.10.

Toute équation de degré 2, $az^2 + bz + c = 0$ ($a \neq 0$) a deux solutions complexes distinctes ou confondues.

Preuve. En utilisant la forme réduite d'un polynôme de degré 2 :

$$az^2 + bz + c = a \left(\left(z + \frac{b}{2a} \right)^2 - \frac{b^2 - 4ac}{4a^2} \right)$$

on est ramené à l'équation $\left(z + \frac{b}{2a} \right)^2 = \frac{b^2 - 4ac}{4a^2}$ qui a deux solutions distinctes (si $b^2 - 4ac \neq 0$) ou confondues (si $b^2 - 4ac = 0$). □

Avec les notations du théorème la quantité $\delta = b^2 - 4ac$ est appelé discriminant de l'équation $ax^2 + bx + c = 0$.

Pour $\delta = 0$, $z_1 = z_2 = -\frac{b}{2a}$ est la seule solution (on dit que c'est une racine double) et pour $\delta \neq 0$, les deux solutions sont $z_1 = \frac{-b - \gamma}{2a}$ et $z_2 = \frac{-b + \gamma}{2a}$ où γ est une racine carrée de δ (i. e. $\gamma^2 = \delta$). Dans les deux cas, on $z_1 + z_2 = -\frac{b}{a}$ et $z_1 z_2 = \frac{b^2 - \gamma^2}{4a^2} = \frac{b^2 - \delta}{4a^2} = \frac{c}{a}$. Réciproquement si z_1, z_2 sont deux nombres complexes tels que $z_1 + z_2 = -\frac{b}{a}$ et $z_1 z_2 = \frac{c}{a}$, on a $z_2 = -\frac{b}{a} - z_1$ et $z_1 \left(-\frac{b}{a} - z_1 \right) = \frac{c}{a}$, c'est-à-dire que z_1 est solution de l'équation $z^2 + \frac{b}{a}z + \frac{c}{a} = 0$, soit de l'équation $az^2 + bz + c = 0$. Comme z_1 et z_2 jouent des rôles symétriques, z_2 est également solution de cette équation. En résumé, on a le résultat suivant.

Théorème 5.11.

Étant donnés des nombres complexes a, b, c avec $a \neq 0$, les nombres complexes z_1 et z_2 sont les racines de l'équation $az^2 + bz + c = 0$ si, et seulement si, on a $z_1 + z_2 = -\frac{b}{a}$ et $z_1 z_2 = \frac{c}{a}$.

Dans le cas où les coefficients a, b, c sont réels, il en est de même de δ et on distingue trois cas de figure :

- soit $\delta = 0$ et $x_1 = -\frac{b}{2a}$ est la seule solution réelle de cette équation ;
- soit $\delta > 0$ et $x_1 = \frac{-b - \sqrt{\delta}}{2a}$, $x_2 = \frac{-b + \sqrt{\delta}}{2a}$ sont les deux solutions réelles de cette équation ;
- soit $\delta < 0$ et $x_1 = \frac{-b - i\sqrt{-\delta}}{2a}$, $x_2 = \frac{-b + i\sqrt{-\delta}}{2a}$ sont les deux solutions complexes non réelles de cette équation.

5.5 Les équations de degré 3 et 4

On s'intéresse tout d'abord aux équations polynomiales de degré 3 :

$$P(z) = z^3 + az^2 + bz + c = 0$$

où a, b, c sont des nombres complexes.

Dans un premier temps, on effectue une translation en vue de supprimer le terme en z^2 de cette équation, c'est-à-dire qu'on cherche $\lambda \in \mathbb{C}$ qui permette de supprimer z^2 dans $P(z - \lambda) = (z - \lambda)^3 + a(z - \lambda)^2 + b(z - \lambda) + c$. En développant, on a $P(z - \lambda) = z^3 + (a - 3\lambda)z^2 + (\lambda^2 - 2a\lambda + b)z + (c - b\lambda + a\lambda^2 - \lambda^3)$. Le choix de $\lambda = \frac{a}{3}$ donne $P(z - \lambda) = z^3 + \left(b - \frac{a^2}{3}\right)z + \left(c - \frac{ab}{3} + \frac{2a^3}{27}\right)$. On est donc ramené à l'équation $Q(z) = z^3 + pz + q = 0$, où on a noté $p = b - \frac{a^2}{3}$ et $q = c - \frac{ab}{3} + \frac{2a^3}{27}$. Si z est solution de $Q(z) = 0$, alors $z - \lambda$ est solution de $P(t) = 0$.

Si $p = 0$, alors les solutions de $Q(z) = 0$ sont les racines cubiques de $-q$.

Si $p \neq 0$, on cherche alors les solutions sous la forme $z = u + v$ en imposant une condition supplémentaire à u et v . En développant :

$$P(u + v) = u^3 + v^3 + (3uv + p)(u + v) + q$$

on est amené à imposer $3uv + p = 0$, ce qui donne le système de deux équations à deux inconnues :

$$\begin{cases} u^3 + v^3 = -q \\ 3uv = -p \end{cases}$$

Les nombres complexes u^3 et v^3 sont alors solutions de :

$$\begin{cases} u^3 + v^3 = -q \\ u^3 v^3 = -\frac{p^3}{27} \end{cases}$$

ce qui revient à dire que ce sont les solutions de l'équation de degré 2 :

$$x^2 + qx - \frac{p^3}{27} = 0$$

Le discriminant de cette équation est $\delta = \frac{4p^3 + 27q^2}{27}$. Notant ω une racine carrée de δ ($\omega^2 = \delta$), on a $u^3 = \frac{-q - \omega}{2}$ et $v^3 = u^3 = \frac{-q + \omega}{2}$. En désignant par w une racine cubique $\frac{-q - \omega}{2}$, les deux autres sont jw et $\bar{j}w$. Enfin la relation $3uv = -p$ avec $p \neq 0$, donne $u \neq 0$, $v \neq 0$ et $v = -\frac{p}{3u}$. On a donc ainsi trouvé trois solutions (u, v) , à savoir :

$$\left(w, -\frac{p}{3w}\right), \left(jw, -\frac{p}{3jw}\right) = \left(jw, -\frac{p\bar{j}}{3w}\right) \text{ et } \left(\bar{j}w, -\frac{p}{3\bar{j}w}\right) = \left(\bar{j}w, -\frac{pj}{3w}\right)$$

ce qui donne trois solutions pour l'équation $Q(z) = 0$:

$$z_1 = w - \frac{p}{3w}, \quad z_2 = jw - \frac{p\bar{j}}{3w}, \quad z_3 = \bar{j}w - \frac{pj}{3w}$$

et on les a toutes.

5.6 Arguments d'un nombre complexe

On suppose connues du cours d'analyse les fonctions trigonométriques $\cos$, $\sin$ et $\tan$ avec leurs principales propriétés. En particulier, les fonctions $\cos$ et $\sin$ sont définies sur $\mathbb{R}$, périodiques de période 2π , la fonction $\cos$ est paire, la fonction $\sin$ est impaire et on a les formules de trigonométrie suivantes valables pour tous réels a, b :

$$\begin{cases} \cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b) \\ \sin(a+b) = \sin(a)\cos(b) + \cos(a)\sin(b) \end{cases}$$

desquelles on déduit les suivantes bien utiles :

$$\begin{cases} \cos(a-b) = \cos(a)\cos(b) + \sin(a)\sin(b) \\ \sin(a-b) = \sin(a)\cos(b) - \cos(a)\sin(b) \\ \cos(2a) = \cos^2(a) - \sin^2(a), \quad \sin(2a) = 2\sin(a)\cos(b) \\ \cos(a)\cos(b) = \frac{\cos(a+b) + \cos(a-b)}{2}, \quad \sin(a)\sin(b) = \frac{\cos(a-b) - \cos(a+b)}{2} \\ \cos(a)\sin(b) = \frac{\sin(a+b) - \sin(a-b)}{2}, \quad \sin(a)\cos(b) = \frac{\sin(a+b) + \sin(a-b)}{2} \end{cases}$$

enfin avec $\cos(0) = 1$, on déduit que $\cos^2(a) + \sin^2(a) = 1$.

La fonction $\tan$ est définie sur $\mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi \mid k \in \mathbb{Z} \right\}$ par $\tan(x) = \frac{\sin(x)}{\cos(x)}$ et elle est impaire et π -périodique.

La fonction $\cos$ réalise une bijection de $[0, \pi]$ sur $[-1, 1]$, la fonction $\sin$ une bijection de $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ sur $[-1, 1]$ et la fonction $\tan$ une bijection de $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ sur $\mathbb{R}$. Les fonctions réciproques de ces fonctions trigonométriques sont notées respectivement $\arccos$, $\arcsin$ et $\arctan$. On a donc :

$$\begin{aligned} (x \in [0, \pi] \text{ et } y = \cos(x)) &\Leftrightarrow (y \in [-1, 1] \text{ et } x = \arccos(y)) \\ \left(x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \text{ et } y = \sin(x)\right) &\Leftrightarrow (y \in [-1, 1] \text{ et } x = \arcsin(y)) \\ \left(x \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\text{ et } y = \tan(x)\right) &\Leftrightarrow (y \in \mathbb{R} \text{ et } x = \arctan(y)) \end{aligned}$$

Le résultat suivant est la base de la définition de l'argument d'un nombre complexe.

Théorème 5.12.

Si z est un nombre complexe de module 1, il existe un unique réel θ dans $[-\pi, \pi[$ tel que $z = \cos(\theta) + i \sin(\theta)$.

Preuve. Le nombre complexe $z = x + iy$ est de module 1 si, et seulement si, $x^2 + y^2 = 1$. En particulier x est dans $[-1, 1]$ et il existe un unique réel $\alpha \in [0, \pi]$ tel que $x = \cos(\alpha)$. Avec $y^2 = 1 - x^2 = \sin^2(\alpha)$, on déduit que $y = \pm \sin(\alpha)$, soit $y = \sin(\pm\alpha)$. En utilisant la parité de la fonction $\cos$, on peut écrire que $x = \cos(\pm\alpha)$ et on aboutit à $(x, y) = (\cos(\theta), \sin(\theta))$ avec $\theta \in [-\pi, \pi[$ (pour $(x, y) = (\cos(\pi), \sin(\pi)) = (-1, 0)$, on écrit $(x, y) = (\cos(-\pi), \sin(-\pi))$).

Si $\theta' \in [-\pi, \pi[$ est une autre solution, de $\cos(\theta) = \cos(\theta')$, on déduit que $\theta' = \pm\theta$. Si $\theta' = \theta$, c'est terminé, sinon $\theta' = -\theta$ et de $\sin(\theta) = \sin(\theta') = -\sin(\theta)$, on déduit que θ vaut 0 ou $-\pi$, 0 étant la seule solution puisque $\theta' = \pi \notin [-\pi, \pi[$. D'où l'unicité. $\square$

Corollaire 5.2. *Pour tout nombre complexe non nul z , il existe un unique réel $\theta \in [-\pi, \pi[$ tel que $\frac{z}{|z|} = \cos(\theta) + i \sin(\theta)$.*

Preuve. On applique le théorème précédent à $\frac{z}{|z|}$ qui est de module égal à 1. $\square$

Définition 5.6. *Avec les notations du corollaire qui précède, on dit que le réel $\theta \in [-\pi, \pi[$ est l'argument principal du nombre complexe non nul z .*

Si $\theta \in [-\pi, \pi[$ est l'argument principal d'un nombre complexe $z \in \mathbb{C}^*$, les seuls réels θ' tels que $\frac{z}{|z|} = \cos(\theta') + i \sin(\theta')$ sont les réels $\theta' = \theta + 2k\pi$, où k est un entier relatif. En effet ces réels conviennent et les égalités $\cos(\theta) = \cos(\theta')$ et $\sin(\theta) = \sin(\theta')$ sont réalisées si, et seulement si, il existe un entier relatif k tel que $\theta' = \theta + 2k\pi$ (on peut trouver un entier k tel que $\theta' - 2k\pi$ soit dans $[-\pi, \pi[$, c'est-à-dire que k est tel que $-\pi \leq \theta' - 2k\pi < \pi$, soit $k \leq \frac{\theta' + \pi}{2\pi} < k + 1$, encore équivalent à $k = \left\lfloor \frac{\theta' + \pi}{2\pi} \right\rfloor$ et $\theta' - 2k\pi$ est l'argument principal de z).

Définition 5.7. *On dit qu'un réel θ est un argument du nombre complexe non nul z si $\frac{z}{|z|} = \cos(\theta) + i \sin(\theta)$.*

Ce qui précède nous dit qu'un nombre complexe non nul admet une infinité d'arguments et que deux tels arguments diffèrent d'un multiple entier de 2π , on dit alors qu'ils sont égaux modulo 2π . On notera $\theta' \equiv \theta \pmod{2\pi}$ pour signifier que les réels θ' et θ sont égaux modulo 2π . Si θ est un argument d'un nombre complexe non nul z , on notera $\arg(z) \equiv \theta \pmod{2\pi}$. La notation $\arg(z)$ signifie qu'on a choisi un argument de z , c'est donc un réel défini modulo 2π . Par abus de langage, on écrira $\theta = \arg(z)$ quand il n'y a pas d'ambiguïté.

Avec le théorème qui suit on donne quelques propriétés des arguments d'un nombre complexe.

Théorème 5.13.

En désignant par z et z' des nombres complexes non nuls, λ un réel non nul et n un entier relatif, on a :

1. $\arg(\bar{z}) \equiv -\arg(z) \pmod{2\pi}$;
2. $\arg(zz') \equiv \arg(z) + \arg(z') \pmod{2\pi}$;

3. $\arg\left(\frac{z}{z'}\right) \equiv \arg(z) - \arg(z') = \arg(\overline{zz'}) \pmod{2\pi}$;
4. $\arg(z^n) \equiv n \arg(z) \pmod{2\pi}$;
5. pour $\lambda > 0$, on a $\arg(\lambda z) \equiv \arg(z) \pmod{2\pi}$ et pour $\lambda < 0$, on a $\arg(\lambda z) \equiv \arg(z) + \pi \pmod{2\pi}$;
6. z est réel si, et seulement si, $\arg(z) \equiv 0 \pmod{\pi}$ (c'est-à-dire que les arguments de z sont de la forme $k\pi$ avec $k \in \mathbb{Z}$, l'argument principal étant 0 ou $-\pi$);
7. z est imaginaire pur si, et seulement si, $\arg(z) \equiv \frac{\pi}{2} \pmod{\pi}$ (c'est-à-dire que les arguments de z sont de la forme $\frac{\pi}{2} + k\pi$ avec $k \in \mathbb{Z}$, l'argument principal étant $-\frac{\pi}{2}$ ou $\frac{\pi}{2}$).

Preuve. Il suffit de considérer le cas des nombres complexes de module 1 par définition des arguments.

1. Pour $z = \cos(\theta) + i \sin(\theta)$, on a $\bar{z} = \cos(\theta) - i \sin(\theta) = \cos(-\theta) + i \sin(-\theta)$ et $-\theta$ est un argument de $\bar{z}$, donc $\arg(\bar{z}) \equiv -\theta \pmod{2\pi}$.
2. Pour $z = \cos(\theta) + i \sin(\theta)$ et $z' = \cos(\theta') + i \sin(\theta')$, on a :

$$\begin{aligned} zz' &= (\cos(\theta) \cos(\theta') - \sin(\theta) \sin(\theta')) + i (\sin(\theta) \cos(\theta') + \cos(\theta) \sin(\theta')) \\ &= \cos(\theta + \theta') + i \sin(\theta + \theta') \end{aligned}$$

et donc $\arg(zz') \equiv \theta + \theta' \pmod{2\pi}$.

3. On a $\arg\left(\frac{z}{z'}\right) \equiv \arg(zz') \equiv \arg(z) + \arg(\bar{z'}) \equiv \arg(z) - \arg(z') \pmod{2\pi}$.
4. Se déduit de ce qui précède.

□

En désignant par ψ l'application qui associe à tout réel θ le nombre complexe $\psi(\theta) = \cos(\theta) + i \sin(\theta)$ on réalise une application surjective de $\mathbb{R}$ sur l'ensemble Γ des nombres complexes de module 1. Cette application n'est pas injective puisque l'égalité $\psi(\theta) = \psi(\theta')$ équivaut à $\theta' \equiv \theta \pmod{2\pi}$. En restriction à $[-\pi, \pi[$ cette application ψ est bijective.

Théorème 5.14.

Avec les notations qui précèdent, on a $\psi(0) = 1$ et $\psi(\theta + \theta') = \psi(\theta) \psi(\theta')$ pour tous réels θ, θ' .

Preuve. On a $\psi(0) = \cos(0) + i \sin(0) = 1$ et :

$$\begin{aligned} \psi(\theta + \theta') &= \cos(\theta + \theta') + i \sin(\theta + \theta') \\ &= (\cos(\theta) \cos(\theta') - \sin(\theta) \sin(\theta')) + i (\sin(\theta) \cos(\theta') + \cos(\theta) \sin(\theta')) \\ &= (\cos(\theta) + i \sin(\theta)) (\cos(\theta') + i \sin(\theta')) = \psi(\theta) \psi(\theta') \end{aligned}$$

□

La fonction ψ vérifie donc la même équation fonctionnelle que la fonction exponentielle réelle. Cette remarque justifie la notation $\psi(\theta) = e^{i\theta}$.

Avec $1 = \psi(0) = \psi(\theta - \theta) = \psi(\theta)\psi(-\theta)$, on déduit que $\frac{1}{\psi(\theta)} = \psi(-\theta) = \overline{\psi(\theta)}$ (ce que l'on savait déjà : l'inverse d'un nombre complexe de module 1 est égal à son conjugué).

On a donc en résumé la notation :

$$\forall \theta \in \mathbb{R}, e^{i\theta} = \cos(\theta) + i \sin(\theta)$$

ce qui définit une fonction 2π -périodique surjective de $\mathbb{R}$ sur l'ensemble Γ des nombres complexes de module 1 avec les propriétés suivantes :

$$\left\{ \begin{array}{l} e^{i \cdot 0} = e^0 = 1 \\ \forall (\theta, \theta') \in \mathbb{R}^2, e^{i(\theta+\theta')} = e^{i\theta} e^{i\theta'} \\ \forall \theta \in \mathbb{R}, \frac{1}{e^{i\theta}} = e^{-i\theta} = \overline{e^{i\theta}} \\ \forall (\theta, \theta') \in \mathbb{R}^2, (e^{i\theta} = e^{i\theta'}) \Leftrightarrow (\exists k \in \mathbb{Z} \mid \theta' = \theta + 2k\pi) \\ \forall \theta \in \mathbb{R}, \cos(\theta) = \operatorname{Re}(e^{i\theta}) = \frac{e^{i\theta} + e^{-i\theta}}{2} \text{ et } \sin(\theta) = \operatorname{Im}(e^{i\theta}) = \frac{e^{i\theta} - e^{-i\theta}}{2i} \end{array} \right.$$

Par récurrence sur $n \geq 0$, on déduit facilement que $(e^{i\theta})^n = e^{in\theta}$. Puis pour $n < 0$ on a $e^{in\theta} = \frac{1}{e^{-in\theta}} = \left(\frac{1}{e^{i\theta}}\right)^{-n} = (e^{-i\theta})^{-n} = e^{in\theta}$, c'est-à-dire que cette formule est valable pour tous les entiers relatifs. Nous verrons un peu plus loin l'intérêt de cette égalité.

On a en particulier les valeurs suivantes $e^{i\pi} = -1$, $e^{i\frac{\pi}{2}} = i$, les égalités $e^{i\theta} = 1$, $e^{i\theta} = -1$ et $e^{i\theta} = i$ étant réalisées respectivement si, et seulement si, $\theta = 2k\pi$, $\theta = (2k+1)\pi$ et $\theta = \frac{\pi}{2} + 2k\pi$, où k est un entier relatif.

Un nombre complexe non nul peut donc s'écrire sous la forme $z = \rho e^{i\theta}$ où ρ est un réel strictement positif uniquement déterminé, c'est le module de z , et θ est un argument de z . Cette écriture est l'écriture polaire (ou trigonométrique) de z .

L'écriture polaire des nombres complexes non nuls peut aussi être utilisée pour résoudre des équations de la forme $a \cos(x) + b \sin(x) = 0$. Pour ce faire, on utilise le résultat suivant.

Théorème 5.15.

Soient a, b des réels non tous deux nuls (c'est-à-dire que $(a, b) \neq (0, 0)$). Il existe un unique couple (ρ, θ) de réels dans $\mathbb{R}^{+,*} \times [-\pi, \pi[$ tel que :

$$\forall x \in \mathbb{R}, a \cos(x) + b \sin(x) = \rho \cos(x - \theta).$$

ρ est le module de $u = a + ib$ et θ son argument principal.

Preuve. Comme $(a, b) \neq (0, 0)$, on a $u = a + ib \neq 0$ et ce nombre complexe s'écrit de manière unique $u = \rho e^{i\theta}$ avec $\rho = |u| = \sqrt{a^2 + b^2} > 0$ et $\theta \in [-\pi, \pi[$. Pour tout réel x , on a alors d'une part :

$$ue^{-ix} = (a + ib)(\cos(x) - i \sin(x)) = a \cos(x) + b \sin(x) + i(b \cos(x) - a \sin(x))$$

et d'autre part $ue^{-ix} = \rho e^{i(\theta-x)} = \rho \cos(\theta-x) + i\rho \sin(\theta-x)$, ce qui donne $a \cos(x) + b \sin(x) = \rho \cos(\theta-x) = \rho \cos(x-\theta)$ et $a \sin(x) - b \cos(x) = \rho \sin(x-\theta)$. Pour ce qui est de l'unicité, supposons que (ρ', θ') soit une autre solution à notre problème. On a alors $\rho \cos(x-\theta) = \rho' \cos(x-\theta')$ pour tout réel x . Prenant $x = \theta$ [resp. $x = \theta'$], on en déduit que $\rho = \rho' \cos(\theta-\theta') \leq \rho'$ [resp. $\rho' = \rho \cos(\theta'-\theta) \leq \rho$] et $\rho = \rho'$. Prenant $x = 0$ [resp. $x = \frac{\pi}{2}$], on a $\cos(\theta) = \cos(\theta')$ [resp. $\cos(\frac{\pi}{2}-\theta) = \sin(\theta) = \cos(\frac{\pi}{2}-\theta') = \sin(\theta')$] avec θ, θ' dans $[-\pi, \pi[$, ce qui équivaut à $\theta = \theta'$. $\square$

Corollaire 5.3. Soient a, b des réels non tous deux nuls. Les solutions de l'équation :

$$a \cos(x) + b \sin(x) = 0 \quad (5.1)$$

sont les réels $x = \theta + \frac{\pi}{2} + k\pi$, où $\theta \in [-\pi, \pi[$ est l'argument principal de $a + ib$ et k un entier relatif.

Preuve. Le théorème précédent nous que l'équation (5.1) est équivalente à $\cos(x-\theta) = 0$, soit à $(x-\theta) \equiv \frac{\pi}{2} \pmod{\pi}$, encore équivalent à dire que $x = \theta + \frac{\pi}{2} + k\pi$ avec $k \in \mathbb{Z}$. $\square$

Les formules suivantes, valables pour $\theta \in \mathbb{R}$ et $n \in \mathbb{Z}$:

$$\cos(\theta) = \operatorname{Re}(e^{i\theta}) = \frac{e^{i\theta} + e^{-i\theta}}{2}, \quad \sin(\theta) = \operatorname{Im}(e^{i\theta}) = \frac{e^{i\theta} - e^{-i\theta}}{2i}$$

(formules d'Euler) et $(\cos(\theta) + i \sin(\theta))^n = (e^{i\theta})^n = e^{in\theta} = \cos(n\theta) + i \sin(n\theta)$ (formule de Moivre) permettent d'obtenir relativement facilement des formules de trigonométrie.

En utilisant la formule du binôme de Newton, la formule de Moivre s'écrit :

$$\cos(n\theta) + i \sin(n\theta) = \sum_{k=0}^n \binom{n}{k} \cos^k(\theta) \sin^{n-k}(\theta) i^{n-k}$$

et l'identification des parties réelles et imaginaires nous permet d'exprimer $\cos(n\theta)$ et $\sin(n\theta)$ comme combinaisons linéaires de puissances de $\cos(\theta)$ et $\sin(\theta)$.

L'utilisation des formules d'Euler et de la formule du binôme de Newton nous permet d'exprimer des puissances $\cos(\theta)$ et $\sin(\theta)$ comme combinaisons linéaires de $\cos(p\theta)$ et $\sin(q\theta)$. On dit qu'on linéarise $\cos^n(\theta)$ ou $\sin^m(\theta)$, n et m étant des entiers naturels. Pour ce faire, on écrit que :

$$\begin{aligned} \cos^n(\theta) &= \frac{1}{2^n} (e^{i\theta} + e^{-i\theta})^n = \frac{1}{2^n} \sum_{k=0}^n \binom{n}{k} e^{ik\theta} e^{-i(n-k)\theta} = \frac{1}{2^n} \sum_{k=0}^n \binom{n}{k} e^{i(2k-n)\theta} \\ &= \frac{1}{2^n} \sum_{k=0}^n \binom{n}{k} \cos((2k-n)\theta) + i \frac{1}{2^n} \sum_{k=0}^n \binom{n}{k} \sin((2k-n)\theta) \end{aligned}$$

et nécessairement :

$$\cos^n(\theta) = \frac{1}{2^n} \sum_{k=0}^n \binom{n}{k} \cos((2k-n)\theta) \quad (5.2)$$

On peut remarquer que l'on a aussi $\sum_{k=0}^n \binom{n}{k} \sin((2k-n)\theta) = 0$. En réalité cette formule est évidente. Par exemple, pour $n = 2p$, le changement d'indice $k = 2p - j$ nous permet d'écrire la deuxième moitié de cette somme sous la forme :

$$\begin{aligned} \sum_{k=p+1}^{2p} \binom{2p}{k} \sin((2k-2p)\theta) &= \sum_{j=0}^{p-1} \binom{2p}{2p-j} \sin((2p-2j)\theta) \\ &= - \sum_{j=0}^p \binom{2p}{j} \sin((2j-2p)\theta) \\ &= - \sum_{k=0}^p \binom{n}{k} \sin((2k-n)\theta) \end{aligned}$$

La vérification étant analogue pour n impair.

La parité de $\cos^p(\theta)$ peut aussi justifier l'absence de termes en $\sin(q\theta)$ dans la formule (5.2).

5.7 Racines n -ièmes d'un nombre complexe

La représentation polaire des nombres complexes nous sera très utile pour résoudre des équations de la forme $x^n = \alpha$.

On rappelle que pour tout entier naturel non nul n , la fonction $f : x \mapsto x^n$ réalise une bijection de $\mathbb{R}^+$ sur lui-même. L'application réciproque de f est notée $x \mapsto \sqrt[n]{x}$ ou $x \mapsto x^{\frac{1}{n}}$ et on l'appelle fonction racine n -ième. Donc pour tout réel positif a , l'unique solution de l'équation $x^n = a$ est le réel positif $\sqrt[n]{a}$.

On rappelle que si $z = re^{it}$ avec $r > 0$ et $t \in \mathbb{R}$, on a pour tout entier relatif n , $z^n = r^n e^{int}$ et pour $z' = r' e^{it'}$ avec $r' > 0$ et $t' \in \mathbb{R}$, l'égalité $z = z'$ est réalisée si, et seulement si, $r = r'$ et $t \equiv t'$ modulo 2π .

Dans le cas où $n = 2$ et $\alpha \neq 0$, on écrit $\alpha = \rho e^{i\theta}$ avec $\rho > 0$ et $\theta \in [-\pi, \pi[$ et on cherche $z = re^{it}$ avec $r > 0$ et $t \in [-\pi, \pi[$ tel que $z^2 = r^2 e^{2it} = \rho e^{i\theta}$, ce qui équivaut à $r^2 = \rho$ et $2t = \theta + 2k\pi$ avec $k \in \mathbb{Z}$. On a donc $r = \sqrt{\rho}$ et $t = \frac{\theta}{2} + k\pi$ avec $k \in \mathbb{Z}$, ce qui donne $z = \sqrt{\rho} e^{i\frac{\theta}{2}} e^{ik\pi}$ et α a deux racines carrées qui sont $z_1 = \sqrt{\rho} e^{i\frac{\theta}{2}}$ et $z_2 = \sqrt{\rho} e^{i\frac{\theta}{2}} e^{i\pi} = -z_1$.

Définition 5.8. Étant donné un nombre complexe α et un entier naturel non nul n , on appelle racine n -ième de α tout nombre complexe z tel que $z^n = \alpha$.

Si $\alpha = 0$, l'équation $z^n = \alpha$ équivaut à $z = 0$, c'est-à-dire que 0 est l'unique racine n -ième de 0. Si $\alpha \neq 0$, une racine n -ième de α est nécessairement non nulle.

Si α est un nombre complexe non nul, il s'écrit alors $\alpha = \rho e^{i\theta}$ avec $\rho > 0$ et $\theta \in [-\pi, \pi[$ (ou $\theta \in \mathbb{R}$ si on se contente d'un quelconque argument de α) et le nombre complexe $z_0 = \sqrt[n]{\rho} e^{i\frac{\theta}{n}}$ nous fournit une solution de l'équation $z^n = \alpha$.

Pour tout autre solution z de cette équation on aura $z^n = z_0^n$, soit $\left(\frac{z}{z_0}\right)^n = 1$ et la connaissance de toutes les racines n -ièmes de 1 nous fournira toutes les racines n -ièmes de α .

Définition 5.9. *Étant donné un entier naturel non nul n , on appelle racine n -ième de l'unité toute racine n -ième de 1.*

Théorème 5.16.

Soit n un entier naturel non nul. Il y a exactement n racines n -ièmes de l'unité qui sont données par :

$$\omega_k = e^{\frac{2ik\pi}{n}} = \cos\left(\frac{2k\pi}{n}\right) + i \sin\left(\frac{2k\pi}{n}\right) \quad (0 \leq k \leq n-1)$$

Preuve. Si $z^n = 1$, on a alors $|z|^n = |z^n| = 1$, donc $|z| = 1$ (c'est l'unique racine n -ième réelle positive de 1) et $z = e^{i\theta}$ avec $\theta \in \mathbb{R}$. L'équation $z^n = 1$ équivaut alors à $e^{in\theta} = 1$, encore équivalent à $n\theta \equiv 0$ modulo 2π . Les racines n -ièmes de l'unité sont donc les nombres complexes $e^{\frac{2ik\pi}{n}}$ où k décrit l'ensemble $\mathbb{Z}$ des entiers relatifs. En effectuant la division euclidienne par n , tout entier k s'écrit $k = qn + r$ avec $0 \leq r \leq n-1$ et $e^{\frac{2ik\pi}{n}} = e^{\frac{2ir\pi}{n}} = \omega_r$. De plus pour j, k entiers compris entre 0 et $n-1$, l'égalité $\omega_j = \omega_k$ est équivalente à $e^{\frac{2i(j-k)\pi}{n}} = 1$, encore équivalent à $\frac{2(j-k)\pi}{n} \equiv 0$ modulo 2π , ce qui revient à dire que $j-k$ est divisible par n , soit $j-k = qn$ et avec $|j-k| \leq n-1$ (puisque j et k sont dans l'intervalle $[0, n-1]$), on déduit que $q = 0$ est la seule possibilité, ce qui signifie que $j = k$. On a donc bien le résultat annoncé. $\square$

Le théorème précédent peut aussi s'énoncer comme suit.

Théorème 5.17.

Pour tout entier naturel non nul n et tout nombre complexe z , on a

$$z^n - 1 = \prod_{k=0}^{n-1} (z - \omega_k), \text{ où les } \omega_k = e^{\frac{2ik\pi}{n}}, \text{ pour } k \text{ compris entre } 0 \text{ et } n-1,$$

sont les racines n -ièmes de l'unité.

Preuve. Le polynôme $P(z) = z^n - 1 - \prod_{k=0}^{n-1} (z - \omega_k)$ est nul ou de degré au plus égal à $n-1$ (le coefficient dominant de $\prod_{k=0}^{n-1} (z - \omega_k)$ est z^n) et s'annule en n points distincts (les ω_k), c'est donc le polynôme nul. $\square$

Exemple 5.1 Les racines cubiques de l'unité sont :

$$\begin{cases} \omega_0 = 1, \\ \omega_1 = e^{\frac{2i\pi}{3}} = \cos\left(\frac{2\pi}{3}\right) + i \sin\left(\frac{2\pi}{3}\right) = -\frac{1}{2} + i\frac{\sqrt{3}}{2} \\ \omega_2 = e^{\frac{4i\pi}{3}} = e^{-\frac{2i\pi}{3}} = \cos\left(\frac{2\pi}{3}\right) - i \sin\left(\frac{2\pi}{3}\right) = -\frac{1}{2} - i\frac{\sqrt{3}}{2} \end{cases}$$

La racine ω_1 est usuellement notée j et $\omega_2 = \bar{j}$.

Corollaire 5.4. Soit n un entier naturel non nul. Tout nombre complexe non nul $\alpha = \rho e^{i\theta}$ a exactement n racines n -ièmes données par :

$$u_k = u_0 \omega_k = \sqrt[n]{\rho} e^{i\frac{\theta}{n}} e^{\frac{2ik\pi}{n}} \quad (0 \leq k \leq n-1)$$

5.8 Exercices

Exercice 5.1. Écrire sous la forme $x + iy$ les nombres complexes suivants :

$$u = \left(\frac{\sqrt{3} - i}{1 + i\sqrt{3}} \right)^{27}, \quad z = \left(\frac{\sqrt{3} + i}{\sqrt{3} - i} + \frac{\sqrt{3} - i}{\sqrt{3} + i} - 1 \right)^{111}, \quad w = \frac{(1 + i\sqrt{3})^4}{(1 + i)^3}$$

Solution. On a $\frac{\sqrt{3} - i}{1 + i\sqrt{3}} = \frac{\sqrt{3} - i}{i(\sqrt{3} - i)} = \frac{1}{i} = -i$, donc $u = -i^{27} = -(i^4)^7 \frac{1}{i} = i$.

On a $\frac{\sqrt{3} + i}{\sqrt{3} - i} + \frac{\sqrt{3} - i}{\sqrt{3} + i} = \frac{(\sqrt{3} + i)^2 + (\sqrt{3} - i)^2}{3 - i^2} = \frac{2(3 + i^2)}{3 - i^2} = 1$, donc $z = 0$.

On a $(1 + i\sqrt{3})^4 = -8(1 + i\sqrt{3})$ et $(1 + i)^3 = -2(1 - i)$, donc $w = 4 \frac{1 + i\sqrt{3}}{1 - i} = 4 \frac{(1 + i\sqrt{3})(1 + i)}{(1 - i)(1 + i)} = 2(1 - \sqrt{3}) + 2i(1 + \sqrt{3})$.

Exercice 5.2. Montrer qu'il n'existe pas de relation d'ordre sur $\mathbb{C}$ qui prolonge la relation $\leq$ de $\mathbb{R}$ et qui soit compatible avec la somme et le produit, c'est-à-dire telle que $a \leq b$ et $c \leq d$ entraîne $a + c \leq b + d$ et $a \leq b$ et $0 \leq c$ entraîne $ac \leq bc$.

Solution. Supposant qu'une telle relation existe. Si $0 \leq i$ [resp. $0 \leq -i$], alors $0 \leq i^2 \leq -1$ [resp. $0 \leq (-i)^2 = (-1)^2 i^2 = -1$] dans $\mathbb{R}$, ce qui est incompatible avec la relation d'ordre sur $\mathbb{R}$.

Exercice 5.3. Soit $j = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$. Calculer j^2 , j^3 , $1 + j + j^2$ et j^n pour tout entier relatif n .

Solution. On a $j^0 = 1$, $j^1 = j$, $j^2 = -\frac{1}{2} - i\frac{\sqrt{3}}{2}$, $j^3 = 1$ et $1 + j + j^2 = 0$. En écrivant n sous la forme $n = 3q + r$ avec $r = 0, 1$ ou 2 , on a :

$$j^n = j^r = \begin{cases} 1 & \text{si } r = 0 \\ j & \text{si } r = 1 \\ j^2 & \text{si } r = 2 \end{cases}$$

Exercice 5.4. Calculer $(1 + i)^n$ pour tout entier naturel n . En déduire les valeurs des sommes $\sum_{j=0}^p \binom{2p}{2j} (-1)^j$ et $\sum_{j=0}^{p-1} \binom{2p}{2j+1} (-1)^j$ pour tout entier naturel non nul p .

Solution. On a $(1 + i)^2 = 2i$, donc $(1 + i)^{2p} = 2^p i^p$ pour tout $p \geq 0$ et on connaît les i^p . Pour les entiers impairs, on a $(1 + i)^{2p+1} = 2^p (1 + i) i^p = 2^p (i^p + i^{p+1})$. On a donc :

$$(1 + i)^{2p} = 2^p i^p = \begin{cases} 2^p & \text{si } p = 4q \\ 2^p i & \text{si } p = 4q + 1 \\ -2^p & \text{si } p = 4q + 2 \\ -2^p i & \text{si } p = 4q + 3 \end{cases}$$

et :

$$(1 + i)^{2p+1} = 2^p (i^p + i^{p+1}) = \begin{cases} 2^p (1 + i) & \text{si } p = 4q \\ 2^p (-1 + i) & \text{si } p = 4q + 1 \\ -2^p (1 + i) & \text{si } p = 4q + 2 \\ 2^p (1 - i) & \text{si } p = 4q + 3 \end{cases}$$

En utilisant la formule du binôme de Newton, on a :

$$\begin{aligned} (1 + i)^{2p} &= \sum_{k=0}^{2p} \binom{2p}{k} i^k = \sum_{j=0}^p \binom{2p}{2j} i^{2j} + \sum_{j=0}^{p-1} \binom{2p}{2j+1} i^{2j+1} \\ &= \sum_{j=0}^p \binom{2p}{2j} (-1)^j + i \sum_{j=0}^{p-1} \binom{2p}{2j+1} (-1)^j \end{aligned}$$

En identifiant les parties réelles et imaginaires, on en déduit que :

$$\begin{aligned} \sum_{j=0}^{4q} \binom{8q}{2j} (-1)^j &= 2^{4q} \text{ et } \sum_{j=0}^{4q-1} \binom{8q}{2j+1} (-1)^j = 0 \\ \sum_{j=0}^{4q+1} \binom{8q+2}{2j} (-1)^j &= 0 \text{ et } \sum_{j=0}^{4q} \binom{8q+2}{2j+1} (-1)^j = 2^{4q+1} \end{aligned}$$

$$\sum_{j=0}^{4q+2} \binom{8q+4}{2j} (-1)^j = -2^{4q+2} \text{ et } \sum_{j=0}^{4q+1} \binom{8q+4}{2j+1} (-1)^j = 0$$

$$\sum_{j=0}^{4q+3} \binom{8q+6}{2j} (-1)^j = 0 \text{ et } \sum_{j=0}^{4q+2} \binom{8q+6}{2j+1} (-1)^j = -2^{4q+3}$$

Exercice 5.5. Montrer que si z est un nombre complexe de module égal à 1, alors $u = i \frac{1+z}{1-z}$ est réel.

Solution. Comme $|z| = 1$, on a $\bar{z} = \frac{1}{z}$ et $\bar{u} = -i \frac{1+\bar{z}}{1-\bar{z}} = -i \frac{1+\frac{1}{z}}{1-\frac{1}{z}} = -i \frac{z+1}{z-1} = u$, ce qui prouve que u est réel.

Exercice 5.6. Soient $z_1, \dots, z_n$ des nombres complexes non nuls deux à deux distincts et tels que $\sum_{k=1}^n z_k = 0$. Montrer qu'il existe deux indices $j \neq k$ compris entre 1 et n tels que $1 \leq \frac{|z_j|}{|z_k|} \leq 2$.

Solution. Quitte à réordonner, on peut supposer que $|z_1| \leq \dots \leq |z_n|$. En supposant que le résultat annoncé est faux, on a $\frac{|z_k|}{|z_{k-1}|} \notin [1, 2]$ pour tout k compris entre 2 et n et comme $\frac{|z_k|}{|z_{k-1}|} \geq 1$, on a nécessairement $\frac{|z_k|}{|z_{k-1}|} > 2$ pour tout k compris entre 2 et n et $|z_n| > 2|z_{n-1}| > 2^2|z_{n-2}| > \dots > 2^{n-1}|z_1|$. Mais l'hypothèse $\sum_{k=1}^n z_k = 0$ nous donne $|z_n| = \left| \sum_{k=1}^{n-1} z_k \right| \leq \sum_{k=1}^{n-1} |z_k| < |z_n| \sum_{k=1}^{n-1} \frac{1}{2^k}$, soit $\sum_{k=1}^{n-1} \frac{1}{2^k} = \frac{1}{2} \frac{1 - \frac{1}{2^{n-1}}}{1 - \frac{1}{2}} = 1 - \frac{1}{2^{n-1}} > 1$, ce qui est impossible.

Exercice 5.7. Déterminer l'ensemble des nombres complexes z tels que $|z - i| = |z - iz| = |z - 1|$.

Solution. On note $z = x + iy$ un nombre complexe. L'égalité $|z - i| = |z - iz|$ équivaut à $x^2 + (y-1)^2 = (x+y)^2 + (y-x)^2$, ou encore à $x^2 - 2y + y^2 + 1 = 2x^2 + 2y^2$, soit à $x^2 + y^2 + 2y - 1 = 0$. L'égalité $|z - i| = |z - 1|$ équivaut à $x^2 + (y-1)^2 = (x-1)^2 + y^2$, encore équivalent à $x = y$. L'équation (8) nous dit alors que x est nécessairement solution de $2x^2 + 2x - 1 = 0$, ce qui donne deux solutions possibles $z_1 = -\frac{1+\sqrt{3}}{2}(1+i)$ et $z_2 = \frac{\sqrt{3}-1}{2}(1+i)$. Réciproquement, on vérifie que ces solutions conviennent bien. Géométriquement, l'ensemble des

points cherché est l'intersection du cercle $\mathcal{C}$ d'équation (2.11) et de la droite $\mathcal{D}$ d'équation $y = x$.

Exercice 5.8. Factoriser dans $\mathbb{R}$ puis dans $\mathbb{C}$, $x^4 + 1$.

Solution. Dans $\mathbb{R}$ on a $x^4 + 1 = (x^2 + 1)^2 - 2x^2 = (x^2 - \sqrt{2}x + 1)(x^2 + \sqrt{2}x + 1)$, les deux polynômes $x^2 \pm \sqrt{2}x + 1$ de discriminant $\delta = -2 < 0$ étant sans racines réelles. Sur $\mathbb{C}$ ces polynômes ont pour racines $\frac{\pm 1 \pm i1}{\sqrt{2}}$, ce qui donne :

$$x^4 + 1 = \left(x - \frac{1-i}{\sqrt{2}}\right) \left(x - \frac{1+i}{\sqrt{2}}\right) \left(x + \frac{1+i}{\sqrt{2}}\right) \left(x + \frac{1-i}{\sqrt{2}}\right)$$

Exercice 5.9. Déterminer les racines carrées complexes de $\alpha = -7 + 24i$.

Solution. En écrivant $z = x + iy$, l'équation $z^2 = \alpha$ équivaut à $x^2 - y^2 = -7$ et $xy = 12$. Considérant que $|z|^2 = x^2 + y^2 = |\alpha| = 25$, on déduit que $x^2 = 9$, donc $x = 3$ et $y = 4$ ou $x = -3$ et $y = -4$. Les deux racines carrées de α sont donc $\pm(3 + 4i)$.

Exercice 5.10. Résoudre dans $\mathbb{C}$ l'équation $z^2 - (1 - 2i)z + 1 - 7i = 0$.

Solution. Cette équation est équivalente à :

$$\left(z - \frac{1-2i}{2}\right)^2 = \frac{(1-2i)^2}{4} - 1 + 7i = \frac{-7+24i}{4}$$

soit à $Z^2 = \alpha = -7+24i$, où on a posé $Z = 2z - 1 + 2i$, ce qui donne $Z = \pm(3 + 4i)$ et $z = \frac{1-2i}{2} \pm \frac{3+4i}{2}$. Les deux solutions complexes de cette équation sont donc $z_1 = 2 + i$, $z_2 = -2 - 3i$.

Exercice 5.11. Résoudre dans $\mathbb{C}$ l'équation $P(z) = z^3 - 3z^2 + 4z - 4 = 0$.

Solution. On élimine tout d'abord le terme en z^2 . On a :

$$\begin{aligned} P(z - \lambda) &= (z - \lambda)^3 - 3(z - \lambda)^2 + 4(z - \lambda) - 4 \\ &= z^3 - 3(\lambda + 1)z^2 + (4 + 6\lambda + 3\lambda^2)z - (\lambda^3 + 3\lambda^2 + 4\lambda + 4) \end{aligned}$$

et $\lambda = -1$ donne $Q(z) = P(z + 1) = z^3 + z - 2$. Cherchant les solutions sous la forme $z = u + v$, on aboutit à $u^3 + v^3 = 2$ et $u^3v^3 = -\frac{1}{27}$, qui nous conduit à résoudre $x^2 - 2x - \frac{1}{27} = 0$ de solutions $u^3 = 1 + \frac{2\sqrt{7}}{3\sqrt{3}}$ et $v^3 = 1 - \frac{2\sqrt{7}}{3\sqrt{3}}$. Ce qui

donne :

$$u \in \left\{ \sqrt[3]{1 + \frac{2\sqrt{7}}{3\sqrt{3}}}, j\sqrt[3]{1 + \frac{2\sqrt{7}}{3\sqrt{3}}}, \bar{j}\sqrt[3]{1 + \frac{2\sqrt{7}}{3\sqrt{3}}} \right\}$$

et $v = -\frac{1}{3u}$ avec $\frac{1}{u} = \frac{1}{\sqrt[3]{1 + \frac{2\sqrt{7}}{3\sqrt{3}}}} = \frac{\sqrt[3]{\frac{2\sqrt{7}}{3\sqrt{3}} - 1}}{\sqrt[3]{\frac{28}{27} - 1}} = 3\sqrt[3]{\frac{2\sqrt{7}}{3\sqrt{3}} - 1}$. D'où les solu-

tions de $Q(z) = 0$:

$$z_1 = \sqrt[3]{\frac{2\sqrt{7}}{3\sqrt{3}} + 1} - \sqrt[3]{\frac{2\sqrt{7}}{3\sqrt{3}} - 1}, \quad z_2 = j\sqrt[3]{\frac{2\sqrt{7}}{3\sqrt{3}} + 1} - \bar{j}\sqrt[3]{\frac{2\sqrt{7}}{3\sqrt{3}} - 1}$$

$$z_3 = \bar{j}\sqrt[3]{\frac{2\sqrt{7}}{3\sqrt{3}} + 1} - j\sqrt[3]{\frac{2\sqrt{7}}{3\sqrt{3}} - 1}$$

Comme 1 est racine évidente de $Q(z) = 1$ et que z_1 est la seule solution réelle, on a nécessairement $\sqrt[3]{\frac{2\sqrt{7}}{3\sqrt{3}} + 1} - \sqrt[3]{\frac{2\sqrt{7}}{3\sqrt{3}} - 1} = 1$, ce qui peut se vérifier en élevant au carré.

Les solutions de $P(z) = 0$ sont alors $z_1 + 1 = 2$ et $z_2 + 1, z_3 + 1$.

Exercice 5.12. Soient z, z' deux nombres complexes non nuls. Montrer que $|z + z'| = |z| + |z'|$ si, et seulement si, $\arg(z) \equiv \arg(z') \pmod{2\pi}$.

Solution. L'égalité $|z + z'| = |z| + |z'|$ est équivalente à $|z + z'|^2 = (|z| + |z'|)^2$ avec :

$$|z + z'|^2 = (z + z')(\bar{z} + \bar{z}') = |z|^2 + |z'|^2 + z\bar{z}' + \bar{z}z'$$

On a donc $|z + z'| = |z| + |z'|$ si, et seulement si, $z\bar{z}' + \bar{z}z' = 2|z||z'|$, ce qui équivaut encore à $\operatorname{Re}(z\bar{z}') = |z||z'|$. En utilisant l'écriture polaire, $z = \rho e^{i\theta}$ et $z' = \rho' e^{i\theta'}$ avec $\rho > 0, \rho' > 0$ et θ, θ' réels, on déduit que $|z + z'| = |z| + |z'|$ si, et seulement si, $\cos(\theta - \theta') = 1$, ce qui équivaut à $\theta - \theta' \equiv 0 \pmod{2\pi}$ et signifie que $\arg(z) \equiv \arg(z') \pmod{2\pi}$.

Exercice 5.13. Démontrer que, pour tout entier naturel r non nul et toute famille $(z_1, \dots, z_r)$ de complexes non nuls, l'égalité $\left| \sum_{k=1}^r z_k \right| = \sum_{k=1}^r |z_k|$ est réalisée si, et seulement si :

$$\forall k \in \mathbb{N}, (2 \leq k \leq r) \Rightarrow (\exists \lambda_k \in]0, +\infty[, z_k = \lambda_k z_1)$$

ce qui revient à dire que tous les z_k ont le même argument modulo 2π .

Solution. Chaque nombre complexe non nul z_k ($1 \leq k \leq r$) peut s'écrire $z_k = \rho_k e^{i\theta_k}$ avec $\rho_k = |z_k| > 0$ et $\theta_k \in]-\pi, \pi[$. On a alors :

$$\begin{cases} \left| \sum_{k=1}^r z_k \right|^2 = \sum_{k=1}^r |z_k|^2 + 2 \sum_{1 \leq j < k \leq r} \rho_j \rho_k \cos(\theta_j - \theta_k) \\ \left(\sum_{k=1}^r |z_k| \right)^2 = \sum_{k=1}^r |z_k|^2 + 2 \sum_{1 \leq j < k \leq r} \rho_j \rho_k \end{cases}$$

et l'égalité $\left| \sum_{k=1}^r z_k \right| = \sum_{k=1}^r |z_k|$ équivaut à $\sum_{1 \leq j < k \leq r} \rho_j \rho_k (1 - \cos(\theta_j - \theta_k)) = 0$. Tous

les termes de cette somme étant positifs ou nuls avec $\rho_j \rho_k > 0$, on en déduit que $\cos(\theta_j - \theta_k) = 1$ avec $\theta_j - \theta_k \in]-2\pi, 2\pi[$ pour $1 \leq j < k \leq r$ (on a $-\pi \leq \theta_j < \pi$ et $-\pi \leq \theta_k < \pi$ donc $-\pi < -\theta_k \leq \pi$ et $-2\pi < \theta_j - \theta_k < 2\pi$), ce qui donne $\theta_j = \theta_k$ et en notant θ cette valeur commune on a $z_k = \rho_k e^{i\theta} = |z_k| e^{i\theta}$ pour tout entier k compris entre 1 et r ou encore :

$$z_k = \frac{|z_k|}{|z_1|} |z_1| e^{i\theta} = \lambda_k z_1 \quad (1 \leq k \leq r)$$

où on a posé $\lambda_k = \frac{|z_k|}{|z_1|}$ pour tout k compris entre 1 et r . Réciproquement si $z_k = \lambda_k z_1$ avec $\lambda_k > 0$ pour tout k compris entre 2 et r et $\lambda_1 = 1$, on a :

$$\left| \sum_{k=1}^r z_k \right| = |z_1| \sum_{k=1}^r \lambda_k = \sum_{k=1}^r \lambda_k |z_1| = \sum_{k=1}^r |z_k|$$

On peut aussi démontrer ce résultat par récurrence sur $r \geq 1$, le cas $r = 2$ correspondant au cas d'égalité dans l'inégalité triangulaire sur $\mathbb{C}$.

Exercice 5.14. Déterminer, pour tout couple de réel (θ, θ') tel que $\cos\left(\frac{\theta - \theta'}{2}\right) \neq 0$, le module et un argument de $e^{i\theta} + e^{i\theta'}$.

Solution. On a $e^{i\theta} + e^{i\theta'} = e^{i\frac{\theta+\theta'}{2}} \left(e^{i\frac{\theta-\theta'}{2}} + e^{-i\frac{\theta-\theta'}{2}} \right) = 2 \cos\left(\frac{\theta - \theta'}{2}\right) e^{i\frac{\theta+\theta'}{2}}$,

donc $e^{i\theta} + e^{i\theta'} \neq 0$ si $\cos\left(\frac{\theta - \theta'}{2}\right) \neq 0$, $|e^{i\theta} + e^{i\theta'}| = 2 \left| \cos\left(\frac{\theta - \theta'}{2}\right) \right|$ et :

$$\arg(e^{i\theta} + e^{i\theta'}) \equiv \begin{cases} \frac{\theta + \theta'}{2} \quad (2\pi) & \text{si } \cos\left(\frac{\theta - \theta'}{2}\right) > 0 \\ \frac{\theta + \theta'}{2} + \pi \quad (2\pi) & \text{si } \cos\left(\frac{\theta - \theta'}{2}\right) < 0 \end{cases}$$

Exercice 5.15. Pour tout réel θ , on désigne par z_θ le nombre complexe $z_\theta = 1 - e^{i\theta}$.

1. Exprimer z_θ en fonction de $\sin\left(\frac{\theta}{2}\right)$ et de $e^{i\frac{\theta}{2}}$.
2. Montrer que pour tout entier naturel n et tout réel $\theta \in \mathbb{R} \setminus 2\pi\mathbb{Z}$, on a
$$\sum_{k=0}^n e^{ik\theta} = e^{in\frac{\theta}{2}} \frac{\sin\left(\frac{n+1}{2}\theta\right)}{\sin\left(\frac{\theta}{2}\right)}.$$
3. En déduire des expressions de $\sum_{k=0}^n \cos(k\theta)$ et de $\sum_{k=1}^n \sin(k\theta)$ pour tout entier naturel non nul n et tout réel θ .

Solution.

1. On a $z_\theta = e^{i\frac{\theta}{2}} \left(e^{-i\frac{\theta}{2}} - e^{i\frac{\theta}{2}} \right) = -2i \sin\left(\frac{\theta}{2}\right) e^{i\frac{\theta}{2}}$. On peut aussi écrire que :

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

2. Comme $e^{i\theta} \neq 1$ pour $\theta \in \mathbb{R} \setminus 2\pi\mathbb{Z}$, on a :

$$\begin{aligned} \sum_{k=0}^n e^{ik\theta} &= \sum_{k=0}^n (e^{i\theta})^k = \frac{1 - e^{i(n+1)\theta}}{1 - e^{i\theta}} = \frac{-2i \sin\left(\frac{n+1}{2}\theta\right) e^{i\frac{n+1}{2}\theta}}{-2i \sin\left(\frac{\theta}{2}\right) e^{i\frac{\theta}{2}}} \\ &= e^{in\frac{\theta}{2}} \frac{\sin\left(\frac{n+1}{2}\theta\right)}{\sin\left(\frac{\theta}{2}\right)} \end{aligned}$$

3. Ce qui donne en identifiant les parties réelles et imaginaires dans l'identité précédente :

$$\sum_{k=0}^n \cos(k\theta) = \cos\left(n\frac{\theta}{2}\right) \frac{\sin\left(\frac{n+1}{2}\theta\right)}{\sin\left(\frac{\theta}{2}\right)} \quad \text{et} \quad \sum_{k=1}^n \sin(k\theta) = \sin\left(n\frac{\theta}{2}\right) \frac{\sin\left(\frac{n+1}{2}\theta\right)}{\sin\left(\frac{\theta}{2}\right)}$$

pour $n \geq 1$ et $\theta \in \mathbb{R} \setminus 2\pi\mathbb{Z}$. Pour $\theta \in 2\pi\mathbb{Z}$, on a $\cos(k\theta) = 1$ et $\sin(k\theta) = 0$ pour tout k , de sorte que $\sum_{k=0}^n \cos(k\theta) = n+1$ et $\sum_{k=1}^n \sin(k\theta) = 0$.

Exercice 5.16. Pour tout réel $\theta \in]-\pi, \pi[$, on désigne par z_θ le nombre complexe $z_\theta = 1 + e^{i\theta}$.

1. Exprimer z_θ en fonction de $\cos\left(\frac{\theta}{2}\right)$ et de $e^{i\frac{\theta}{2}}$.

2. En calculant, pour n entier naturel non nul, z_θ^n de deux manières différentes, montrer que :

$$\sum_{k=0}^n \binom{n}{k} \cos(k\theta) = 2^n \cos^n\left(\frac{\theta}{2}\right) \cos\left(n\frac{\theta}{2}\right)$$

et :

$$\sum_{k=1}^n \binom{n}{k} \sin(k\theta) = 2^n \cos^n\left(\frac{\theta}{2}\right) \sin\left(n\frac{\theta}{2}\right)$$

Solution.

1. On a $z_\theta = e^{i\frac{\theta}{2}} (e^{-i\frac{\theta}{2}} + e^{i\frac{\theta}{2}}) = 2 \cos\left(\frac{\theta}{2}\right) e^{i\frac{\theta}{2}}$. On peut aussi écrire que :

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

2. On a d'une part $z_\theta^n = 2^n \cos^n\left(\frac{\theta}{2}\right) e^{i\frac{n\theta}{2}}$ et d'autre part, en utilisant la formule du binôme de Newton, $z_\theta^n = (1 + e^{i\theta})^n = \sum_{k=0}^n \binom{n}{k} e^{ik\theta}$, ce qui donne en identifiant les parties réelles et imaginaires :

$$\begin{aligned} \sum_{k=0}^n \binom{n}{k} \cos(k\theta) &= 2^n \cos^n\left(\frac{\theta}{2}\right) \cos\left(n\frac{\theta}{2}\right) \\ \sum_{k=1}^n \binom{n}{k} \sin(k\theta) &= 2^n \cos^n\left(\frac{\theta}{2}\right) \sin\left(n\frac{\theta}{2}\right) \end{aligned}$$

Exercice 5.17. On désigne par $D(0, 1) = \{z \in \mathbb{C} \mid |z| \leq 1\}$ le disque unité fermé du plan complexe. Montrer que l'application :

$$\begin{aligned} \varphi : \mathbb{C} \setminus D(0, 1) &\rightarrow \mathbb{C} \setminus [-1, 1] \\ z &\mapsto \frac{1}{2} \left(z + \frac{1}{z} \right) \end{aligned}$$

est bijective.

Solution. Tout $z \in \mathbb{C} \setminus D(0, 1)$ s'écrit $z = \rho e^{i\theta}$ avec $\rho > 1$ et :

$$\varphi(z) = \frac{1}{2} \left(\rho e^{i\theta} + \frac{1}{\rho} e^{-i\theta} \right) = \frac{1}{2} \left(\left(\rho + \frac{1}{\rho} \right) \cos(\theta) + i \left(\rho - \frac{1}{\rho} \right) \sin(\theta) \right)$$

Si $\varphi(z)$ est réel, nécessairement $\sin(\theta) = 0$, donc $\cos(\theta) = \pm 1$ et :

$$|\varphi(z)| = \frac{\rho^2 + 1}{2\rho} = \frac{(\rho - 1)^2 + 2\rho}{2\rho} > \frac{2\rho}{2\rho} = 1$$

On a donc bien $\varphi(z) \in \mathbb{C} \setminus [-1, 1]$ pour tout $z \in \mathbb{C} \setminus D(0, 1)$. Il s'agit maintenant de montrer que pour tout $Z \in \mathbb{C} \setminus [-1, 1]$ l'équation $\frac{1}{2} \left(z + \frac{1}{z} \right) = Z$ a une unique solution $z \in \mathbb{C} \setminus D(0, 1)$. Cette équation est équivalente à $z^2 - 2Zz + 1 = 0$. Le discriminant réduit de cette équation est $\delta' = Z^2 - 1 \neq 0$, ce qui donne deux solutions distinctes $z_1 = \rho_1 e^{i\theta_1}$ et $z_2 = \rho_2 e^{i\theta_2}$ dans $\mathbb{C}$. De $z_1 z_2 = 1$, on déduit que $\rho_1 \rho_2 e^{i(\theta_1 + \theta_2)} = 1$, donc que $e^{i(\theta_1 + \theta_2)}$ est réel positif et $\theta_1 + \theta_2 \equiv 0$ modulo 2π . On a donc $z_1 = \rho_1 e^{i\theta_1}$ et $z_2 = \frac{1}{\rho_1} e^{-i\theta_1}$. Si $\rho_1 = 1$, on a alors :

$$2Z = z_1 + z_2 = e^{i\theta_1} + e^{-i\theta_1} = 2\cos(\theta_1)$$

et $Z = \cos(\theta_1) \in [-1, 1]$, ce qui n'est pas. On a donc $\rho_1 \neq 1$, de sorte que $z_1 \in \mathbb{C} \setminus D(0, 1)$ et $z_2 \notin \mathbb{C} \setminus D(0, 1)$ pour $\rho_1 > 1$, ou $z_2 \in \mathbb{C} \setminus D(0, 1)$ et $z_1 \notin \mathbb{C} \setminus D(0, 1)$ pour $\rho_1 < 1$. Il y a donc une unique racine dans $\mathbb{C} \setminus D(0, 1)$.

Exercice 5.18. Exprimer $\cos(4\theta)$ et $\sin(4\theta)$ comme combinaisons linéaires de puissances de $\cos(\theta)$ et $\sin(\theta)$.

Solution. On a :

$$\begin{aligned} \cos(4\theta) + i \sin(4\theta) &= e^{i4\theta} = (e^{i\theta})^4 = (\cos(\theta) + i \sin(\theta))^4 \\ &= \cos^4(\theta) + 4\cos^3(\theta) \sin(\theta) i - 6\cos^2(\theta) \sin^2(\theta) \\ &\quad - 4\cos(\theta) \sin^3(\theta) i + \sin^4(\theta) \end{aligned}$$

et en conséquence :

$$\begin{aligned} \cos(4\theta) &= \cos^4(\theta) - 6\cos^2(\theta) \sin^2(\theta) + \sin^4(\theta) \\ &= (\cos^2(\theta) + \sin^2(\theta))^2 - 8\cos^2(\theta) \sin^2(\theta) = 1 - 8\cos^2(\theta) \sin^2(\theta) \end{aligned}$$

et :

$$\sin(4\theta) = 4(\cos^3(\theta) \sin(\theta) - \cos(\theta) \sin^3(\theta)) = 4\cos(\theta) \sin(\theta) (\cos^2(\theta) - \sin^2(\theta))$$

Exercice 5.19. Linéariser $\cos^4(\theta)$ et $\sin^4(\theta)$.

Solution. On a :

$$\begin{aligned} \cos^4(\theta) &= \frac{1}{16} (e^{i\theta} + e^{-i\theta})^4 \\ &= \frac{1}{16} (e^{4i\theta} + 4e^{3i\theta} e^{-i\theta} + 6e^{2i\theta} e^{-2i\theta} + 4e^{i\theta} e^{-3i\theta} + e^{-4i\theta}) \\ &= \frac{1}{16} (e^{4i\theta} + e^{-4i\theta} + 4(e^{2i\theta} + e^{-2i\theta}) + 6) = \frac{1}{8} \cos(4\theta) + \frac{1}{2} \cos(2\theta) + \frac{3}{8} \end{aligned}$$

et :

$$\begin{aligned}\sin^4(\theta) &= \frac{1}{16} (e^{i\theta} - e^{-i\theta})^4 \\ &= \frac{1}{16} (e^{4i\theta} - 4e^{3i\theta}e^{-i\theta} + 6e^{2i\theta}e^{-2i\theta} - 4e^{i\theta}e^{-3i\theta} + e^{-4i\theta}) \\ &= \frac{1}{16} (e^{4i\theta} + e^{-4i\theta} - 4(e^{2i\theta} + e^{-2i\theta}) + 6) = \frac{1}{8} \cos(4\theta) - \frac{1}{2} \cos(2\theta) + \frac{3}{8}\end{aligned}$$

Exercice 5.20. Déterminer, pour n entier naturel non nul, toutes les racines n -ièmes de -1 .

Solution. Il s'agit de résoudre l'équation $z^n = -1 = e^{i\pi}$. Les solutions de cette équation sont les $u_k = e^{(2k+1)\frac{i\pi}{n}}$, où $0 \leq k \leq n-1$.

Exercice 5.21. En notant, pour n entier naturel non nul, $(\omega_k)_{0 \leq k \leq n-1}$ la suite de toutes les racines n -ièmes de l'unité, montrer que pour $n \geq 2$, on a $\sum_{k=0}^{n-1} \omega_k = 0$ et $\prod_{k=0}^{n-1} \omega_k = (-1)^{n-1}$.

Solution. On a $\sum_{k=0}^{n-1} \omega_k = \sum_{k=0}^{n-1} \omega_1^k = \frac{1 - \omega_1^n}{1 - \omega_1} = 0$ (pour $n \geq 2$, on a bien $\omega_1 = e^{\frac{2i\pi}{n}} \neq 1$) et :

$$\begin{aligned}\prod_{k=0}^{n-1} \omega_k &= \prod_{k=0}^{n-1} \omega_1^k = \omega_1^{1+2+\dots+(n-1)} = \omega_1^{\frac{n(n-1)}{2}} \\ &= \left(e^{\frac{2i\pi}{n}}\right)^{\frac{n(n-1)}{2}} = e^{\frac{n(n-1)}{2} \frac{2i\pi}{n}} = e^{i(n-1)\pi} = (e^{i\pi})^{n-1} = (-1)^{n-1}\end{aligned}$$

($m = \frac{n(n-1)}{2}$ étant entier, on a bien $(e^{i\theta})^m = e^{im\theta}$). De la première identité, on déduit que $\sum_{k=0}^{n-1} \cos\left(\frac{2k\pi}{n}\right) = \sum_{k=0}^{n-1} \sin\left(\frac{2k\pi}{n}\right) = 0$.

Exercice 5.22. Montrer que, pour tout entier $n \geq 1$, l'ensemble des racines $2n$ -ièmes de l'unité est aussi donnée par :

$$\Gamma_{2n} = \{-1, 1\} \cup \left\{e^{\frac{ik\pi}{n}} \mid 1 \leq k \leq n-1\right\} \cup \left\{e^{\frac{-ik\pi}{n}} \mid 1 \leq k \leq n-1\right\}$$

En déduire que, pour tout nombre complexe z , on a :

$$z^{2n} - 1 = (z^2 - 1) \prod_{k=1}^{n-1} \left(z^2 - 2 \cos\left(\frac{k\pi}{n}\right) z + 1\right)$$

Solution. Ces racines $2n$ -ièmes sont les :

$$\omega_k = e^{\frac{2ik\pi}{2n}} = e^{\frac{ik\pi}{n}} \quad (0 \leq k \leq 2n-1)$$

Pour $k=0$, on a $\omega_0 = 1$, pour $k=n$, on a $\omega_n = e^{i\pi} = -1$ et pour $k=2n-j$ compris entre $n+1$ et $2n-1$, on a $\omega_k = e^{\frac{i\{2n-j\}\pi}{n}} = e^{\frac{-ij\pi}{n}}$, ce qui donne le résultat attendu. On a donc, pour tout nombre complexe z :

$$\begin{aligned} z^{2n} - 1 &= (z-1)(z+1) \prod_{k=1}^{n-1} \left(z - e^{i\frac{k\pi}{n}} \right) \left(z - e^{-i\frac{k\pi}{n}} \right) \\ &= (z^2 - 1) \prod_{k=1}^{n-1} \left(z^2 - \left(e^{i\frac{k\pi}{n}} + e^{-i\frac{k\pi}{n}} \right) z + 1 \right) \\ &= (z^2 - 1) \prod_{k=1}^{n-1} \left(z^2 - 2 \cos \left(\frac{k\pi}{n} \right) z + 1 \right) \end{aligned}$$

Exercice 5.23. Résoudre dans $\mathbb{C}$ l'équation $z^8 + z^4 + 1 = 0$.

Solution. Si z est solution de cette équation, alors $t = z^4$ est solution de $t^2 + t + 1 = 0$, ce qui donne $t \neq 1$ et $\frac{t^3 - 1}{t - 1} = 0$, donc $t = j = e^{\frac{2i\pi}{3}}$ ou $t = \bar{j} = j^2$. Il s'agit alors de calculer les racines quatrièmes de j et de $\bar{j}$, ces racines sont les :

$$z_k = e^{i(\frac{\pi}{6} + k\frac{\pi}{2})} \text{ et } \bar{z}_k = e^{-i(\frac{\pi}{6} + k\frac{\pi}{2})} \quad (0 \leq k \leq 3)$$

soit $z_0 = e^{i\frac{\pi}{6}} = \frac{\sqrt{3}}{2} + \frac{i}{2}$, $z_1 = e^{2i\frac{\pi}{3}} = j = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$, $z_2 = e^{7i\frac{\pi}{6}} = -z_1$, $z_3 = e^{5i\frac{\pi}{6}} = -z_0$ et leurs conjugués. On peut aussi procéder comme suit. Si z est solution de cette équation, il est alors non nul et $z^4 + 1 + \frac{1}{z^4} = 0$, soit $\left(z^2 + \frac{1}{z^2} \right)^2 = 1$, donc $z^2 + \frac{1}{z^2} = \pm 1$, soit $\left(z + \frac{1}{z} \right)^2 - 2 = \pm 1$, c'est-à-dire $\left(z + \frac{1}{z} \right)^2 = 1$ ou $\left(z + \frac{1}{z} \right)^2 = 3$. Si $\left(z + \frac{1}{z} \right)^2 = 1$, on a alors $z + \frac{1}{z} = \pm 1$, soit $z^2 \pm z + 1 = 0$ avec $z \neq \pm 1$ ou encore $\frac{z^3 \pm 1}{z \pm 1} = 0$, ce qui donne les 4 solutions, $j, \bar{j}, -j, -\bar{j}$. Si $\left(z + \frac{1}{z} \right)^2 = 3$, on a alors $z + \frac{1}{z} = \pm\sqrt{3}$, soit $z^2 \pm \sqrt{3}z + 1 = 0$ ou encore $\left(z \pm \frac{\sqrt{3}}{2} \right)^2 = -\frac{1}{4}$, ce qui donne les 4 autres solutions, $z_0 = \frac{\sqrt{3}}{2} + \frac{i}{2}$, $\bar{z}_0, -z_0, -\bar{z}_0$.

Chapitre 6

Espaces vectoriels réels ou complexes

On notera $\mathbb{K}$ l'ensemble des nombres réels ou complexes, en précisant quand cela sera nécessaire s'il s'agit de $\mathbb{R}$ ou $\mathbb{C}$. Par scalaire on entend réel ou complexe.

6.1 L'espace vectoriel $\mathbb{K}^n$

On peut utiliser les nombres réels $x \in \mathbb{R}$ pour représenter tous les points d'une droite, les couples de réels $(x, y) \in \mathbb{R}^2$ pour représenter tous les points d'un plan et les triplets de réels $(x, y, z) \in \mathbb{R}^3$ pour représenter tous les points d'un espace.

De manière plus générale, étant donné un entier naturel non nul n , on appelle vecteur tout élément du produit cartésien $\mathbb{K}^n = \mathbb{K} \times \mathbb{K} \times \cdots \times \mathbb{K}$ (répété n fois). Ces vecteurs sont des listes ordonnées de n scalaires $x_1, x_2, \dots, x_n$ et seront notés

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ ou plus simplement } x = (x_i)_{1 \leq i \leq n} \text{ et on dit que les } x_i \text{ sont les}$$

composantes du vecteur x . La représentation sous forme de vecteurs colonnes sera justifiée plus loin par l'utilisation du calcul matriciel.

On peut naturellement munir cet ensemble d'une opération interne d'addition notée $+$ et définie par :

$$\forall x = (x_i)_{1 \leq i \leq n} \in \mathbb{K}^n, \forall y = (y_i)_{1 \leq i \leq n} \in \mathbb{K}^n, x + y = (x_i + y_i)_{1 \leq i \leq n}$$

Des propriétés de l'addition des scalaires, on déduit facilement que cette opération d'addition vérifie les propriétés suivantes :

- (i) elle est commutative, ce qui signifie que l'on a $x + y = y + x$ pour tous vecteurs x et y ;
- (ii) elle est associative, ce qui signifie que l'on a $x + (y + z) = (x + y) + z$ pour tous vecteurs x, y et z ;

(iii) le vecteur nul $0 = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ est un élément neutre pour cette addition, ce qui

signifie que pour tout vecteur x , on a $x + 0 = 0 + x = x$;

(iv) pour tout vecteur $x = (x_i)_{1 \leq i \leq n} \in \mathbb{K}^n$, le vecteur $x' = (-x_i)_{1 \leq i \leq n}$ est tel que $x + x' = x' + x = 0$, on dit que x' est un opposé de x et on le note $-x$.

Tout cela se résume en disant que l'ensemble $\mathbb{K}^n$ muni de l'addition, que l'on note $(\mathbb{K}^n, +)$, est un groupe commutatif

De même, on peut naturellement munir $\mathbb{K}^n$ d'une multiplication externe définie par :

$$\forall \lambda \in \mathbb{K}, \forall x = (x_i)_{1 \leq i \leq n} \in \mathbb{K}^n, \lambda \cdot x = (\lambda \cdot x_i)_{1 \leq i \leq n}$$

On dit que cette opération est externe car elle associe à un scalaire λ (en dehors de $\mathbb{K}^n$) et à un élément x de $\mathbb{K}^n$ un élément $\lambda \cdot x$ de $\mathbb{K}^n$.

On écrira plus simplement λx pour $\lambda \cdot x$.

Là encore des propriétés de l'addition et de la multiplication des scalaires, on déduit que cette opération externe vérifie les propriétés suivantes :

- (v) pour tout scalaire λ et tous vecteurs x et y , on a $\lambda(x + y) = \lambda x + \lambda y$;
- (vi) pour tous scalaires λ, μ et tout vecteur x , on a $(\lambda + \mu)x = \lambda x + \mu x$;
- (vii) pour tous scalaires λ, μ et tout vecteur x , on a $\lambda(\mu x) = (\lambda\mu)x$;
- (viii) pour tout vecteur x , on a $1 \cdot x = x$.

Tout cela se résume en disant que l'ensemble $\mathbb{K}^n$ muni de cette addition interne et de cette multiplication externe, que l'on note $(\mathbb{K}^n, +, \cdot)$, est un $\mathbb{K}$ -espace vectoriel, ou simplement un espace vectoriel, le corps $\mathbb{K}$ étant sous-entendu.

6.2 Définition d'un espace vectoriel réel ou complexe

De manière plus générale, on donne la définition suivante.

Définition 6.1. On appelle $\mathbb{K}$ -espace vectoriel tout ensemble non vide E muni d'une addition interne $(x, y) \in E \times E \mapsto x + y \in E$ et d'une multiplication externe $(\lambda, x) \in \mathbb{K} \times E \mapsto \lambda x \in E$ vérifiant les propriétés (i) à (viii) précédentes.

Pour simplifier, on dira espace vectoriel pour $\mathbb{K}$ -espace vectoriel. Quand cela sera nécessaire, on précisera espace vectoriel réel (*i. e.* pour $\mathbb{K} = \mathbb{R}$) ou espace vectoriel complexe (*i. e.* pour $\mathbb{K} = \mathbb{C}$). Les éléments d'un espace vectoriel sont appelés vecteurs.

Dans un espace vectoriel l'élément neutre pour l'addition est noté 0 et on dit que c'est le vecteur nul et le symétrique d'un vecteur x est noté $-x$ et on dit que c'est l'opposé de x . Pour x, y dans E , la somme $x + (-y)$ est simplement notée $x - y$.

Exemples 6.1

1. L'ensemble $\mathbb{C}$ des nombres complexes est un espace vectoriel réel et aussi un espace vectoriel complexe.
2. On rappelle qu'une suite réelle est une application u définie sur $\mathbb{N}$ et à valeurs réelles. L'ensemble de toutes les suites réelles est un espace vectoriel réel.
3. Plus généralement, étant donné une partie I non vide de $\mathbb{R}$, l'ensemble E de toutes les applications de I dans $\mathbb{R}$ (resp. dans $\mathbb{C}$) est un espace vectoriel réel (resp. complexe).
4. L'ensemble noté $\mathbb{R}[x]$ des fonctions polynomiales réelles (on dira plus simplement polynômes réel), c'est-à-dire l'ensemble des fonctions P définies par
$$P(x) = \sum_{k=0}^n a_k x^k$$
 pour tout réel x , où les coefficients a_k sont réels, est un espace vectoriel réel. Même chose sur $\mathbb{C}$.
5. L'ensemble $\mathcal{F}$ des polynômes trigonométriques, c'est-à-dire l'ensemble des fonctions P définies par
$$P(x) = \sum_{k=0}^n (a_k \cos(kx) + b_k \sin(kx))$$
 pour tout réel x , où les coefficients a_k et b_k sont réels, est un espace vectoriel réel.

Définition 6.2. Soient E un espace vectoriel, n un entier naturel non nul et $x, y, x_1, \dots, x_n$ des éléments de E . On dit que y est colinéaire à x s'il existe un scalaire λ tel que $y = \lambda x$. Plus généralement, on dit que y est combinaison linéaire de $x_1, \dots, x_n$ s'il existe des scalaires $\lambda_1, \dots, \lambda_n$ tels que
$$y = \sum_{k=1}^n \lambda_k x_k.$$

On peut remarquer que, par définition, un espace vectoriel est stable par combinaison linéaire, c'est-à-dire que si $x_1, \dots, x_n$ sont dans E et $\lambda_1, \dots, \lambda_n$ dans $\mathbb{K}$, alors la combinaison linéaire $\sum_{k=1}^n \lambda_k x_k$ est encore dans E .

En s'inspirant de la construction l'espace vectoriel $\mathbb{K}^n$ comme produit cartésien de p exemplaires de l'espace vectoriel $\mathbb{K}$, on vérifie facilement que le produit cartésien $F = F_1 \times \dots \times F_p$ de p espaces vectoriels $F_1, \dots, F_p$ est naturellement muni d'une structure d'espace vectoriel avec les lois définies par :

$$\begin{cases} x + y = (x_1 + y_1, \dots, x_p + y_p) \\ \lambda x = (\lambda x_1, \dots, \lambda x_p) \end{cases}$$

où $x = (x_1, \dots, x_p)$, $y = (y_1, \dots, y_p)$ sont deux éléments de F et λ un scalaire.

6.3 Sous-espaces vectoriels

Définition 6.3. Soit E un espace vectoriel. On dit qu'une partie F de E est un sous-espace vectoriel de E si :

1. le vecteur 0 est dans F ;
2. pour tous vecteurs x, y dans F et tout scalaire λ , les vecteurs $x + y$ et λx sont dans F .

L'appellation sous-espace vectoriel est justifiée par le résultat suivant.

Théorème 6.1.

Tout sous-espace vectoriel d'un espace vectoriel est un espace vectoriel.

Preuve. Soit F un sous-espace vectoriel d'un espace vectoriel E . Comme F contient 0 , il est non vide. Le deuxième point de la définition nous dit que l'addition des vecteurs et la multiplication d'un vecteur par un scalaire restreintes à F y définissent bien respectivement une opération interne et externe. L'addition des vecteurs qui est commutative sur E l'est en particulier sur F . L'élément neutre 0 pour l'addition est bien dans F . Tout vecteur $x \in F$ admet un opposé $-x \in E$ et en écrivant que $-x = (-1)x$, on voit que $-x$ est bien dans F . Les propriétés (v) à (viii) vérifiées dans E le sont en particulier dans F . $\square$

De manière équivalente, on peut dire qu'une partie F d'un espace vectoriel E est un sous-espace vectoriel si, et seulement si, F est non vide et pour tous vecteurs x, y dans F , tous scalaires λ, μ , le vecteur $\lambda x + \mu y$ est dans F .

De manière plus générale, un sous-espace vectoriel d'un espace vectoriel est une partie non vide stable par combinaison linéaire.

Exemples 6.2

1. Si E un espace vectoriel, alors $\{0\}$ et E sont des sous-espaces vectoriels de E .
2. $\mathbb{R}$ et l'ensemble des imaginaires purs sont des sous-espaces vectoriels réels de $\mathbb{C}$.

Définition 6.4. Dans l'espace $\mathbb{K}^n$, où n est un entier naturel non nul, on appelle droite vectorielle tout sous-ensemble de la forme $D = \{\lambda a \mid \lambda \in \mathbb{K}\}$, où a est un vecteur non nul donné.

On note $D = \mathbb{K}a$ une telle droite. On a donc, en notant $a = (a_k)_{1 \leq k \leq n}$, pour tout $x = (x_k)_{1 \leq k \leq n} \in \mathbb{K}^n$:

$$(x \in D) \Leftrightarrow (\exists \lambda \in \mathbb{K} \mid \forall k \in \{1, 2, \dots, n\}, x_k = \lambda a_k)$$

Une telle représentation est appelée représentation paramétrique de la droite D . Cette définition correspond bien à la notion de droite vectorielle du plan $\mathbb{R}^2$ ou de l'espace $\mathbb{R}^3$ étudiée en Lycée.

On vérifie facilement qu'une droite de $\mathbb{K}^n$ est un sous-espace vectoriel.

Pour $n = 1$, $D = \mathbb{K}$ est la seule droite vectorielle puisque, pour tout $a \neq 0$, tout scalaire x peut s'écrire $x = \frac{x}{a}a = \lambda a$, donc $\mathbb{K} \subset \mathbb{K}a \subset \mathbb{K}$ et $\mathbb{K} = \mathbb{K}a$.

Définition 6.5. Dans l'espace $\mathbb{K}^n$, où $n \geq 2$, on appelle *plan vectoriel* tout sous-ensemble de la forme $P = \{\lambda a + \mu b \mid (\lambda, \mu) \in \mathbb{K}^2\}$, où a et b sont deux vecteurs non colinéaires donnés.

On note $P = \mathbb{K}a \oplus \mathbb{K}b$ un tel plan. On a donc, en notant $a = (a_k)_{1 \leq k \leq n}$ et $b = (b_k)_{1 \leq k \leq n}$, pour tout $x = (x_k)_{1 \leq k \leq n} \in \mathbb{K}^n$:

$$(x \in P) \Leftrightarrow (\exists (\lambda, \mu) \in \mathbb{K} \mid \forall k \in \{1, 2, \dots, n\}, x_k = \lambda a_k + \mu b_k)$$

Une telle représentation est appelée représentation paramétrique du plan P . Là encore cette définition correspond bien à la notion de plan vectoriel de l'espace $\mathbb{R}^3$ étudiée en Lycée. On vérifie facilement qu'un plan de $\mathbb{K}^n$ est un sous-espace vectoriel.

Une partie finie d'un espace vectoriel E distincte de $\{0\}$ n'est pas un sous-espace vectoriel, mais à partir d'une partie finie de vecteurs de E , on peut engendrer un sous-espace vectoriel en s'inspirant des définitions de droites et plans.

Théorème 6.2.

Soient E un espace vectoriel et $x_1, \dots, x_n$ des éléments de E . L'ensemble F de toutes les combinaisons linéaires de $x_1, \dots, x_n$ est un sous-espace vectoriel de E .

Preuve. L'ensemble F contient $0 = \sum_{k=1}^n 0 \cdot x_k$ et pour $x = \sum_{k=1}^n \lambda_k x_k$, $y = \sum_{k=1}^n \mu_k x_k$ dans F et λ, μ dans $\mathbb{K}$, on a $\lambda x + \mu y = \sum_{k=1}^n (\lambda \lambda_k + \mu \mu_k) x_k \in F$, donc F est bien un sous-espace vectoriel de E . $\square$

Définition 6.6. Avec les notations du théorème précédent, on dit que l'ensemble F est le sous-espace vectoriel de E engendré par $x_1, \dots, x_n$ et on le note $F = \langle x_1, \dots, x_n \rangle$, ou $F = \text{Vect} \{x_1, \dots, x_n\}$ ou encore $F = \sum_{k=1}^n \mathbb{K}x_k$.

Si tous les x_k sont nuls, on a alors $F = \{0\}$.

Exemple 6.1 Dans l'espace $\mathbb{K}[x]$ des fonctions polynomiales, pour tout entier naturel non nul n , le sous-espace vectoriel engendré par $1, x, \dots, x^n$ est formé de l'ensemble des polynômes de degré au plus égal à n , on le note $\mathbb{K}_n[x]$ (le cas $n = 0$ correspond aux polynômes constants).

De manière un peu plus générale, on peut définir le sous-espace vectoriel d'un espace vectoriel E engendré par une famille X non vide de E (non nécessairement

finie) comme l'ensemble $F = \text{Vect}(X)$ (ou $F = \langle X \rangle$) de toutes les combinaisons linéaires d'éléments de X . Un vecteur x de E est donc dans $\text{Vect}(X)$ si, et seulement si, il existe un entier $p \geq 1$, des vecteurs $x_1, \dots, x_p$ dans X et des scalaires

$$\lambda_1, \dots, \lambda_p \text{ tels que } x = \sum_{k=1}^p \lambda_k x_k.$$

Le théorème qui suit nous donne deux définitions équivalentes de $\text{Vect}(X)$.

Théorème 6.3.

Si X est une partie non vide d'un espace vectoriel E , $\text{Vect}(X)$ est l'intersection de tous les sous-espaces vectoriels de E qui contiennent X . C'est aussi le plus petit (pour l'ordre défini par l'inclusion) sous-espace vectoriel de E qui contient X , c'est-à-dire que $\text{Vect}(X)$ contient X et est contenu dans tout sous-espace vectoriel de E qui contient X .

On peut aussi définir des sous-espaces vectoriels de $\mathbb{K}^n$ en utilisant des équations linéaires (on verra plus loin, avec la notion de base d'un espace vectoriel, que cela est encore possible pour n'importe quel espace vectoriel).

Théorème 6.4.

Étant donné un entier naturel non nul n et n scalaires non tous nuls $a_1, \dots, a_n$, l'ensemble $F = \left\{ x = (x_i)_{1 \leq i \leq n} \in \mathbb{K}^n \mid \sum_{k=1}^n a_k x_k = 0 \right\}$ est un sous-espace vectoriel de $\mathbb{K}^n$.

Preuve. Il suffit de vérifier. □

En fait si tous les a_k sont nuls, F est encore défini et c'est $\mathbb{K}^n$ tout entier.

Pour $n = 1$, on a $a_1 \neq 0$ et cet espace F est réduit à $\{0\}$. Pour $n = 2$, on a $(a_1, a_2) \neq (0, 0)$ et supposant par exemple que $a_2 \neq 0$, l'équation $a_1 x_1 + a_2 x_2 = 0$ équivaut à $x_2 = -\frac{a_1}{a_2} x_1$, ce qui signifie que F est l'ensemble des vecteurs de la

forme $x = x_1 \begin{pmatrix} 1 \\ -\frac{a_1}{a_2} \end{pmatrix}$, où x_1 décrit $\mathbb{K}$, ce qui équivaut encore à dire que F est

l'ensemble des vecteurs de la forme $x = \frac{x_1}{a_2} \begin{pmatrix} a_2 \\ -a_1 \end{pmatrix} = \lambda \begin{pmatrix} a_2 \\ -a_1 \end{pmatrix}$, où λ décrit $\mathbb{K}$. En définitive F est la droite engendrée par $\begin{pmatrix} a_2 \\ -a_1 \end{pmatrix}$.

Pour $n = 3$, on a $(a_1, a_2, a_3) \neq (0, 0, 0)$ et supposant par exemple que $a_3 \neq 0$, l'équation $a_1 x_1 + a_2 x_2 + a_3 x_3 = 0$ équivaut à $x_3 = -\frac{a_1}{a_3} x_1 - \frac{a_2}{a_3} x_2$, ce qui signifie

que F est l'ensemble des vecteurs de la forme $x = x_1 \begin{pmatrix} 1 \\ 0 \\ -\frac{a_1}{a_3} \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ -\frac{a_2}{a_3} \end{pmatrix}$, où (x_1, x_2) décrit $\mathbb{K}^2$, ce qui équivaut encore à dire que F est l'ensemble des vecteurs

de la forme :

$$x = \frac{x_1}{a_3} \begin{pmatrix} a_3 \\ 0 \\ -a_1 \end{pmatrix} + \frac{x_2}{a_3} \begin{pmatrix} 0 \\ a_3 \\ -a_2 \end{pmatrix} = \lambda \begin{pmatrix} a_3 \\ 0 \\ -a_1 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ a_3 \\ -a_3 \end{pmatrix}$$

où (λ, μ) décrit $\mathbb{K}^2$, les vecteurs $\begin{pmatrix} a_3 \\ 0 \\ -a_1 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ a_3 \\ -a_3 \end{pmatrix}$ étant non colinéaires puisque $a_3 \neq 0$. En définitive F est le plan engendré par $\begin{pmatrix} a_3 \\ 0 \\ -a_1 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ a_3 \\ -a_3 \end{pmatrix}$.

De manière générale, on donne la définition suivante.

Définition 6.7. *Étant donné un entier naturel non nul n , on appelle hyperplan vectoriel de $\mathbb{K}^n$ tout sous-espace vectoriel de la forme :*

$$H = \left\{ x = (x_i)_{1 \leq i \leq n} \in \mathbb{K}^n \mid \sum_{k=1}^n a_k x_k = 0 \right\}$$

où les scalaires $a_1, \dots, a_n$ ne sont pas tous nuls.

On dit aussi que H est l'hyperplan d'équation $\sum_{k=1}^n a_k x_k = 0$.

De manière un peu plus générale, on appelle hyperplan affine de $\mathbb{K}^n$ tout sous-ensemble de $\mathbb{K}^n$ de la forme $H = \left\{ x = (x_i)_{1 \leq i \leq n} \in \mathbb{K}^n \mid \sum_{k=1}^n a_k x_k = \lambda \right\}$, où λ est un scalaire et les scalaires $a_1, \dots, a_n$ ne sont pas tous nuls. On dit aussi que H est l'hyperplan d'équation $\sum_{k=1}^n a_k x_k = \lambda$.

Pour $n = 2$, un hyperplan est une droite affine et pour $n = 3$, un hyperplan est un plan affine.

Dans l'espace $\mathbb{R}^3$ l'intersection de deux plans vectoriels distincts est une droite. Plus généralement, on a le résultat suivant.

Théorème 6.5.

Étant donnés deux entiers naturels non nuls n et p , des scalaires $a_{1,1}, \dots, a_{1,n}, \dots, a_{p,1}, \dots, a_{p,n}$, l'ensemble :

$$F = \left\{ x = (x_i)_{1 \leq i \leq n} \in \mathbb{K}^n \mid \sum_{k=1}^n a_{i,k} x_k = 0 \ (1 \leq i \leq p) \right\}$$

est un sous-espace vectoriel de $\mathbb{K}^n$.

On dit que F est le sous-espace vectoriels de $\mathbb{K}^n$ d'équations linéaires :

$$\sum_{k=1}^n a_{1,k}x_k = 0, \dots, \sum_{k=1}^n a_{p,k}x_k = 0$$

6.4 Applications linéaires

Pour ce paragraphe, on désigne par E et F deux espaces vectoriels. On note 0 le vecteur nul de E et aussi celui de F . En toute rigueur il faudrait noter 0_E et 0_F ces vecteurs nuls, mais en fonction du contexte on sait en général de quel vecteur nul il s'agit.

Définition 6.8. On dit qu'une application u de E dans F est linéaire (ou que c'est un morphisme d'espaces vectoriels) si pour tous vecteurs x, y de E et tout scalaire λ , on a $u(x + y) = u(x) + u(y)$ et $u(\lambda x) = \lambda u(x)$.

Une application linéaire $u : E \rightarrow F$ transforme 0_E en 0_F et l'opposé de x dans E en l'opposé de $u(x)$ dans F . En effet, on a $u(0) = u(0 + 0) = u(0) + u(0)$, donc $u(0) = 0$ et $0 = u(0) = u(x + (-x)) = u(x) + u(-x)$ et donc $u(-x) = -u(x)$.

Exemples 6.3

1. Pour tout scalaire λ , l'application $u : x \mapsto \lambda x$ est linéaire de E dans E . On dit que c'est l'homothétie de rapport λ . Pour $\lambda = 1$, cette application est l'application identité et on la note Id_E ou Id .
2. Pour tout entier j compris entre 1 et n , l'application u définie sur $\mathbb{K}^n$ par $u(x) = x_j$, si $x = (x_i)_{1 \leq i \leq n}$ est linéaire de E dans $\mathbb{K}$. On dit que c'est la j -ième projection canonique.
3. Une translation de vecteur non nul $x \mapsto x + a$, définie de E dans E , n'est pas linéaire.
4. Étant donné un intervalle réel I , la dérivation $f \mapsto f'$ est linéaire de l'espace vectoriel des fonctions dérivables de I dans $\mathbb{R}$ dans l'espace vectoriel des fonctions de I dans $\mathbb{R}$.

On notera $\mathcal{L}(E, F)$ l'ensemble de toutes les applications linéaires de E dans F . Dans le cas où $F = E$, on note plus simplement $\mathcal{L}(E)$ pour $\mathcal{L}(E, E)$. Les éléments de $\mathcal{L}(E)$ sont aussi appelés endomorphismes de E .

Si u et v sont deux applications linéaires de E dans F , $u + v$ est l'application définie sur E par :

$$\forall x \in E, (u + v)(x) = u(x) + v(x)$$

et pour tout scalaire λ , λu est l'application définie sur E par :

$$\forall x \in E, (\lambda u)(x) = \lambda u(x)$$

Il est facile de vérifier que $u + v$ et λu sont aussi des applications linéaires de E dans F . On a donc ainsi défini une addition interne sur $\mathcal{L}(E, F)$ et une multiplication externe. Le résultat qui suit se démontre alors facilement.

Théorème 6.6.

L'ensemble $\mathcal{L}(E, F)$ de toutes les applications linéaires de E dans F muni de ces deux opérations est un espace vectoriel.

La composition des applications permet aussi de construire des applications linéaires à partir d'applications linéaires données.

Théorème 6.7.

Soient E, F, G des espaces vectoriels, u une application linéaire de E dans F et v une application linéaire de F dans G . La composée $v \circ u$ est alors une application linéaire de E dans G .

Preuve. Il suffit de vérifier. □

Dans le cas où $F = E$, on vérifie facilement que $(\mathcal{L}(E), \circ)$ est anneau unitaire non commutatif pour $n \geq 2$. Le fait que $\mathcal{L}(E)$ soit un espace vectoriel et un anneau unitaire, se traduit en disant que c'est une algèbre.

Théorème 6.8.

Soit $u : E \rightarrow F$ une application linéaire. Pour tout $n \in \mathbb{N}^*$, tous vecteurs $x_1, \dots, x_n$ de E et tous scalaires $\lambda_1, \dots, \lambda_n$, on a :

$$u \left(\sum_{k=1}^n \lambda_k x_k \right) = \sum_{k=1}^n \lambda_k u(x_k)$$

Preuve. Pour $n = 1$, on a bien $u(\lambda x) = \lambda u(x)$ pour tout scalaire λ et tout vecteur x par définition d'une application linéaire. Pour $n = 2$, toujours par définition d'une application linéaire, on a pour tous vecteurs x, y et tous scalaires λ, μ :

$$u(\lambda x + \mu y) = u(\lambda x) + u(\mu y) = \lambda u(x) + \mu u(y).$$

Supposant le résultat acquis au rang $n \geq 2$, on se donne $n+1$ vecteurs $x_1, \dots, x_n, x_{n+1}$ et $n+1$ scalaires $\lambda_1, \dots, \lambda_n, \lambda_{n+1}$ et on a :

$$\begin{aligned} u \left(\sum_{k=1}^{n+1} \lambda_k x_k \right) &= u \left(\sum_{k=1}^n \lambda_k x_k \right) + u(\lambda_{n+1} x_{n+1}) \\ &= \sum_{k=1}^n \lambda_k u(x_k) + \lambda_{n+1} u(x_{n+1}) = \sum_{k=1}^{n+1} \lambda_k u(x_k) \end{aligned}$$

□

Définition 6.9. Soit u une application linéaire de E dans F . Le noyau de u est l'ensemble $\ker(u) = \{x \in E \mid u(x) = 0\}$ et l'image de u l'ensemble $\text{Im}(u) = \{u(x) \mid x \in E\}$.

Théorème 6.9.

Le noyau d'une application linéaire u de E dans F est un sous-espace vectoriel de E et son image un sous-espace vectoriel de F .

Preuve. On a vu que $\ker(u)$ contient 0 et pour x, y dans $\ker(u)$, λ, μ dans $\mathbb{K}$, on a $u(\lambda x + \mu y) = \lambda u(x) + \mu u(y) = 0$, ce qui signifie que $\lambda x + \mu y \in \ker(u)$. Donc $\ker(u)$ est bien un sous-espace vectoriel de E . De manière analogue, en utilisant la linéarité de u , on montre que $\text{Im}(u)$ est un sous-espace vectoriel de F . $\square$

Théorème 6.10.

Soit u une application linéaire de E dans F .

1. *L'application u est injective si, et seulement si, $\ker(u)$ est réduit à $\{0\}$.*
2. *L'application u est surjective si, et seulement si, $\text{Im}(u) = F$.*

Preuve.

1. Supposons u injective. Pour tout $x \in \ker(u)$, on a $u(x) = u(0)$ et nécessairement $x = 0$ puisque u est injective. Donc $\ker(u) = \{0\}$. Réciproquement, supposons que $\ker(u) = \{0\}$. Si x, y dans E sont tels que $u(x) = u(y)$, on a alors $u(x - y) = u(x) - u(y) = 0$, c'est-à-dire que $x - y \in \ker(u)$ et donc $x = y$.
2. Ce résultat est en fait valable pour toute application de E dans F (la linéarité de u n'intervient pas ici).

$\square$

Les applications linéaires de E dans $\mathbb{K}$ ont un statut particulier.

Définition 6.10. *On appelle forme linéaire sur E toute application linéaire de E dans $\mathbb{K}$.*

Exemple 6.2 *Étant donné un segment $I = [a, b]$ non réduit à un point, l'application $f \mapsto \int_a^b f(x) dx$ est une forme linéaire sur l'espace vectoriel des fonctions continues de I dans $\mathbb{R}$.*

On rappelle que si u est une bijection de E sur F , elle admet alors une application réciproque notée u^{-1} et définie par :

$$(y \in F \text{ et } x = u^{-1}(y)) \Leftrightarrow (x \in E \text{ et } y = u(x))$$

Cette application est aussi l'unique application de F dans E telle que $u \circ u^{-1} = \text{Id}_F$ et $u^{-1} \circ u = \text{Id}_E$.

Dans le cas où u est linéaire, il en est de même de u^{-1} . En effet si y, y' sont dans F , ils s'écrivent alors de manière unique $y = u(x)$, $y' = u(x')$ et on a :

$$u^{-1}(y + y') = u^{-1}(u(x) + u(x')) = u^{-1}(u(x + x')) = x + x' = u^{-1}(y) + u^{-1}(y')$$

et $u^{-1}(\lambda y) = u^{-1}(\lambda u(x)) = u^{-1}(u(\lambda x)) = \lambda x = \lambda u^{-1}(y)$ pour tout scalaire λ .

Définition 6.11. On appelle isomorphisme de E sur F toute application linéaire bijective de E sur F .

Dans le cas où $E = F$, un isomorphisme de E sur E est un automorphisme de E . On note $GL(E)$ l'ensemble de tous les automorphismes de E et on dit que $GL(E)$ est le groupe linéaire de E .

6.5 Base canonique de $\mathbb{K}^n$ et expression matricielle des applications linéaires de $\mathbb{K}^n$ dans $\mathbb{K}^m$

Tout vecteur $x = (x_i)_{1 \leq i \leq n}$ de $\mathbb{K}^n$ s'écrit $\sum_{k=1}^n x_k e_k$, où on a noté, pour tout entier k compris entre 1 et n , e_k le vecteur dont toutes les composantes sont nulles sauf la k -ième qui vaut 1. L'espace vectoriel $\mathbb{K}^n$ est donc engendré par les vecteurs $e_1, \dots, e_n$. On dit que le système $(e_1, \dots, e_n)$, que l'on note plus simplement $(e_k)_{1 \leq k \leq n}$, est un système générateur de $\mathbb{K}^n$. De plus par définition du produit cartésien $\mathbb{K}^n$ une telle écriture est unique, ce qui se traduit en disant que le système $(e_k)_{1 \leq k \leq n}$ est un système libre. On dit que $(e_k)_{1 \leq k \leq n}$ est la base canonique de $\mathbb{K}^n$. Nous définirons plus loin la notion de base d'un espace vectoriel.

Si m est un autre entier naturel non nul, on note $(f_k)_{1 \leq k \leq m}$ la base canonique de $\mathbb{K}^m$.

Étant donnée une application linéaire u de E dans F , on a pour tout vecteur $x = \sum_{j=1}^n x_j e_j$ de $\mathbb{K}^n$:

$$u(x) = u\left(\sum_{j=1}^n x_j e_j\right) = \sum_{j=1}^n x_j u(e_j)$$

et chacun des vecteurs $u(e_j)$, pour j compris entre 1 et n , étant dans $\mathbb{K}^m$, il s'écrit

$$u(e_j) = \sum_{i=1}^m a_{ij} f_i. \text{ On a donc en définitive :}$$

$$u(x) = \sum_{j=1}^n x_j \sum_{i=1}^m a_{ij} f_i = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j \right) f_i$$

ce qui signifie que les composantes du vecteurs $u(x)$ sont données par :

$$y_i = \sum_{j=1}^n a_{ij} x_j \quad (1 \leq i \leq m) \quad (6.1)$$

Par exemple pour $n = m = 2$, on a $u(e_1) = a_{11}f_1 + a_{21}f_2$, $u(e_2) = a_{12}f_1 + a_{22}f_2$ et $u(x) = y_1f_1 + y_2f_2$ avec :

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 \\ y_2 = a_{21}x_1 + a_{22}x_2 \end{cases} \quad (6.2)$$

L'application linéaire u est donc uniquement déterminée par les 4 scalaires a_{11}, a_{12}, a_{21} et a_{22} . On stocke ces scalaires dans un tableau à 2 lignes et 2 colonnes $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, où la première colonne $\begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$ est le vecteur $u(e_1)$ et la deuxième colonne $\begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$ le vecteur $u(e_2)$. Un tel tableau est appelé matrice à 2 lignes et 2 colonnes ou plus simplement matrice 2×2 .

On traduit les deux égalités de (6.2) en utilisant le produit matriciel :

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

De manière générale, une application linéaire u de $\mathbb{K}^n$ dans $\mathbb{K}^m$ est donc uniquement déterminée par la matrice A à m lignes et n colonnes :

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

où la colonne numéro j , pour $1 \leq j \leq n$, est le vecteur $u(e_j) = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$. Les

égalités (6.1) se traduisent alors par le produit matriciel :

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$$

qui s'effectue comme suit :

$$y_i = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \sum_{j=1}^n a_{ij} x_j$$

pour tout i compris entre 1 et m . Tout cela est compacté en :

$$(y = u(x)) \Leftrightarrow (y = Ax)$$

et on dit que A est la matrice de u dans les bases canoniques de $\mathbb{K}^n$ et $\mathbb{K}^m$. Si $n = m$, on dit simplement que A est la matrice de $u \in \mathcal{L}(\mathbb{K}^n)$ dans la base canonique de $\mathbb{K}^n$.

6.6 Matrices réelles ou complexes

On note $\mathcal{M}_{m,n}(\mathbb{K})$ l'ensemble de toutes les matrices à m lignes et n colonnes et à coefficients dans $\mathbb{K}$.

Une matrice $\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in \mathcal{M}_{m,n}(\mathbb{K})$ est notée $A = ((a_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$,

le premier indice i étant le numéro de ligne et le deuxième j , le numéro de colonne.

Pour $n = m$, on note $\mathcal{M}_n(\mathbb{K})$ l'ensemble $\mathcal{M}_{n,n}(\mathbb{K})$ et dit que c'est l'ensemble des matrices carrées d'ordre n à coefficients dans $\mathbb{K}$.

6.6.1 Opérations sur les matrices

Théorème 6.11.

Si u et v sont deux applications linéaires de $\mathbb{K}^n$ dans $\mathbb{K}^m$ ayant pour matrices respectives $A = ((a_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ et $B = ((b_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ dans les bases canoniques, alors l'application linéaire $u + v$ a pour matrice dans ces bases canoniques, la matrice $((a_{i,j} + b_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$. Et pour tout scalaire λ , la matrice de λu dans les bases canoniques est la matrice $((\lambda a_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$.

Preuve. Résulte de :

$$(u + v)(e_j) = u(e_j) + v(e_j) = \sum_{i=1}^m a_{ij} f_i + \sum_{i=1}^m b_{ij} f_i = \sum_{i=1}^m (a_{ij} + b_{ij}) f_i$$

et de :

$$(\lambda u)(e_j) = \lambda u(e_j) = \lambda \sum_{i=1}^m a_{ij} f_i = \sum_{i=1}^m \lambda a_{ij} f_i$$

pour $1 \leq j \leq n$. □

On définit donc naturellement la somme de deux matrices $n \times m$ et le produit d'une telle matrice par un scalaire par :

$$A + B = ((a_{i,j} + b_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \text{ et } \lambda A = ((\lambda a_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$$

en utilisant les notations précédentes.

On vérifie alors facilement le résultat suivant.

Théorème 6.12.

L'ensemble $\mathcal{M}_{m,n}(\mathbb{K})$ des matrices à m lignes et n colonnes est un espace vectoriel.

On note 0 la matrice nulle, c'est-à-dire l'élément de $\mathcal{M}_{m,n}(\mathbb{K})$ dont toutes les composantes sont nulles et pour toute matrice $A = ((a_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$, on note $-A = ((-a_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ l'opposé de A .

En explicitant la matrice d'une composée de deux applications linéaires on définira le produit de deux matrices.

On se donne donc une application linéaire v de $\mathbb{K}^n$ dans $\mathbb{K}^m$ et une application linéaire u de $\mathbb{K}^m$ dans $\mathbb{K}^r$ de matrices respectives $B = ((b_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ dans $\mathcal{M}_{m,n}(\mathbb{K})$ et $A = ((a_{i,j}))_{\substack{1 \leq i \leq r \\ 1 \leq j \leq m}}$ dans $\mathcal{M}_{r,m}(\mathbb{K})$.

On note toujours $(e_k)_{1 \leq k \leq n}$ la base canonique de $\mathbb{K}^n$, $(f_k)_{1 \leq k \leq m}$ celle de $\mathbb{K}^m$ et $(g_k)_{1 \leq k \leq r}$ est celle de $\mathbb{K}^r$.

La matrice $C = ((c_{i,j}))_{\substack{1 \leq i \leq r \\ 1 \leq j \leq n}} \in \mathcal{M}_{r,n}(\mathbb{K})$ de $u \circ v \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^r)$ est obtenu en calculant les composantes des vecteurs $u \circ v(e_j)$, pour j compris entre 1 et n , dans la base $(g_k)_{1 \leq k \leq r}$.

On a :

$$u \circ v(e_j) = u(v(e_j)) = u\left(\sum_{k=1}^m b_{kj} f_k\right) = \sum_{k=1}^m b_{kj} u(f_k)$$

avec $u(f_k) = \sum_{i=1}^r a_{ik} g_i$ pour k compris entre 1 et m , ce qui donne :

$$u \circ v(e_j) = \sum_{k=1}^m b_{kj} \left(\sum_{i=1}^r a_{ik} g_i \right) = \sum_{i=1}^r \left(\sum_{k=1}^m a_{ik} b_{kj} \right) g_i$$

et signifie que les coefficients de la matrice C sont donnés par :

$$c_{ij} = \sum_{k=1}^m a_{ik} b_{kj} \quad (1 \leq i \leq r, 1 \leq j \leq n)$$

Au vu de ce résultat, on donne la définition suivante.

Définition 6.12. Pour toutes matrices $A = ((a_{i,j}))_{\substack{1 \leq i \leq r \\ 1 \leq j \leq m}}$ dans $\mathcal{M}_{r,m}(\mathbb{K})$ et $B = ((b_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ dans $\mathcal{M}_{m,n}(\mathbb{K})$, le produit AB de A par B est la matrice $C = ((c_{i,j}))_{\substack{1 \leq i \leq r \\ 1 \leq j \leq n}}$ de $\mathcal{M}_{r,n}(\mathbb{K})$ définie par :

$$c_{ij} = \sum_{k=1}^m a_{ik} b_{kj} \quad (1 \leq i \leq r, 1 \leq j \leq n)$$

Et nous venons de montrer le résultat suivant.

Théorème 6.13.

Si $v : \mathbb{K}^n \rightarrow \mathbb{K}^m$ est une application linéaire de matrice $B = ((b_{ij}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ et $u : \mathbb{K}^m \rightarrow \mathbb{K}^r$ une application linéaire de matrice $A = ((a_{ij}))_{\substack{1 \leq i \leq r \\ 1 \leq j \leq m}}$ dans les bases canoniques, alors la matrice dans les bases canoniques de l'application linéaire $u \circ v$ de $\mathbb{K}^n$ dans $\mathbb{K}^r$ est la matrice produit $C = AB$.

Il est important de remarquer que l'on ne peut définir le produit AB que si le nombre de colonnes de la matrice A est égal au nombre de lignes de la matrice B , ce produit est donc défini de $\mathcal{M}_{r,m}(\mathbb{K}) \times \mathcal{M}_{m,n}(\mathbb{K})$ dans $\mathcal{M}_{r,n}(\mathbb{K})$.

L'opération de multiplication des matrices est une opération interne sur l'espace $\mathcal{M}_n(\mathbb{K})$ des matrices carrées d'ordre n vérifiant les propriétés suivante :

- elle est associative, c'est-à-dire que pour toutes matrices A, B, C dans $\mathcal{M}_n(\mathbb{K})$, on a $A(BC) = (AB)C$;
- elle est distributive par rapport à l'addition, c'est-à-dire que pour toutes matrices A, B, C dans $\mathcal{M}_n(\mathbb{K})$, on a $A(B + C) = AB + AC$;
- la matrice $I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$ est l'élément neutre pour ce produit, c'est-à-dire que $A \cdot I_n = I_n \cdot A = A$ pour toute matrice A dans $\mathcal{M}_n(\mathbb{K})$.

Ces propriétés ajoutées à celle de l'addition des matrices se traduisent en disant que $(\mathcal{M}_n(\mathbb{K}), +, \cdot)$ est un anneau unitaire. Sachant que c'est aussi un espace vectoriel, c'est une algèbre.

L'associativité du produit matriciel dans $\mathcal{M}_n(\mathbb{K})$ permet de définir les puissances successives d'une matrice A par la relation de récurrence :

$$\begin{cases} A^0 = I_n \\ \forall p \in \mathbb{N}, A^{p+1} = A^p A = A A^p \end{cases}$$

Le produit des matrices dans $\mathcal{M}_n(\mathbb{K})$ n'est pas commutatif. Par exemple pour $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$, on a $AB = \begin{pmatrix} 5 & 11 \\ 11 & 25 \end{pmatrix} \neq BA = \begin{pmatrix} 10 & 14 \\ 14 & 20 \end{pmatrix}$.

On dira que deux matrices A et B dans $\mathcal{M}_n(\mathbb{K})$ commutent si $AB = BA$.

6.6.2 Matrices inversibles

Définition 6.13. *On dit qu'une matrice A dans $\mathcal{M}_n(\mathbb{K})$ est inversible s'il existe une matrice A' dans $\mathcal{M}_n(\mathbb{K})$ telle que $AA' = A'A = I_n$. On dit alors que A' est un inverse de A .*

Si une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est inversible son inverse est alors unique et on le note A^{-1} .

Théorème 6.14.

Si $A \in \mathcal{M}_n(\mathbb{K})$ est inversible, alors A^{-1} est aussi inversible et on a $(A^{-1})^{-1} = A$. Le produit de deux matrices inversibles A et B est inversible et $(AB)^{-1} = B^{-1}A^{-1}$.

Preuve. Le premier point résulte de la définition et le deuxième de :

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$$

et :

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = B^{-1}I_nB = B^{-1}B = I_n$$

□

Par récurrence, on déduit que le produit de p matrices inversibles $A_1, \dots, A_p$ est inversible avec $(A_1 \cdots A_p)^{-1} = A_p^{-1} \cdots A_1^{-1}$.

Théorème 6.15.

Un endomorphisme u de $\mathbb{K}^n$ est bijectif si, et seulement si, sa matrice A dans la base canonique est inversible et dans ce cas A^{-1} est la matrice de u^{-1} dans la base canonique.

Preuve. Supposons u bijectif et notons A' la matrice de u^{-1} dans la base canonique de $\mathbb{K}^n$. De $u \circ u^{-1} = u^{-1} \circ u = Id$, on déduit que $AA' = A'A = I_n$, ce qui signifie que A est inversible d'inverse A' . Réciproquement supposons A inversible et désignons par u' l'endomorphisme de $\mathbb{K}^n$ de matrice A^{-1} dans la base canonique. La matrice de $u \circ u'$ [resp. de $u' \circ u$] est $AA^{-1} = I_n$ [resp. $A^{-1}A = I_n$], donc $u \circ u' = Id$ [resp. de $u' \circ u = Id$] et u est surjective [resp. injective]. L'endomorphisme u est donc bijectif. □

Théorème 6.16.

Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est inversible si, et seulement si, l'unique solution du système linéaire $Ax = 0$ est $x = 0$.

Preuve. Si A est inversible et x est solution de $Ax = 0$, on a alors $A^{-1}Ax = x = 0$, donc 0 est l'unique solution de $Ax = 0$. Pour la réciproque, on procède par récurrence sur $n \geq 1$.

On désigne par $(e_i)_{1 \leq i \leq n}$ la base canonique de $\mathbb{K}^n$. Pour $n = 1$, $A = (a)$ est inversible si, et seulement si, $a \neq 0$ et le résultat est évident. Supposons le résultat acquis pour $n - 1 \geq 1$ et soit $A \in \mathcal{M}_n(\mathbb{K})$ telle que $x = 0$ soit l'unique solution du système linéaire $Ax = 0$. La première colonne de A n'est pas nulle puisque c'est $C_1 = Ae_1$ avec $e_1 \neq 0$, il existe donc un indice j tel que $a_{j1} \neq 0$.

Montrer que A est inversible équivaut à montrer que $\frac{1}{a_{j1}}A$ est inversible, ce qui nous ramène à $a_{j1} = 1$. Si $a_{11} = 0$, alors $j \geq 2$ et il est équivalent de montrer que la matrice $P_j(1)A$ (notations de l'exercice 6.10) est inversible, ce qui nous ramène à $a_{11} = 1$ ($P_j(1)A$ se déduit de A en ajoutant la ligne j à la ligne 1). Il est encore

équivalent de montrer que $AP_2(-a_{12})$ est inversible, ce qui ramène à $a_{12} = 0$ et multipliant à droite par $P_j(-a_{1j})$ pour $2 \leq j \leq n$, on se ramène à $a_{1j} = 0$. En résumé, il suffit de considérer le cas où $A = \begin{pmatrix} 1 & 0 \\ c & B \end{pmatrix}$ avec $0 \in \mathcal{M}_{1,n-1}(\mathbb{K})$, $c \in \mathcal{M}_{n-1,1}(\mathbb{K})$ et $B \in \mathcal{M}_{n-1}(\mathbb{K})$. Si $x' \in \mathbb{K}^{n-1} \setminus \{0\}$ est solution de $Bx' = 0$, alors $x = \begin{pmatrix} 0 \\ x' \end{pmatrix} \in \mathbb{K}^n \setminus \{0\}$ est solution de $Ax = 0$, ce qui contredit l'hypothèse de départ. Donc $x' = 0$ est l'unique solution de $Bx' = 0$ et B est inversible. En posant $A' = \begin{pmatrix} 1 & 0 \\ d & B^{-1} \end{pmatrix}$ où $d \in \mathcal{M}_{n-1,1}(\mathbb{K})$ est à préciser, on a :

$$AA' = \begin{pmatrix} 1 & 0 \\ c & B \end{pmatrix} \begin{pmatrix} 1 & 0 \\ d & B^{-1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ c + Bd & I_{n-1} \end{pmatrix}$$

et :

$$A'A = \begin{pmatrix} 1 & 0 \\ d & B^{-1} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ c & B \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ d + B^{-1}c & I_{n-1} \end{pmatrix}$$

soit $AA' = A'A = I_n$ en prenant $d = -B^{-1}c$. La matrice A est donc inversible. $\square$

Corollaire 6.1. *Un endomorphisme u de $\mathbb{K}^n$ est bijectif si, et seulement si, son noyau est réduit à $\{0\}$.*

Preuve. Si A est la matrice canoniquement associée à u , on a $u(x) = Ax$ et $\ker(u) = \{0\}$ équivaut à dire que 0 est l'unique solution de $Ax = 0$, ce qui équivaut à A inversible encore équivaut à dire que u est un isomorphisme. $\square$

Corollaire 6.2. *Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est inversible d'inverse A' dans $\mathcal{M}_n(\mathbb{K})$ si, et seulement si, on a $A'A = I_n$ [resp. $AA' = I_n$].*

Preuve. Supposons que $A'A = I_n$. Si x est solution de $Ax = 0$, on a alors $x = I_n x = A'(Ax) = A'0 = 0$ et A est inversible avec $A^{-1} = (A'A)A^{-1} = A'$. Si $AA' = I_n$, la matrice A' est alors inversible d'inverse A , donc $A = (A')^{-1}$ est inversible d'inverse A' . $\square$

Montrer qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est inversible et calculer son inverse revient à montrer que pour tout vecteur $y \in \mathbb{K}^n$ le système linéaire de n équation à n inconnues $Ax = y$ a une unique solution et à exprimer cette solution x en fonction de y .

Théorème 6.17.

Une matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{K})$ est inversible si, et seulement si, on a $ad - bc \neq 0$ et dans ce cas, on a $A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

Preuve. Pour tous scalaires a, b, c, d , on a $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = (ad - bc) I_2$.
Si $ad - bc \neq 0$, cela s'écrit $AA' = I_2$ avec $A' = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$, ce qui signifie que A est inversible d'inverse A' . Réciproquement si A est inversible, on a :

$$\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = A^{-1} \left(A \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \right) = (ad - bc) A^{-1}$$

donc $ad - bc \neq 0$ (puisque $A \neq 0$) et $A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. □

6.6.3 Déterminant d'une matrice d'ordre 2

Définition 6.14. Le déterminant d'une matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{K})$ est le scalaire $\det(A) = ad - bc$.

Ce déterminant est aussi noté $\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$.

Le théorème 6.17 nous dit qu'une matrice $A \in \mathcal{M}_2(\mathbb{K})$ est inversible si, et seulement si, son déterminant est non nul. Nous généraliserons plus loin cette définition du déterminant.

Avec le théorème qui suit on résume les propriétés fondamentales du déterminant des matrices d'ordre 2.

Théorème 6.18.

On désigne par A, B des matrices d'ordre 2.

1. $\det(I_2) = 1$.
2. Pour tout scalaire λ , on a, $\det(\lambda A) = \lambda^2 \det(A)$.
3. $\det(AB) = \det(A) \det(B)$.
4. Si A est inversible, alors $\det(A^{-1}) = \frac{1}{\det(A)}$.
5. Si l'une des lignes [resp. des colonnes] de A est nulle, alors $\det(A) = 0$.
6. Si A' est la matrice déduite de A en permutant les deux lignes [resp. les deux colonnes], alors $\det(A') = -\det(A)$.

Preuve. On note $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ et $B = \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}$.

1. Il suffit de vérifier.
2. On a $\lambda A = \begin{pmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{pmatrix}$ et $\det(\lambda A) = \lambda^2 ad - bc = \lambda^2 \det(A)$.

3. On a $AB = \begin{pmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{pmatrix}$ et :

$$\begin{aligned} \det(AB) &= (aa' + bc')(cb' + dd') - (ab' + bd')(ca' + dc') \\ &= aa'cb' + aa'dd' + bc'cb' + bc'dd' - ab'ca' - ab'dc' - bd'ca' - bd'dc' \\ &= aa'dd' + bc'cb' - ab'dc' - bd'ca' \\ &= ad(a'd' - b'c') - bc(a'd' - b'c') \\ &= (ad - bc)(a'd' - b'c') = \det(A) \det(B) \end{aligned}$$

4. Pour A inversible, on $AA^{-1} = I_2$ et $\det(A) \det(A^{-1}) = \det(AA^{-1}) = \det(I_2) = 1$, ce qui donne $\det(A^{-1}) = \frac{1}{\det(A)}$.

5. Résulte de la définition.

6. On a $A' = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$ [resp. $A' = \begin{pmatrix} c & d \\ a & b \end{pmatrix}$] et $\det(A') = cb - ad = -\det(A)$.

□

6.6.4 Transposée d'une matrice

Définition 6.15. La transposée d'une matrice $A = ((a_{ij}))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$ à n lignes et m colonnes est la matrice à m lignes et n colonnes ${}^tA = ((a'_{ij}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ où $a'_{ij} = a_{ji}$.

Le transposé d'un vecteur colonne $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ (matrice à n lignes et 1 colonne) est le vecteur ligne ${}^tX = (x_1, x_2, \dots, x_n)$.

En représentant $A = ((a_{ij}))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$ sous forme de lignes $A = \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix}$ [resp. de colonnes $M = (C_1, C_2, \dots, C_m)$] où :

$$L_i = (a_{i1}, a_{i2}, \dots, a_{im}) \text{ [resp. } C_j = \begin{pmatrix} a_{1,j} \\ a_{2,j} \\ \vdots \\ a_{n,j} \end{pmatrix}]$$

est la ligne numéro i [resp. la colonne numéro j] de M , on a :

$${}^tA = ({}^tL_1, {}^tL_2, \dots, {}^tL_n) \text{ [resp. } {}^tA = \begin{pmatrix} {}^tC_1 \\ {}^tC_2 \\ \vdots \\ {}^tC_m \end{pmatrix}]$$

Exemple 6.3 La transposée d'une matrice carrée triangulaire supérieure [resp. inférieure] est triangulaire inférieure [resp. supérieure].

Définition 6.16. On dit qu'une matrice carrée $A = ((a_{ij}))_{1 \leq i, j \leq n}$ est symétrique si elle est égale à sa transposée, ce qui revient à dire que $a_{ij} = a_{ji}$ pour tous i, j compris entre 1 et n .

Définition 6.17. On dit qu'une matrice carrée $A = ((a_{ij}))_{1 \leq i, j \leq n}$ est antisymétrique si ${}^tA = -A$, ce qui revient à dire que $a_{ij} = -a_{ji}$ pour tous i, j compris entre 1 et n .

Une matrice anti-symétrique a tous ses termes diagonaux nuls. En effet, pour tout i compris entre 1 et n , on $a_{ii} = -a_{ii}$ et en conséquence, $a_{ii} = 0$.

Théorème 6.19.

Pour toutes matrices $A = ((a_{ij}))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$ et $B = ((b_{ij}))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$ dans $\mathcal{M}_{n,m}(\mathbb{K})$ et tout scalaire λ , on a :

$${}^t({}^tA) = A, {}^t(A+B) = {}^tA + {}^tB, {}^t(\lambda A) = \lambda {}^tA$$

Pour toutes matrices $A = ((a_{ij}))_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}}$ dans $\mathcal{M}_{p,n}(\mathbb{K})$ et $B = ((b_{ij}))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$ dans $\mathcal{M}_{n,m}(\mathbb{K})$, on a ${}^t(AB) = {}^tB {}^tA$. Si $A \in \mathcal{M}_n(\mathbb{K})$ est inversible, alors ${}^tA$ est aussi inversible et on a $({}^tA)^{-1} = {}^t(A^{-1})$.

Preuve. On a ${}^tA = ((a'_{ij}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ avec $a'_{ij} = a_{ji}$ et ${}^t({}^tA) = {}^tA' = ((a''_{ij}))_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}}$ avec $a''_{ij} = a'_{ji} = a_{ij}$. Donc ${}^t({}^tA) = A$. Les résultats sur les combinaisons linéaires

de matrices sont évidents. Les coefficients de $C = AB$ sont les $c_{ij} = \sum_{k=1}^n a_{ik}b_{kj}$ et

ceux de ${}^tC$ les $c'_{ij} = c_{ji} = \sum_{k=1}^n a_{jk}b_{ki} = \sum_{k=1}^n b'_{ik}a'_{kj}$ et on reconnaît là les coefficients de ${}^tB {}^tA$. Si $A \in \mathcal{M}_n(\mathbb{K})$ est inversible, on a alors :

$$I_n = {}^tI_n = {}^t(AA^{-1}) = {}^t(A^{-1}) {}^tA$$

ce qui signifie que ${}^tA$ est inversible avec $({}^tA)^{-1} = {}^t(A^{-1})$. $\square$

Cette notion de matrice transposée nous sera utile lors de l'étude des formes bilinéaires et quadratiques.

6.6.5 Trace d'une matrice carrée

Définition 6.18. La trace de $A = ((a_{i,j}))_{1 \leq i,j \leq n} \in \mathcal{M}_n(\mathbb{K})$ est le scalaire $\text{Tr}(A) = \sum_{i=1}^n a_{ii}$ (somme des termes diagonaux).

Exemple 6.4 La trace de la matrice identité I_n est $\text{Tr}(I_n) = n$.

Théorème 6.20.

L'application trace est linéaire de $\mathcal{M}_n(\mathbb{K})$ dans $\mathbb{K}$ (c'est une forme linéaire) et on a $\text{Tr}(AB) = \text{Tr}(BA)$ pour toutes matrices A, B dans $\mathcal{M}_n(\mathbb{K})$.

Preuve. Il suffit de vérifier. □

6.7 Systèmes d'équations linéaires

On se donne $A = ((a_{ij}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathcal{M}_{m,n}(\mathbb{K})$, un vecteur $b = (b_i)_{1 \leq i \leq m} \in \mathbb{K}^m$ et on s'intéresse au système linéaire $Ax = b$ d'inconnue $x = (x_i)_{1 \leq i \leq n}$ dans $\mathbb{K}^n$. Un tel système s'écrit :

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

Pour $b = 0$, un tel système a au moins $x = 0$ comme solution. Le système $Ax = 0$ est le système homogène associé au système $Ax = b$.

En notant, pour j compris entre 1 et n , $C_j = (a_{ij})_{1 \leq i \leq m}$ la colonne numéro j de la matrice A , résoudre le système $Ax = b$ revient à trouver tous les scalaires $x_1, \dots, x_n$ tels que $x_1C_1 + x_2C_2 + \cdots + x_nC_n = b$. Dans le cas où le nombre d'inconnues n est strictement supérieur au nombre d'équations m , le système homogène $Ax = 0$ a une infinité de solutions.

Théorème 6.21.

Pour toute matrice $A = ((a_{ij}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in \mathcal{M}_{m,n}(\mathbb{K})$ avec $n > m$, le système homogène $Ax = 0$ a une infinité de solutions.

Preuve. On sait déjà que $x = 0$ est solution. On désigne par $B = \begin{pmatrix} A \\ 0_{n-m,n} \end{pmatrix}$ la matrice carrée d'ordre n ayant pour m premières lignes celles de A , les suivantes étant nulles. Pour toute matrice $B' \in \mathcal{M}_n(\mathbb{K})$ la matrice BB' a sa dernière ligne

nulle et ne peut donc également I_n , ce qui signifie que la matrice B n'est pas inversible. Il existe donc $x \in \mathbb{K}^n \setminus \{0\}$ tel que $Bx = \begin{pmatrix} Ax \\ 0_{n-m,n} \end{pmatrix} = 0$ (théorème 6.16). Le vecteur x est donc solution non nulle de $Ax = 0$ et la droite dirigée par x nous donne une infinité de solutions de ce système linéaire. $\square$

Théorème 6.22.

Soit A dans $\mathcal{M}_n(\mathbb{K})$. Le système $Ax = b$ a une unique solution dans $\mathbb{K}^n$ pour tout vecteur b , si, et seulement si, la matrice A est inversible.

Preuve. Supposons A inversible. Le vecteur $A^{-1}b$ est solution de $Ax = b$ et si x est solution de ce système, on a alors $A^{-1}(Ax) = A^{-1}b$, soit $x = A^{-1}b$. Notre système a bien une unique solution. Réciproquement, supposons que, pour tout $b \in \mathbb{K}^n$ le système $Ax = b$ a unique solution. En désignant par $(e_j)_{1 \leq j \leq n}$ la base canonique de $\mathbb{K}^n$ et, pour tout j compris entre 1 et n , par C_j la solution de $Ax = e_j$, la matrice $A' = (C_1, \dots, C_n)$ est telle que :

$$AA' = (AC_1, \dots, AC_n) = (e_1, \dots, e_n) = I_n$$

ce qui signifie que A est inversible d'inverse A' . $\square$

6.8 Sommes et sommes directes de sous-espaces vectoriels

On se donne pour ce paragraphe un espace vectoriel E .

Définition 6.19. Soient F et G deux sous-espaces vectoriels de E . On dit que E est somme des espaces F et G si l'application :

$$\begin{array}{ccc} \varphi & F \times G & \rightarrow E \\ & (x, y) & \mapsto x + y \end{array}$$

est surjective et on note alors $E = F + G$. Si cette application est bijective, on dit que E est somme directe des espaces F et G et on note $E = F \oplus G$.

Dire que $E = F + G$ signifie donc que l'on peut écrire tout vecteur $x \in E$ sous la forme $x = y + z$, où $y \in F$ et $z \in G$ et dire que $E = F \oplus G$ signifie donc que l'on peut écrire de manière unique tout vecteur $x \in E$ sous la forme $x = y + z$, où $y \in F$ et $z \in G$.

Théorème 6.23.

Soient F et G deux sous-espaces vectoriels de E . On a $E = F \oplus G$ si, et seulement si, $E = F + G$ et $F \cap G = \{0\}$.

Preuve. Supposons que $E = F \oplus G$, on a alors $E = F + G$ et tout $x \in F \cap G$ s'écrit $x = x + 0 = 0 + x$ avec $(x, 0)$ et $(0, x)$ dans $F \times G$, ce qui impose $x = 0$

puisque φ est bijective. Réciproquement supposons que $E = F + G$ et $F \cap G = \{0\}$. Si $x \in E$ s'écrit $x = y + z = y' + z'$, avec y, y' dans F et z, z' dans G , on a alors $y - y' = z' - z \in F \cap G$, donc $y - y' = z - z' = 0$ et $(y, z) = (y', z')$. La somme est donc directe. $\square$

Définition 6.20. On dit que deux sous-espaces vectoriels F et G de E sont *supplémentaires*, si $E = F \oplus G$. On dit aussi que F est un *supplémentaire* de G ou que G est un *supplémentaire* de F dans E .

E est l'unique supplémentaire de $\{0\}$, mais un sous-espace vectoriel F de E distinct de $\{0\}$ et de E admet une infinité de supplémentaires. Il suffit de considérer deux droites de $\mathbb{R}^2$ dirigées par deux vecteurs non colinéaires pour s'en convaincre.

On peut définir la somme ou la somme directe de p sous-espaces de E comme suit.

Définition 6.21. Soient $p \geq 2$ un entier et $F_1, \dots, F_p$ des sous-espaces vectoriels de E . On dit que E est *somme* des espaces $F_1, \dots, F_p$ si l'application :

$$\begin{aligned} \varphi \quad F_1 \times \dots \times F_p &\rightarrow E \\ (x_1, \dots, x_p) &\mapsto x_1 + \dots + x_p \end{aligned}$$

est surjective et on note alors $E = F_1 + \dots + F_p$ ou de manière plus compacte

$E = \sum_{k=1}^p F_k$. Si cette application est bijective, on dit que E est *somme directe*

des espaces $F_1, \dots, F_p$ et on note $E = F_1 \oplus \dots \oplus F_p$ ou $E = \bigoplus_{k=1}^p F_k$.

En fait la somme de deux sous-espaces vectoriels F et G peut se définir par :

$$F + G = \{y + z \mid y \in F \text{ et } z \in G\}$$

Il est facile de vérifier que $F + G$ est un sous-espace vectoriel de E . Ce sous-espace n'est en général pas égal à E .

On peut aussi définir $F + G$ comme le sous-espace vectoriel de E engendré par la réunion $F \cup G$ (qui en général n'est pas un espace vectoriel).

Théorème 6.24.

Si F, G sont deux sous-espaces vectoriels de E , alors la somme $F + G$ est le sous-espace vectoriel de E engendré par $F \cup G$.

Preuve. Dire que $x \in \text{Vect}(F \cup G)$ équivaut à dire qu'il s'écrit $x = \sum_{k=1}^p \lambda_k x_k$ où les x_k sont des éléments de $F \cup G$ et les λ_k des scalaires. En séparant les x_k qui sont dans F de ceux qui sont dans G , cette somme peut s'écrire $x = y + z$ avec $y \in F$ et $z \in G$, ce qui signifie que $x \in F + G$. Réciproquement $x \in F + G$ s'écrit $x = y + z$ avec $(y, z) \in F \times G$ et il est bien dans $\text{Vect}(F \cup G)$. $\square$

De manière un peu plus générale, on a le résultat suivant.

Théorème 6.25.

Si $F_1, \dots, F_p$ sont des sous-espaces vectoriels de E , alors la somme $\sum_{k=1}^p F_k$ est le sous-espace vectoriel de E engendré par $\bigcup_{k=1}^p F_k$.

6.9 Exercices

Exercice 6.1. Soit E un $\mathbb{K}$ -espace vectoriel.

1. Montrer que l'égalité $\lambda x = 0$ où λ est un scalaire et x un vecteur est équivalente à $\lambda = 0$ ou $x = 0$.
2. Montrer que dans un espace vectoriel E , on a $(-1)x = -x$ pour tout vecteur x .

Solution.

1. Pour tout $x \in E$, on a $0 \cdot x + x = 0 \cdot x + 1 \cdot x = (0 + 1)x = 1 \cdot x = x$ et simplifiant par x (ce qui revient à ajouter $-x$ aux deux membres de cette égalité), on aboutit à $0 \cdot x = 0$. Pour tout $\lambda \in \mathbb{K}$, on a $\lambda \cdot 0 = \lambda \cdot (0 + 0) = \lambda \cdot 0 + \lambda \cdot 0$ et simplifiant par $\lambda \cdot 0$, on aboutit à $\lambda \cdot 0 = 0$. Supposons que $\lambda x = 0$. Si $\lambda = 0$ c'est terminé, sinon λ est inversible dans $\mathbb{K}$ et $x = 1 \cdot x = \left(\frac{1}{\lambda}\lambda\right)x = \frac{1}{\lambda}(\lambda x) = \frac{1}{\lambda}0 = 0$.
2. Pour tout vecteur x , on a $x + (-1)x = 1 \cdot x + (-1)x = (1 - 1)x = 0 \cdot x = 0$ et $(-1)x = -x$ puisque l'opposé de x est unique.

Exercice 6.2. Montrer que l'intersection de deux sous-espaces vectoriels d'un espace vectoriel E est un sous-espace vectoriel. Qu'en est-il de la réunion ?

Solution. Soient F, G deux sous-espaces vectoriels de E . L'intersection $H = F \cap G$ contient 0 puisqu'ils sont dans F et G et pour tous x, y dans H , λ, μ dans $\mathbb{K}$, le vecteur $\lambda x + \mu y$ est dans F et G , donc dans H . En définitive, H est un sous-espace vectoriel de E . En général la réunion de deux sous-espaces vectoriels de E n'est pas un sous-espace vectoriel. Par exemple dans $\mathbb{R}^2$, les ensembles F et G définis par $F = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \mid x \in \mathbb{R} \right\}$ (l'axe des x) et $G = \left\{ \begin{pmatrix} 0 \\ y \end{pmatrix} \mid y \in \mathbb{R} \right\}$ (l'axe des y) sont des sous-espaces vectoriels, mais pas $F \cup G$ puisque $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ sont dans cette réunion, mais pas leur somme $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Exercice 6.3. Soient F, G deux sous-espaces vectoriels de E . Montrer que $F \cup G$ est un sous-espace vectoriel de E si, et seulement si, $F \subset G$ ou $G \subset F$.

Solution. Si $F \subset G$ [resp. $G \subset F$], on a alors $F \cup G = G$ [resp. $F \cup G = F$] et c'est un sous-espace vectoriel de E . Réciproquement supposons que $F \cup G$ soit un sous-espace vectoriel de E . Si $F \not\subset G$ et $G \not\subset F$, il existe alors $x \in F \setminus G$ et $y \in G \setminus F$, donc comme x et y sont dans $F \cup G$, il en est de même de $x + y$, mais $x + y \in F$ entraîne $y = (x + y) - x \in F$, ce qui n'est pas et $x + y \in G$ entraîne $x = (x + y) - y \in G$, ce qui n'est pas, il y a donc une impossibilité et on a nécessairement $F \subset G$ ou $G \subset F$.

Exercice 6.4. On désigne par $\mathcal{F}$ l'espace vectoriel des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ et par $\mathcal{P}$ [resp. $\mathcal{I}$] le sous-ensemble de $\mathcal{F}$ formés de toutes les fonctions paires [resp. impaires] de $\mathbb{R}$ dans $\mathbb{R}$.

1. Montrer que $\mathcal{P}$ et $\mathcal{I}$ sont des sous-espaces vectoriels de $\mathcal{F}$.
2. Calculer $\mathcal{P} \cap \mathcal{I}$.
3. Montrer que toute fonction $f \in \mathcal{F}$, s'écrit de manière unique comme somme d'une fonction paire et d'une fonction impaire. Ce résultat se traduit en disant que $\mathcal{F}$ est somme directe des sous-espaces $\mathcal{P}$ et $\mathcal{I}$ et on note $\mathcal{F} = \mathcal{P} \oplus \mathcal{I}$.

Solution.

1. La fonction nulle est à la fois paire et impaire donc dans $\mathcal{P}$ et dans $\mathcal{I}$. Si f, g sont deux fonctions paires [resp. impaires], il en est alors de même de $f + g$ et de λf pour tout réel λ . Les ensembles $\mathcal{P}$ et $\mathcal{I}$ sont donc bien des sous-espaces vectoriels de $\mathcal{F}$.
2. Dire qu'une fonction f est dans $\mathcal{P} \cap \mathcal{I}$ signifie qu'elle est à la fois paire et impaire et donc que pour tout réel x , on a $f(x) = f(-x) = -f(x)$, ce qui revient à dire que $f(x) = 0$. On a donc $\mathcal{P} \cap \mathcal{I} = \{0\}$.
3. Pour $f \in \mathcal{F}$, la fonction g [resp. h] définie sur $\mathbb{R}$ par $g(x) = \frac{f(x) + f(-x)}{2}$ [resp. $h(x) = \frac{f(x) - f(-x)}{2}$] est paire [resp. impaire] et $f = g + h$. Si (g', h') est un autre couple dans $\mathcal{P} \times \mathcal{I}$ tel que $f = g' + h'$, la fonction $g - g' = h' - h$ est dans $\mathcal{P} \cap \mathcal{I}$ donc nulle et $g = g'$, $h = h'$. Une telle écriture est donc unique. Par exemple si f est la fonction exp, les fonctions g et h sont les fonctions hyperboliques ch et sh définies par $\text{ch}(x) = \frac{e^x + e^{-x}}{2}$ et $\text{sh}(x) = \frac{e^x - e^{-x}}{2}$.

Exercice 6.5. Montrer qu'une forme linéaire φ sur E non identiquement nulle est surjective.

Solution. Dire que $\varphi \neq 0$ signifie qu'il existe $x_0 \in E$ tel que $\lambda = \varphi(x_0) \neq 0$. Pour tout scalaire y , on peut alors écrire $y = \frac{y}{\lambda} \lambda = \frac{y}{\lambda} \varphi(x_0) = \varphi\left(\frac{y}{\lambda} x_0\right)$, soit $y = \varphi(x)$ avec $x = \frac{y}{\lambda} x_0 \in E$, ce qui signifie que φ est surjective.

Exercice 6.6. Soit u une application linéaire de E dans E (i. e. un endomorphisme de E). Montrer que $\text{Im}(u) \subset \ker(u) \Leftrightarrow u \circ u = 0$.

Solution. Si $\text{Im}(u) \subset \ker(u)$, on a alors $u(x) \in \ker(u)$ pour tout $x \in E$ et $u(u(x)) = 0$, ce qui signifie que $u \circ u = 0$. Réciproquement si $u \circ u = 0$, pour tout $y = u(x) \in \text{Im}(u)$, on a $u(y) = u(u(x)) = u \circ u(x) = 0$, ce qui signifie que $y \in \ker(u)$ et $\text{Im}(u) \subset \ker(u)$.

Exercice 6.7. Soit $u \in \mathcal{L}(\mathbb{K}^3)$ de matrice $A = \begin{pmatrix} 4 & 2 & -4 \\ -6 & -4 & 6 \\ -1 & -1 & 1 \end{pmatrix}$ dans la base canonique (e_1, e_2, e_3) . Déterminer le noyau u .

Solution. L'image du vecteur $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ est le vecteur :

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 4 & 2 & -4 \\ -6 & -4 & 6 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4x_1 + 2x_2 - 4x_3 \\ -6x_1 - 4x_2 + 6x_3 \\ -x_1 - x_2 + x_3 \end{pmatrix}$$

Dire que $x \in \mathbb{K}^3$ est dans le noyau de u équivaut à dire que ses composantes sont solutions du système linéaire de 3 équations à 3 inconnues :

$$\begin{cases} 4x_1 + 2x_2 - 4x_3 = 0 \\ -6x_1 - 4x_2 + 6x_3 = 0 \\ -x_1 - x_2 + x_3 = 0 \end{cases}$$

L'équation (3) donne $x_3 = x_1 + x_2$ qui reporté dans la première donne $x_2 = 0$ et $x_1 = x_3$. Réciproquement tout vecteur vérifiant ces conditions est solution du système linéaire. Le noyau de u est donc :

$$\ker(u) = \left\{ \begin{pmatrix} x_1 \\ 0 \\ x_1 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = x_1 (e_1 + e_3) \mid x \in \mathbb{K} \right\}$$

c'est donc la droite vectorielle engendré par le vecteur $e_1 + e_3$.

Exercice 6.8. Calculer A^n pour tout $n \in \mathbb{N}$, où $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

Solution. On a $A^0 = I_3$, $A^1 = A$ et $A^2 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$, $A^3 = \begin{pmatrix} 1 & 3 & 6 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$.

En supposant que, pour $n \geq 1$, on a $A^n = \begin{pmatrix} 1 & n & a_n \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}$, où a_n est un entier à déterminer, on a :

$$\begin{aligned} A^{n+1} &= \begin{pmatrix} 1 & n & a_n \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & n+1 & n+a_n+1 \\ 0 & 1 & n+1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & n & a_{n+1} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

avec $a_{n+1} = a_n + n + 1$. La suite $(a_n)_{n \geq 1}$ est donc définie par la relation de récurrence :

$$\begin{cases} a_1 = 1 \\ \forall n \geq 1, a_{n+1} = a_n + n + 1 \end{cases}$$

ce qui donne $a_n = 1 + 2 + \dots + n = \frac{n(n+1)}{2}$ (vérification par récurrence sur $n \geq 1$). On a donc, pour tout $n \geq 0$, $A^n = \begin{pmatrix} 1 & n & \frac{n(n+1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}$.

Exercice 6.9. Montrer que si la matrice $A \in \mathcal{M}_n(\mathbb{K})$ a une colonne [resp. une ligne] nulle, alors elle n'est pas inversible.

Solution. En notant $C_1, \dots, C_n$ [resp. $L_1, \dots, L_n$] les colonne [resp. lignes] de A , on a pour toute matrice $A' \in \mathcal{M}_n(\mathbb{K})$, $A'A = (A'C_1, \dots, A'C_n)$ et pour $C_j = 0$, la colonne j de $A'A$ est nulle, ce qui interdit l'égalité $A'A = I_n$. Pour ce qui est

des lignes, on écrit que $AA' = \begin{pmatrix} L_1 A' \\ \vdots \\ L_n A' \end{pmatrix}$.

Exercice 6.10. On désigne par $(e_i)_{1 \leq i \leq n}$ la base canonique de $\mathbb{K}^n$ avec $n \geq 2$ et pour tout entier j compris entre 1 et n , par M_j la matrice $M_j = (0, \dots, 0, e_1, 0, \dots, 0)$, la colonne e_1 étant en position j . Montrer que pour

tout scalaire λ et tout entier j compris entre 2 et n , la matrice $P_j(\lambda) = I_n + \lambda M_j$ est inversible et déterminer son inverse.

Solution. Soit, pour j et λ fixés, u l'endomorphisme de $\mathbb{K}^n$ canoniquement associé à $P_j(\lambda)$. Il est défini par :

$$u(e_k) = \begin{cases} \lambda e_1 + e_j & \text{si } k = j \\ e_k & \text{si } k \neq j \end{cases}$$

Cet endomorphisme est inversible d'inverse u' défini par :

$$u'(e_k) = \begin{cases} -\lambda e_1 + e_j & \text{si } k = j \\ e_k & \text{si } k \neq j \end{cases}$$

donc $P_j(\lambda)$ est inversible d'inverse $P_j(-\lambda)$.

Exercice 6.11. Montrer que la matrice $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ est inversible et déterminer son inverse.

Solution. Il s'agit de résoudre, pour $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ donné dans $\mathbb{K}^2$, le système :

$$\begin{cases} x_1 + 2x_2 = y_1 \\ 3x_1 + 4x_2 = y_2 \end{cases}$$

Multipliant la première équation par 3 et retranchant la deuxième équation au résultat obtenu, on a $2x_2 = 3y_1 - y_2$, soit $x_2 = \frac{3}{2}y_1 - \frac{1}{2}y_2$ qui reporté dans la première équation donne $x_1 = -y_1 + y_2$. On a donc :

$$\begin{cases} x_1 = -2y_1 + y_2 \\ x_2 = \frac{3}{2}y_1 - \frac{1}{2}y_2 \end{cases}$$

ce qui signifie que A est inversible et que $A^{-1} = \frac{1}{2} \begin{pmatrix} -4 & 2 \\ 3 & -1 \end{pmatrix}$.

Exercice 6.12. Soit $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$. Pour tout $n \in \mathbb{N}^*$, on désigne par A^n la matrice $A \cdot A \cdots A$, ce produit ayant n termes. On convient que $A^0 = I_2$, où $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

1. Calculer A^2 et A^3 .

2. Montrer que pour tout entier $n \geq 1$, la matrice A^n est de la forme

$$A^n = \begin{pmatrix} a_n + (-1)^n & a_n \\ a_n & a_n + (-1)^n \end{pmatrix}, \text{ où } a_n \text{ est un entier.}$$

3. Calculer $A^2 - 2A - 3I_2$.

4. Montrer que A est inversible et calculer A^{-1} .

5. Montrer que pour tout entier $n \geq 2$, il existe un polynôme Q_n et deux entiers α_n et β_n tels que :

$$X^n = Q_n(X) (X^2 - 2X - 3) + \alpha_n X + \beta_n. \quad (6.3)$$

En évaluant (6.3) en -1 et 3 déterminer α_n et β_n .

6. En déduire A^n pour tout $n \geq 2$.

Solution.

1. On a $A^2 = \begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}$, $A^3 = \begin{pmatrix} 13 & 14 \\ 14 & 13 \end{pmatrix}$.

2. C'est vrai pour $n = 1$ avec $a_1 = 2$ et en supposant le résultat acquis pour n , on a :

$$\begin{aligned} A^{n+1} &= \begin{pmatrix} a_n + (-1)^n & a_n \\ a_n & a_n + (-1)^n \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 3a_n + (-1)^n & 3a_n + 2(-1)^n \\ 3a_n + 2(-1)^n & 3a_n + (-1)^n \end{pmatrix} \\ &= \begin{pmatrix} 3a_n + 2(-1)^n + (-1)^{n+1} & 3a_n + 2(-1)^n \\ 3a_n + 2(-1)^n & 3a_n + 2(-1)^n + (-1)^{n+1} \end{pmatrix} \end{aligned}$$

puisque $2(-1)^n + (-1)^{n+1} = (-1)^n (2 - 1) = (-1)^n$.

3. On a $A^2 - 2A - 3I_2 = 0$.

4. On a $A^{-1} = \frac{1}{3} \begin{pmatrix} -1 & 2 \\ 2 & -1 \end{pmatrix}$ par calcul direct ou avec la question précédente,
 $A^{-1} = \frac{1}{3} (A - 2I_2)$.

5. Pour $n = 2$, on a $X^2 = (X^2 - 2X - 3) + 2X + 3$, donc $Q_2 = 1$ et $(\alpha_2, \beta_2) = (2, 3)$. Supposant le résultat acquis pour $n \geq 2$, on a :

$$\begin{aligned} X^{n+1} &= XQ_n(X) (X^2 - 2X - 3) + \alpha_n X^2 + \beta_n X \\ &= XQ_n(X) (X^2 - 2X - 3) + \alpha_n ((X^2 - 2X - 3) + 2X + 3) + \beta_n X \\ &= (\alpha_n + XQ_n)(X) (X^2 - 2X - 3) + (2\alpha_n + \beta_n) X + 3\alpha_n \end{aligned}$$

soit le résultat au rang $n+1$. Les racines de $X^2 - 2X - 3$ sont -1 et 3 , donc $-\alpha_n + \beta_n = (-1)^n$, $3\alpha_n + \beta_n = 3^n$ et résolvant ce système, on a $\alpha_n = \frac{3^n - (-1)^n}{4}$ et $\beta_n = \frac{3^n + 3(-1)^n}{4}$.

6. On a :

$$\begin{aligned}
 A^n &= \alpha_n A + \beta_n I_2 = \frac{1}{4} ((3^n - (-1)^n) A + (3^n + 3(-1)^n) I_2) \\
 &= \frac{1}{4} \begin{pmatrix} 3^n - (-1)^n + 3^n + 3(-1)^n & 2(3^n - (-1)^n) \\ 2(3^n - (-1)^n) & 3^n - (-1)^n + 3^n + 3(-1)^n \end{pmatrix} \\
 &= \frac{1}{2} \begin{pmatrix} 3^n + (-1)^n & 3^n - (-1)^n \\ 3^n - (-1)^n & 3^n + (-1)^n \end{pmatrix} = \begin{pmatrix} a_n + (-1)^n & a_n \\ a_n & a_n + (-1)^n \end{pmatrix}
 \end{aligned}$$

avec $a_n = \frac{3^n - (-1)^n}{2}$ et $a_n + (-1)^n = \frac{3^n + (-1)^n}{2}$.

Exercice 6.13. Montrer que si φ est une forme linéaire non nulle sur E , il existe alors un vecteur non nul a dans E tel que $E = \ker(\varphi) \oplus \mathbb{K}a$.

Solution. La forme linéaire φ étant non nulle, on peut trouver un vecteur a dans E tel que $\varphi(a) \neq 0$. Ce vecteur a est nécessairement non nul. Pour tout vecteur x dans E , le vecteur $h = x - \frac{\varphi(x)}{\varphi(a)}a$ est dans le noyau de φ et en écrivant que $x = h + \frac{\varphi(x)}{\varphi(a)}a$ on déduit que $E = \ker(\varphi) + \mathbb{K}a$. Si x est dans $\ker(\varphi) \cap \mathbb{K}a$ on a alors $x = \lambda a$ et $\lambda\varphi(a) = \varphi(x) = 0$ avec $\varphi(a) \neq 0$ ce qui entraîne $\lambda = 0$ et $x = 0$. On a donc $\ker(\varphi) \cap \mathbb{K}a = \{0\}$ et $E = \ker(\varphi) \oplus \mathbb{K}a$.

Chapitre 7

Espaces vectoriels réels ou complexes de dimension finie

On note toujours $\mathbb{K}$ le corps de réels ou des complexes.

7.1 Systèmes libres, systèmes générateurs et bases

Nous avons déjà rencontré et utilisé la base canonique de $\mathbb{K}^n$. Nous allons donner une définition précise de cette notion dans le cadre des espaces vectoriels réels ou complexes, ce qui nous mènera à la notion de dimension.

On se donne un espace vectoriel E .

Définition 7.1. Soit $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une famille (ou un système) de n vecteurs de E , où n est un entier naturel non nul. On dit que $\mathcal{B}$ est :

- une famille libre, ou que les vecteurs $e_1, \dots, e_n$ sont linéairement indépendants, si pour toute famille $(\lambda_i)_{1 \leq i \leq n}$ l'égalité $\sum_{i=1}^n \lambda_i e_i = 0$ est réalisée si, et seulement si, tous les λ_i sont nuls ;
- une famille liée, si ce n'est pas une famille libre (i. e. il existe des scalaires $\lambda_1, \dots, \lambda_n$ non tous nuls tels que $\sum_{i=1}^n \lambda_i e_i = 0$) ;
- une famille génératrice, si pour tout vecteur $x \in E$, il existe des scalaires $\lambda_1, \dots, \lambda_n$ tels que $x = \sum_{i=1}^n \lambda_i e_i$;
- une base de E , si elle est libre et génératrice.

Dire que $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ est une famille génératrice de E équivaut à dire que $E = \text{vect}(\mathcal{B})$ (l'espace vectoriel engendré par $\mathcal{B}$).

Avec le théorème qui suit, on résume quelques propriétés des familles libres ou liées.

Théorème 7.1.

Soit $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une famille de n vecteurs de E , où n est un entier naturel non nul.

1. Si $n = 1$, dire que $\mathcal{B}$ est libre [resp. liée] signifie que $e_1 \neq 0$ [resp. $e_1 = 0$].
2. Si $\mathcal{B}$ est libre, alors tous les vecteurs e_i sont non nuls.
3. Si l'un des e_i est nul, alors $\mathcal{B}$ est liée.
4. Si $\mathcal{B}$ contient une famille liée, elle est elle-même liée.
5. Si $\mathcal{B}$ est contenue dans une partie libre, elle est elle-même libre.
6. Si $\mathcal{B}$ est liée, l'un des vecteurs e_j est combinaison linéaire des autres.
7. Si $\mathcal{B}$ est une base de E , alors tout vecteur x de E s'écrit de manière unique comme combinaison linéaire des vecteurs $e_1, \dots, e_n$.

Preuve. Résultent des définitions. □

L'utilisation des déterminants, définis pour l'instant dans le seul cas des matrices d'ordre 2, nous donne un moyen élémentaire de vérifier que deux vecteurs de $\mathbb{K}^2$ sont linéairement indépendants.

Théorème 7.2.

Les vecteurs $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ et $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ sont linéairement indépendants si, et seulement si, on a $\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} \neq 0$.

Preuve. Il revient au même de montrer que x et y sont liés si, et seulement si, $\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = 0$. Supposons le système (x, y) lié. On a alors $y = \lambda x$ ou $x = \lambda y$ pour un scalaire λ et, par exemple dans le premier cas :

$$\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = \begin{vmatrix} x_1 & \lambda x_1 \\ x_2 & \lambda x_2 \end{vmatrix} = \lambda (x_1 x_2 - x_1 x_2) = 0.$$

Réciproquement, on suppose que $\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = 0$. Si $x = 0$ ou $y = 0$, le système (x, y) est alors lié. Si x et y sont non nuls, en supposant que y_2 est non nul, l'égalité $\begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = x_1 y_2 - y_1 x_2 = 0$ entraîne (c'est même équivalent) :

$$y_2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - x_2 \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

avec $(y_2, -x_2) \neq (0, 0)$, ce qui signifie que x et y sont liés. Si $y_2 = 0$, on a alors $y_1 \neq 0$ et on écrit que l'égalité $x_1 y_2 - x_2 y_1 = 0$ entraîne :

$$x_1 \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - y_1 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

avec $(x_1, -y_1) \neq (0, 0)$. $\square$

Si $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ est une base de E , alors tout vecteur $x \in E$ s'écrit $x = \sum_{i=1}^n \lambda_i e_i$

et les scalaires λ_i qui sont uniquement déterminés sont appelés les composantes ou coordonnées de x dans la base $\mathcal{B}$. Réciproquement une telle famille $\mathcal{B}$ vérifiant cette propriété est une base de E .

La base canonique de $\mathbb{K}^n$ est bien une base au sens de la définition qu'on a donné.

Exemple 7.1 Dans l'espace $\mathbb{R}_n[x]$ [resp. $\mathbb{C}_n[x]$] des fonctions polynomiales réelles [resp. complexes] de degré au plus égal à n , la famille de polynômes $(1, x, \dots, x^n)$ est une base puisque tout polynôme dans $\mathbb{R}_n[x]$ [resp. $\mathbb{C}_n[x]$] s'écrit sous la forme

$P = \sum_{k=0}^n a_k x^k$, les réels [resp. complexes] a_k étant uniquement déterminés. On dit que cette base est la base canonique de $\mathbb{R}_n[x]$ [resp. $\mathbb{C}_n[x]$].

On se limitera dans ce chapitre aux familles libres ou génératrices qui sont finies. Mais en réalité, on est rapidement amené à considérer des familles qui peuvent être infinies. On donne donc les définitions suivantes (qui ne seront pas utilisées au niveau élémentaire où se situe ce cours).

Définition 7.2. Soit $\mathcal{B} = (e_i)_{i \in I}$ une famille de vecteurs de E , où I est un ensemble non vide quelconque (fini ou infini) d'indices. On dit que $\mathcal{B}$ est :

- une famille libre, ou que les vecteurs e_i , pour $i \in I$ sont linéairement indépendants, si toute sous-famille finie de $\mathcal{B}$ est libre, ce qui signifie que pour tout sous-ensemble non vide et fini J de I , une combinaison linéaire $\sum_{j \in J} \lambda_j e_j$, où les λ_j pour $j \in J$ sont des scalaires, est nulle si, et seulement si, tous ces λ_j sont nuls ;
- une famille liée, si ce n'est pas une famille libre (i. e. il existe une partie fini J de I et des scalaires λ_j où j décrit J qui sont non tous nuls tels que $\sum_{j \in J} \lambda_j e_j = 0$) ;
- une famille génératrice, si l'espace vectoriel engendré par $\mathcal{B}$ est l'espace E tout entier, ce qui signifie que pour tout vecteur $x \in E$, il existe une partie finie J de I et des scalaires λ_j où j décrit J tels que $x = \sum_{j \in J} \lambda_j e_j$;
- une base de E , si elle est libre et génératrice.

7.2 Espaces vectoriels de dimension finie

Le théorème qui suit est essentiel pour définir la notion de dimension finie.

Théorème 7.3.

Si un espace vectoriel E admet une famille génératrice $\mathcal{B}$ formée de $n \geq 1$ éléments, alors toute famille libre dans E a au plus n éléments (ce qui équivaut à dire qu'un système de plus de $n + 1$ vecteurs est lié).

Preuve. Soit $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une famille génératrice de E . Si $n = 1$, alors pour tout couple (x, y) de vecteurs non nuls de E on peut trouver deux scalaires non nuls λ et μ tels que $x = \lambda e_1$ et $y = \mu e_1$ et on a la combinaison linéaire nulle $\mu x - \lambda y = 0$ avec μ et $-\lambda$ non nuls, ce qui signifie que le système (x, y) est lié. Il ne peut donc exister de famille libre à 2 éléments dans E et a fortiori il ne peut en exister à plus de 2 éléments. Supposons le résultat acquis au rang $n - 1 \geq 1$, c'est-à-dire que dans tout espace vectoriel F admettant un système générateur de $n - 1$ vecteurs une famille de plus de n vecteurs est liée. Supposons que E soit un espace vectoriel admettant une famille génératrice à n éléments. Supposons qu'il existe une famille libre ayant $m \geq n + 1$ éléments. On peut extraire de cette famille une famille libre à $n + 1$ éléments puisque toute sous-famille d'une famille libre est libre. Soit $\mathcal{L} = (f_i)_{1 \leq i \leq n+1}$ une telle famille. Comme $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ est génératrice, il existe des scalaires a_{ij} tels que :

$$\begin{cases} f_1 = a_{11}e_1 + \cdots + a_{1n}e_n \\ \vdots \\ f_{n+1} = a_{n+1,1}e_1 + \cdots + a_{n+1,n}e_n \end{cases}$$

Si tous les $a_{i,n}$ sont nuls alors les f_i sont dans l'espace vectoriel F engendré par les $n - 1$ vecteurs $e_1, \dots, e_{n-1}$ et en conséquence liés (hypothèse de récurrence), en contradiction avec $\mathcal{L}$ libre. Il existe donc un indice i compris entre 1 et $n + 1$ tel que $a_{i,n} \neq 0$ et en changeant au besoin la numérotation des éléments de $\mathcal{L}$ on peut supposer que $i = n + 1$. Les n vecteurs :

$$\begin{cases} g_1 = a_{n+1,n}f_1 - a_{1n}f_{n+1} \\ \vdots \\ g_n = a_{n+1,n}f_1 - a_{nn}f_{n+1} \end{cases}$$

sont dans l'espace vectoriel F engendré par les $n - 1$ vecteurs $e_1, \dots, e_{n-1}$ (on a annulé les composantes en e_{n+1}) et en conséquence liés (hypothèse de récurrence), c'est-à-dire qu'il existe des scalaires $\lambda_1, \dots, \lambda_n$ non tous nuls tels que $\lambda_1 g_1 + \cdots + \lambda_n g_n = 0$, ce qui entraîne que :

$$a_{n+1,n}(\lambda_1 f_1 + \cdots + \lambda_n f_n) - (\lambda_1 a_{1n} + \cdots + \lambda_n a_{nn})f_{n+1} = 0$$

les scalaires $a_{n+1,n}\lambda_1, \dots, a_{n+1,n}\lambda_n$ n'étant pas tous nuls, ce qui nous dit encore que les f_i sont liés et est en contradiction avec $\mathcal{L}$ libre. Il est donc impossible de trouver un tel système $\mathcal{L}$ libre. $\square$

Définition 7.3. On dit qu'un espace vectoriel est de dimension finie s'il est réduit à $\{0\}$ ou s'il est différent de $\{0\}$ et admet une base formée d'un nombre fini de vecteurs. Dans le cas contraire, on dit qu'il est de dimension infinie.

On déduit alors du théorème précédent le suivant.

Théorème 7.4.

Si E un espace vectoriel non réduit à $\{0\}$ et de dimension finie, alors toutes les bases ont le même nombre d'éléments.

Preuve. Si $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ et $\mathcal{B}' = (e'_i)_{1 \leq i \leq n'}$ sont deux bases de l'espace vectoriel E , ce sont alors deux familles génératrices et libres et le théorème précédent nous dit que $n' \leq n$ et $n \leq n'$, soit $n = n'$. $\square$

On peut alors donner la définition suivante.

Définition 7.4. Si E est un espace vectoriel non réduit à $\{0\}$ et de dimension finie, alors sa dimension est le nombre de l'une quelconque de ses bases. On note $\dim_{\mathbb{K}}(E)$ (ou simplement $\dim(E)$) cette dimension.

Par convention, on dira que l'espace vectoriel $\{0\}$ est de dimension 0. Un espace vectoriel E est donc de dimension 0 si, et seulement si, il est réduit à $\{0\}$.

Dans le cas général on peut montrer, mais cela dépasse le niveau de ce cours d'introduction, que tout espace vectoriel admet une base (finie ou infinie).

On appelle droite tout espace vectoriel de dimension 1 et plan tout espace vectoriel de dimension 2.

Si $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ est une base E , on a alors $E = \text{vect}(\mathcal{B})$.

Exemples 7.1

1. Bien entendu l'espace $\mathbb{K}^n$ est de dimension n .
2. L'ensemble $\mathbb{C}$ des nombres complexes est un espace vectoriel réel de dimension 2 et un espace vectoriel complexe de dimension 1.
3. Pour tout entier naturel n , l'espace $\mathbb{K}_n[x]$ des fonctions polynomiales de degré au plus égal à n est de dimension $n + 1$.
4. Pour tous entiers naturels non nuls n et m , l'espace $\mathcal{M}_{m,n}(\mathbb{K})$ des matrices à m lignes et n colonnes est un espace vectoriel de dimension $m \cdot n$. En particulier l'espace $\mathcal{M}_n(\mathbb{K})$ des matrices carrées d'ordre n est de dimension n^2 . La base canonique de $\mathcal{M}_{m,n}(\mathbb{K})$ est $(E_{i,j})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ où $E_{i,j}$ est la matrice dont tous les coefficients sont nuls sauf celui en position (i, j) .
5. L'espace des fonctions définies sur un intervalle réel I et à valeurs réelles est de dimension infinie (l'exercice 7.2 nous montre qu'on peut trouver des familles libres ayant une infinité d'éléments). Il admet des bases, mais il n'est pas possible d'en expliciter une.

Le théorème qui suit explique l'importance de l'espace vectoriel $\mathbb{K}^n$.

Théorème 7.5.

Un espace vectoriel de dimension n est isomorphe à $\mathbb{K}^n$.

Preuve. C'est le choix d'une base $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ de E qui nous permet de définir un isomorphisme de E sur $\mathbb{K}^n$. En effet l'unicité de l'écriture de tout vecteur x de

E sous la forme $x = \sum_{k=1}^n \lambda_k e_k$ se traduit en disant que l'application :

$$\begin{array}{ccc} E & \rightarrow & \mathbb{K}^n \\ x = \sum_{k=1}^n \lambda_k e_k & \mapsto & (\lambda_1, \lambda_2, \dots, \lambda_n) \end{array}$$

est bijective et il est facile de vérifier que cette application est linéaire. $\square$

Théorème 7.6.

Soit E un espace vectoriel de dimension $n \geq 1$.

1. *Une famille libre dans E a au plus n élément et c'est une base si, et seulement si, elle a exactement n éléments.*
2. *Une famille génératrice dans E a au moins n élément et c'est une base si, et seulement si, elle a exactement n éléments.*

Preuve. Le cas $n = 1$ est laissé au lecteur et on suppose que $n \geq 2$.

1. Le théorème 7.3 nous dit qu'une famille libre dans E a au plus n élément et si c'est une base, elle a obligatoirement n éléments. Il reste à montrer qu'une famille libre de n éléments est une base. Pour ce faire il suffit de montrer qu'elle est génératrice. Notons $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une telle famille libre. Pour tout vecteur $x \in E$ la famille $\mathcal{B} \cup \{x\}$ est liée puisque formée de $n+1$ éléments, il existe donc des scalaires $\lambda, \lambda_1, \dots, \lambda_n$ non tous nuls tels que $\lambda x + \sum_{i=1}^n \lambda_i e_i = 0$. Si $\lambda = 0$, on

a alors $\sum_{i=1}^n \lambda_i e_i = 0$ et tous les λ_i sont nuls puisque $\mathcal{B}$ est libre, ce qui n'est pas

possible. On a donc $\lambda \neq 0$ et $x = -\sum_{i=1}^n \frac{\lambda_i}{\lambda} e_i$. On a donc ainsi montré que $\mathcal{B}$ est génératrice et que c'est une base.

2. Le théorème 7.3 nous dit qu'une famille génératrice dans E a au moins n élément et si c'est une base, elle a obligatoirement n éléments. Il reste à montrer qu'une famille génératrice de n éléments est une base. Pour ce faire il suffit de montrer qu'elle est libre. Notons $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une telle famille génératrice. Si cette famille est liée, l'un des e_i , disons e_n , est combinaison linéaire des autres et la famille $(e_i)_{1 \leq i \leq n-1}$ est génératrice, ce qui est en contradiction avec le théorème 7.3. La famille $\mathcal{B}$ est donc génératrice et c'est une base de E . $\square$

On retient de ces résultats, que pour montrer qu'une famille finie $\mathcal{B}$ de vecteurs est une base de E , on peut procéder comme suit :

- si on ne connaît pas la dimension de E , on montre que $\mathcal{B}$ est génératrice et libre ;
- si on sait que E est de dimension n , on vérifie que $\mathcal{B}$ a exactement n éléments et on montre que $\mathcal{B}$ est libre ou qu'elle est génératrice (il est inutile de montrer les deux points).

On a défini un espace vectoriel de dimension finie comme un espace vectoriel admettant une base finie. Le théorème qui suit nous dit qu'on peut aussi définir un espace vectoriel de dimension finie comme un espace vectoriel admettant une famille génératrice finie.

Théorème 7.7.

Soit E un espace vectoriel admettant une famille génératrice finie. De cette famille on peut extraire une base et E est de dimension finie.

Preuve. Soit $\mathcal{G} = (u_i)_{1 \leq i \leq p}$ une famille génératrice de E . Si cette famille est libre, elle constitue alors une base de E et E est de dimension p . Sinon, cette famille est liée et l'un de ses éléments, disons u_p est combinaison linéaire des autres (en changeant la numérotation des éléments de $\mathcal{G}$ on peut toujours se ramener à ce cas de figure), ce qui implique que la famille $\mathcal{G}' = (u_i)_{1 \leq i \leq p-1}$ est encore génératrice. Si cette famille est libre, c'est alors une base et E est de dimension finie, sinon on recommence. En un nombre fini de telles opérations on construit ainsi une base de E formée de $n \leq p$ éléments. $\square$

Théorème 7.8. Base incomplète

Soit E un espace vectoriel de dimension $n \geq 1$. Toute famille libre à p éléments dans E (nécessairement $1 \leq p \leq n$) peut se compléter en une base.

Preuve. Soient $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une base de E et $\mathcal{L} = (u_i)_{1 \leq i \leq p}$ une famille libre dans E . On sait déjà que $p \leq n$. Si $p = n$, $\mathcal{L}$ est alors une base. Supposons que $p < n$. Il existe alors un vecteur e_k dans $\mathcal{B}$ tel que $\mathcal{L}' = \mathcal{L} \cup \{e_k\}$ soit libre. En effet si un tel système est lié pour tout entier k compris entre 1 et n , il existe alors, pour chaque entier k , des scalaires $\lambda_1, \dots, \lambda_p, \lambda_{p+1}$ non tous nuls tels que

$$\sum_{i=1}^p \lambda_i u_i + \lambda_{p+1} e_k = 0. \text{ Si } \lambda_{p+1} = 0, \text{ on a alors } \sum_{i=1}^p \lambda_i u_i = 0 \text{ et tous les } \lambda_i \text{ sont nuls}$$

puisque $\mathcal{L}$ est libre. On a donc $\lambda_{p+1} \neq 0$ et $e_k = -\sum_{i=1}^p \frac{\lambda_i}{\lambda_{p+1}} u_i$. En conséquence,

tous les vecteurs de la base $\mathcal{B}$ sont combinaisons linéaires des éléments de $\mathcal{L}$ et $\mathcal{L}$ est alors générateur de E , ce qui est impossible pour $p < n$. Le système $\mathcal{L}'$ est donc libre. Si $p+1 = n$, c'est une base et sinon on recommence. On arrive ainsi à compléter $\mathcal{L}$ en une base de E au bout d'un nombre fini d'opérations. $\square$

Les deux corollaires qui suivent sont des résultats importants à retenir.

Corollaire 7.1. *Soit E un espace vectoriel de dimension n . Tout sous-espace vectoriel F de E est de dimension finie $m \leq n$ et $m = n$ si, et seulement si, $F = E$.*

Preuve. Si $F = \{0\}$, il est alors de dimension $0 \leq n$ et $n = 0$ équivaut à $F = E$. On suppose que $F \neq \{0\}$ (donc $n \geq 1$). Montrons tout d'abord que F est de dimension finie. Comme $n + 1$ vecteurs de F sont nécessairement liés, on peut définir l'entier m comme le plus grand entier pour lequel on peut trouver m vecteurs de F linéairement indépendants. On a $m \geq 1$ puisque $F \neq \{0\}$ et $m \leq n$ d'après le théorème 7.3. Soient donc $(f_i)_{1 \leq i \leq m}$ une famille libre dans F . Pour tout vecteur $x \in F$, la famille $(f_1, \dots, f_m, x)$ est liée et x est combinaison linéaire des f_i puisque $(f_i)_{1 \leq i \leq m}$ est libre. La famille $(f_i)_{1 \leq i \leq m}$ est donc une base de F et cet espace est de dimension finie $m \leq n$. Si $m = n$, une base de F est aussi une base de E et $F = E$. La réciproque est évidente. $\square$

Corollaire 7.2. *Tout sous-espace-vectoriel F d'un espace vectoriel E de dimension finie admet des supplémentaires et pour tout supplémentaire G de F dans E , on :*

$$\dim(E) = \dim(F) + \dim(G) \quad (7.1)$$

Preuve. On suppose que F est un sous-espace vectoriel strict de E , c'est-à-dire que $F \neq \{0\}$ et $F \neq E$. On a alors $1 \leq p = \dim(F) \leq n - 1$. Une base $\mathcal{L} = (u_i)_{1 \leq i \leq p}$ de F se complète en une base $\mathcal{B} = (u_i)_{1 \leq i \leq n}$ et on vérifie facilement que le sous-espace vectoriel G de E engendré par $\mathcal{L}' = (u_i)_{p+1 \leq i \leq n}$ est un supplémentaire de F . Pour cet espace G , on a $\dim(G) = n - p = \dim(E) - \dim(F)$. Réciproquement si $E = F \oplus G$, on vérifie facilement que la réunion d'une base de F et d'une base de G nous fournit une base de E , ce qui implique que $\dim(E) = \dim(F) + \dim(G)$. $\square$

De manière plus générale, on peut montrer que si E est un espace vectoriel de dimension finie ou non, alors tout sous-espace vectoriel de E admet un supplémentaire dans E .

L'égalité (7.1) pour $E = F \oplus G$ peut se généraliser.

Théorème 7.9.

Soit E un espace vectoriel de dimension n . Si E est somme directe de $p \geq 2$ sous-espaces stricts $F_1, \dots, F_p$, soit $E = \bigoplus_{k=1}^p F_k$, alors en désignant, pour tout k compris entre 1 et p , par $\mathcal{B}_k$ une base de F_k , la famille $\bigcup_{k=1}^p \mathcal{B}_k$ est une base de E et $\dim(E) = \sum_{k=1}^p \dim(F_k)$.

Preuve. On vient de voir que le résultat est vrai pour $p = 2$ et une récurrence nous montre qu'il est vrai pour tout $p \geq 2$. $\square$

Dans le cas de la somme, non nécessairement directe, de deux sous-espaces d'un espace de dimension finie, on a le résultat suivant.

Théorème 7.10.

Soient E un espace vectoriel de dimension finie et F, G deux sous-espaces vectoriels de E . On a $\dim(F + G) + \dim(F \cap G) = \dim(F) + \dim(G)$.

Preuve. Comme $F \cap G$ est un sous-espace de G , il admet un supplémentaire H dans G , soit $G = (F \cap G) \oplus H$. Ce sous-espace H de G est aussi un sous-espace de $F + G$. En fait H est un supplémentaire de F dans $F + G$. En effet, on a $F + H \subset F + G$ puisque $H \subset G$ et tout $x \in F + G$ s'écrit $x = y + z$ avec $y \in F$ et $z \in G = (F \cap G) \oplus H$, donc $z = z_1 + z_2$ avec $z_1 \in F \cap G \subset F$ et $z_2 \in H$, ce qui donne $x = (y + z_1) + z_2 \in F + H$. On a donc $F + G \subset F + H$ et l'égalité $F + G = F + H$. Si maintenant x est dans $F \cap H$, il est dans $F \cap G$ puisque $H \subset G$, donc dans $(F \cap G) \cap H = \{0\}$. On a donc $F \cap H = \{0\}$ et $F + G = F \oplus H$. Il en résulte que $\dim(F + G) = \dim(F) + \dim(H) = \dim(F) + \dim(G) - \dim(F \cap G)$.

□

Ce résultat a pour conséquence le résultat suivant très utile pour montrer qu'un espace est somme directe de deux sous-espaces.

Théorème 7.11.

Soient E un espace vectoriel de dimension finie et F, G deux sous-espaces vectoriels de E . On a :

$$\begin{aligned} (E = F \oplus G) &\Leftrightarrow (E = F + G \text{ et } \dim(E) = \dim(F) + \dim(G)) \\ &\Leftrightarrow (F \cap G = \{0\} \text{ et } \dim(E) = \dim(F) + \dim(G)) \end{aligned}$$

Preuve. On sait déjà que si $E = F \oplus G$, alors $E = F + G$, $F \cap G = \{0\}$ et $\dim(E) = \dim(F) + \dim(G)$. Supposons que $E = F + G$ et $\dim(E) = \dim(F) + \dim(G)$. Le théorème précédent nous dit alors que $\dim(F \cap G) = 0$, soit que $F \cap G = \{0\}$ et on a $E = F \oplus G$. De même si $F \cap G = \{0\}$ et $\dim(E) = \dim(F) + \dim(G)$, on a alors $F + G = F \oplus G$ et cet espace a même dimension que E , donc $E = F \oplus G$.

□

Les propriétés des applications linéaires injectives, surjectives ou bijectives décrites par le théorème qui suit sont importantes.

Théorème 7.12.

Soient E, F deux espaces vectoriels de dimension finie et u une application linéaire de E dans F .

1. *Si u est injective, elle transforme alors tout système libre de E en un système libre de F et $\dim(E) \leq \dim(F)$.*
2. *Si u est surjective, elle transforme alors tout système générateur de E en un système générateur de F et $\dim(F) \leq \dim(E)$.*

3. Si u est bijective, elle transforme alors toute base de E en une base de F et $\dim(E) = \dim(F)$ (deux espaces vectoriels de dimension finie isomorphes ont la même dimension).
4. Si $\dim(E) = \dim(F)$, alors :

$$(u \text{ bijective}) \Leftrightarrow (u \text{ injective}) \Leftrightarrow (u \text{ surjective})$$

Preuve.

1. Soit $\mathcal{L} = (x_i)_{1 \leq i \leq p}$ un système libre dans E . Si $\sum_{i=1}^p \lambda_i u(x_i) = 0$, on a alors du fait de la linéarité de u , $u\left(\sum_{i=1}^p \lambda_i x_i\right) = 0$, ce qui signifie que $\sum_{i=1}^p \lambda_i x_i$ est dans le noyau de u , donc nul puisque u est injective, ce qui équivaut à la nullité de tous les coefficients λ_i puisque $\mathcal{L}$ est libre. La famille $u(\mathcal{L})$ est donc libre. Prenant pour $\mathcal{L}$ une base de E , elle est formée de $n = \dim(E)$ éléments et $u(\mathcal{L})$ est libre à n éléments dans F , donc $n \leq m = \dim(F)$.
2. Soit $\mathcal{L} = (x_i)_{1 \leq i \leq p}$ un système générateur de E . Comme u est surjective tout vecteur y de F s'écrit $y = u(x)$ avec x dans E qui s'écrit $x = \sum_{i=1}^p \lambda_i x_i$, ce qui donne $y = \sum_{i=1}^p \lambda_i u(x_i)$. Le système $u(\mathcal{L})$ est donc générateur de F . Prenant pour $\mathcal{L}$ une base de E , elle est formée de n éléments et $u(\mathcal{L})$ est générateur de F à n éléments, donc $n \geq m$.
3. et 4. Résultent des deux points précédents.

□

Le théorème 7.5 et le point 3. du théorème précédent nous disent que deux espaces vectoriels de dimension finie ont même dimension si, et seulement si, ils sont isomorphes.

En vue de généraliser le théorème 7.11, on utilisera le résultat suivant.

Lemme 7.1 Si $F_1, \dots, F_p$ sont des espaces vectoriels de dimension finie, il en est de même de l'espace produit $F = F_1 \times \dots \times F_p$ et on a $\dim(F) = \sum_{k=1}^p \dim(F_k)$.

Preuve. Pour $p = 1$, il n'y a rien à montrer. En procédant par récurrence sur $p \geq 2$, il suffit de montrer le résultat pour $p = 2$. Pour ce faire, on vérifie que si $\mathcal{B}_1 = (e_i)_{1 \leq i \leq n}$ est une base de F_1 et $\mathcal{B}_2 = (f_j)_{1 \leq j \leq m}$ une base de F_2 , alors la famille $\mathcal{B} = \{(e_i, 0) \mid 1 \leq i \leq n\} \cup \{(0, f_j) \mid 1 \leq j \leq m\}$ est une base de $F = F_1 \times F_2$. En effet tout vecteur $z = (x, y)$ dans F s'écrit :

$$z = \left(\sum_{i=1}^n \lambda_i e_i, \sum_{j=1}^m \mu_j f_j \right) = \sum_{i=1}^n \lambda_i (e_i, 0) + \sum_{j=1}^m \mu_j (0, f_j)$$

donc $\mathcal{B}$ engendre F et l'égalité $\sum_{i=1}^n \lambda_i (e_i, 0) + \sum_{j=1}^m \mu_j (0, f_j) = 0$ est équivalente à $\left(\sum_{i=1}^n \lambda_i e_i, \sum_{j=1}^m \mu_j f_j \right) = (0, 0)$, soit à $\sum_{i=1}^n \lambda_i e_i = 0$ et $\sum_{j=1}^m \mu_j f_j = 0$ qui impose $\lambda_i = 0$ pour tout i compris entre 1 et n et $\mu_j = 0$ pour tout j compris entre 1 et m . La famille $\mathcal{B}$ est donc libre et c'est une base de F . L'espace F est donc de dimension finie égale au nombre d'éléments de $\mathcal{B}$, soit à $n + m$. $\square$

Théorème 7.13.

Si $F_1, \dots, F_p$ sont des sous-espaces vectoriels d'un espace vectoriel E de dimension finie, on a alors $E = \bigoplus_{k=1}^n F_k$ si, et seulement si, $E = \sum_{k=1}^p F_k$ et $\dim(E) = \sum_{k=1}^p \dim(F_k)$.

Preuve. On sait déjà que la condition est nécessaire. Dire que $E = \sum_{k=1}^p F_k$, équivaut à dire que l'application linéaire :

$$\begin{array}{ccc} \varphi : F = F_1 \times \dots \times F_p & \rightarrow & E \\ (x_1, \dots, x_p) & \mapsto & x_1 + \dots + x_p \end{array}$$

est surjective et si de plus $\dim(F) = \dim(E)$, cette application est en fait une bijection ce qui signifie que $E = \bigoplus_{k=1}^n F_k$. $\square$

7.3 Rang d'un système de vecteurs ou d'une application linéaire

On désigne par E un espace vectoriel de dimension finie.

On rappelle que le sous-espace vectoriel $F = \text{Vect}\{x_1, \dots, x_p\}$ de E engendré par des vecteurs $x_1, \dots, x_p$ de E est l'ensemble de toutes les combinaisons linéaires de ces vecteurs, soit $F = \left\{ x = \sum_{k=1}^p \lambda_k x_k \mid (\lambda_1, \lambda_2, \dots, \lambda_p) \in \mathbb{K}^p \right\}$.

Définition 7.5. *Le rang de la famille $\{x_1, \dots, x_p\}$ de vecteurs de E est la dimension de l'espace vectoriel engendré par ces vecteurs. On le note $\text{rg}(x_1, \dots, x_p)$.*

Théorème 7.14.

Le rang d'une famille $\{x_1, \dots, x_p\}$ de p vecteurs de E est au maximum égal à p et ce rang vaut p si, et seulement si, cette famille est libre.

Preuve. Si le système $\mathcal{L} = \{x_1, \dots, x_p\}$ est libre, il constitue alors une base de $F = \text{Vect}(\mathcal{L})$ et $\text{rg}(\mathcal{L}) = \dim(F) = p$. Réciproquement si le rang vaut p , la famille $\mathcal{L}$ est génératrice de F avec p éléments, c'est donc une base de F et en conséquence une famille libre. $\square$

Si E est de dimension n , le rang d'une famille de vecteurs de E est au plus égal à n . Si ce rang vaut n , on a alors $\text{Vect}\{x_1, \dots, x_p\} = E$ et $\{x_1, \dots, x_p\}$ est un système générateur de E . Dans le cas où $p = n$, c'est une base.

Définition 7.6. Soient E, F deux espaces vectoriels de dimension finie et u une application linéaire de E dans F . Le rang de u est la dimension de $\text{Im}(u)$. On le note $\text{rg}(u)$.

En désignant par $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une base de E , l'image de u est le sous-espace vectoriel de F engendré par $\{u(e_1), \dots, u(e_n)\}$ et $\text{rg}(u) = \text{rg}(u(e_1), \dots, u(e_n))$.

Comme $\text{Im}(u)$ est un sous-espace vectoriel de F , on a $\text{rg}(u) \leq \dim(F)$ et comme $\text{rg}(u) = \text{rg}(u(e_1), \dots, u(e_n))$, on a aussi $\text{rg}(u) \leq \dim(E)$. On a donc $\text{rg}(u) \leq \min(\dim(E), \dim(F))$. Plus précisément, on a les résultats suivants.

Théorème 7.15.

Soient E, F deux espaces vectoriels de dimension finie et u une application linéaire de E dans F . On a $\text{rg}(u) = \dim(F)$ si, et seulement si, u est surjective.

Preuve. Dire que u est surjective équivaut à dire que $\text{Im}(u) = F$, ce qui est encore équivalent à $\text{rg}(u) = \dim(\text{Im}(u)) = \dim(F)$ puisque $\text{Im}(u)$ est un sous-espace vectoriel de F . $\square$

Théorème 7.16. Théorème du rang

Soient E, F deux espaces vectoriels de dimension finie et u une application linéaire de E dans F . On a $\dim(E) = \dim(\ker(u)) + \text{rg}(u)$.

Preuve. Soit H un supplémentaire de $\ker(u)$ dans E et v la restriction de u à H , c'est-à-dire l'application v définie sur H par :

$$\forall x \in H, v(x) = u(x)$$

Le noyau de cette application est $\ker(v) = H \cap \ker(u) = \{0\}$, ce qui signifie que v est injective de H dans F et réalise une bijection de H dans $\text{Im}(v)$. En écrivant tout vecteur y de $\text{Im}(u)$ sous la forme $y = u(x)$ avec $x \in E$ qui s'écrit $x = x_1 + x_2$ où $x_1 \in \ker(u)$ et $x_2 \in H$, on a $y = u(x_1) + u(x_2) = v(x_2)$, c'est-à-dire que y est dans $\text{Im}(v)$. On a donc $\text{Im}(u) \subset \text{Im}(v)$ et comme $\text{Im}(v) \subset \text{Im}(u)$, on a en fait

$\text{Im}(v) = \text{Im}(u)$ et v réalise un isomorphisme de H sur $\text{Im}(u)$. Il en résulte que :

$$\text{rg}(u) = \dim(\text{Im}(u)) = \dim(H) = \dim(E) - \dim(\ker(u))$$

□

En montrant le théorème du rang, on a en fait montré que $\text{Im}(u)$ est isomorphe à un supplémentaire de $\ker(u)$ dans E .

Corollaire 7.3. Soient E, F deux espaces vectoriels de dimension finie et u une application linéaire de E dans F . On a :

1. $\text{rg}(u) \leq \min(\dim(E), \dim(F))$;
2. $\text{rg}(u) = \dim(E)$ si, et seulement si, u est injective.

Preuve.

1. De la formule du rang, on déduit que $\text{rg}(u) \leq \dim(E)$ et $\text{rg}(u) \leq \dim(F)$ par définition.
2. Si $\text{rg}(u) = \dim(E)$, la formule du rang nous dit que $\dim(\ker(u)) = 0$, soit que $\ker(u) = \{0\}$ et u est injective. Réciproquement si u est injective, on a $\ker(u) = \{0\}$ et $\text{rg}(u) = \dim(E)$.

□

7.4 Expression matricielle des applications linéaires

Nous avons déjà introduit les matrices en utilisant les bases canoniques de $\mathbb{K}^n$ et $\mathbb{K}^m$. En fait tout peut être repris dans le cadre des espaces vectoriels de dimension fini en utilisant des bases de ces espaces, les démonstrations étant identiques à celles du paragraphe 6.5.

On se donne un espace vectoriel E de dimension n , une base $\mathcal{B} = (e_k)_{1 \leq k \leq n}$ de E , un espace vectoriel F de dimension m , une base $\mathcal{B}' = (f_k)_{1 \leq k \leq m}$ de F et une application linéaire u de E dans F .

Pour tout entier j compris entre 1 et n , il existe des scalaires a_{ij} tels que :

$$u(e_j) = \sum_{i=1}^m a_{ij} f_i$$

et pour tout vecteur $x = \sum_{j=1}^n x_j e_j$ dans E , on a :

$$u(x) = \sum_{j=1}^n x_j \sum_{i=1}^m a_{ij} f_i = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j \right) f_i$$

c'est-à-dire que les composantes dans la base $\mathcal{B}'$ de $u(x)$ sont les :

$$y_i = \sum_{j=1}^n a_{ij} x_j \quad (1 \leq i \leq m)$$

Définition 7.7. Avec les notations qui précèdent, on dit que la matrice $A = ((a_{ij}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ est la matrice de u dans les bases $\mathcal{B}$ et $\mathcal{B}'$. Dans le cas où $E = F$ et $\mathcal{B} = \mathcal{B}'$, on dira simplement que A est la matrice de u dans la base $\mathcal{B}$.

Pour tout j compris entre 1 et n , la colonne numéro j de la matrice A est le vecteur $C_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$ formé des composantes de $u(e_j)$ dans la base $\mathcal{B}'$. L'égalité $y = u(x)$ se traduit alors par le produit matriciel :

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = AX$$

où les x_j , pour j compris entre 1 et n sont les composantes de x dans la base $\mathcal{B}$ et les y_i , pour i compris entre 1 et m celles de $u(x)$ dans la base $\mathcal{B}'$.

Comme dans le cas de $\mathbb{K}^n$ et $\mathbb{K}^m$, on a le résultat suivant.

Théorème 7.17.

Si u et v sont deux applications linéaires de E dans F de matrices respectives $A = ((a_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ et $B = ((b_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ dans les bases $\mathcal{B}$ et $\mathcal{B}'$ alors, pour tous scalaires λ, μ , l'application linéaire $\lambda u + \mu v$ a pour matrice dans ces bases la matrice $\lambda A + \mu B = ((\lambda a_{i,j} + \mu b_{i,j}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$.

Ce résultat peut se traduire en disant que l'application qui associe à toute application linéaire $u \in \mathcal{L}(E, F)$ sa matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$ dans les bases $\mathcal{B}$ et $\mathcal{B}'$ est linéaire. Plus précisément, on a le résultat suivant.

Théorème 7.18.

Étant données une base $\mathcal{B} = (e_k)_{1 \leq k \leq n}$ de E et une base $\mathcal{B}' = (f_k)_{1 \leq k \leq m}$ de F , l'application φ qui associe à toute application linéaire $u \in \mathcal{L}(E, F)$ sa matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$ dans les bases $\mathcal{B}$ et $\mathcal{B}'$ est un isomorphisme de $\mathcal{L}(E, F)$ sur $\mathcal{M}_{m,n}(\mathbb{K})$.

Preuve. On vient de voir que φ est linéaire et cette application est bijective du fait qu'une application linéaire $u \in \mathcal{L}(E, F)$ est uniquement déterminée par sa matrice dans les bases $\mathcal{B}$ et $\mathcal{B}'$. $\square$

En particulier, pour toute matrice $A \in \mathcal{M}_{m,n}(\mathbb{K})$ il existe une unique application linéaire $u \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ ayant pour matrice A dans les bases canoniques de $\mathbb{K}^n$ et $\mathbb{K}^m$. On dira que u est l'application linéaire canoniquement associée à la matrice A .

Pour ce qui est de la matrice d'une composée d'applications linéaires nous avons le résultat suivant où G est un espace vectoriel de dimension r et $\mathcal{B}'' = (g_k)_{1 \leq k \leq r}$ une base de G .

Théorème 7.19.

Si v est une application linéaire de E dans F de matrice $B \in \mathcal{M}_{m,n}(\mathbb{K})$ dans les bases $\mathcal{B}$ et $\mathcal{B}'$ et u une application linéaire de F dans G de matrice $A \in \mathcal{M}_{r,m}(\mathbb{K})$ dans les bases $\mathcal{B}'$ et $\mathcal{B}''$, alors la matrice dans les bases $\mathcal{B}$ et $\mathcal{B}''$ de l'application linéaire $u \circ v$ de E dans G est la matrice produit $C = AB \in \mathcal{M}_{r,n}(\mathbb{K})$.

On en déduit le résultat suivant, où $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de E .

Théorème 7.20.

Un endomorphisme u de E est bijectif si, et seulement si, sa matrice A dans les bases $\mathcal{B}$ et $\mathcal{B}'$ est inversible et dans ce cas A^{-1} est la matrice de u^{-1} dans les bases $\mathcal{B}'$ et $\mathcal{B}$.

En regardant une matrice comme un ensemble de vecteurs colonnes, on peut donner la définition suivante.

Définition 7.8. Soit $A = ((a_{ij}))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ une matrice dans $\mathcal{M}_{m,n}(\mathbb{K})$. En désignant, pour tout j compris entre 1 et n , par $C_j = (a_{ij})_{1 \leq i \leq n}$ le vecteur de $\mathbb{K}^n$ représentant la colonne numéro j de A , le rang de A est le rang de la famille $(C_1, C_2, \dots, C_n)$ de vecteurs de $\mathbb{K}^n$.

Théorème 7.21.

Le rang d'une matrice $A \in \mathcal{M}_{r,m}(\mathbb{K})$ est égal au rang de l'application linéaire $u \in \mathcal{L}(\mathbb{K}^n, \mathbb{K}^m)$ canoniquement associé à la matrice A .

Preuve. En désignant par $\mathcal{B} = (e_k)_{1 \leq k \leq n}$ une base de $\mathbb{K}^n$ et par $\mathcal{B}' = (f_k)_{1 \leq k \leq m}$ une base de $\mathbb{K}^m$, pour tout j compris entre 1 et n , la colonne numéro j de A est $C_j = u(e_j)$ et on a :

$$\text{rg}(A) = \text{rg}(C_1, C_2, \dots, C_n) = \text{rg}(u(e_1), u(e_2), \dots, u(e_n)) = \text{rg}(u)$$

□

Théorème 7.22.

Si $u \in \mathcal{L}(E, F)$ a pour matrice $A \in \mathcal{M}_{r,m}(\mathbb{K})$ dans les bases $\mathcal{B}$ et $\mathcal{B}'$ (toujours avec les notations du début de ce paragraphe), alors le rang de u est égal au rang de A .

Preuve. Dire que $x = \sum_{j=1}^n x_j e_j$ est dans le noyau de u équivaut à dire que $u(x) = 0$, ce qui se traduit par $AX = 0$, où $X = (x_j)_{1 \leq j \leq n}$ est le vecteur de $\mathbb{K}^n$ formé des composantes de x dans la base $\mathcal{B}$. En désignant par v l'application linéaire canoniquement associée à A , on a $v(X) = AX = 0$, c'est-à-dire que X est dans le noyau de v . Réciproquement si $X \in \ker(v)$, on a $AX = 0$, ce qui équivaut à $u(x) = 0$, soit $x \in \ker(u)$. L'application $x \mapsto X$ réalise donc un isomorphisme de $\ker(u)$ sur $\ker(v)$ et $\dim(\ker(u)) = \dim(\ker(v))$, ce qui équivaut à $\operatorname{rg}(u) = \operatorname{rg}(v)$ en utilisant le théorème du rang. Et donc $\operatorname{rg}(u) = \operatorname{rg}(v) = \operatorname{rg}(A)$. $\square$

7.5 Formules de changement de base

On se donne un espace vectoriel E de dimension n et deux bases de E , $\mathcal{B}_1 = (e_k)_{1 \leq k \leq n}$ et $\mathcal{B}_2 = (e'_k)_{1 \leq k \leq n}$. Une question naturelle est de savoir comment passer des composantes d'un vecteur de E d'une base à l'autre.

Pour tout vecteur $x = \sum_{j=1}^n x_j e_j = \sum_{j=1}^n x'_j e'_j$ dans E , on note $X = (x_j)_{1 \leq j \leq n}$ et $X' = (x'_j)_{1 \leq j \leq n}$ les vecteurs de $\mathbb{K}^n$ respectivement formés des composantes de x dans les bases $\mathcal{B}_1$ et $\mathcal{B}_2$. Comme, pour j compris entre 1 et n le vecteur e'_j est dans E , il s'écrit $e'_j = \sum_{i=1}^n p_{ij} e_i$, où les p_{ij} sont des scalaires uniquement déterminés. On a alors :

$$x = \sum_{i=1}^n x_i e_i = \sum_{j=1}^n x'_j e'_j = \sum_{j=1}^n x'_j \left(\sum_{i=1}^n p_{ij} e_i \right) = \sum_{i=1}^n \left(\sum_{j=1}^n p_{ij} x'_j \right) e_i$$

et avec l'unicité de l'écriture de x dans la base $\mathcal{B}_1$, on déduit que :

$$x_i = \sum_{j=1}^n p_{ij} x'_j \quad (1 \leq i \leq n)$$

ce qui peut se traduire par l'égalité matricielle $X = PX'$, où P est la matrice :

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}$$

ayant pour colonne j le vecteur de $\mathbb{K}^n$ formé des composantes de e'_j dans la base $\mathcal{B}_1$.

On peut retenir cette formule sous la forme $X_{\mathcal{B}_1} = P_{\mathcal{B}_1 \rightarrow \mathcal{B}_2} X_{\mathcal{B}_2}$ avec des notations évidentes.

Définition 7.9. Avec les notations qui précèdent, on dit que P est la matrice de passage de la base $\mathcal{B}_1$ à la base $\mathcal{B}_2$.

Théorème 7.23.

Avec les notations qui précèdent, la matrice de passage P de $\mathcal{B}_1$ à $\mathcal{B}_2$ est inversible et son inverse est la matrice de passage de $\mathcal{B}_2$ à $\mathcal{B}_1$.

Preuve. Il suffit de remarquer que la matrice P de $\mathcal{B}_1$ à $\mathcal{B}_2$ est la matrice dans les bases $\mathcal{B}_2$ et $\mathcal{B}_1$ de l'identité de E . En effet, pour tout j compris entre 1 et n on a $Id_E(e'_j) = e'_j = \sum_{i=1}^n p_{ij}e_i$. Cette matrice est donc inversible et son inverse est la matrice dans les bases $\mathcal{B}_1$ et $\mathcal{B}_2$ de $(Id_E)^{-1} = Id_E$ c'est-à-dire la matrice de passage de $\mathcal{B}_2$ à $\mathcal{B}_1$. $\square$

On a donc $(P_{\mathcal{B}_1 \rightarrow \mathcal{B}_2})^{-1} = P_{\mathcal{B}_2 \rightarrow \mathcal{B}_1}$. Le calcul de l'inverse de la matrice de passage P de $\mathcal{B}_1$ à $\mathcal{B}_2$ peut se calculer en résolvant le système linéaire aux inconnues e_i :

$$e'_j = \sum_{i=1}^n p_{ij}e_i \quad (1 \leq i \leq n)$$

On est maintenant en mesure d'exprimer la matrice d'un endomorphisme u de E dans la base $\mathcal{B}_2$ en fonction de sa matrice dans la base $\mathcal{B}_1$.

Théorème 7.24.

Si u est un endomorphisme de E de matrice $A = ((a_{ij}))_{1 \leq i, j \leq n}$ dans la base $\mathcal{B}_1$ et de matrice $A' = ((a'_{ij}))_{1 \leq i, j \leq n}$ dans la base $\mathcal{B}_2$, on a alors $A' = P^{-1}AP$, où P est la matrice de passage de $\mathcal{B}_1$ à $\mathcal{B}_2$.

Preuve. PA' est la matrice de $Id_E \circ u = u$ dans les bases $\mathcal{B}_2$ et $\mathcal{B}_1$ et AP est aussi la matrice de $u \circ Id_E = u$ dans les bases $\mathcal{B}_2$ et $\mathcal{B}_1$. On a donc $PA' = AP$, ce qui équivaut à $A' = P^{-1}AP$. $\square$

La formule de changement de base peut se retenir sous la forme :

$$A_{\mathcal{B}_2} = P_{\mathcal{B}_2 \rightarrow \mathcal{B}_1} A_{\mathcal{B}_1} P_{\mathcal{B}_1 \rightarrow \mathcal{B}_2}$$

Dans le cas où la matrice A' de u dans la base $\mathcal{B}_2$ a une forme plus simple que celle de A , on peut l'utiliser pour calculer les puissances de u ou de A . L'idée étant que l'égalité $A' = P^{-1}AP$ entraîne pour tout entier naturel p , on a $(A')^p = P^{-1}A^pP$ ou encore $A^p = P(A')^pP^{-1}$. La vérification étant immédiate par récurrence sur $p \geq 0$ (toujours avec la convention $A^0 = I_n$).

Si par exemple, A' est diagonale, c'est-à-dire de la forme $A' = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$,

on a alors $(A')^p = \begin{pmatrix} \lambda_1^p & 0 & \cdots & 0 \\ 0 & \lambda_2^p & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n^p \end{pmatrix}$ pour tout $p \geq 0$. Pour $p \geq 1$, l'endo-

morphisme u^p est défini par la formule de récurrence $u^p = u^{p-1} \circ u$ avec $u^0 = Id_E$. Si A est la matrice de u dans la base $\mathcal{B}_1$, alors celle de A^p dans cette même base est A^p .

7.6 Exercices

Exercice 7.1. Montrer que la famille $\mathcal{B} = (L_0, L_1, L_2)$ de polynômes de $\mathbb{K}_2[x]$ définie par :

$$L_0(x) = (x-1)(x-2), \quad L_1(x) = x(x-2), \quad L_2(x) = x(x-1)$$

forme une base de $\mathbb{K}_2[x]$. En déduire que pour tout polynôme P dans $\mathbb{K}_2[x]$

$$\text{on a } \int_0^2 P(t) dt = \frac{P(0) + 4P(1) + P(2)}{3} \quad (\text{formule des trois niveaux}).$$

Solution. Supposons que $\lambda_0 P_0 + \lambda_1 P_1 + \lambda_2 P_2 = 0$. Prenant les valeurs successives $x = 0, 1, 2$, on déduit que $\lambda_0 = \lambda_1 = \lambda_2 = 0$. Le système $\mathcal{B}$ est donc libre. Étant donné $P = ax^2 + bx + c$, on cherche des scalaires $\lambda_0, \lambda_1, \lambda_2$ tels que :

$$ax^2 + bx + c = \lambda_0 P_0 + \lambda_1 P_1 + \lambda_2 P_2.$$

Les valeurs $x = 0, 1, 2$ nous donne $2\lambda_0 = c$, $-\lambda_1 = a + b + c$ et $2\lambda_2 = 4a + 2b + c$, ce qui détermine de manière unique les scalaires $\lambda_0, \lambda_1, \lambda_2$. Le système $\mathcal{B}$ est donc générateur et c'est une base de $\mathbb{K}_2[x]$. Tout polynôme P dans $\mathbb{K}_2[x]$ s'écrit donc de manière unique $P = \lambda_0 P_0 + \lambda_1 P_1 + \lambda_2 P_2$ avec $\lambda_0 = \frac{P(0)}{2}$, $\lambda_1 = -P(1)$ et $\lambda_2 = \frac{P(2)}{2}$. On a alors :

$$\begin{aligned} \int_0^2 P(t) dt &= \frac{P(0)}{2} \int_0^2 L_0(t) dt - P(1) \int_0^2 L_1(t) dt + \frac{P(2)}{2} \int_0^2 L_2(t) dt \\ &= \frac{P(0) + 4P(1) + P(2)}{3} \end{aligned}$$

Exercice 7.2. On désigne par E l'espace vectoriel des applications de $\mathbb{R}$ dans $\mathbb{R}$. Montrer, pour tous réels $0 < a_1 < a_2 < \dots < a_n$, la famille :

$$\mathcal{L} = \{f_{a_k} : x \mapsto \sin(a_k x) \mid 1 \leq k \leq n\}$$

est libre dans E (ce qui peut se traduire en disant que la famille de fonctions $\{f_a : x \mapsto \sin(ax) \mid a \in \mathbb{R}^{+,*}\}$ est libre dans E).

Solution. Pour $n = 1$, la fonction $f_a : x \mapsto \sin(ax)$ n'est pas la fonction nulle, donc (f_a) est libre dans E . Supposons le résultat acquis au rang $n - 1 \geq 1$ et soient

$0 < a_1 < a_2 < \dots < a_n$, $\lambda_1, \lambda_2, \dots, \lambda_n$ des réels tels que $\sum_{k=1}^n \lambda_k \sin(a_k x) = 0$ pour

tout $x \in \mathbb{R}$. En dérivant deux fois, on a $\sum_{k=1}^n \lambda_k a_k^2 \sin(a_k x) = 0$ pour tout $x \in \mathbb{R}$. Il

en résulte que :

$$\forall x \in \mathbb{R}, \sum_{k=1}^{n-1} \lambda_k (a_k^2 - a_n^2) \sin(a_k x) = 0$$

et l'hypothèse de récurrence nous dit que $\lambda_k (a_k^2 - a_n^2) = 0$ pour tout k compris entre 1 et $n - 1$, ce qui équivaut à dire que $\lambda_k = 0$ pour tout k compris entre 1 et $n - 1$ puisque $a_k^2 \neq a_n^2$ pour $k \neq n$. Il reste alors $\lambda_n f_{a_n} = 0$ dans E et $\lambda_n = 0$. On a donc ainsi montré que la famille $(f_{a_k})_{1 \leq k \leq n}$ est libre dans E .

Exercice 7.3. Pour tout $n \in \mathbb{N}^*$, on désigne par $\mathbb{R}_n[x]$ l'espace vectoriel des fonctions polynomiales à coefficients réels de degré au plus égal à n et on considère l'application $u : P \in \mathbb{R}_n[x] \mapsto xP' \in \mathbb{R}_n[x]$, où on a noté, pour toute fonction polynomiale $P \in \mathbb{R}_n[x]$, P' le polynôme dérivé de P .

1. Montrer que u est une application linéaire de E dans E .
2. Donner la matrice de u dans la base canonique $\mathcal{B} = (1, x, \dots, x^n)$ de E .
3. L'application u est-elle injective ?
4. L'application u est-elle surjective ?
5. Calculer le noyau, l'image et le rang de u .
6. Soit F le sous-espace-vectoriel de E engendré par $(x, x^2, \dots, x^n)$. Montrer que l'application u est bijective de F sur F .

Solution.

1. Pour $P \in E$, on a $P' \in \mathbb{R}_{n-1}[x]$ et $xP' \in \mathbb{R}_n[x] = E$, donc u va bien de E dans E . Pour P, Q dans E et λ, μ dans $\mathbb{R}$, on a :

$$u(\lambda P + \mu Q) = x(\lambda P + \mu Q)' = \lambda(xP') + \mu(xQ') = \lambda u(P) + \mu u(Q)$$

donc u est linéaire.

2. On a $u(1) = 0$ et $u(x^k) = kx^k$ pour k entier compris entre 1. La matrice de u

dans $\mathcal{B}$ est donc $A = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & n-1 & 0 \\ 0 & 0 & 0 & \cdots & 0 & n \end{pmatrix}.$

3. On a $u(1) = 0$ avec $1 \neq 0$, donc u non injective.
4. Pour tout polynôme $P \in E$, on a $u(P)(0) = 0$ car $u(P) = xP'$, donc 1 n'est pas dans l'image de u et u est non surjective.
5. On a $P \in \ker(u)$ si, et seulement si, $xP' = 0$, ce qui équivaut encore à $P' = 0$, soit P constant. Donc $\ker(u)$ est la droite dirigée par le polynôme constant 1. En utilisant le théorème du rang, on déduit que $\text{rg}(u) = \dim(E) - 1 = n$. L'image de u est contenu dans l'espace $x\mathbb{R}_{n-1}[x]$ des polynômes de $\mathbb{R}_n[x]$ multiples de x . Réciproquement tout polynôme dans $x\mathbb{R}_{n-1}[x]$ s'écrit xQ avec $Q \in \mathbb{R}_{n-1}[x]$ et désignant par $P \in \mathbb{R}_n[x]$ une primitive de Q , on a $xQ = xP' = u(P)$. Donc $\text{Im}(u) = x\mathbb{R}_{n-1}[x]$. On retrouve le rang de u :

$$\text{rg}(u) = \dim(x\mathbb{R}_{n-1}[x]) = \dim(\mathbb{R}_{n-1}[x]) = n.$$

6. On remarque que $F = x\mathbb{R}_{n-1}[x] = \text{Im}(u)$. Donc la restriction v de u à F est un endomorphisme de F . Son noyau est $\ker(v) = F \cap \ker(u) = \{0\}$, il en résulte que u est un isomorphisme.

Exercice 7.4. Soit E un espace vectoriel de dimension finie et $u \in \mathcal{L}(E)$.

1. Montrer que $\text{Im}(u) = \text{Im}(u^2) \Leftrightarrow \ker(u) = \ker(u^2)$, où $u^2 = u \circ u$.
2. Montrer que :

$$(\text{Im}(u) = \text{Im}(u^2)) \Leftrightarrow (E = \ker(u) \oplus \text{Im}(u)) \Leftrightarrow (\ker(u) = \ker(u^2))$$

Solution.

1. On a toujours $\text{Im}(u^2) \subset \text{Im}(u)$, $\ker(u) \subset \ker(u^2)$, donc :

$$(\text{Im}(u) = \text{Im}(u^2)) \Leftrightarrow (\text{rg}(u) = \text{rg}(u^2))$$

et :

$$(\ker(u) = \ker(u^2)) \Leftrightarrow (\dim(\ker(u)) = \dim(\ker(u^2)))$$

D'autre part, le théorème du rang nous dit que :

$$\dim(E) = \dim(\ker(u)) + \text{rg}(u) = \dim(\ker(u^2)) + \text{rg}(u^2)$$

ce qui permet de déduire que :

$$(\text{Im}(u) = \text{Im}(u^2)) \Leftrightarrow (\ker(u) = \ker(u^2))$$

2. Il suffit de montrer que :

$$(\operatorname{Im}(u) = \operatorname{Im}(u^2)) \Leftrightarrow (E = \ker(u) \oplus \operatorname{Im}(u))$$

Si $\operatorname{Im}(u) = \operatorname{Im}(u^2)$, alors pour tout x dans E , il existe y dans E tel que $u(x) = u^2(y)$, donc $x - u(y) \in \ker(u)$ et $x = (x - u(y)) + u(y) \in \ker(u) + \operatorname{Im}(u)$. On a donc $E = \ker(u) + \operatorname{Im}(u)$ et avec le théorème du rang, on déduit que $E = \ker(u) \oplus \operatorname{Im}(u)$. Si $E = \ker(u) \oplus \operatorname{Im}(u)$ alors tout $x \in \ker(u^2)$ s'écrit $x = x_1 + u(x_2)$ avec $u(x_1) = 0$ et $0 = u^2(x) = u^3(x_2)$ entraîne que $u^2(x_2) \in \ker(u) \cap \operatorname{Im}(u) = \{0\}$, donc $u(x_2) \in \ker(u) \cap \operatorname{Im}(u) = \{0\}$ et $x = x_1 \in \ker(u)$. On a donc $\ker(u^2) \subset \ker(u)$ et $\ker(u) = \ker(u^2)$, ce qui équivaut à $\operatorname{Im}(u) = \operatorname{Im}(u^2)$.

Exercice 7.5. On désigne par $\mathcal{B}_1 = (e_1, e_2, e_3)$ la base canonique de $\mathbb{K}^3$.

1. Montrer que $\mathcal{B}_2 = (e'_1 = e_1 - e_2, e'_2 = 2e_2 + e_3, e'_3 = e_1 + e_3)$ est une base de $\mathbb{K}^3$.
2. Calculer l'inverse de la matrice de passage P de $\mathcal{B}_1$ à $\mathcal{B}_2$.

Solution.

1. L'égalité $\lambda_1 e'_1 + \lambda_2 e'_2 + \lambda_3 e'_3 = 0$ équivaut à :

$$\begin{cases} \lambda_1 + \lambda_3 = 0 \\ -\lambda_1 + 2\lambda_2 = 0 \\ \lambda_2 + \lambda_3 = 0 \end{cases}$$

En additionnant les deux premières équations, on a $\lambda_3 + 2\lambda_2 = 0$ qui reporté dans la troisième donne $\lambda_2 = 0$. Les équations 2 et 3 donnent alors $\lambda_1 = \lambda_3 = 0$. La famille $\mathcal{B}_2$ est donc libre dans $\mathbb{K}^3$ et c'est une base puisqu'elle a trois éléments.

2. La matrice de passage de $\mathcal{B}_1$ à $\mathcal{B}_2$ est la matrice $P = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix}$. Avec $e'_1 = e_1 - e_2$, $e'_2 = 2e_2 + e_3$ et $e'_3 = e_1 + e_3$ on déduit que $e_1 = 2e'_1 + e'_2 - e'_3$, $e_2 = e'_1 + e'_2 - e'_3$ et $e_3 = -2e'_1 - e'_2 + 2e'_3$, ce qui signifie que $P^{-1} = \begin{pmatrix} 2 & 1 & -2 \\ 1 & 1 & -1 \\ -1 & -1 & 2 \end{pmatrix}$.

Exercice 7.6. On désigne par $\mathcal{B}_1$ et $\mathcal{B}_2$ les bases de $\mathbb{K}^3$ de l'exercice 7.5

et par u l'endomorphisme de $\mathbb{K}^3$ de matrice $A = \begin{pmatrix} 4 & 2 & -4 \\ -6 & -4 & 6 \\ -1 & -1 & 1 \end{pmatrix}$ dans la base canonique $\mathcal{B}_1$.

1. Déterminer la matrice de u dans la base $\mathcal{B}_2$.
2. Calculer, pour tout $p \in \mathbb{N}$, la matrice de u^p dans la base canonique de $\mathbb{K}^3$.

Solution.

1. La matrice de u dans la base $\mathcal{B}_2$ est :

$$\begin{aligned} A' &= P^{-1}AP = \begin{pmatrix} 2 & 1 & -2 \\ 1 & 1 & -1 \\ -1 & -1 & 2 \end{pmatrix} \begin{pmatrix} 4 & 2 & -4 \\ -6 & -4 & 6 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

2. La matrice de u^p dans la base $\mathcal{B}_2$ est $(A')^p = \begin{pmatrix} 2^p & 0 & 0 \\ 0 & (-1)^p & 0 \\ 0 & 0 & 0 \end{pmatrix}$ et dans la base $\mathcal{B}_1$ c'est :

$$\begin{aligned} A^p &= P(A')^p P^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2^p & 0 & 0 \\ 0 & (-1)^p & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 & -2 \\ 1 & 1 & -1 \\ -1 & -1 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 2^{p+1} & 2^p & -2^{p+1} \\ 2(-1)^p - 2^{p+1} & 2(-1)^p - 2^p & 2^{p+1} - 2(-1)^p \\ (-1)^p & (-1)^p & (-1)^{p+1} \end{pmatrix} \end{aligned}$$

Chapitre 8

Opérations élémentaires et déterminants

On note toujours $\mathbb{K}$ le corps de réels ou des complexes.

On se donne un entier $n \geq 1$ et $\mathcal{M}_n(\mathbb{K})$ désigne l'espace vectoriel des matrices carrées d'ordre n à coefficients dans $\mathbb{K}$.

Pour i, j entiers compris entre 1 et n , on note E_{ij} la matrice dont tous les coefficients sont nuls sauf celui d'indice (i, j) (ligne i et colonne j) qui vaut 1. On rappelle que la famille $(E_{ij})_{1 \leq i, j \leq n}$ est une base de $\mathcal{M}_n(\mathbb{K})$ qui est donc de dimension n^2 .

Pour toute matrice $A = ((a_{ij}))_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbb{K})$, on note pour tout entier i compris entre 1 et n , $L_i = (a_{i1}, a_{i2}, \dots, a_{in})$ sa ligne numéro i (c'est une matrice à

une ligne et n colonnes) et pour tout entier j compris entre 1 et n , $C_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$

sa colonne numéro j (c'est une matrice à n lignes et une colonne). On écrira

$$A = \begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix} \text{ ou } A = \begin{pmatrix} C_1 & \cdots & C_n \end{pmatrix}.$$

Définition 8.1. On dit qu'une matrice $A = ((a_{ij}))_{1 \leq i, j \leq n} \in \mathcal{M}_n(\mathbb{K})$ est triangulaire inférieure [resp. supérieure] si $a_{ij} = 0$ pour $1 \leq i < j \leq n$ [resp. pour $1 \leq j < i \leq n$].

Une matrice triangulaire inférieure est donc de la forme

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

et une matrice triangulaire supérieure de la forme

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}.$$

Définition 8.2. Une matrice diagonale est une matrice triangulaire inférieure et supérieure.

Une matrice diagonale est donc de la forme

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}.$$

8.1 Opérations élémentaires. Matrices de dilatation et de transvection

On suppose que $n \geq 2$.

On appelle matrice déduite de A par opération élémentaire sur les lignes de A

toute matrice de la forme $A_i(\lambda) = \begin{pmatrix} L_1 \\ \vdots \\ L_{i-1} \\ \lambda L_i \\ L_{i+1} \\ \vdots \\ L_n \end{pmatrix}$ avec $1 \leq i \leq n$ et $\lambda \in \mathbb{K}^*$, c'est-

à-dire que la matrice $A_i(\lambda)$ est déduite de la matrice A en multipliant sa ligne

numéro i par λ ou de la forme $A_{ij}(\lambda) = \begin{pmatrix} L_1 \\ \vdots \\ L_{i-1} \\ L_i + \lambda L_j \\ L_{i+1} \\ \vdots \\ L_n \end{pmatrix}$ avec $1 \leq i \neq j \leq n$ et

$\lambda \in \mathbb{K}$, c'est-à-dire que la matrice $A_{ij}(\lambda)$ est déduite de la matrice A en ajoutant à la ligne numéro i la ligne numéro j multipliée par λ .

On appelle matrice déduite de A par opération élémentaire sur les colonnes de A toute matrice de la forme :

$$A'_j(\lambda) = \begin{pmatrix} C_1 & \cdots & C_{j-1} & \lambda C_j & C_{j+1} & \cdots & C_n \end{pmatrix}$$

avec $1 \leq j \leq n$ et $\lambda \in \mathbb{K}^*$, c'est-à-dire que la matrice $A'_i(\lambda)$ est déduite de la matrice A en multipliant sa colonne numéro j par λ ou de la forme :

$$A'_{ij}(\lambda) = \begin{pmatrix} C_1 & \cdots & C_{j-1} & C_j + \lambda C_i & C_{j+1} & \cdots & C_n \end{pmatrix}$$

$1 \leq i \neq j \leq n$ et $\lambda \in \mathbb{K}$, c'est-à-dire que la matrice $A'_{ij}(\lambda)$ est déduite de la matrice A en ajoutant à la colonne numéro j la colonne numéro i multipliée par λ .

Définition 8.3. On appelle matrice de transvection toute matrice de la forme $T_{ij}(\lambda) = I_n + \lambda E_{ij}$ avec $1 \leq i \neq j \leq n$ et $\lambda \in \mathbb{K}$.

Une matrice de transvection $T_{ij}(\lambda)$ est donc une matrice triangulaire dont tous les termes diagonaux valent 1 et de termes hors de la diagonale tous nuls sauf celui d'indice (i, j) (i. e. en ligne i et colonne j) qui vaut λ .

Définition 8.4. On appelle matrice de dilatation toute matrice de la forme $D_i(\lambda) = I_n + (\lambda - 1) E_{ii}$ avec $1 \leq i \leq n$ et $\lambda \in \mathbb{K}^*$.

Une matrice de dilatation $D_i(\lambda)$ est donc diagonale de termes diagonaux tous égaux à 1 sauf le numéro i qui vaut λ .

Théorème 8.1.

Avec les notations qui précèdent on a :

$$\begin{aligned} A_i(\lambda) &= D_i(\lambda) A, \quad A_{ij}(\lambda) = T_{ij}(\lambda) A \\ A'_j(\lambda) &= A D_j(\lambda), \quad A'_{ij}(\lambda) = A T_{ij}(\lambda) \end{aligned}$$

C'est-à-dire que :

1. la multiplication à gauche par une matrice de dilatation $D_i(\lambda)$ a pour effet de multiplier la ligne i par λ ;
2. la multiplication à droite par une matrice de dilatation $D_j(\lambda)$ a pour effet de multiplier la colonne j par λ ;
3. la multiplication à gauche par une matrice de transvection $T_{ij}(\lambda)$ a pour effet de remplacer la ligne L_i par $L_i + \lambda L_j$;
4. la multiplication à droite par une matrice de transvection $T_{ij}(\lambda)$ a pour effet de remplacer la colonne C_j par $C_j + \lambda C_i$.

Preuve. Le coefficient d'indice (p, q) du produit de matrices $D_i(\lambda) A$ est obtenu en faisant le produit de la ligne p de $D_i(\lambda)$ par la colonne q de A , ce qui donne en notant $\alpha_{p,q}$ ce coefficient :

$$\alpha_{p,q} = \begin{cases} a_{p,q} & \text{si } 1 \leq p \neq i \leq n, 1 \leq q \leq n \\ \lambda a_{iq} & \text{si } p = i, 1 \leq q \leq n \end{cases}$$

On a donc bien $A_i(\lambda) = D_i(\lambda)A$. Les autres égalités se montrent de façon analogue. $\square$

Ce résultat justifie la définition suivante.

Définition 8.5. On appelle *matrice élémentaire* une matrice de dilatation ou de transvection.

Lemme 8.1 Une matrice élémentaire est inversible avec $T_{ij}(\lambda)^{-1} = T_{ij}(-\lambda)$ pour une matrice de transvection et $D_i(\lambda)^{-1} = D_i\left(\frac{1}{\lambda}\right)$ pour une matrice de dilatation.

Preuve. Pour λ, μ dans $\mathbb{K}$ et $i \neq j$ compris entre 1 et n , la matrice $T_{ij}(\lambda)T_{ij}(\mu)$ se déduit de $T_{ij}(\mu)$ en ajoutant à sa ligne i sa ligne j multipliée par λ , ce qui donne la matrice $T_{ij}(\lambda + \mu)$. Prenant $\mu = -\lambda$, on a $T_{ij}(\lambda)T_{ij}(-\lambda) = T_0 = I_n$, ce qui signifie que $T_{ij}(\lambda)$ est inversible d'inverse $T_{ij}(-\lambda)$.

Le deuxième résultat est évident. $\square$

Théorème 8.2.

Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ (où $n \geq 2$) est inversible si, et seulement si, elle est produit de matrices élémentaires. Précisément si $A \in \mathcal{M}_n(\mathbb{K})$ est inversible, il existe alors des matrices de transvection $P_1, \dots, P_r$ et $Q_1, \dots, Q_s$ et une matrice de dilatation $D_n(\lambda)$ telles que :

$$A = P_1 \cdots P_r D_n(\lambda) Q_1 \cdots Q_s$$

Preuve. Le cas $n = 2$ est traité avec l'exercice 8.1. On le suppose vrai pour toutes les matrices inversibles d'ordre $n - 1 \geq 2$ et on se donne une matrice inversible A d'ordre n . On se ramène tout d'abord par opération élémentaire au cas où $a_{21} \neq 0$. Si $a_{21} = 0$, comme A est inversible, sa colonne 1 n'est pas nulle (cette colonne est Ae_1 où e_1 est le premier vecteur de base canonique et $x = 0$ est l'unique solution de $Ax = 0$), il existe donc un indice $i \in \{1, 3, \dots, n\}$ tel que $a_{i1} \neq 0$ et la matrice $T_{2i}(1)A$ (déduite de A en ajoutant la ligne i à la ligne 2) est telle que son coefficient d'indice $(2, 1)$ est non nul. Une fois ramené à $a_{21} \neq 0$, on se ramène à $a_{11} = 1$ en remplaçant la première ligne L_1 par $L_1 + \lambda L_2$ (multiplication à gauche par $T_{12}(\lambda)$) où le scalaire λ est choisi tel que $a_{11} + \lambda a_{21} = 1$. Ensuite, pour tout $i \in \{2, 3, \dots, n\}$, en remplaçant la ligne L_i par $L_i - a_{i1}L_1$ (multiplication à gauche par $T_{i1}(-a_{i1})$), on annule le coefficient d'indice $(i, 1)$. On peut donc trouver des matrices de transvection $P_1, \dots, P_k$ telles que :

$$P_k \cdots P_1 A = \begin{pmatrix} 1 & \alpha_{12} & \cdots & \alpha_{1n} \\ 0 & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix}$$

De manière analogue, en multipliant à droite par des matrices de transvection, $Q_1, \dots, Q_m$, on obtient :

$$P_k \cdots P_1 A Q_1 \cdots Q_m = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \beta_{22} & \cdots & \beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \beta_{n2} & \cdots & \beta_{nn} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & B & \\ 0 & & & \end{pmatrix}$$

On peut alors conclure en appliquant l'hypothèse de récurrence à la matrice B qui est d'ordre $n - 1$ et inversible. En effet, si B n'est pas inversible, il existe $x' \neq 0$ dans $\mathbb{K}^{n-1}$ tel que $Bx' = 0$, donc $x = \begin{pmatrix} 0 \\ x' \end{pmatrix} \in \mathbb{K}^n$ est non nul solution de $P_k \cdots P_1 A Q_1 \cdots Q_m x = 0$ qui équivaut à $Ay = 0$ avec $y = Q_1 \cdots Q_m x \neq 0$ puisque les matrices P_i et Q_j sont inversibles, en contradiction avec A inversible. $\square$

Nous verrons plus loin (paragraphe 8) que, comme dans le cas $n = 2$, le scalaire λ qui intervient dans le théorème précédent est uniquement déterminé par la matrice A , c'est son déterminant.

Pour $n = 1$, le résultat est encore vrai avec $A = (a) = D(a)$.

8.2 Déterminants des matrices carrées

Nous avons déjà défini le déterminant d'une matrice d'ordre deux, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ par $\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$ (définition 6.6.3).

Une matrice d'ordre 1 étant tout simplement un réel ou un complexe, son déterminant est lui même.

Le déterminant d'une matrice carrée $A = ((a_{ij}))_{1 \leq i, j \leq n}$ d'ordre $n \geq 3$ peut se définir par récurrence comme suit :

$$\det(A) = \sum_{i=1}^n (-1)^{i+1} a_{i,1} \det(A_{i,1})$$

où $A_{i,1}$ est, pour i compris entre 1 et n , la matrice d'ordre $n - 1$ déduite de A en supprimant la première colonne et la ligne numéro i . Dans cette expression, on dit qu'on développe le déterminant suivant la première colonne.

$$\text{On note } \det(A) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

Les lignes d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ étant notées $L_1, L_2, \dots, L_n$, on écrira aussi $\det(A) = \det \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix}$.

Exemple 8.1 Pour $n = 3$ et $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$, on a $\det(A) = \begin{vmatrix} 5 & 6 \\ 8 & 9 \end{vmatrix} - 4 \begin{vmatrix} 2 & 3 \\ 8 & 9 \end{vmatrix} + 7 \begin{vmatrix} 2 & 3 \\ 5 & 6 \end{vmatrix} = 0$.

Théorème 8.3.

Si $A_i(\lambda)$ est la matrice déduite de $A \in \mathcal{M}_n(\mathbb{K})$ en multipliant sa ligne i par un scalaire λ , on a alors $\det(A_i(\lambda)) = \lambda \det(A)$, soit :

$$\det \begin{pmatrix} L_1 \\ \vdots \\ L_{i-1} \\ \lambda L_i \\ L_{i+1} \\ \vdots \\ L_n \end{pmatrix} = \lambda \det \begin{pmatrix} L_1 \\ \vdots \\ L_{i-1} \\ L_i \\ L_{i+1} \\ \vdots \\ L_n \end{pmatrix}$$

Preuve. Pour $n = 1$ c'est clair et pour $n = 2$, on a $\begin{vmatrix} \lambda a & \lambda b \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ \lambda c & \lambda d \end{vmatrix} = \lambda(ad - bc) = \lambda \det(A)$. Supposons le résultat acquis pour les matrices d'ordre $n - 1 \geq 2$. Soient A d'ordre n et $A' = A_i(\lambda)$ déduite de A en multipliant sa ligne i par λ . On a alors :

$$\det(A') = (-1)^{i+1} \lambda a_{i,1} \det(A_{i,1}) + \sum_{\substack{k=1 \\ k \neq i}}^n (-1)^{k+1} a_{k,1} \det(A'_{k,1})$$

la matrice $A'_{k,1}$, pour $k \neq i$, étant déduite de $A_{k,1}$ en multipliant sa ligne i par λ . On a donc $\det(A'_{k,1}) = \lambda \det(A_{k,1})$ pour $k \neq i$ et $\det(A') = \lambda \det(A)$. $\square$

Corollaire 8.1. Si $A \in \mathcal{M}_n(\mathbb{K})$ a une ligne nulle, on a alors $\det(A) = 0$.

Preuve. Supposons que la ligne i de A soit nulle. En désignant par $A' = A_i(\lambda)$ la matrice déduite de A en multipliant sa ligne i par $\lambda = 0$, on a $A' = A$ et $\det(A') = \det(A) = 0$ et $\det(A') = 0 \det(A) = 0$. $\square$

Corollaire 8.2. Pour tout $A \in \mathcal{M}_n(\mathbb{K})$ et tout scalaire λ , on a $\det(\lambda A) = \lambda^n \det(A)$.

Preuve. En utilisant n fois le théorème précédent, on a :

$$\det(\lambda A) = \det \begin{pmatrix} \lambda L_1 \\ \lambda L_2 \\ \vdots \\ \lambda L_n \end{pmatrix} = \lambda \det \begin{pmatrix} L_1 \\ \lambda L_2 \\ \vdots \\ \lambda L_n \end{pmatrix} = \dots = \lambda^n \det \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_n \end{pmatrix}$$

$\square$

Théorème 8.4.

Le déterminant d'une matrice triangulaire est égale au produit de ses termes diagonaux, soit $\det(A) = \prod_{i=1}^n a_{ii}$.

Preuve. Considérons tout d'abord le cas des matrices triangulaires inférieures. Pour $n = 1$ c'est clair et pour $n = 2$, on a $\begin{vmatrix} a & 0 \\ c & d \end{vmatrix} = ad - 0 \cdot c = ad$. Supposons le résultat acquis pour les matrices triangulaires inférieures d'ordre $n - 1 \geq 2$ et soit A triangulaire inférieure d'ordre n . La matrice A_{11} est triangulaire inférieure de diagonale $a_{22}, \dots, a_{nn}$ et pour i compris entre 2 et n , la matrice A_{i1} est telle que sa première ligne est nulle, elle est donc de déterminant nul et :

$$\det(A) = a_{1,1} \det(A_{1,1}) = \prod_{i=1}^n a_{ii}$$

Pour le cas des matrices triangulaires supérieures, le cas $n = 1$ est encore évident et le cas $n = 2$ se vérifie par le calcul. Supposant le résultat acquis au rang $n - 1 \geq 2$, pour A triangulaire supérieure d'ordre n , La matrice A_{11} est triangulaire supérieure de diagonale $a_{22}, \dots, a_{nn}$ et pour i compris entre 2 et n , les coefficients a_{i1} sont nuls de sorte que $\det(A) = a_{1,1} \det(A_{1,1}) = \prod_{i=1}^n a_{ii}$. $\square$

De manière plus générale, on peut s'intéresser au calcul du déterminant d'une matrice triangulaire par blocs.

Lemme 8.2 *Pour toute matrice $M = \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} \in \mathcal{M}_n(\mathbb{K})$, où $A \in \mathcal{M}_p(\mathbb{K})$, $B \in \mathcal{M}_{p \times q}(\mathbb{K})$, $D \in \mathcal{M}_q(\mathbb{K})$ et $p + q = n$, on a $\det(M) = \det(A) \det(D)$.*

Preuve. Pour $p = 1$, on a $A = (a_{11})$, $M = \begin{pmatrix} a_{11} & B \\ 0 & D \end{pmatrix}$ où B est une matrice à une ligne et $\det(M) = a_{11} \det(D) = \det(A) \det(D)$. Supposant le résultat acquis au rang $p - 1 \geq 1$, on a $\det(M) = \sum_{i=1}^p (-1)^{i+1} a_{i,1} \det(M_{i,1})$, où $M_{i1} = \begin{pmatrix} A_{i1} & B_{i1} \\ 0 & D \end{pmatrix} \in \mathcal{M}_{n-1}(\mathbb{K})$ avec $A \in \mathcal{M}_{p-1}(\mathbb{K})$, ce qui nous donne en exploitant l'hypothèse de récurrence :

$$\det(M) = \left(\sum_{i=1}^p (-1)^{i+1} a_{i,1} \det(A_{i,1}) \right) \det(D) = \det(A) \det(D)$$

$\square$

Théorème 8.5.

$$\text{Si } A = \begin{pmatrix} A_1 & B_{12} & \cdots & B_{1r} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & B_{r-1,r-1} \\ 0 & \cdots & 0 & A_r \end{pmatrix}, \text{ où les } A_k \text{ sont des matrices carrées, on a alors } \det(A) = \prod_{k=1}^r \det(A_k).$$

Preuve. Se déduit de ce qui précède par récurrence sur $r \geq 1$. $\square$

Exemples 8.1

1. Si $A = I_n$ est la matrice identité, on a alors $\det(I_n) = 1$.
2. Si $A = D_i(\lambda)$ est une matrice de dilatation, on a alors $\det(D_i(\lambda)) = \lambda$.
3. Si $A = T_{ij}(\lambda)$ est une matrice de transvection, on a alors $\det(T_{ij}(\lambda)) = 1$.

Théorème 8.6.

Soient A, A', A'' des matrices de lignes respectives L_i, L'_i, L''_i (pour i compris entre 1 et n) telles que $L_i = L'_i = L''_i$ pour $i \neq k$ et $L'_k = L_k + L'_k$ où k est un indice compris entre 1 et n . On a $\det(A'') = \det(A) + \det(A')$, soit :

$$\det \begin{pmatrix} L_1 \\ \vdots \\ L_{k-1} \\ L_k + L'_k \\ L_{k+1} \\ \vdots \\ L_n \end{pmatrix} = \det \begin{pmatrix} L_1 \\ \vdots \\ L_{k-1} \\ L_k \\ L_{k+1} \\ \vdots \\ L_n \end{pmatrix} + \det \begin{pmatrix} L_1 \\ \vdots \\ L_{k-1} \\ L'_k \\ L_{k+1} \\ \vdots \\ L_n \end{pmatrix}$$

Preuve. Pour $n = 1$ c'est clair et pour $n = 2$, il suffit de calculer. Supposons le résultat acquis pour les matrices d'ordre $n - 1 \geq 2$. Soient A, A', A'' d'ordre n vérifiant les conditions du théorème. On a alors :

$$\det(A'') = (-1)^{k+1} (a_{k,1} + a'_{k,1}) \det(A''_{k,1}) + \sum_{\substack{i=1 \\ i \neq k}}^n (-1)^{i+1} a_{i,1} \det(A''_{i,1})$$

avec $A''_{k,1} = A_{k,1} = A'_{k,1}$ et pour $i \neq k$, les matrices $A_{i,1}, A'_{i,1}, A''_{i,1}$ vérifiant les hypothèses du théorème au rang $n - 1$ (avec des notations évidentes), donc :

$$\begin{aligned} \det(A'') &= (-1)^{k+1} (a_{k,1} \det(A_{k,1}) + a'_{k,1} \det(A'_{k,1})) \\ &\quad + \sum_{\substack{i=1 \\ i \neq k}}^n (-1)^{i+1} a_{i,1} \det(A_{i,1}) + \sum_{\substack{i=1 \\ i \neq k}}^n (-1)^{i+1} a'_{i,1} \det(A'_{i,1}) \\ &= \det(A) + \det(A'). \end{aligned}$$

$\square$

Les théorèmes 8.3 et 8.6 se traduisent en disant que le déterminant est linéaire par rapport à chaque ligne.

Théorème 8.7.

Si A' est la matrice déduite de $A \in \mathcal{M}_n(\mathbb{K})$ en permutant deux lignes, on a alors $\det(A') = -\det(A)$, soit :

$$\det \begin{pmatrix} \vdots \\ L_i \\ \vdots \\ L_j \\ \vdots \end{pmatrix} = -\det \begin{pmatrix} \vdots \\ L_j \\ \vdots \\ L_i \\ \vdots \end{pmatrix}$$

où les pointillés indiquent les lignes inchangées.

Preuve. Pour $n = 2$, il suffit de calculer. Supposons le résultat acquis pour les matrices d'ordre $n-1 \geq 2$. La permutation de deux lignes se faisant avec un nombre impair de permutations de deux lignes successives (par exemple la permutation $(2, 4)$ se fait par les trois permutations $(2, 3, 4) \rightarrow (3, 2, 4) \rightarrow (3, 4, 2) \rightarrow (4, 3, 2)$), il suffit de considérer le cas où $j = i + 1$ (montrer ce point rigoureusement). On se donne donc A d'ordre n et A' est déduite de $A \in \mathcal{M}_n(\mathbb{K})$ en permutant les lignes i et $i + 1$. Pour $k \neq i$ et $k \neq i + 1$, on a $a'_{k,1} = a_{k,1}$ et $\det(A'_{k,1}) = -\det(A_{k,1})$ par hypothèse de récurrence, et avec $a'_{i,1} = a_{(i+1),1}$, $a'_{(i+1),1} = a_{i,1}$, $A'_{i,1} = A_{(i+1),1}$, $A'_{(i+1),1} = A_{i,1}$, on déduit que :

$$\begin{aligned} \det(A') &= (-1)^{i+1} a_{(i+1),1} \det(A_{(i+1),1}) + (-1)^i a_{i,1} \det(A_{i,1}) \\ &\quad - \sum_{\substack{k=1 \\ k \neq i, k \neq i+1}}^n (-1)^{k+1} a_{k,1} \det(A_{k,1}) = -\det(A) \end{aligned}$$

□

Le résultat précédent se traduit en disant que le déterminant est une forme alternée sur les lignes.

Corollaire 8.3. *Si la matrice $A \in \mathcal{M}_n(\mathbb{K})$ a deux lignes identiques, alors $\det(A) = 0$.*

Preuve. Si $L_i = L_j$ avec $i \neq j$, alors matrice A' déduite de A en permutant ces deux lignes est égale à A et $\det(A) = -\det(A)$, donc $\det(A) = 0$. □

Corollaire 8.4. *On ne change pas la valeur d'un déterminant si on ajoute à une ligne une combinaison linéaire des autres lignes.*

Preuve. Il suffit de montrer le résultat quand on ajoute à la ligne L_i la ligne L_j multipliée par un scalaire λ où $i \neq j$. Dans ce cas, on a :

$$\det \begin{pmatrix} \vdots \\ L_i + \lambda L_j \\ \vdots \\ L_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ L_j \\ \vdots \\ L_j \\ \vdots \end{pmatrix} + \lambda \det \begin{pmatrix} \vdots \\ L_j \\ \vdots \\ L_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ L_j \\ \vdots \\ L_j \\ \vdots \end{pmatrix}$$

où les pointillés indiquent les lignes inchangées. $\square$

En effectuant des opérations élémentaires sur les lignes d'une matrice A , on peut se ramener à une matrice triangulaire supérieure de même déterminant que celui de A .

Les théorèmes 8.3, 8.6 et le corollaire 8.1 se traduisent aussi par le résultat suivant.

Corollaire 8.5. *Pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$, toute matrice de dilatation $D_i(\lambda)$ et toute matrice de transvection $T_{ij}(\lambda)$, on a :*

$$\begin{cases} \det(D_i(\lambda) A) = \det(D_i(\lambda)) \det(A) = \lambda \det(A) \\ \det(T_{ij}(\lambda) A) = \det(T_{ij}(\lambda)) \det(A) = \det(A) \end{cases}$$

En utilisant le théorème 8.2, on obtient le résultat suivant.

Théorème 8.8.

Pour toute matrice inversible A et toute matrice B dans $\mathcal{M}_n(\mathbb{K})$, on a $\det(AB) = \det(A) \det(B)$.

Preuve. La matrice A étant inversible s'écrit $A = P_1 \cdots P_r D_n(\lambda) Q_1 \cdots Q_s$, où les matrices P_i et Q_j sont des matrices de transvection et la matrice $D_n(\lambda)$ une matrice de dilatation (théorème 8.2). Une utilisation répétée du corollaire précédent nous donne $\det(A) = \det(D_n(\lambda)) = \lambda$ et pour toute matrice B :

$$\det(AB) = \det(D_n(\lambda)) \det(B) = \det(A) \det(B)$$

$\square$

Le résultat précédent est en fait valable pour toutes matrices A et B . Le cas où la matrice A n'est pas inversible se traite en utilisant le résultat suivant.

Théorème 8.9.

Une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est inversible si, et seulement si, son déterminant est non nul et dans ce cas, on a $\det(A^{-1}) = \frac{1}{\det(A)}$.

Preuve. Si A est inversible d'inverse A^{-1} , on $AA^{-1} = I_n$ et le théorème précédent nous dit que $\det(A) \det(A^{-1}) = \det(I_n) = 1$. On a donc $\det(A) \neq 0$ et

$\det(A^{-1}) = \frac{1}{\det(A)}$. La réciproque se démontre par récurrence sur $n \geq 1$. Pour $n = 1$, le résultat est évident car $\det(a) = a$ pour tout scalaire a . Supposons le résultat acquis pour les matrices d'ordre $n - 1 \geq 1$ et soit A d'ordre n telle que $\det(A) \neq 0$. La première colonne de A est nécessairement non nulle (définition du déterminant) et on peut reprendre la démonstration du théorème 8.2 pour trouver des matrices de transvection $P_1, \dots, P_k$ telles que :

$$P_k \cdots P_1 A = \begin{pmatrix} 1 & \alpha_{12} & \cdots & \alpha_{1n} \\ 0 & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix} = \begin{pmatrix} 1 & \alpha \\ 0 & B \end{pmatrix}$$

où α est un vecteur ligne à $n - 1$ composantes et B une matrice carrée d'ordre $n - 1$. Comme les matrices P_k sont inversibles, on a :

$$\det(A) = \det(P_k \cdots P_1 A) = \det(B)$$

et $\det(B) \neq 0$. La matrice B est donc inversible, ce qui implique que A est aussi inversible. En effet si $Ax = 0$, en notant $x = \begin{pmatrix} x_1 \\ x' \end{pmatrix}$ avec $x_1 \in \mathbb{K}$ et $x' \in \mathbb{K}^{n-1}$, on a $x_1 + \alpha x' = 0$ et $Bx' = 0$, ce qui entraîne $x' = 0$ et $x_1 = 0$, soit $x = 0$. La matrice A est donc inversible. $\square$

Théorème 8.10.

Pour toutes matrices A et B dans $\mathcal{M}_n(\mathbb{K})$, on a :

$$\det(AB) = \det(BA) = \det(A) \det(B)$$

Preuve. Il reste à traiter le cas où la matrice A n'est pas inversible. Dans ce cas la matrice AB ne peut être inversible (sinon, en notant C l'inverse de AB , on a $(AB)C = I_n$, soit $A(BC) = I_n$ et A est inversible) et on a $0 = \det(AB) = \det(A) \det(B) = 0 \cdot \det(B)$. L'égalité $\det(BA) = \det(B) \det(A)$ donne $\det(AB) = \det(BA)$. $\square$

On peut remarquer que $\det(AB) = \det(BA)$ alors qu'en général $AB \neq BA$.

La multiplication à droite par une matrice élémentaire se traduisant par une action particulière sur les colonnes, on déduit de ce théorème et du théorème 8.1 les propriétés suivantes du déterminant.

Corollaire 8.6. Si $A'_j(\lambda)$ est la matrice déduite de $A \in \mathcal{M}_n(\mathbb{K})$ en multipliant sa colonne j par un scalaire λ , on a alors $\det(A'_j(\lambda)) = \lambda \det(A)$. Si $A \in \mathcal{M}_n(\mathbb{K})$ a une colonne nulle, alors $\det(A) = 0$.

Pour l'instant, le déterminant d'une matrice se calcule en utilisant la première colonne de cette dernière. En réalité, on peut aussi utiliser la première ligne et nous en déduisons que cette première ligne ou colonne peut être remplacée par n'importe quelle autre. Précisément, on a les résultats suivants.

Théorème 8.11.

Pour toute matrice A dans $\mathcal{M}_n(\mathbb{K})$, on a $\det({}^t A) = \det(A)$.

Preuve. Comme d'habitude c'est trivial pour $n = 1$. On suppose donc que $n \geq 2$. Si A n'est pas inversible, il en est de même de sa transposée (en effet si ${}^t A$ est inversible, il en est de même de $A = {}^t({}^t A)$ – théorème 6.19 –) et on a alors $\det({}^t A) = \det(A) = 0$. Si A est inversible, elle s'écrit $A = P_1 \cdots P_r D_n(\lambda) Q_1 \cdots Q_s$, où les P_i, Q_j sont des matrices de transvection et $\lambda = \det(A)$, ce qui donne :

$${}^t A = {}^t Q_s \cdots {}^t Q_1 {}^t D_n(\lambda) {}^t P_r \cdots {}^t P_1$$

les transposées de matrices élémentaires étant des matrices élémentaires de même type avec ${}^t D_n(\lambda) = D_n(\lambda)$, ce qui donne $\det({}^t A) = \det(D_n(\lambda)) = \lambda = \det(A)$.
□

De ce résultat, on déduit le développement du déterminant suivant la première ligne (pour $n \geq 2$) :

$$\det(A) = \det({}^t A) = \sum_{j=1}^n (-1)^{j+1} a_{1,j} \det(A_{1,j})$$

où $A_{1,j}$ est la matrice carrée d'ordre $n - 1$ déduite de A en supprimant la ligne 1 et la colonne j .

On en déduit alors les propriétés suivantes relatives aux colonnes, ces propriétés étant analogues à celles obtenues pour les lignes.

Corollaire 8.7. Si A' est la matrice déduite de $A \in \mathcal{M}_n(\mathbb{K})$ en permutant deux colonnes, on a alors $\det(A') = -\det(A)$. Soient A, A', A'' des matrices de lignes respectives C_j, C'_j, C''_j (pour j compris entre 1 et n) telles que $C_j = C'_j = C''_j$ pour $j \neq k$ et $C''_k = C_k + C'_k$ où k est un indice compris entre 1 et n . On a $\det(A'') = \det(A) + \det(A')$. On ne change pas la valeur d'un déterminant si on ajoute à une ligne une combinaison linéaire des autres lignes.

Ce corollaire se traduit en disant que le déterminant est linéaire par rapport à chaque colonne et que c'est une forme alternée sur les colonnes.

En général, pour calculer un déterminant, on essaiera d'effectuer des opérations élémentaires sur les lignes ou les colonnes dans le but de faire apparaître un maximum de coefficients nuls, ce qui facilitera le calcul du déterminant de la matrice obtenue.

De tout ce qui précède, on déduit les différentes formes de développement d'un déterminant suivant une ligne ou une colonne (pour $n \geq 2$).

Théorème 8.12.

Pour toute matrice $A \in \mathcal{M}_n(\mathbb{K})$, on a :

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{i,j} \det(A_{i,j}) \quad (1 \leq j \leq n)$$

(développement suivant la colonne j) et

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{i,j} \det(A_{i,j}) \quad (1 \leq i \leq n)$$

(développement suivant la ligne i) où $A_{i,j}$ est la matrice carrée d'ordre $n-1$ déduite de A en supprimant la ligne i et la colonne j .

Preuve. Pour $j = 1$, c'est la définition première du déterminant et pour $i = 1$ c'est une conséquence immédiate de $\det({}^t A) = \det(A)$. Fixons la colonne $j \geq 2$ et notons $(e_i)_{1 \leq i \leq n}$ la base canonique de $\mathbb{K}^n$. La colonne C_j s'écrit $C_j = \sum_{i=1}^n a_{ij} e_i$ et en utilisant la linéarité du déterminant par rapport à la j -ième colonne, on a $\det(A) = \sum_{i=1}^n a_{ij} \det(B_{i,j})$, où $B_{i,j}$ est la matrice déduite de A en remplaçant C_j par e_i . En permutant la colonne j avec la colonne $j-1$, puis $j-1$ avec $j-2, \dots, 2$ avec 1 et ensuite la ligne i avec la ligne $i-1$, $i-1$ avec $i-2, \dots, 2$ avec 1 (on fait rien pour $i = 1$) on aboutit à :

$$\begin{aligned} \det(B_{i,j}) &= (-1)^{i+j} \begin{vmatrix} 1 & a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ 0 & a_{21} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix} \\ &= (-1)^{i+j} \det(A_{i,j}) \end{aligned}$$

et on a le résultat annoncé. On procède de manière analogue pour la deuxième formule. $\square$

Avec les notations du théorème, on dit que $\det(A_{i,j})$ est le mineur d'indice (i, j) de la matrice A et que $(-1)^{i+j} \det(A_{i,j})$ est le cofacteur d'indice (i, j) de A .

8.3 Déterminant d'une famille de vecteurs

Étant donnée une famille $(x_j)_{1 \leq j \leq n}$ de n vecteurs de $\mathbb{K}^n$, on définit le déterminant de cette famille comme le déterminant de la matrice A dont les colonnes sont

formées de ces vecteurs. En notant, pour j compris entre 1 et n , $x_j = (x_{i,j})_{1 \leq i \leq n}$ (vecteur colonne), on a donc $\det(x_1, \dots, x_n) = \det((x_{i,j}))_{1 \leq i,j \leq n}$.

Du théorème 8.9 on déduit le résultat suivant bien utile pour vérifier qu'un système de n vecteurs dans $\mathbb{K}^n$ est libre et donc forme une base.

Théorème 8.13.

Une famille $(x_j)_{1 \leq j \leq n}$ de n vecteurs de $\mathbb{K}^n$ est libre si, et seulement si, son déterminant est non nul.

Preuve. En utilisant les notations qui précèdent, on note $P = ((x_{i,j}))_{1 \leq i,j \leq n}$. Dire que le système $(x_j)_{1 \leq j \leq n}$ est libre équivaut à dire que l'unique solution $\lambda \in \mathbb{K}^n$ du système linéaire $\sum_{j=1}^n \lambda_j x_j = 0$ est $\lambda = 0$, ce système s'écrivant aussi $P\lambda = 0$,

cela revient à dire que la matrice P est inversible, ce qui est encore équivalent à $\det(P) \neq 0$. $\square$

Si E un espace vectoriel réel ou complexe de dimension $n \geq 1$ et $\mathcal{B}$ une base E , en notant pour tout vecteur $x \in E$, X le vecteur colonne de $\mathbb{K}^n$ formé des composantes de x dans la base $\mathcal{B}$, on définit le déterminant d'une famille $(x_j)_{1 \leq j \leq n}$ de n vecteurs de E dans la base $\mathcal{B}$ par $\det_{\mathcal{B}}(x_1, \dots, x_n) = \det(X_1, \dots, X_n)$.

Ce déterminant dépend du choix d'une base de E .

Théorème 8.14.

Si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de E , alors pour tout n -uplet $(x_1, x_2, \dots, x_n)$ de vecteurs de E , on a $\det_{\mathcal{B}}(x_1, x_2, \dots, x_n) = \det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}'}(x_1, x_2, \dots, x_n)$.

Preuve. En désignant pour tout vecteur x de E par X [resp. X'] le vecteur colonne de $\mathbb{K}^n$ formé des composantes de x dans la base $\mathcal{B}$ [resp. $\mathcal{B}'$] et par P la matrice de passage de $\mathcal{B}$ à $\mathcal{B}'$, on a $X = PX'$ et :

$$\begin{aligned} \det_{\mathcal{B}}(x_1, x_2, \dots, x_n) &= \det(X_1, X_2, \dots, X_n) = \det(PX'_1, PX'_2, \dots, PX'_n) \\ &= \det(P \cdot (X'_1, X'_2, \dots, X'_n)) = \det(P) \det((X'_1, X'_2, \dots, X'_n)) \\ &= \det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}'}(x_1, x_2, \dots, x_n) \end{aligned}$$

$\square$

8.4 Déterminant d'un endomorphisme

On désigne par E un espace vectoriel réel ou complexe de dimension $n \geq 1$.

Si $\mathcal{B}_1$ et $\mathcal{B}_2$ sont deux bases de E , u un endomorphisme de E , A_1 la matrice de u dans la base $\mathcal{B}_1$ et A_2 sa matrice dans la base $\mathcal{B}_2$, on sait alors que $A_2 = P^{-1}AP$ où P est la matrice de passage de $\mathcal{B}_1$ à $\mathcal{B}_2$. Il en résulte alors que :

$$\begin{aligned} \det(A_2) &= \det(P^{-1}AP) = \det(P^{-1}) \det(A_1) \det(P) \\ &= \frac{1}{\det(P)} \det(A_1) \det(P) = \det(A_1) \end{aligned}$$

C'est-à-dire que ces déterminants ne dépendent pas du choix d'une base. On peut alors donner la définition suivante.

Définition 8.6. *Le déterminant d'un endomorphisme u de E est le déterminant de la matrice de u dans une base de E .*

Du théorème 7.19, on déduit le résultat suivant.

Théorème 8.15.

Si u, v sont deux endomorphismes de E , on a alors :

$$\det(u \circ v) = \det(v \circ u) = \det(u) \det(v)$$

Et du théorème 7.20, on déduit le résultat suivant.

Théorème 8.16.

Un endomorphisme u de E est inversible si, et seulement si, son déterminant est non nul.

8.5 Exercices

Exercice 8.1. Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ une matrice inversible.

1. On suppose que $c \neq 0$. Déterminer des scalaires $\lambda_1, \lambda_2, \lambda_3$ tels que

$$A_1 = T_{12}(\lambda_1) A = \begin{pmatrix} 1 & b_1 \\ c_1 & d_1 \end{pmatrix}, \quad A_2 = T_{21}(\lambda_2) A_1 = \begin{pmatrix} 1 & b_2 \\ 0 & d_2 \end{pmatrix}$$

et $A_3 = A_2 T_{12}(\lambda_3) = \begin{pmatrix} 1 & 0 \\ 0 & d_1 \end{pmatrix}$. En déduire qu'il existe des matrices de transvection P_1, P_2, Q_1 et une matrice de dilatation D telles que $A = P_1 P_2 D Q_1$.

2. Donner un résultat analogue dans le cas où $c = 0$.

Solution.

1. Pour tout scalaire λ_1 , on a :

$$T_{12}(\lambda_1) A = \begin{pmatrix} 1 & \lambda_1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a + c\lambda_1 & b + d\lambda_1 \\ c & d \end{pmatrix}$$

et prenant λ_1 tel que $a + c\lambda_1 = 1$, soit $\lambda_1 = \frac{1-a}{c}$, on a :

$$T_{12}(\lambda_1) A = \begin{pmatrix} 1 & \frac{d - \det(A)}{c} \\ c & d \end{pmatrix}$$

Pour tout scalaire λ_2 , on a :

$$T_{21}(\lambda_2)A_1 = \begin{pmatrix} 1 & 0 \\ \lambda_2 & 1 \end{pmatrix} \begin{pmatrix} 1 & b_1 \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & b_1 \\ c + \lambda_2 & d + b_1\lambda_2 \end{pmatrix}$$

et prenant $\lambda_2 = -c$, on a :

$$T_{21}(\lambda_2)A_1 = \begin{pmatrix} 1 & b_1 \\ 0 & d - cb_1 \end{pmatrix} = \begin{pmatrix} 1 & \frac{d - \det(A)}{c} \\ 0 & \det(A) \end{pmatrix}$$

Pour tout scalaire λ_3 , on a :

$$A_2T_{12}(\lambda_3) = \begin{pmatrix} 1 & b_2 \\ 0 & d_2 \end{pmatrix} \begin{pmatrix} 1 & \lambda_3 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & b_2 + \lambda_3 \\ 0 & d_2 \end{pmatrix}$$

et prenant $\lambda_3 = -b_2$, on a $A_2T_{12}(\lambda_3) = \begin{pmatrix} 1 & 0 \\ 0 & d_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \det(A) \end{pmatrix}$. On a donc en définitive $T_{21}(\lambda_2)T_{12}(\lambda_1)AT_{12}(\lambda_3) = D(\det(A))$ et utilisant le fait que les matrices de transvections sont inversibles, on déduit que :

$$A = T_{12}(-\lambda_1)T_{21}(-\lambda_2)D(\det(A))T_{12}(-\lambda_3)$$

$$\text{où } \lambda_1 = \frac{1-a}{c}, \lambda_2 = -c \text{ et } \lambda_3 = \frac{\det(A) - d}{c}.$$

2. Si $c = 0$, on a nécessairement $a \neq 0$ puisque A est inversible et :

$$T_{21}(1)A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} = \begin{pmatrix} a & b \\ a & b + d \end{pmatrix}$$

ce qui nous ramène au cas précédent et donne $T_{21}(1)A = P_1P_2DQ_1$, ou encore $A = T_{21}(-1)P_1P_2DQ_1$.

Exercice 8.2. Calculer le déterminant de $A = \begin{pmatrix} 5 & 4 & 2 & 1 \\ 2 & 3 & 1 & -2 \\ -5 & -7 & -3 & 9 \\ 1 & -2 & -1 & 4 \end{pmatrix}$ en effectuant des opérations élémentaires.

Solution. Les opérations $L_2 \rightarrow L_2 - \frac{2}{5}L_1$, $L_3 \rightarrow L_3 + L_1$, $L_4 \rightarrow L_4 - \frac{1}{5}L_1$ donnent :

$$\begin{aligned} \det(A) &= \begin{vmatrix} 5 & 4 & 2 & 1 \\ 0 & \frac{7}{5} & \frac{1}{5} & -\frac{12}{5} \\ 0 & -3 & -1 & 10 \\ 0 & -\frac{14}{5} & -\frac{7}{5} & \frac{19}{5} \end{vmatrix} = 5 \begin{vmatrix} \frac{7}{5} & \frac{1}{5} & -\frac{12}{5} \\ -3 & -1 & 10 \\ -\frac{14}{5} & -\frac{7}{5} & \frac{19}{5} \end{vmatrix} \\ &= 5 \frac{1}{5} \frac{1}{5} \begin{vmatrix} 7 & 1 & -12 \\ -3 & -1 & 10 \\ -14 & -7 & 19 \end{vmatrix} = \frac{1}{5} \begin{vmatrix} 7 & 1 & -12 \\ -3 & -1 & 10 \\ -14 & -7 & 19 \end{vmatrix} \end{aligned}$$

Puis les opérations $L_2 \rightarrow L_2 + \frac{3}{7}L_1$, $L_3 \rightarrow L_3 + \frac{14}{7}L_1 = L_3 + 2L_1$ donnent :

$$\det(A) = \frac{1}{5} \begin{vmatrix} 7 & 1 & -12 \\ 0 & -\frac{4}{7} & \frac{34}{7} \\ 0 & -5 & -5 \end{vmatrix} = \frac{1}{5} \cdot 7 \cdot \frac{2}{7} \cdot 5 \begin{vmatrix} -2 & 17 \\ -1 & -1 \end{vmatrix} = 2 \cdot 19 = 38$$

Exercice 8.3. Développer le déterminant de la matrice suivante sous la forme d'un produit de facteurs linéaires en x :

$$A(x) = \begin{pmatrix} x+2 & 2x+3 & 3x+4 \\ 2x+3 & 3x+4 & 4x+5 \\ 3x+5 & 5x+8 & 10x+17 \end{pmatrix}$$

Solution. Les opérations $L_3 \rightarrow L_3 - L_2$, $L_2 \rightarrow L_2 - L_1$ (dans l'ordre indiqué) donnent :

$$\det(A(x)) = \begin{vmatrix} x+2 & 2x+3 & 3x+4 \\ x+1 & x+1 & x+1 \\ x+2 & 2x+4 & 6x+12 \end{vmatrix} = (x+1)(x+2) \begin{vmatrix} x+2 & 2x+3 & 3x+4 \\ 1 & 1 & 1 \\ 1 & 2 & 6 \end{vmatrix}$$

puis $L_3 \rightarrow L_3 - L_2$ donne :

$$\begin{aligned} \det(A(x)) &= (x+1)(x+2) \begin{vmatrix} x+2 & 2x+3 & 3x+4 \\ 1 & 1 & 1 \\ 0 & 1 & 5 \end{vmatrix} \\ &= (x+1)(x+2) \left((x+2) \begin{vmatrix} 1 & 1 \\ 1 & 5 \end{vmatrix} - \begin{vmatrix} 2x+3 & 3x+4 \\ 1 & 5 \end{vmatrix} \right) \\ &= (x+1)(x+2) (4(x+2) - (7x+11)) \\ &= -(x+1)(x+2)(3x+3) = -3(x+1)^2(x+2) \end{aligned}$$

Exercice 8.4. Soient α, β deux scalaires et $A(\alpha, \beta) = ((a_{ij}))_{1 \leq i, j \leq n}$ la matrice d'ordre $n \geq 3$ définie par :

$$\forall i \in \{1, 2, \dots, n\}, \quad \begin{cases} a_{ii} = \beta \\ a_{ij} = \alpha \text{ si } j \in \{1, 2, \dots, n\} - \{i\} \end{cases}$$

Calculer $\Delta(\alpha, \beta) = \det(A(\alpha, \beta))$.

Solution. La matrice $A(\alpha, \beta)$ est de la forme $A(\alpha, \beta) = \begin{pmatrix} \beta & \alpha & \alpha & \cdots & \alpha \\ \alpha & \beta & \alpha & \cdots & \alpha \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \alpha & \cdots & \alpha & \beta & \alpha \\ \alpha & \cdots & \alpha & \alpha & \beta \end{pmatrix}$.

Si $\alpha = 0$, la matrice est diagonale et on a $\Delta(0, \beta) = \beta^n$. On suppose que $\alpha \neq 0$.

En ajoutant les lignes 2 à n à la première ligne on a :

$$\Delta(\alpha, \beta) = (\beta + (n-1)\alpha) \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ \alpha & \beta & \alpha & \cdots & \alpha \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \alpha & \cdots & \alpha & \beta & \alpha \\ \alpha & \cdots & \alpha & \alpha & \beta \end{vmatrix}$$

Puis en retranchant la première ligne multipliée par α aux lignes 2 à n on obtient :

$$\begin{aligned} \Delta(\alpha, \beta) &= (\beta + (n-1)\alpha) \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & \beta - \alpha & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \beta - \alpha & 0 \\ 0 & \cdots & 0 & 0 & \beta - \alpha \end{vmatrix} \\ &= (\beta + (n-1)\alpha)(\beta - \alpha)^{n-1} \end{aligned}$$

Exercice 8.5. Sachant que 1700, 1020, 1122 et 1309 sont divisibles par

$$17, \text{ montrer sans le calculer que } D = \begin{vmatrix} 1 & 7 & 0 & 0 \\ 1 & 0 & 2 & 0 \\ 1 & 1 & 2 & 2 \\ 1 & 3 & 0 & 9 \end{vmatrix} \text{ est divisible par } 17.$$

Solution. On ne change pas la valeur de ce déterminant si on remplace la colonne 4 par $C_4 + 10C_3 + 10^2C_2 + 10^3C_1$, ce qui donne :

$$D = \begin{vmatrix} 1 & 7 & 0 & 1700 \\ 1 & 0 & 2 & 1020 \\ 1 & 1 & 2 & 1122 \\ 1 & 3 & 0 & 1309 \end{vmatrix} = 17 \begin{vmatrix} 1 & 7 & 0 & 100 \\ 1 & 0 & 2 & 60 \\ 1 & 1 & 2 & 66 \\ 1 & 3 & 0 & 77 \end{vmatrix} = 17q$$

avec q entier puisque tous les coefficients du déterminant sont entiers.

Exercice 8.6. Soient $n \geq 2$ un entier et $\alpha_1, \alpha_2, \dots, \alpha_n$ des scalaires. Calculer le déterminant $\Delta(\alpha_1, \dots, \alpha_n)$ de la matrice :

$$V(\alpha_1, \dots, \alpha_n) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \cdots & \alpha_n^{n-1} \end{pmatrix}$$

Une telle matrice est dite de Vandermonde. À quelle condition une telle matrice est-elle inversible ?

Solution. Pour $n = 2$, on a $\Delta(\alpha_1, \alpha_2) = \alpha_2 - \alpha_1$. Le calcul de $\Delta(\alpha_1, \dots, \alpha_n)$ se fait par récurrence sur $n \geq 2$. En retranchant, pour $i = n, n-1, \dots, 2$ à la ligne i la ligne $i-1$ multipliée par α_1 , on obtient :

$$\begin{aligned} \Delta(\alpha_1, \dots, \alpha_n) &= \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & \alpha_2 - \alpha_1 & \cdots & \alpha_n - \alpha_1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \alpha_2^{n-2}(\alpha_2 - \alpha_1) & \cdots & \alpha_n^{n-2}(\alpha_n - \alpha_1) \end{vmatrix} \\ &= \begin{vmatrix} \alpha_2 - \alpha_1 & \alpha_3 - \alpha_1 & \cdots & \alpha_n - \alpha_1 \\ \alpha_2(\alpha_2 - \alpha_1) & \alpha_3(\alpha_3 - \alpha_1) & \cdots & \alpha_n(\alpha_n - \alpha_1) \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_2^{n-2}(\alpha_2 - \alpha_1) & \alpha_3^{n-2}(\alpha_3 - \alpha_1) & \cdots & \alpha_n^{n-2}(\alpha_n - \alpha_1) \end{vmatrix} \end{aligned}$$

soit :

$$\begin{aligned} \Delta(\alpha_1, \dots, \alpha_n) &= \left(\prod_{k=2}^n (\alpha_k - \alpha_1) \right) \begin{vmatrix} 1 & \cdots & 1 \\ \alpha_2 & \cdots & \alpha_n \\ \vdots & \ddots & \vdots \\ \alpha_2^{n-2} & \cdots & \alpha_n^{n-2} \end{vmatrix} \\ &= \left(\prod_{k=2}^n (\alpha_k - \alpha_1) \right) \Delta(\alpha_2, \dots, \alpha_n) \end{aligned}$$

et par récurrence, on a :

$$\det(A_n) = \prod_{k=2}^n (\alpha_k - \alpha_1) \prod_{2 \leq i < j \leq n} (\alpha_j - \alpha_i) = \prod_{1 \leq i < j \leq n} (\alpha_j - \alpha_i)$$

Cette matrice est inversible si, et seulement si, les α_i sont deux à deux distincts.

Exercice 8.7. Calculer le déterminant de $A_n = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 2 & 2^2 & \cdots & 2^n \\ \vdots & \vdots & \ddots & \vdots \\ n & n^2 & \cdots & n^n \end{pmatrix}$.

Solution. On a :

$$\begin{aligned} \det(A_n) &= n! \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & n & \cdots & n^{n-1} \end{vmatrix} = n! \Delta(1, 2, \dots, n) \\ &= n! \prod_{1 \leq j < i \leq n} (i - j) = n! \prod_{i=2}^n (i-1)! = \prod_{i=2}^n i! \end{aligned}$$

Exercice 8.8. Soit $A_n = \begin{pmatrix} a_1 & c_1 & 0 & \cdots & 0 \\ b_2 & a_2 & c_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & b_{n-1} & a_{n-1} & c_{n-1} \\ 0 & \cdots & 0 & b_n & a_n \end{pmatrix}$ une matrice

tridiagonale d'ordre $n \geq 3$ à coefficients réels ou complexes. Pour tout entier k compris entre 1 et n , on désigne par D_k le déterminant de la matrice d'ordre k formée des k premières lignes et k premières colonnes de A_n (les D_k sont les déterminants extraits principaux de A_n).

1. Exprimer, pour tout k compris entre 3 et n , D_k en fonction de D_{k-1} et D_{k-2} .
2. Calculer le déterminant de :

$$A_n = \begin{pmatrix} 2 & 1 & 0 & \cdots & 0 \\ 2^2 & 5 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & (n-1)^2 & 2n-1 & 1 \\ 0 & \cdots & 0 & n^2 & 2n+1 \end{pmatrix}$$

Solution.

1. En développant D_k suivant la dernière ligne on a :

$$\begin{aligned} D_k &= a_k \begin{vmatrix} a_1 & c_1 & 0 & \cdots & 0 \\ b_2 & a_2 & c_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & b_{k-2} & a_{k-2} & c_{k-2} \\ 0 & \cdots & 0 & b_{k-1} & a_{k-1} \end{vmatrix} - b_k \begin{vmatrix} a_1 & c_1 & 0 & \cdots & 0 \\ b_2 & a_2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & c_{k-3} & 0 \\ \vdots & \ddots & b_{k-2} & a_{k-2} & 0 \\ 0 & \cdots & 0 & b_{k-1} & c_{k-1} \end{vmatrix} \\ &= a_k D_{k-1} - b_k c_{k-1} D_{k-2}. \end{aligned}$$

Ce qui donne, avec les valeurs initiales $D_1 = a_1$ et $D_2 = a_1 a_2 - b_2 c_1$, un algorithme de calcul de D_n .

2. On a $D_n = (2n+1) D_{n-1} - n^2 D_{n-2}$ avec les valeurs initiales $D_1 = 2$, $D_2 = 6$. En calculant D_3 et D_4 on conjecture que $D_n = (n+1)!$. Ce qui se montre par récurrence sur $n \geq 2$. C'est vrai pour $n = 2$ et le supposant acquis jusqu'au rang $n-1 \geq 2$, on a $D_n = (2n+1)n! - n^2(n-1)! = (n+1)!$

Chapitre 9

Polynômes à coefficients réels ou complexes

$\mathbb{K}$ désigne le corps de réels ou des complexes

Pour tout couple d'entiers $(n, k) \in \mathbb{N}^2$, on note :

$$\delta_{n,k} = \begin{cases} 1 & \text{si } k = n \\ 0 & \text{si } k \neq n \end{cases}$$

(symboles de Kronecker).

9.1 L'algèbre $\mathbb{K}[X]$ des polynômes à coefficients dans $\mathbb{K}$

On désigne par $\mathbb{K}^{\mathbb{N}}$ l'ensemble des suite $a = (a_k)_{k \in \mathbb{N}}$ d'éléments de $\mathbb{K}$. On écrira aussi une telle suite sous la forme $a = (a_0, a_1, \dots, a_n, \dots)$. Muni des opérations d'addition et de multiplication externe définies par :

$$a + b = (a_k + b_k)_{k \in \mathbb{N}} \text{ et } \lambda a = (\lambda a_k)_{k \in \mathbb{N}}$$

(où a, b sont dans $\mathbb{K}^{\mathbb{N}}$ et λ dans $\mathbb{K}$), cet ensemble est un $\mathbb{K}$ -espace vectoriel.

On définit une multiplication interne sur $\mathbb{K}^{\mathbb{N}}$ en posant pour $a = (a_k)_{k \in \mathbb{N}}$, $b = (b_k)_{k \in \mathbb{N}}$ dans $\mathbb{K}^{\mathbb{N}}$, $ab = (c_k)_{k \in \mathbb{N}}$, où les c_k sont définis par :

$$\forall k \in \mathbb{N}, c_k = \sum_{j=0}^k a_j b_{k-j} = \sum_{j=0}^k a_{k-j} b_j = \sum_{\substack{0 \leq i, j \leq k \\ i+j=k}} a_i b_j$$

Muni de cette multiplication, $\mathbb{K}^{\mathbb{N}}$ est un anneau commutatif, unitaire, le neutre pour le produit étant $P_0 = (\delta_{0,k})_{k \in \mathbb{N}} = (1, 0, 0, \dots, 0, \dots)$.

Le produit que nous avons défini dans $\mathbb{K}^{\mathbb{N}}$ n'est pas le produit usuelle des suites ($ab = (a_k b_k)_{k \in \mathbb{N}}$), il s'agit en fait du produit de Cauchy des séries formelles.

Le fait que $\mathbb{K}^{\mathbb{N}}$ soit muni d'une structure d'espace vectoriel et d'une structure d'anneau commutatif unitaire se traduit en disant que $\mathbb{K}^{\mathbb{N}}$ est une $\mathbb{K}$ -algèbre commutative et unitaire.

Définition 9.1. Le support d'une suite $a = (a_k)_{k \in \mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$ est l'ensemble $\text{Supp}(a) = \{k \in \mathbb{N} \mid a_k \neq 0\}$, ce support pouvant être vide.

Définition 9.2. On appelle polynôme à coefficients dans $\mathbb{K}$ toute suite $P = (a_k)_{k \in \mathbb{N}} \in \mathbb{K}^{\mathbb{N}}$ de support fini.

Un polynôme est donc une suite d'éléments de $\mathbb{K}$ dont tous les termes, sauf un nombre fini d'entre eux, sont nuls. On peut aussi dire qu'un polynôme est une suite nulle à partir d'un certain rang, c'est-à-dire telle qu'il existe un entier $m \in \mathbb{N}$ tel que $a_k = 0$ pour tout $k \geq m$.

On note 0 le polynôme nul, c'est-à-dire celui dont tous les termes sont nuls et 1 le polynôme $P_0 = (\delta_{0,k})_{k \in \mathbb{N}}$.

On note, temporairement, $\mathbb{K}^{(\mathbb{N})}$ l'ensemble de ces polynômes.

Si $P = (a_k)_{k \in \mathbb{N}}$ est un polynôme non nul, il existe alors un nombre fini d'indices $k \in \mathbb{N}$ tels que $a_k \neq 0$ et on peut noter :

$$\deg(P) = \max \{k \in \mathbb{N} \mid a_k \neq 0\}, \quad \text{val}(P) = \min \{k \in \mathbb{N} \mid a_k \neq 0\}$$

Définition 9.3. Avec les notations qui précèdent, on dit que $\deg(P)$ est le degré de P et $\text{val}(P)$ la valuation de P .

Par convention, on pose $\deg(0) = -\infty$ et $\text{val}(0) = +\infty$.

Un polynôme non nul de degré $n \geq 0$ peut être noté $P = (a_0, \dots, a_n, 0, \dots, 0, \dots)$ et on dit que $a_n \neq 0$ est le coefficient dominant de P .

Théorème 9.1.

Pour tous P, Q dans $\mathbb{K}^{(\mathbb{N})}$, on a $\deg(P + Q) \leq \max(\deg(P), \deg(Q))$ avec égalité si $\deg(P) \neq \deg(Q)$, $\text{val}(P + Q) \geq \min(\text{val}(P), \text{val}(Q))$ avec égalité si $\text{val}(P) \neq \text{val}(Q)$, $\deg(PQ) = \deg(P) + \deg(Q)$ et $\text{val}(PQ) = \text{val}(P) + \text{val}(Q)$.

Preuve. Pour $P = 0$, on a $\deg(P) = -\infty$ et $P + Q = Q$, $PQ = 0$, donc $\deg(P + Q) = \deg(Q) = \max(\deg(P), \deg(Q))$ et $\deg(PQ) = -\infty = \deg(P) + \deg(Q)$. Comme P et Q jouent des rôles symétriques, on a un résultat analogue dans le cas où $Q = 0$.

Pour $P = (a_0, \dots, a_n, 0, \dots, 0, \dots)$ et $Q = (b_0, \dots, b_m, 0, \dots, 0, \dots)$ non nuls dans $\mathbb{K}^{(\mathbb{N})}$ avec $n = \deg(P) \leq m = \deg(Q)$, on a :

$$P + Q = (a_0 + b_0, \dots, a_m + b_m, 0, \dots, 0, \dots)$$

(on a $a_k = 0$ pour $k \geq n + 1$) avec $a_m + b_m = b_m \neq 0$ si $n < m$, donc $P + Q$ est de degré au plus égal à m , l'égalité étant réalisée pour $n < m$, et :

$$PQ = (c_0, \dots, c_{n+m}, 0, \dots, 0, \dots)$$

avec $c_{n+m} = a_n b_m \neq 0$, donc PQ est de degré $n + m$. On procède de manière analogue pour la valuation. $\square$

Avec la démonstration qui précède, on a aussi prouvé que l'anneau $\mathbb{K}^{(\mathbb{N})}$ est intègre (on a vu que pour P, Q non nuls, PQ est de degré $nm \neq -\infty$, donc $PQ \neq 0$).

Théorème 9.2.

$\mathbb{K}^{(\mathbb{N})}$ est une sous-algèbre de $\mathbb{K}^{\mathbb{N}}$ (i. e. un sous-anneau et un sous-espace vectoriel) et la famille $(P_n)_{n \in \mathbb{N}}$ de polynômes définie par :

$$\forall n \in \mathbb{N}, P_n = (\delta_{n,k})_{k \in \mathbb{N}} = (0, \dots, 0, 1, 0, \dots, 0, \dots)$$

en est une base.

Preuve. Le polynôme nul 0 et le polynôme $1 = P_0$ sont dans $\mathbb{K}^{(\mathbb{N})}$ et on vérifie facilement que si P, Q sont dans $\mathbb{K}^{(\mathbb{N})}$ et λ dans $\mathbb{K}$, alors $P + \lambda Q$ et PQ sont dans $\mathbb{K}^{(\mathbb{N})}$. Donc $\mathbb{K}^{(\mathbb{N})}$ est une sous-algèbre unitaire de $\mathbb{K}^{\mathbb{N}}$. Si $P = (a_k)_{k \in \mathbb{N}}$ est un polynôme

non nul de degré n , il s'écrit $P = \sum_{k=0}^n a_k P_k$, donc la famille $(P_n)_{n \in \mathbb{N}}$ engendre $\mathbb{K}^{(\mathbb{N})}$.

Pour tout entier n , l'égalité $\sum_{k=0}^n a_k P_k = 0$ équivaut à $(a_0, \dots, a_n, 0, \dots, 0, \dots) = 0$, qui est encore équivalent à $a_k = 0$ pour tout k compris entre 0 et n . La famille $(P_n)_{n \in \mathbb{N}}$ est donc libre et c'est une base de $\mathbb{K}^{(\mathbb{N})}$. $\square$

L'application $a_0 \in \mathbb{K} \mapsto (a_0, 0, \dots, 0, \dots) = a_0 P_0 \in \mathbb{K}^{(\mathbb{N})}$ réalise un morphisme d'anneaux injectif de $\mathbb{K}$ dans $\mathbb{K}^{(\mathbb{N})}$, ce qui permet d'identifier $\mathbb{K}$ au sous-anneau de $\mathbb{K}^{(\mathbb{N})}$ formé du polynôme nul et des polynômes non nuls de degré 0. Un polynôme $a_0 P_0$, identifié à $a_0 \in \mathbb{K}$, est appelé polynôme constant.

On a noté $1 = P_0 = (1, 0, \dots, 0, \dots)$ le neutre de $\mathbb{K}^{(\mathbb{N})}$ pour le produit et en notant $X = P_1 = (0, 1, 0, \dots, 0, \dots)$, on a $X^n = P_n$ pour tout $n \in \mathbb{N}^*$. En effet, c'est vrai pour $n = 1$ et supposant le résultat acquis pour $n \geq 1$, on a :

$$X^{n+1} = X \cdot X^n = P_1 P_n = (\delta_{1,k}) (\delta_{n,k}) = (c_k)_{k \in \mathbb{N}}$$

avec $c_{n+1} = \sum_{j=0}^{n+1} \delta_{1,j} \delta_{n,n+1-j} = \delta_{n,n} = 1$ et pour $k \neq n+1$, on a $c_k = \sum_{j=0}^k \delta_{1,j} \delta_{n,k-j} = \delta_{n,k-1} = 0$, soit $c_k = \delta_{n+1,k}$ pour tout k et $X^{n+1} = P_{n+1}$.

En définitive, un polynôme non nul de degré n s'écrit $P = \sum_{k=0}^n a_k X^k$, en convenant que $X^0 = 1$. On notera aussi $P(X) = \sum_{k=0}^n a_k X^k$. On dit alors que X est une indéterminée et on note $\mathbb{K}[X]$ l'algèbre des polynômes à coefficients dans $\mathbb{K}$.

Un monôme est un polynôme de la forme λX^n où $\lambda \in \mathbb{K}$ et $n \in \mathbb{N}$ (pour $n = 0$, il s'agit d'un polynôme constant).

Pour tout entier naturel n , on note $\mathbb{K}_n[X]$ le sous-ensemble de $\mathbb{K}[X]$ formé des polynômes de degré au plus égal à n . On vérifie facilement que $\mathbb{K}_n[X]$ est un

sous-espace vectoriel de $\mathbb{K}[X]$. Ce sous-espace $\mathbb{K}_n[X]$ est de dimension $n+1$, la famille $(X^k)_{0 \leq k \leq n}$ en étant une base (la base canonique de $\mathbb{K}_n[X]$).

Un polynôme $P(X) = \sum_{k=0}^n a_k X^k$ de degré $n \geq 0$ est dit unitaire (ou normalisé) si $a_n = 1$.

Théorème 9.3.

Les éléments inversibles de l'anneau $\mathbb{K}[X]$ sont les polynômes constants non nuls.

Preuve. Si $P \in \mathbb{K}[X]$ est inversible, il existe alors un polynôme $Q \in \mathbb{K}[X]$ tel que $PQ = 1$ et on a $\deg(PQ) = \deg(P) + \deg(Q) = \deg(1) = 0$, dans $\mathbb{N}$, ce qui impose $\deg(P) = \deg(Q) = 0$. Donc P est constant non nul. Réciproquement un polynôme $P = a_0$ avec $a_0 \in \mathbb{K}^*$ est inversible d'inverse $\frac{1}{a_0}$. $\square$

On peut définir sur $\mathbb{K}[X]$ une autre opération interne qui est la composition de deux polynômes.

Définition 9.4. Soient $P = \sum_{k=0}^n a_k X^k$ et Q deux polynômes $\mathbb{K}[X]$. La composée des polynômes P et Q est le polynôme $P \circ Q = \sum_{k=0}^n a_k Q^k$ avec la convention que $Q^0 = 1$ pour tout polynôme Q .

On note aussi $P(Q)$ pour $P \circ Q$. Pour $Q = X$, on retrouve le polynôme $P(X)$ (ce qui justifie la notation $P(X)$). Pour $Q = X - \alpha$, où $\alpha \in \mathbb{K}$, on a le polynôme $P(X - \alpha) = \sum_{k=0}^n a_k (X - \alpha)^k$ qui est de même degré que P .

Théorème 9.4.

Si P, Q sont deux polynômes non nuls dans $\mathbb{K}[X]$, on a alors $\deg(P \circ Q) = \deg(P) \deg(Q)$.

Preuve. Pour $P = \sum_{k=0}^n a_k X^k$ de degré $n \geq 0$ et $Q = \sum_{j=0}^m b_j X^j$ de degré $m \geq 0$, le polynôme Q^k est de degré mk , donc $a_n Q^n$ est de degré nm et les $a_k Q^k$ sont de degré inférieur ou égal à $(n-1)m < nm$ pour tout k compris entre 0 et $n-1$. D'où le résultat. $\square$

Définition 9.5. On dit qu'une famille de polynômes non nuls $(P_k)_{k \in \mathbb{N}}$ est étagée en degrés [resp. en valuations] si on a $\deg(P_k) < \deg(P_{k+1})$ [resp. $\text{val}(P_k) < \text{val}(P_{k+1})$] pour tout $k \in \mathbb{N}$ et qu'elle est échelonnée en degrés

[resp. en valuations] si on a $\deg(P_k) = k$ [resp. $\text{val}(P_k) = k$] pour tout $k \in \mathbb{N}$.

Une famille de polynômes échelonnée en degrés [resp. en valuations] est étagée en degrés [resp. en valuations].

Théorème 9.5.

Une famille $(P_k)_{k \in \mathbb{N}}$ de polynômes non nuls étagée en degrés [resp. en valuations] est libre dans $\mathbb{K}[X]$.

Preuve. Soit $(P_k)_{k \in \mathbb{N}}$ une famille de polynômes étagée en degrés. On vérifie par récurrence sur $n \geq 0$, que chaque famille $(P_k)_{0 \leq k \leq n}$ est libre. Pour $n = 0$, P_0 est non nul, donc (P_0) est libre. Supposons le résultat acquis pour $n - 1 \geq 0$ et soit

$(\lambda_k)_{0 \leq k \leq n}$ des scalaires tels que $\sum_{k=0}^n \lambda_k P_k = 0$. On a alors $\lambda_n P_n = -\sum_{k=0}^{n-1} \lambda_k P_k$. Si

$\lambda_n \neq 0$, on a alors $\deg(P_n) = \deg\left(\sum_{k=0}^{n-1} \lambda_k P_k\right)$, ce qui est en contradiction avec

$\deg(P_k) < \deg(P_n)$ pour tout k compris entre 0 et $n - 1$. On a donc $\lambda_n = 0$

et $\sum_{k=0}^{n-1} \lambda_k P_k = 0$, ce qui impose $\lambda_k = 0$ pour tout k compris entre 0 et $n - 1$

puisque $(P_k)_{0 \leq k \leq n-1}$ est libre par hypothèse de récurrence. On procède de manière analogue pour une famille de polynômes étagée en valuations. $\square$

Corollaire 9.1. Une famille $(P_k)_{k \in \mathbb{N}}$ [resp. $(P_k)_{0 \leq k \leq n}$] de polynômes non nuls échelonnée en degrés [resp. en valuations] est une base de $\mathbb{K}[X]$ [resp. de $\mathbb{K}_n[X]$].

Preuve. La famille $(P_k)_{0 \leq k \leq n}$ est libre formée de $n + 1$ polynômes dans $\mathbb{K}_n[X]$, c'est donc une base. Comme $\mathbb{K}[X] = \bigcup_{n \geq 0} \mathbb{K}_n[X]$, on en déduit que $(P_k)_{k \in \mathbb{N}}$ est une base de $\mathbb{K}[X]$. $\square$

Exemple 9.1 Pour tout scalaire $a \in \mathbb{K}$, la famille $\left((X - a)^k\right)_{k \in \mathbb{N}}$ est une base de $\mathbb{K}[X]$ puisqu'elle est échelonnée en degrés.

9.2 Division euclidienne dans $\mathbb{K}[X]$

Comme pour l'anneau $\mathbb{Z}$ des entiers relatifs, on dispose dans l'anneau $\mathbb{K}[X]$ d'une division euclidienne qui nous permettra de définir les notions de pgcd et de ppcm.

Théorème 9.6.

Pour tout couple (A, B) de polynômes à coefficients dans $\mathbb{K}$ avec $B \neq 0$, il existe un unique couple (Q, R) de polynômes tel que $A = BQ + R$ et $\deg(R) < \deg(B)$.

Preuve. On montre tout d'abord l'existence du quotient Q et du reste R . Pour $A = 0$, le couple $(Q, R) = (0, 0)$ convient. Pour $A \neq 0$, on procède par récurrence sur $n = \deg(A) \geq 0$ (à B fixé). Si $\deg(A) = 0$, on a alors $A = a_0 \in \mathbb{K}^*$. Pour $B = b_0 \in \mathbb{K}^*$, on écrit que $a_0 = b_0 \frac{a_0}{b_0} + 0$ et le couple $(Q, R) = \left(\frac{a_0}{b_0}, 0\right)$ convient. Pour B non constant, le couple $(Q, R) = (0, a_0)$ convient. Supposons le résultat acquis pour les polynômes de degré au plus égal à $n - 1 \geq 0$ et soit A de degré n . Si $\deg(A) < \deg(B)$, le couple $(Q, R) = (0, A)$ convient. Sinon, on a $A(X) = \sum_{k=0}^n a_k X^k$ et $B(X) = \sum_{k=0}^m b_k X^k$ avec $n \geq m$ et $a_n \neq 0$, $b_m \neq 0$. On note $A_1(X) = A(X) - \frac{a_n}{b_m} X^{n-m} B(X)$. Si $A_1 = 0$, on a alors une division euclidienne $A = BQ + R$ avec $Q = \frac{a_n}{b_m} X^{n-m}$ et $R = 0$. Sinon, on a $\deg(A_1) < n = \deg(A)$ et l'hypothèse de récurrence nous dit qu'il existe deux polynômes Q_1, R_1 tels que $A_1 = BQ_1 + R_1$ et $\deg(R_1) < \deg(B)$. Il en résulte que $A = BQ + R$ avec $Q = Q_1 + \frac{a_n}{b_m} X^{n-m}$ et $R = R_1$. Si on a une autre division euclidienne $A = BQ_1 + R_1$ avec $\deg(R_1) < \deg(B)$, on a alors $R - R_1 = B(Q_1 - Q)$. Si $Q_1 \neq Q$, on a alors $R_1 \neq R$ ($B \neq 0$ et $\mathbb{K}[X]$ est intègre) et $\deg(R - R_1) \geq \deg(B)$, ce qui n'est pas possible puisque R et R_1 sont de degré strictement inférieur à celui de B (éventuellement égal à $-\infty$). On a donc $Q = Q_1$ et $R = R_1$. Le reste et le quotient dans une telle division euclidienne sont donc uniques. $\square$

Avec l'exercice 9.1, on donne une démonstration très simple du théorème de division euclidienne qui utilise l'algèbre linéaire.

Définition 9.6. On dit qu'un polynôme non nul B divise un polynôme A dans $\mathbb{K}[X]$, s'il existe un polynôme Q tel que $A = BQ$.

Le polynôme B divise le polynôme A si, et seulement si, le reste dans la division euclidienne de A par B est nul (unicité du quotient et du reste).

Lemme 9.1 Soient A, B non nuls dans $\mathbb{K}[X]$. On a A divise B et B divise A si, et seulement si, il existe $\lambda \in \mathbb{K}^*$ tel que $B = \lambda A$.

Preuve. Si A divise B et B divise A , il existe alors deux polynômes Q_1, Q_2 tel que $B = Q_1 A$ et $A = Q_2 B$, donc $B = Q_1 Q_2 B$, soit $(1 - Q_1 Q_2) B = 0$ et $Q_1 Q_2 = 1$ puisque $B \neq 0$ et $\mathbb{K}[X]$ est intègre. On a donc $Q_1 = \lambda \in \mathbb{K}^*$ et $B = \lambda A$. La réciproque est évidente. $\square$

Définition 9.7. On dit que deux polynômes A, B non nuls dans $\mathbb{K}[X]$ sont associés, si il existe $\lambda \in \mathbb{K}^*$ tel que $B = \lambda A$.

Définition 9.8. Un idéal de $\mathbb{K}[X]$ est un sous-ensemble I de $\mathbb{K}[X]$ tel que :

$$\begin{cases} I \text{ est un sous-groupe de } (\mathbb{K}[X], +) \\ \forall (P, Q) \in I \times \mathbb{K}[X], PQ \in I \end{cases}$$

De manière plus générale, on peut définir la notion d'idéal d'un anneau.

Exemples 9.1

1. Si A est un anneau et $\varphi : \mathbb{K}[X] \rightarrow A$ est un morphisme d'anneaux, alors son noyau $\ker(\varphi)$ est un idéal de $\mathbb{K}[X]$. En effet, comme φ est en particulier un morphisme de groupes additifs, $\ker(\varphi)$ est un sous-groupe de $(\mathbb{K}[X], +)$. Pour $P \in \ker(\varphi)$ et $Q \in \mathbb{K}[X]$, on a $\varphi(PQ) = \varphi(P)\varphi(Q) = 0$, donc PQ est dans $\ker(\varphi)$. Cet exemple nous sera utile pour définir le polynôme minimal d'un endomorphisme ou d'une matrice.
2. Pour tout $P \in \mathbb{K}[X]$, l'ensemble $(P) = \{PQ \mid q \in \mathbb{K}[X]\}$ des multiples de P dans $\mathbb{K}[X]$ est un idéal. On dit que c'est l'idéal engendré par P et on le note aussi $P \cdot \mathbb{K}[X]$. Un tel idéal est dit principal. Plus généralement, pour toute famille $(P_k)_{1 \leq k \leq r}$ d'éléments de $\mathbb{K}[X]$, l'ensemble :

$$(P_1, \dots, P_r) = \left\{ \sum_{k=1}^r Q_k P_k \mid (Q_k)_{1 \leq k \leq r} \in \mathbb{K}[X]^r \right\}$$

est un idéal. On dit que c'est l'idéal engendré par $P_1, \dots, P_r$. On le note aussi $\sum_{k=1}^r P_k \cdot \mathbb{K}[X]$.

Lemme 9.2 Soient A, B dans $\mathbb{K}[X]$. Le polynôme A divise B si, et seulement si, on a $(B) \subset (A)$. En les supposant non nuls, les polynômes A et B sont associés si, et seulement si, on a $(B) = (A)$.

Preuve. Si A divise B , il existe alors $Q \in \mathbb{K}[X]$ tel que $B = QA$, donc tout multiple de B est multiple de A , c'est-à-dire que $(B) \subset (A)$. Réciproquement, si $(B) \subset (A)$, on a alors $B \in (A)$, soit $B = QA$ avec $Q \in \mathbb{K}[X]$. Le deuxième point s'en déduit immédiatement. $\square$

Théorème 9.7.

Pour tout idéal I de $\mathbb{K}[X]$ non réduit à $\{0\}$, il existe un unique polynôme unitaire A tel que $I = (A)$.

Preuve. Soit I un idéal de $\mathbb{K}[X]$. Si $I = \{0\}$, il est principal engendré par 0. On suppose que $I \neq \{0\}$ et on pose $n_0 = \min \{\deg(P) \mid P \in I \setminus \{0\}\}$ (plus petit élément d'une partie non vide de $\mathbb{N}$). En désignant par A un polynôme unitaire dans I de degré n_0 , la division euclidienne de $P \in I$ par A nous donne $P = QA + R$ avec $R = 0$ puisque A est de degré minimal dans $I \setminus \{0\}$, donc $P = QA$ et $I \subset (A)$. Comme par ailleurs $(A) \subset I$ puisque I est un idéal, on a $I = (A)$. Si B est un autre générateur de I , on a B qui divise A et A qui divise B , donc $B = \lambda A$ avec $\lambda \in \mathbb{K}^*$ et $\lambda = 1$ si B est unitaire. Le générateur unitaire de I est donc unique. $\square$

Le théorème précédent se traduit en disant que l'anneau $\mathbb{K}[X]$ est principal.

Théorème 9.8.

Soient $r \in \mathbb{N} \setminus \{0, 1\}$ et $A_1, \dots, A_r$ dans $\mathbb{K}[X]$ non tous nuls. Il existe un unique polynôme unitaire $\Delta \in \mathbb{K}[X]$ tel que $(A_1, \dots, A_r) = (\Delta)$. Il existe des polynômes $U_1, \dots, U_r$ tels que $\Delta = \sum_{k=1}^r U_k A_k$ et Δ est aussi caractérisé par la propriété suivante :

$$\begin{cases} \forall k \in \{1, \dots, r\}, \Delta \text{ divise } A_k \\ \text{tout diviseur commun à } A_1, \dots, A_r \text{ divise } \Delta \end{cases} \quad (9.1)$$

Preuve. L'existence et l'unicité de Δ se déduisent du fait que $(A_1, \dots, A_r)$ est un idéal de l'anneau principal $\mathbb{K}[X]$. Comme $\Delta \in (\Delta) = (A_1, \dots, A_r)$, il existe $(U_1, \dots, U_r) \in (\mathbb{K}[X])^r$ tel que $\Delta = \sum_{k=1}^r U_k A_k$. De $(A_k) \subset (\Delta)$ pour tout k compris entre 1 et r , on déduit que Δ divise A_k . Si $D \in \mathbb{K}[X]$ est un diviseur commun aux A_k , il divise aussi $\Delta = \sum_{k=1}^r U_k A_k$. Réciproquement si D est un polynôme unitaire vérifiant (9.1), il est alors divisible par Δ et Δ divise D , donc $D = \lambda \Delta$ avec $\lambda \in \mathbb{K}^*$ et $\lambda = 1$ si D est unitaire. $\square$

On peut donc donner la définition suivante.

Définition 9.9. Soient $r \in \mathbb{N} \setminus \{0, 1\}$ et $A_1, \dots, A_r$ dans $\mathbb{K}[X]$ non tous nuls. On appelle plus grand commun diviseur de $A_1, \dots, A_r$ le polynôme unitaire Δ qui vérifie $(A_1, \dots, A_r) = (\Delta)$ ou (9.1). On note $\text{pgcd}(A_1, \dots, A_r)$ ou $A_1 \wedge \dots \wedge A_r$ ce polynôme.

La relation $\Delta = \sum_{k=1}^r U_k A_k$ est l'identité de Bézout.

Un polynôme unitaire D qui s'écrit $D = \sum_{k=1}^r U_k A_k$ n'est pas nécessairement le $\text{pgcd } \Delta$ de $A_1, \dots, A_r$. Tout ce que l'on peut dire c'est que $D \in (A_1, \dots, A_r) = (\Delta)$ et donc $D = Q\Delta$. Mais si $D = 1$, on a alors nécessairement $\Delta = 1$ et dans ce cas D est le pgcd . Nous reviendrons sur ce cas particulier au paragraphe suivant.

On vérifie facilement que le pgcd satisfait les propriétés énoncées dans le théorème qui suit.

Théorème 9.9.

Soient P, Q, R dans $\mathbb{K}[X]$ et $\lambda \in \mathbb{K}$, avec P et λ non nuls. En désignant par α le coefficient dominant de P , on a :

- $P \wedge 0 = \frac{1}{\alpha} P$; $P \wedge \lambda = 1$;
- $P \wedge Q = Q \wedge P$ (commutativité du pgcd) ;

- $P \wedge Q = \frac{1}{\alpha}P$ si, et seulement si, P divise Q ;
- $(PQ) \wedge (PR) = \frac{1}{\alpha}P(Q \wedge R)$;
- pour Q non nul, $P \wedge (Q \wedge R) = (P \wedge Q) \wedge R$ (associativité du pgcd).

Preuve. Laisée au lecteur. $\square$

Comme pour l'anneau $\mathbb{Z}$ des entiers relatifs, on dispose de l'algorithme d'Euclide pour obtenir le pgcd de deux polynômes A, B non nuls, les justifications étant analogues. Cet algorithme permet également de déterminer des polynômes U, V tels que $AU + BV = A \wedge B$. Tout est basé sur le résultat suivant.

Lemme 9.3 Soient A, B deux polynômes non nuls et R le reste dans la division euclidienne de A par B . On a $A \wedge B = B \wedge R$.

Preuve. On a $(A, B) = (BQ + R, B) \subset (B, R)$ et $(R, B) = (A - BQ, B) \subset (A, B)$ d'où l'égalité $(A, B) = (B, R)$ et $A \wedge B = B \wedge R$. $\square$

Le principe de l'algorithme d'Euclide est le suivant pour A, B non nuls tels que $\deg(A) \geq \deg(B)$ (A et B jouent des rôles symétriques). On note $R_{-1} = A$, $R_0 = B$ et on désigne par R_1 le reste dans la division euclidienne de $R_{-1} = A$ par $R_0 = B$. Si $R_1 = 0$, c'est alors terminé : $A \wedge B = R_0 = B$, à un facteur multiplicatif non nul près, en réalité $A \wedge B = \frac{1}{\beta}B$ où β est le coefficient dominant de B . Sinon, on a $\deg(R_1) < \deg(R_0)$ et d'après le lemme précédent $A \wedge B = R_0 \wedge R_1$. On désigne alors par R_2 le reste dans la division euclidienne de R_0 par R_1 . Si $R_2 = 0$, c'est alors terminé : $A \wedge B = R_0 \wedge R_1 = R_1$, à un facteur multiplicatif non nul près. Sinon, on continue avec $R_0 \wedge R_1 = R_1 \wedge R_2$. On définit donc ainsi une suite $(R_n)_{n \geq 0}$ de polynômes par :

- $R_0 = B$; R_1 est le reste dans la division euclidienne de A par B ; on a donc $R_1 = 0$ ou $0 \leq \deg(R_1) < \deg(R_0)$;
- pour $n \geq 2$, si $R_{n-1} = 0$ alors $R_n = 0$, sinon R_n est le reste dans la division euclidienne de R_{n-2} par R_{n-1} et on a $R_n = 0$ ou $0 \leq \deg(R_n) < \deg(R_{n-1})$.

Il existe alors un entier $p \geq 1$ tel que $R_p = 0$, $0 \leq \deg(R_{p-1}) < \dots < \deg(R_1) < \deg(R_0)$ et $A \wedge B = R_0 \wedge R_1 = \dots = R_{p-1} \wedge R_p = R_{p-1}$ (à un facteur multiplicatif non nul près), c'est-à-dire que $A \wedge B$ est, à un facteur multiplicatif non nul près, le dernier reste non nul dans cette suite finie de divisions euclidiennes. On a en fait construit avec l'algorithme d'Euclide, dans le cas où B ne divise pas A , deux suites de polynômes $(R_n)_{0 \leq n \leq p}$ et $(Q_n)_{1 \leq n \leq p}$ de la manière suivante :

$$\begin{cases} A = Q_1 R_0 + R_1 & (R_1 \neq 0 \text{ et } \deg(R_1) < \deg(R_0)) \\ R_0 = Q_2 R_1 + R_2 & (R_2 \neq 0 \text{ et } \deg(R_2) < \deg(R_1)) \\ \vdots \\ R_{p-3} = Q_{p-1} R_{p-2} + R_{p-1} & (R_{p-1} \neq 0 \text{ et } \deg(R_{p-1}) < \deg(R_{p-2})) \\ R_{p-2} = Q_p R_{p-1} + R_p & (R_p = 0) \end{cases}$$

On vérifie alors, par récurrence finie sur $k \in \{0, 1, \dots, p-1\}$, qu'il existe deux polynômes U_k et V_k tels que $R_k = AU_k + BV_k$. En particulier pour $k = p-1$

on a $A \wedge B = R_{p-1} = AU_{p-1} + BV_{p-1} = AU + BV$. Un tel couple (U, V) n'est pas unique puisque si (U_0, V_0) est une solution, pour tout polynôme Q , le couple $(U_0 + QB, V_0 - QA)$ est aussi solution. On a en effet :

$$A(U_0 + QB) + B(V_0 - QA) = AU_0 + BV_0 = A \wedge B$$

9.3 Polynômes premiers entre eux

Définition 9.10. Soient $r \in \mathbb{N} \setminus \{0, 1\}$ et $A_1, \dots, A_r$ des polynômes dans $\mathbb{K}[X]$ tous non nuls. On dit que $A_1, \dots, A_r$ sont premiers entre eux dans leur ensemble si leur pgcd vaut 1. Pour $r = 2$, on dira simplement que A_1 et A_2 sont premiers entre eux.

Dire que $A_1, \dots, A_r$ sont premiers entre eux dans leur ensemble signifie donc que $(A_1, \dots, A_r) = (1) = \mathbb{K}[X]$.

Exemple 9.2 Pour $\alpha \neq \beta$ dans $\mathbb{K}$, les polynômes $X - \alpha$ et $X - \beta$ sont premiers entre eux. En effet, $\beta - \alpha = (X - \alpha) - (X - \beta) \in (X - \alpha, X - \beta)$ avec $\beta - \alpha \neq 0$.

Des polynômes deux à deux premiers entre eux sont premiers entre eux dans leur ensemble puisque $\mathbb{K}[X] = (A_i, A_j) \subset (A_1, \dots, A_r)$ pour $1 \leq i \neq j \leq r$, mais la réciproque est fautive. Par exemple, pour les polynômes $X, X + 1, X(X + 1)$, on a $X \wedge (X + 1) \wedge X(X + 1) = (X \wedge (X + 1)) \wedge X(X + 1) = 1 \wedge X(X + 1) = 1$ et $(X + 1) \wedge X(X + 1) = X + 1 \neq 1$.

Théorème 9.10. Bézout

Soient $A_1, \dots, A_r$ des polynômes dans $\mathbb{K}[X]$ tous non nuls avec $r \geq 2$. Ces polynômes sont premiers entre eux dans leur ensemble si, et seulement si, il existe des polynômes $U_1, \dots, U_r$ tel que $\sum_{k=1}^r U_k A_k = 1$.

Preuve. On sait déjà, par définition du pgcd, que la condition est nécessaire. Réciproquement s'il existe des polynômes $U_1, \dots, U_r$ tel que $\sum_{k=1}^r U_k A_k = 1$, on a alors $1 \in (A_1, \dots, A_r)$, donc $\mathbb{K}[X] = (1) \subset (A_1, \dots, A_r)$ et l'égalité. $\square$

Corollaire 9.2. Soient $A_1, \dots, A_r$ des polynômes non nuls avec $r \geq 2$ et $\Delta = A_1 \wedge \dots \wedge A_r$. Il existe des polynômes, $B_1, \dots, B_r$ premiers entre eux dans leur ensemble, tels que $A_k = \Delta B_k$ pour tout k compris entre 1 et r .

Preuve. Comme, pour tout k compris entre 1 et r , Δ divise A_k , il existe un polynôme B_k tel que $A_k = \Delta B_k$ et on a $\Delta = \sum_{k=1}^r U_k A_k = \Delta \sum_{k=1}^r U_k B_k$ avec $\Delta \neq 0$,

ce qui entraîne $\sum_{k=1}^r U_k B_k = 1$, ce qui signifie que $B_1, \dots, B_r$ premiers entre eux dans leur ensemble. $\square$

Corollaire 9.3. Gauss. Soient A, B, C trois polynômes non nuls. Si A divise BC et est premier avec B , alors A divise C .

Preuve. Comme A et B sont premiers entre eux, il existe deux polynômes U, V tels que $AU + BV = 1$ et pour tout polynôme C , on a $ACU + BCV = C$, de sorte que si A divise BC , il va diviser C . $\square$

Corollaire 9.4. Soient A, B, C trois polynômes non nuls. Si C est premier avec A alors $A \wedge B = A \wedge (BC)$ (le pgcd de deux polynômes est inchangé si on multiplie l'un d'eux par un polynôme premier avec l'autre).

Preuve. Soient $\Delta = A \wedge B$ et $\Gamma = A \wedge (BC)$. Comme Δ divise A et B , il divise A et BC , donc leur pgcd Γ . De $AU + BV = 1$, on déduit que $ABU + BCV = B$ et Γ qui divise A et BC va diviser A et B , donc leur pgcd Δ . En définitive $\Delta = \Gamma$. $\square$

Corollaire 9.5. Soient $r \in \mathbb{N} \setminus \{0, 1\}$ et $A_1, \dots, A_r, C$ des polynômes non nuls. Si C est premier avec chacun des A_k , pour tout k compris entre 1 et r , il est alors premier avec leur produit $\prod_{k=1}^r A_k$.

Preuve. En utilisant le corollaire précédent, on a :

$$C \wedge \prod_{k=1}^r A_k = C \wedge \left(A_1 \prod_{k=2}^r A_k \right) = C \wedge \prod_{k=2}^r A_k$$

puisque A_1 est premier avec C et par récurrence finie, on déduit que :

$$C \wedge \prod_{k=1}^r A_k = C \wedge \prod_{k=2}^r A_k = C \wedge \prod_{k=3}^r A_k = \dots = C \wedge A_r = 1$$

puisque chaque A_k , pour k compris entre 1 et r , est premier avec C . On peut aussi utiliser le théorème de Bézout. On a $A_k U_k + C V_k = 1$ pour tout k compris entre 1 et r , donc en faisant le produit $\prod_{k=1}^r (A_k U_k + C V_k) = \left(\prod_{k=1}^r A_k \right) U + C V = 1$ avec

$$U = \prod_{k=1}^r U_k, \text{ ce qui signifie que } C \text{ est premier avec } \prod_{k=1}^r A_k. \quad \square$$

Exemple 9.3 Si A et B sont premiers entre eux, alors A^n et B^m sont premiers entre eux pour tous n, m entiers naturels non nuls. En effet, comme A est premier avec B , il est premier avec B^m d'après le corollaire précédent et le même corollaire nous dit que B^m est premier avec A^n . En particulier, pour $\alpha \neq \beta$ dans $\mathbb{K}$ et n, m dans $\mathbb{N}^*$, les polynôme $(X - \alpha)^n$ et $(X - \beta)^m$ sont premiers entre eux.

9.4 Fonctions polynomiales, racines des polynômes

À tout polynôme $P(X) = \sum_{k=0}^n a_k X^k \in \mathbb{K}[X]$, on associe la fonction polynomiale $\tilde{P} : \mathbb{K} \rightarrow \mathbb{K}$ définie par $\tilde{P}(x) = \sum_{k=0}^n a_k x^k$ pour tout $x \in \mathbb{K}$. Nous verrons, en utilisant le théorème de division euclidienne, que ce morphisme est injectif.

Pour $\alpha \in \mathbb{K}$, on note $P(\alpha) = \sum_{k=0}^n a_k \alpha^k$ l'évaluation de la fonction polynôme $\tilde{P}$ en α (ce qu'il faudrait noter en toute rigueur $\tilde{P}(\alpha)$). Cette évaluation en α est en fait la composition de P avec le polynôme constant α , qui est bien un polynôme constant.

$P(0) = a_0$ est le coefficient constant de P .

Lemme 9.4 *Pour tout $\alpha \in \mathbb{K}$, l'application $\varphi_\alpha : P \mapsto P(\alpha)$ réalise un morphisme d'algèbres (i. e. un morphisme d'anneaux et d'espaces vectoriels) de $\mathbb{K}[X]$ dans $\mathbb{K}$.*

Preuve. On vérifie tout d'abord que φ_α est linéaire. Pour $P(X) = \sum_{k=0}^n a_k X^k$,

$Q(X) = \sum_{k=0}^m b_k X^k$ et $\lambda \in \mathbb{K}$, on a :

$$(\lambda P + Q)(\alpha) = \sum_{k=0}^{\max(n,m)} (\lambda a_k + b_k) \alpha^k = \lambda P(\alpha) + Q(\alpha)$$

donc $\lambda P + Q$ est linéaire. Pour le produit, on a $PQ = \sum_{k=0}^n a_k X^k Q$ et avec la linéarité de φ_α , on déduit que :

$$(PQ)(\alpha) = \sum_{k=0}^n a_k \varphi_\alpha(X^k Q) = \sum_{k=0}^n a_k \sum_{j=0}^m b_j \alpha^{k+j} = \sum_{\substack{0 \leq k \leq n \\ 0 \leq j \leq m}} a_k b_j \alpha^{k+j} = P(\alpha) Q(\alpha)$$

□

La dernière égalité dans la démonstration précédente justifie la façon de définir le produit de deux polynômes.

Définition 9.11. Soit $P \in \mathbb{K}[X]$. On dit que $\alpha \in \mathbb{K}$ est racine de P , si $P(\alpha) = 0$.

Un polynôme constant non nul n'a pas de racines et le polynôme nul a tous les points de $\mathbb{K}$ comme racines.

Pour $P \in \mathbb{K}[X]$ et $\alpha \in \mathbb{K}$, le reste dans la division euclidienne par $X - \alpha$ est un polynôme de degré 0 ou $-\infty$, c'est-à-dire une constante λ . On a donc

$P(X) = Q(X)(X - \alpha) + \lambda$ et l'évaluation en α nous donne $\lambda = P(\alpha)$. On en déduit que α est racine de P si, et seulement si, $X - \alpha$ divise P .

Définition 9.12. Soient $P \in \mathbb{K}[X] \setminus \mathbb{K}$ un polynôme non constant, $\alpha \in \mathbb{K}$ et $m \in \mathbb{N}^*$. On dit que α est racine d'ordre (ou de multiplicité) m de P si $(X - \alpha)^m$ divise P et $(X - \alpha)^{m+1}$ ne divise pas P .

Si α est racine de $P \in \mathbb{K}[X] \setminus \mathbb{K}$, sa multiplicité est alors comprise entre 1 et $\deg(P)$. Pour $m = 1$, on dit que α est une racine simple, pour $m = 2$, qu'elle est racine double et pour $m \geq 2$, qu'elle est racine multiple. Si α n'est pas racine de P , on peut dire que c'est une racine de multiplicité nulle.

Théorème 9.11.

Soient $P \in \mathbb{K}[X] \setminus \mathbb{K}$, $\alpha_1, \dots, \alpha_r$ dans $\mathbb{K}$ deux à deux distincts et $m_1, \dots, m_r$ des entiers naturels non nuls. Les deux assertions suivantes sont équivalentes :

1. pour tout k compris entre 1 et r , α_k est racine de P de multiplicité m_k ;
2. il existe un polynôme $Q \in \mathbb{K}[X]$ tel que $P(X) = Q(X) \prod_{k=1}^r (X - \alpha_k)^{m_k}$ et $Q(\alpha_k) \neq 0$ pour tout k compris entre 1 et r .

Preuve. Pour $r = 1$ c'est vrai. Supposons le résultat acquis pour $r - 1 \geq 1$. Si, pour tout k compris entre 1 et r , α_k est racine de P de multiplicité m_k , l'hypothèse de récurrence nous dit alors qu'il existe un polynôme Q tel que $P(X) =$

$Q(X) \prod_{k=1}^{r-1} (X - \alpha_k)^{m_k}$ et $Q(\alpha_k) \neq 0$ pour tout k compris entre 1 et $r - 1$. Avec

$P(\alpha_r) = Q(\alpha_r) \prod_{k=1}^{r-1} (\alpha_r - \alpha_k)^{m_k} = 0$ et $\prod_{k=1}^{r-1} (\alpha_r - \alpha_k)^{m_k} \neq 0$, on déduit que

$Q(\alpha_r) = 0$, donc α_r est racine de Q et désignant par $m \geq 1$ sa multiplicité, on a $Q(X) = Q_1(X)(X - \alpha_r)^m$ avec $Q_1(\alpha_r) \neq 0$, donc :

$$P(X) = Q_1(X)(X - \alpha_r)^m \prod_{k=1}^{r-1} (X - \alpha_k)^{m_k} = Q_2(X)(X - \alpha_r)^m$$

avec $Q_2(\alpha_r) = Q_1(\alpha_r) \prod_{k=1}^{r-1} (\alpha_r - \alpha_k)^{m_k} \neq 0$, ce qui signifie que m est la multipli-

cité de α_r , soit que $m = m_r$ et $P(X) = Q_1(X) \prod_{k=1}^r (X - \alpha_k)^{m_k}$ avec $Q_1(\alpha_k) \neq 0$

pour tout k compris entre 1 et r . Réciproquement, s'il existe un polynôme Q tel que $P(X) = Q(X) \prod_{k=1}^r (X - \alpha_k)^{m_k}$ et $Q(\alpha_k) \neq 0$ pour tout k compris entre 1 et r , on

a alors $P(X) = Q_1(X) \prod_{k=1}^{r-1} (X - \alpha_k)^{m_k}$ avec $Q_1(\alpha_k) = Q(\alpha_k)(\alpha_k - \alpha_r)^{m_r} \neq 0$

pour tout k compris entre 1 et $r - 1$, donc α_k est racine de P de multiplicité m_k , pour tout k compris entre 1 et $r - 1$. L'égalité $P(X) = Q_2(X)(X - \alpha_r)^{m_r}$ avec $Q_2(\alpha_r) = Q(\alpha_r) \prod_{k=1}^{r-1} (\alpha_r - \alpha_k)^{m_k} \neq 0$ nous dit que α_r est racine de P de multiplicité m_r . $\square$

Corollaire 9.6. Si $P \in \mathbb{K}[X] \setminus \mathbb{K}$ admet $r \geq 1$ racines distinctes $\alpha_1, \dots, \alpha_r$ dans $\mathbb{K}$ de multiplicités respectives $m_1, \dots, m_r$, on a alors $\deg(P) \geq \sum_{k=1}^r m_k$.

Preuve. Conséquence immédiate du théorème précédent. $\square$

Corollaire 9.7. Un polynôme $P \in \mathbb{K}[X]$ de degré $n \geq 1$ admet au plus n racines distinctes ou confondues dans $\mathbb{K}$.

Preuve. Conséquence immédiate du théorème précédent. $\square$

Du corollaire précédent, on déduit que si P et Q dans $\mathbb{K}_n[X]$ coïncident en au moins $n + 1$ points distincts, ils sont alors égaux. Le corollaire précédent peut aussi s'exprimer en disant que si $P \in \mathbb{K}_n[X]$ admet $r \geq 1$ racines distinctes $\alpha_1, \dots, \alpha_r$ dans $\mathbb{K}$ de multiplicités respectives $m_1, \dots, m_r$ avec $\sum_{i=1}^r m_i > n$, c'est alors le polynôme nul.

Définition 9.13. On dit qu'un polynôme $P \in \mathbb{K}[X]$ est scindé sur $\mathbb{K}$, s'il est constant ou de degré $n \geq 1$ et admet $r \geq 1$ racines distinctes $\alpha_1, \dots, \alpha_r$ dans $\mathbb{K}$ de multiplicités respectives $m_1, \dots, m_r$ avec $\sum_{i=1}^r m_i = n$. Dans le cas où tous les m_i sont égaux à 1, on dit que le polynôme est scindé à racines simples.

Un polynôme scindé non constant est donc de la forme $P(X) = \lambda \prod_{k=1}^r (X - \alpha_k)^{m_k}$, où λ est une constante non nulle, les α_k sont des scalaires deux à deux distincts et les m_k des entiers naturels non nuls.

On peut montrer le théorème suivant que nous admettrons.

Théorème 9.12. d'Alembert-Gauss

Toute équation polynomiale à coefficients complexes de degré non nul n admet n racines complexes distinctes ou confondues.

On peut donner plusieurs démonstrations du théorème de d'Alembert-Gauss mais toutes utilisent des outils d'algèbre ou d'analyse plus sophistiqués que le contenu de ce chapitre.

9.5 Dérivation des polynômes. Formule de Taylor

Définition 9.14. Soit P un polynôme dans $\mathbb{K}[X]$. Le polynôme dérivé de P est le polynôme P' défini par :

$$P'(X) = \begin{cases} 0 & \text{si } P = a_0 \text{ est constant} \\ \sum_{k=1}^n k a_k X^{k-1} & \text{si } P(X) = \sum_{k=0}^n a_k X^k \text{ est de degré } n \geq 1 \end{cases}$$

On a en particulier $(X^k)' = kX^{k-1}$ pour tout entier $k \geq 1$ et, pour tout polynôme P de degré $n \geq 1$, le polynôme P' est de degré $n - 1$.

Théorème 9.13.

L'application $P \mapsto P'$ est linéaire de $\mathbb{K}[X]$ dans $\mathbb{K}[X]$, de $\mathbb{K}_p[X]$ dans $\mathbb{K}_{p-1}[X]$ pour $p \geq 1$ et pour tous polynômes P, Q dans $\mathbb{K}[X]$, on a :

$$(PQ)' = P'Q + PQ', \quad (P \circ Q)' = (P' \circ Q)Q'$$

Preuve. La linéarité se vérifie facilement. Pour $P = a_0 \in \mathbb{K}$ et $Q \in \mathbb{K}[X]$, on a $P' = 0$ et $(PQ)' = a_0 Q' = P'Q + PQ'$. Comme P et Q jouent des rôles symétriques, on a un résultat analogue pour $P \in \mathbb{K}[X]$ et Q constant. Pour $P(X) = \sum_{k=0}^n a_k X^k$ de degré $n \geq 1$ dans $\mathbb{K}[X]$ et $Q(X) = X^q$ avec $q \in \mathbb{N}^*$, on a :

$$(PQ)' = \sum_{k=0}^n a_k (k+q) X^{k+q-1} = X^q \sum_{k=0}^n k a_k X^{k-1} + q X^{q-1} \sum_{k=0}^n a_k X^k = P'Q + PQ'$$

Par linéarité, on en déduit que $(PQ)' = P'Q + PQ'$ pour tous polynômes P, Q . De cette formule, on déduit par récurrence sur $k \geq 1$ que $(Q^k)' = kQ^{k-1}Q'$ pour tout polynôme Q . En effet, c'est vrai pour $k = 1$ et supposant le résultat acquis pour $k \geq 1$, on a :

$$(Q^{k+1})' = (QQ^k)' = Q(Q^k)' + Q'Q^k = kQQ^{k-1}Q' + Q'Q^k = (k+1)Q^kQ'$$

Pour $P = a_0 \in \mathbb{K}$ et $Q \in \mathbb{K}[X]$, on a $(P \circ Q)' = (a_0)' = 0 = (P' \circ Q)Q'$. Pour $P(X) = \sum_{k=0}^n a_k X^k$ de degré $n \geq 1$ et $Q \in \mathbb{K}[X]$, on a :

$$(P \circ Q)' = \sum_{k=0}^n a_k (Q^k)' = \left(\sum_{k=0}^n k a_k Q^{k-1} \right) Q' = (P' \circ Q)Q'$$

□

On définit les dérivées successives d'un polynôme P , en les notant $P^{(k)}$, par la relation de récurrence $P^{(0)} = P$ et $P^{(k)} = (P^{(k-1)})'$ pour tout $k \geq 1$. En

particulier, on a :

$$(X^k)^{(p)} = \begin{cases} \frac{k!}{(k-p)!} X^{k-p} & \text{si } 0 \leq p \leq k \\ 0 & \text{si } p \geq k+1 \end{cases}$$

Théorème 9.14. Leibniz

Pour tous polynômes P, Q et tout entier naturel n , on a :

$$(PQ)^{(n)} = \sum_{k=0}^n \binom{n}{k} P^{(k)} Q^{(n-k)} = \sum_{k=0}^n \binom{n}{k} P^{(n-k)} Q^{(k)}$$

Preuve. On procède par récurrence sur $n \geq 0$, comme pour la formule du binôme de Newton. $\square$

Théorème 9.15. Taylor

Pour tout polynôme $P(X) = \sum_{k=0}^n a_k X^k$ et tout scalaire $a \in \mathbb{K}$, on a :

$$P(X) = \sum_{k=0}^n \frac{P^{(k)}(a)}{k!} (X-a)^k$$

Preuve. Par linéarité, il suffit de montrer le résultat pour $P(X) = X^p$, où $p \in \mathbb{N}$. Dans ce cas, on a $P(X) = (X-a+a)^p = \sum_{k=0}^p \binom{p}{k} a^{p-k} (X-a)^k$ et pour k compris entre 0 et p , $P^{(k)}(a) = \frac{p!}{(p-k)!} a^{p-k} = k! \binom{p}{k} a^{p-k}$, d'où le résultat. $\square$

Exemples 9.2

1. Pour tout $a \in \mathbb{K}$, la famille $\left((X-a)^k\right)_{k \in \mathbb{N}}$ est une base de $\mathbb{K}[X]$ puisqu'elle est échelonnée en degrés. L'écriture de $P \in \mathbb{K}[X]$ dans cette base est donnée par la formule de Taylor, ce qui peut se retrouver comme suit : pour $P \in \mathbb{K}_n[X]$, on a $P = \sum_{k=0}^n \lambda_k (X-a)^k$ et par dérivation $P^{(j)} = \sum_{k=j}^n \lambda_k \frac{k!}{(k-j)!} (X-a)^{k-j}$, ce qui nous donne en évaluant en a , $P^{(j)}(a) = j! \lambda_j$ et $\lambda_j = \frac{P^{(j)}(a)}{j!}$.
2. Pour tout polynôme $P \in \mathbb{K}[X]$ de degré n , la famille $(P^{(k)})_{0 \leq k \leq n}$ est étagée en degrés et c'est une base de $\mathbb{K}_n[X]$.

Théorème 9.16.

Soient $P \in \mathbb{K}[X] \setminus \{0\}$, $\alpha \in \mathbb{K}$ et $m \in \mathbb{N}^*$. Les assertions suivantes sont équivalentes :

1. α est racine d'ordre $m \geq 1$ de P ;
2. il existe un polynôme $Q \in \mathbb{K}[X]$ tel que $P(X) = Q(X)(X - \alpha)^m$ et $Q(\alpha) \neq 0$;
3. $P^{(k)}(\alpha) = 0$ pour k compris entre 0 et $m - 1$ et $P^{(m)}(\alpha) \neq 0$.

Preuve. (1) $\Rightarrow$ (2) Supposons que α soit racine d'ordre $m \geq 1$ de P . Il existe alors un polynôme Q tel que $P(X) = Q(X)(X - \alpha)^m$ et $Q(\alpha) \neq 0$ puisque $(X - \alpha)^{m+1}$ ne divise pas P .

(2) $\Rightarrow$ (3) Supposons que $P(X) = Q(X)(X - \alpha)^m$ avec $Q(\alpha) \neq 0$. On a alors, pour k compris entre 0 et m :

$$\begin{aligned} P^{(k)}(X) &= \sum_{j=0}^k \binom{k}{j} ((X - \alpha)^m)^{(j)} Q^{(k-j)}(X) \\ &= \sum_{j=0}^k \binom{k}{j} \frac{m!}{(m-j)!} (X - \alpha)^{m-j} Q^{(k-j)}(X) \end{aligned}$$

donc $P^{(k)}(\alpha) = 0$ pour k compris entre 0 et $m - 1$ et $P^{(m)}(\alpha) = m!Q(\alpha) \neq 0$.

(3) $\Rightarrow$ (1) Supposons que $P^{(k)}(\alpha) = 0$ pour k compris entre 0 et $m - 1$ et $P^{(m)}(\alpha) \neq 0$. En utilisant la formule de Taylor pour les polynômes, on a :

$$P(X) = \sum_{k=m}^n \frac{P^{(k)}(\alpha)}{k!} (X - \alpha)^k = Q(X)(X - \alpha)^m$$

ou encore $P(X) = Q(X)(X - \alpha)^m$ avec $Q(\alpha) = \frac{P^{(m)}(\alpha)}{m!} \neq 0$, ce qui signifie que α est racine d'ordre m de P . $\square$

9.6 Polynômes irréductibles

Définition 9.15. Un polynôme $P \in \mathbb{K}[X] \setminus \{0\}$ est dit irréductible s'il est non constant (ou de manière équivalente non inversible) et n'est divisible que par les constantes non nulles ou les polynômes λP avec $\lambda \in \mathbb{K}^*$.

Un polynôme non constant P est réductible s'il admet un diviseur P_1 tel que $1 \leq \deg(P_1) < \deg(P)$, ce qui impose que P est degré au moins 2 et s'écrit $P = P_1 Q$, où les polynômes P_1 et Q sont non constants.

Exemples 9.3

1. Un polynôme de degré 1, $P(X) = aX + b$ avec $a \neq 0$, est irréductible.

2. Un polynôme de degré 2 est réductible dans $\mathbb{K}[X]$ si, et seulement si, il admet une racine double ou deux racines simples dans $\mathbb{K}$. En effet, il est clair que la condition est suffisante et pour $P \in \mathbb{K}[X]$ de degré 2 réductible, on a $P = P_1 Q$ avec P_1 et Q de degré 1, soit $P_1(X) = a(X - \alpha)$ et $Q(X) = b(X - \beta)$, donc $P(X) = \gamma(X - \alpha)(X - \beta)$ et le résultat suit.
3. $P(X) = X^2 - 2$ est réductible dans $\mathbb{R}[X]$ avec $-\sqrt{2}$ et $\sqrt{2}$ comme racines simples, mais il est irréductible dans $\mathbb{Q}[X]$ du fait que l'équation $x^2 - 2 = 0$ n'a pas de racines dans $\mathbb{Q}$ (si $\alpha = \frac{p}{q} \in \mathbb{Q}$ est une telle racine avec $p \in \mathbb{Z}^*$ premier avec $q \in \mathbb{N}^*$, on a alors $p^2 = 2q^2$, donc 2 divise p , soit $p = 2r$ et $2r^2 = q^2$ nous dit que q est aussi pair, ce qui contredit $p \wedge q = 1$).
4. Un polynôme de degré au moins égal à 2 et irréductible dans $\mathbb{K}[X]$ n'a pas de racines dans $\mathbb{K}$ (sinon il serait divisible par un polynôme de degré 1).

Théorème 9.17. Euclide

Soient $P \in \mathbb{K}[X]$ un polynôme irréductible. Si P divise un produit $\prod_{k=1}^r A_k$ de $r \geq 2$ polynôme non nuls, il divise alors l'un des A_k .

Preuve. Supposons que P irréductible dans $\mathbb{K}[X]$ divise le produit $A_1 A_2$, où A_1, A_2 sont deux polynômes non nuls. Quitte à diviser par le coefficient dominant, on peut supposer que P est unitaire. Le pgcd Δ de P et A_1 étant un diviseur de P , il est égal à 1 ou à P , puisque P est irréductible. Dans le cas où $\Delta = P$, P divise A_1 . Dans le cas où $P = 1$, il existe deux polynômes U, V tels que $UP + VA_1 = 1$ et P divise $A_2 = UPA_2 + VA_1 A_2$. Supposons le résultat acquis pour $r - 1 \geq 2$ et soit P irréductible qui divise $\prod_{k=1}^r A_k$. Si P divise A_r , c'est terminé, sinon il divise

$\prod_{k=1}^{r-1} A_k$ (cas $r = 2$) et l'hypothèse de récurrence permet de conclure. $\square$

Théorème 9.18.

Tout polynôme non constant $P \in \mathbb{K}[X]$ est produit de polynômes irréductibles et une telle décomposition est unique à l'ordre près des facteurs, ce qui signifie qu'il existe une constante $\lambda \in \mathbb{K}^*$, un entier $r \geq 1$, des polynômes unitaires $P_1, \dots, P_r$ deux à deux distincts et irréductibles et des entiers naturels non nuls $m_1, \dots, m_r$ tels que $P = \lambda \prod_{k=1}^r P_k^{m_k}$. L'unicité signifie que

si on a une autre décomposition du même type $P = \mu \prod_{k=1}^s Q_k^{n_k}$, on a alors $\mu = \lambda$, $s = r$ et il existe une permutation $\sigma \in \mathcal{S}_r$ telle que $Q_k = P_{\sigma(k)}$ et $n_k = m_{\sigma(k)}$ pour tout k compris entre 1 et r .

Preuve. Pour l'existence d'une telle décomposition en facteurs irréductibles, on procède par récurrence sur le degré $n \geq 1$ de P . Pour $n = 1$, on a vu que le

polynôme P est irréductible. Supposons le résultat acquis pour tous les polynômes de degré k compris entre 1 et $n-1 \geq 1$ et soit $P \in \mathbb{K}[X]$ de degré n . Si P est irréductible, c'est terminé, sinon il s'écrit $P = AB$ avec A, B non constants et de degré compris entre 1 et $n-1$. Il suffit d'utiliser l'hypothèse de récurrence pour A et B .

Pour l'unicité à l'ordre près de cette décomposition, on procède également par récurrence sur le degré $n \geq 1$ de P . Pour $n = 1$, on a $P(X) = \lambda(X - \alpha)$ avec $\lambda \in \mathbb{K}^*$ et $X - \alpha$ irréductible uniquement déterminés. Supposons le résultat acquis pour tous les polynômes de degré k compris entre 1 et $n-1 \geq 1$ et supposons que

$P \in \mathbb{K}[X]$ de degré n ait deux décompositions, soit $P = \lambda \prod_{k=1}^r P_k^{m_k} = \mu \prod_{k=1}^s Q_k^{n_k}$.

L'identification des coefficients dominants nous dit que $\lambda = \mu$ est le coefficient dominant de P . Comme Q_1 irréductible divise le produit $\prod_{k=1}^r P_k^{m_k}$, le lemme d'Eu-

clide nous dit qu'il divise l'un des P_j , il est donc égal à ce P_j puisque ces deux polynômes sont irréductibles. Supposons que $n_1 \leq m_j$. En divisant par $Q_1^{n_1}$, on a

$Q_1^{m_j - n_1} \prod_{\substack{k=1 \\ k \neq j}}^r P_k^{m_k} = \prod_{k=2}^s Q_k^{n_k}$ et nécessairement $m_j = n_1$ (sinon le lemme d'Euclide

nous dit Q_1 est égal à l'un des Q_k avec $k \geq 2$, ce qui n'est pas). On note $j = \sigma(1)$

et on a $\prod_{\substack{k=1 \\ k \neq j}}^r P_k^{m_k} = \prod_{k=2}^s Q_k^{n_k}$. L'hypothèse de récurrence permet alors de conclure. $\square$

Corollaire 9.8. *Tout polynôme non constant $P \in \mathbb{K}[X]$ admet au moins un diviseur irréductible.*

Pour la suite de ce paragraphe, on s'intéresse aux polynômes irréductibles dans $\mathbb{R}[X]$. On sait déjà que les polynômes de degré 1 sont irréductibles. Un polynôme à coefficients réels étant aussi un élément de $\mathbb{C}[X]$, il est scindé dans cet anneau de polynômes.

Lemme 9.5 *Soit $P(X) = aX^2 + bX + c$ un polynôme de degré 2 dans $\mathbb{R}[X]$. Ce polynôme est irréductible si, et seulement si, $\delta = b^2 - 4ac < 0$.*

Preuve. On a $P(X) = a \left(\left(X + \frac{b}{2a} \right)^2 - \frac{b^2 - 4ac}{4a^2} \right) = a \left(\left(X + \frac{b}{2a} \right)^2 - \frac{\delta}{4a^2} \right)$.

Si $\delta = 0$, on a alors $P(X) = a \left(X + \frac{b}{2a} \right)^2$ qui est réductible. Si $\delta > 0$, on a alors

$P(X) = a \left(X + \frac{b}{2a} - \frac{\sqrt{\delta}}{2a} \right) \left(X + \frac{b}{2a} + \frac{\sqrt{\delta}}{2a} \right)$ qui est réductible. Si $\delta < 0$, on

a alors $P(x) > 0$ pour tout $x \in \mathbb{R}$ et P n'a pas de racines réelles, il est donc irréductible. $\square$

Lemme 9.6 *Soit P un polynôme non constant dans $\mathbb{R}[X]$. Si $\alpha \in \mathbb{C}$ est une racine de P de multiplicité $m \geq 1$, alors son conjugué $\bar{\alpha}$ est aussi racine de P avec la même multiplicité.*

Preuve. Dans $\mathbb{C}[X]$, on a $P^{(k)}(\alpha) = 0$ pour tout entier k compris entre 1 et $m-1$ et $P^{(m)}(\alpha) \neq 0$. Comme P et tous ses polynômes dérivés sont à coefficients réels, on a $\overline{P^{(k)}(\alpha)} = P^{(k)}(\bar{\alpha}) = 0$ pour tout entier k compris entre 1 et $m-1$ et $\overline{P^{(m)}(\alpha)} = P^{(m)}(\bar{\alpha}) \neq 0$, ce qui revient à dire que $\bar{\alpha}$ est racine de P avec la même multiplicité m . $\square$

Théorème 9.19.

Les polynômes réels irréductibles sont les polynômes de degré 1 et les polynômes de degré 2, $P(X) = aX^2 + bX + c$ tels que $b^2 - 4ac < 0$.

Preuve. On sait déjà que ces polynômes sont irréductibles. Si P de degré supérieur ou égal à 2 dans $\mathbb{R}[X]$ est irréductible, il n'a pas de racines réelles et sa décomposition en facteurs irréductibles dans $\mathbb{C}[X]$ est nécessairement de la forme

$P(X) = \prod_{k=1}^r (X - \alpha_k)^{m_k} (X - \bar{\alpha}_k)^{m_k}$, où $r \geq 1$, les α_k sont complexes non réels et les m_k entiers naturels non nuls. Comme :

$$(X - \alpha_k)^{m_k} (X - \bar{\alpha}_k)^{m_k} = \left(X^2 - 2 \operatorname{Re}(\alpha_k) X + |\alpha_k|^2 \right)^{m_k} \in \mathbb{R}[X]$$

on a nécessairement $r = m_1 = 1$ et $P(X) = X^2 - 2 \operatorname{Re}(\alpha_1) X + |\alpha_1|^2$ avec $(\operatorname{Re}(\alpha_1))^2 - |\alpha_1|^2 = (\operatorname{Im}(\alpha_1))^2 > 0$ puisque α_1 est non réel. $\square$

Le théorème de décomposition en facteurs irréductibles prend donc la forme suivante dans $\mathbb{R}[X]$.

Théorème 9.20.

Tout polynôme non constant $P \in \mathbb{R}[X]$ est produit de polynômes irréductibles de degré 1 ou 2, c'est-à-dire qu'il existe une constante $\lambda \in \mathbb{R}^*$, deux entiers naturels r et s , des réels $a_1, \dots, a_r$ deux à deux distincts, des entiers naturels non nuls $n_1, \dots, n_r$, des couples de réels $(b_1, c_1), \dots, (b_s, c_s)$ deux à deux distincts tels que $b_k^2 - 4c_k < 0$ pour tout k et des entiers naturels non nuls $m_1, \dots, m_s$ tels que :

$$P = \lambda \prod_{k=1}^r (X - a_k)^{n_k} \prod_{k=1}^s (X^2 + b_k X + c_k)^{m_k}$$

une telle décomposition est unique à l'ordre près des facteurs.

Le théorème précédent est utile pour la décomposition en éléments simples des fractions rationnelles réelles.

La décomposition en produit de polynômes irréductibles peut être utilisée pour calculer le pgcd de deux polynômes non constants. Tout est basé sur le lemme qui suit, conséquence de l'unicité de la décomposition en produit de polynômes irréductibles.

Lemme 9.7 Soient P, Q deux polynômes unitaires non constants et $P = \prod_{k=1}^r P_k^{m_k}$,

$Q = \prod_{k=1}^r P_k^{n_k}$ leurs décompositions en produit de polynômes irréductibles avec les P_k unitaires irréductibles deux à deux distincts et les n_k, m_k entiers naturels (certains de ces entiers pouvant être nuls). On a alors :

$$(P \text{ divise } Q) \Leftrightarrow (\forall k \in \{1, \dots, r\}, m_k \leq n_k)$$

Preuve. Dire que P divise Q équivaut à dire qu'il existe un polynôme unitaire Q_1 tel que $Q = PQ_1$ et en écrivant $Q_1 = \prod_{k=1}^r P_k^{s_k}$ où les s_k sont des entiers naturels, on a $Q = P_1^{n_1} \dots P_r^{n_r} = P_1^{m_1+s_1} \dots P_r^{m_r+s_r}$. L'unicité de la décomposition en facteurs irréductibles de Q nous dit alors que $n_k = m_k + s_k \geq m_k$ pour tout k . La réciproque est évidente. $\square$

Théorème 9.21.

Avec les notations du lemme précédent, on a $P \wedge Q = \prod_{k=1}^r P_k^{\min(m_k, n_k)}$.

Preuve. Le polynôme $\Delta = \prod_{k=1}^r P_k^{\min(m_k, n_k)}$ divise P et Q d'après le lemme précédent. Si D est un diviseur de P et Q , il s'écrit sous la forme $D = \prod_{k=1}^r P_k^{s_k}$ où les s_k sont des entiers naturels tels que $s_k \leq m_k$ et $s_k \leq n_k$ pour tout k compris entre 1 et r , on a donc $s_k \leq \min(m_k, n_k)$ pour tout k compris entre 1 et r et D divise Δ . Donc Δ est bien le pgcd de P et Q . $\square$

De manière plus générale si $P_1, \dots, P_n$ sont polynômes unitaires non constants et $P_j = \prod_{k=1}^r P_k^{m_{j,k}}$ la décomposition en produit de polynômes irréductibles de cha-

cun de ces polynômes, on a $P_1 \wedge \dots \wedge P_n = \prod_{k=1}^r P_k^{\min(m_{1,k}, \dots, m_{n,k})}$.

9.7 Décomposition des fractions rationnelles en éléments simples

Les résultats de ce paragraphe seront utilisés pour le calcul des primitives des fonctions rationnelles (voir le cours d'analyse).

Une fonction rationnelle est une fonction $x \mapsto f(x) = \frac{P(x)}{Q(x)}$, où P, Q sont deux fonctions polynomiales à coefficients complexes avec $Q \neq 0$, ces polynômes P et Q étant premiers entre eux. Une telle fonction est définie sur $\mathbb{R}$ privé des racines réelles de Q .

Définition 9.16. On dit qu'un nombre complexe α est un pôle d'ordre m de $f = \frac{P}{Q}$, s'il est racine d'ordre m de Q .

Lemme 9.8 Soient $f = \frac{P}{Q}$ une fraction rationnelle et $\alpha \in \mathbb{C}$ un pôle d'ordre $m \geq 1$ de f . Il existe des nombres complexes $\lambda_1, \dots, \lambda_m$, avec $\lambda_m \neq 0$ et une fonction rationnelle f_1 n'admettant pas α pour pôle, uniquement déterminés, tels que $f(X) = \sum_{k=1}^m \frac{\lambda_k}{(X - \alpha)^k} + f_1(X)$.

Preuve. On a $Q(X) = (X - \alpha)^m Q_1(X)$, où $Q_1 \in \mathbb{C}[X]$ est tel que $Q(\alpha) \neq 0$. Les polynômes Q_1 et $(X - \alpha)^m$ sont donc premiers entre eux et le théorème de Bézout nous dit qu'il existe deux polynômes U, V tels que :

$$U(X)(X - \alpha)^m + V(X)Q_1(X) = 1$$

On peut alors écrire que :

$$\begin{aligned} f(x) &= \frac{P(x)}{Q(x)} = \frac{P(x)(U(X)(X - \alpha)^m + V(X)Q_1(X))}{(X - \alpha)^m Q_1(X)} \\ &= \frac{V(X)}{(X - \alpha)^m} + \frac{P(x)U(X)}{Q_1(X)} = \frac{V(X)}{(X - \alpha)^m} + g_1(X) \end{aligned}$$

où $g_1 = \frac{P \cdot U}{Q_1}$ est une fonction rationnelle n'admettant pas α pour pôle.

En utilisant la formule de Taylor pour les polynômes (ce qui revient à utiliser $((X - \alpha)^k)_{k \in \mathbb{N}}$ comme base de $\mathbb{C}[X]$), on peut écrire que :

$$V(X) = \sum_{j=0}^{m-1} a_j (X - \alpha)^j + (X - \alpha)^m V_1(X)$$

où les a_j sont des nombres complexes et V_1 est un polynôme, ce qui nous donne :

$$\begin{aligned} \frac{V(X)}{(X - \alpha)^m} &= \sum_{j=0}^{m-1} \frac{a_j}{(X - \alpha)^{m-j}} + V_1(X) \\ &= \sum_{k=1}^m \frac{a_{m-k}}{(X - \alpha)^k} + V_1(X) = \sum_{k=1}^m \frac{\lambda_k}{(X - \alpha)^k} + V_1(X) \end{aligned}$$

et $f(X) = \sum_{k=1}^m \frac{\lambda_k}{(X - \alpha)^k} + f_1(X)$, où $f_1 = g_1 + V_1$ est une fonction rationnelle n'admettant pas α pour pôle et $\lambda_m = a_0 = V(\alpha) \neq 0$ (car $V(\alpha)Q_1(\alpha) = 1$).

Pour montrer l'unicité d'une telle décomposition, il nous suffit de montrer que l'égalité $\sum_{k=1}^n \frac{\lambda_k}{(X - \alpha)^k} + f_1(X) = 0$, où les λ_k sont des nombres complexes et f_1

une fonction rationnelle n'admettant pas α pour pôle équivalent à $f_1 = 0$ et $\lambda_k = 0$ pour tout k . Supposons que les λ_k ne sont pas tous nuls et soit p le plus grand entier compris entre 1 et m tel que $\lambda_p \neq 0$. En multipliant par $(X - \alpha)^p$, on a :

$$\lambda_p + \sum_{k=1}^{p-1} \lambda_k (X - \alpha)^{p-k} + (X - \alpha)^p f_1(X) = 0$$

et l'évaluation en α aboutit à $\lambda_p = 0$ ($f_1(\alpha)$ est bien défini puisque α n'est pas pôle de f_1), ce qui est contradictoire. $\square$

Avec les notations de la démonstration qui précède, on a :

$$(X - \alpha)^m f(X) = \frac{P(x)}{Q_1(x)} = \lambda_m + (X - \alpha) g(X)$$

où g est une fonction rationnelle n'admettant pas α pour pôle et l'évaluation en α nous donne $\lambda_m = \frac{P(\alpha)}{Q_1(\alpha)}$. Pour un pôle simple (i. e. $m = 1$), cela donne

$$f(X) = \frac{\lambda_1}{X - \alpha} + f_1(X) \text{ avec } \lambda_1 = \frac{P(\alpha)}{Q_1(\alpha)} = \frac{P(\alpha)}{Q'(\alpha)}.$$

Théorème 9.22.

Soient $f = \frac{P}{Q}$ une fraction rationnelle (les polynômes P et Q étant premiers entre eux) et $\alpha_1, \dots, \alpha_p$ les pôles de f d'ordres respectifs $m_1, \dots, m_p$. Il existe des nombres complexes $(\lambda_{j,k})_{\substack{1 \leq j \leq p \\ 1 \leq k \leq m_j}}$, avec $\lambda_{j,m_j} \neq 0$ pour tout j compris entre 1 et p , et une fonction polynomiale R , uniquement déterminés, tels que :

$$f(X) = \sum_{j=1}^p \sum_{k=1}^{m_j} \frac{\lambda_{j,k}}{(X - \alpha_j)^k} + R(X) \quad (9.2)$$

Preuve. Pour l'existence d'une telle décomposition, on procède par récurrence sur le nombre $p \geq 0$ de pôles de f . Pour $p = 0$, f n'a pas de pôles complexes, ce qui signifie que Q n'a pas de racines et c'est donc un polynôme constant non nul (on est sur $\mathbb{C}$ qui est algébriquement clos) et $f = R$ est une fonction polynomiale. Supposons le résultat acquis pour les fonctions rationnelles ayant $p - 1 \geq 0$ pôles complexes et soit f ayant p pôles complexes $\alpha_1, \dots, \alpha_p$ d'ordres respectifs $m_1, \dots, m_p$. Le théorème précédent nous assure de la décomposition :

$$f(X) = \frac{P(X)}{Q(X)} = \sum_{k=1}^{m_p} \frac{\lambda_{p,k}}{(X - \alpha_p)^k} + f_1(X)$$

où $P(X)$ est premier avec $Q(X) = \prod_{j=1}^p (X - \alpha_j)^{m_j}$ et f_1 est une fonction rationnelle n'admettant pas α_p pour pôle. On a alors :

$$f_1(X) = \frac{1}{(X - \alpha_p)^{m_p}} \left(\frac{P(X)}{Q_1(X)} - \sum_{k=1}^{m_p} \lambda_{p,k} (X - \alpha_p)^{m_p-k} \right) = \frac{1}{(X - \alpha_p)^{m_p}} \frac{P_1(X)}{Q_1(X)}$$

où $Q_1(X) = \prod_{j=1}^{p-1} (X - \alpha_j)^{m_j}$ et $P_1(X) = P(X) - Q_1(X) \sum_{k=1}^{m_p} \lambda_{p,k} (X - \alpha_p)^{m_p-k}$.

Comme α_p n'est pas pôle de f_1 , le polynôme P_1 est multiple de $(X - \alpha_p)^{m_p}$, donc $f_1(X) = \frac{P_2(X)}{Q_1(X)}$ avec $P_2(\alpha_j) = \frac{P_1(\alpha_j)}{(\alpha_j - \alpha_p)^{m_p}} = \frac{P(\alpha_j)}{(\alpha_j - \alpha_p)^{m_p}} \neq 0$ pour tout j compris entre 1 et $p-1$, ce qui signifie que P_2 est premier avec Q_1 . La fonction rationnelle f_1 a donc $p-1$ pôles $\alpha_1, \dots, \alpha_{p-1}$ d'ordres respectifs $m_1, \dots, m_{p-1}$, ce qui permet de conclure avec l'hypothèse de récurrence.

Pour l'unicité, il suffit de montrer que l'égalité $\sum_{j=1}^p \sum_{k=1}^{m_j} \frac{\lambda_{j,k}}{(X - \alpha_j)^k} + R(X) = 0$, où $(\lambda_{j,k})_{\substack{1 \leq j \leq p \\ 1 \leq k \leq m_j}}$ est une suite de nombres complexes et R est une fonction polynomiale impose $R = 0$ et tous les $\lambda_{j,k}$ nuls. En réduisant au même dénominateur, on a l'égalité polynomiale :

$$\sum_{j=1}^p \sum_{k=1}^{m_j} \lambda_{j,k} (X - \alpha_j)^{m_j-k} = - \prod_{j=1}^p (X - \alpha_j)^{m_j} R(X)$$

Si $R \neq 0$ le polynôme du second membre est de degré au moins égal à $\sum_{j=1}^p m_j$ et celui du premier membre de degré inférieur ou égal à $\max_{1 \leq j \leq p} (m_j - 1)$, ce qui donne une impossibilité. On a donc $R = 0$ et tous les $\lambda_{j,k}$ sont nuls. $\square$

Définition 9.17. Avec les notations du théorème précédent, on dit que la décomposition (9.2) est la décomposition en éléments simples de f et, pour tout j compris entre 1 et p , on dit que la fraction rationnelle $\sum_{k=1}^{m_j} \frac{\lambda_{j,k}}{(X - \alpha_j)^k}$ est la partie polaire de f associée au pôle α_j .

Dans une telle décomposition en éléments simples, le polynôme R est le quotient dans la division euclidienne de P par Q . En effet, on a :

$$P(X) = \sum_{j=1}^p \sum_{k=1}^{m_j} \frac{\lambda_{j,k}}{(X - \alpha_j)^k} Q(X) + Q(X) R(X)$$

avec $\frac{Q(X)}{(X - \alpha_j)^k} = (X - \alpha_j)^{m_j-k} \prod_{\substack{i=1 \\ i \neq j}}^p (X - \alpha_i)^{m_i}$ de degré strictement inférieur à $\deg(Q)$.

Si la fraction rationnelle f est à coefficients réels et si α est un pôle complexe non réel d'ordre m , son conjugué $\bar{\alpha}$ est aussi pôle d'ordre m de f .

En notant, pour $f \in \mathbb{R}(X)$, $x_1, \dots, x_r$ les pôles réels de f d'ordres respectifs $m_1, \dots, m_r$ (quand il en existe) et $\alpha_1, \bar{\alpha}_1, \dots, \alpha_s, \bar{\alpha}_s$ les pôles complexes non réels deux à deux distincts f (quand il en existe), où α_k et $\bar{\alpha}_k$ sont d'ordre n_k , on a la

décomposition en éléments simples dans $\mathbb{C}(X)$:

$$f(X) = \sum_{j=1}^r \sum_{k=1}^{m_j} \frac{\lambda_{j,k}}{(X - x_j)^k} + \sum_{j=1}^s \sum_{k=1}^{n_j} \left(\frac{\mu_{j,k}}{(X - \alpha_j)^k} + \frac{\mu'_{j,k}}{(X - \bar{\alpha}_j)^k} \right) + R(X)$$

Comme R est le quotient dans la division euclidienne de P par Q , ces polynômes étant à coefficients réels, on a $R \in \mathbb{R}[X]$ et comme $f(X) = \bar{f}(X)$, on a :

$$\begin{aligned} f(X) &= \sum_{j=1}^r \sum_{k=1}^{m_j} \frac{\overline{\lambda_{j,k}}}{(X - \bar{x}_j)^k} + \sum_{j=1}^s \sum_{k=1}^{n_j} \left(\frac{\overline{\mu_{j,k}}}{(X - \bar{\alpha}_j)^k} + \frac{\overline{\mu'_{j,k}}}{(X - \alpha_j)^k} \right) + \bar{R}(X) \\ &= \sum_{j=1}^r \sum_{k=1}^{m_j} \frac{\overline{\lambda_{j,k}}}{(X - x_j)^k} + \sum_{j=1}^s \sum_{k=1}^{n_j} \left(\frac{\overline{\mu_{j,k}}}{(X - \bar{\alpha}_j)^k} + \frac{\overline{\mu'_{j,k}}}{(X - \alpha_j)^k} \right) + R(X) \end{aligned}$$

et par unicité de la décomposition en éléments simples, on déduit que $\overline{\lambda_{j,k}} = \lambda_{j,k}$, ce qui signifie que les $\lambda_{j,k}$ sont réels et $\mu'_{j,k} = \overline{\mu_{j,k}}$, ce qui nous donne dans $\mathbb{R}(X)$:

$$\begin{aligned} f(X) &= \sum_{j=1}^r \sum_{k=1}^{m_j} \frac{\lambda_{j,k}}{(X - x_j)^k} + \sum_{j=1}^s \sum_{k=1}^{n_j} \left(\frac{\mu_{j,k}}{(X - \alpha_j)^k} + \frac{\overline{\mu_{j,k}}}{(X - \bar{\alpha}_j)^k} \right) + R(X) \\ &= \sum_{j=1}^r \sum_{k=1}^{m_j} \frac{\lambda_{j,k}}{(X - x_j)^k} + \sum_{j=1}^s \sum_{k=1}^{n_j} \left(\frac{\mu_{j,k}(X - \bar{\alpha}_j)^k + \overline{\mu_{j,k}}(X - \alpha_j)^k}{|X - \alpha_j|^{2k}} \right) + R(X) \end{aligned}$$

En notant $\alpha_j = a_j + ib_j$, où $a_j = \operatorname{Re}(\alpha_j)$ et $b_j = \operatorname{Im}(\alpha_j)$, on a :

$$|X - \alpha_j|^2 = (X - a_j)^2 + b_j^2 = X^2 + \beta_j X + \gamma_j$$

avec $\beta_j^2 - 4\gamma_j < 0$.

Lemme 9.9 Soit $A(X) = X^2 + \beta X + \gamma$ un polynôme réel irréductible (i. e. tel que $\beta^2 - 4\gamma < 0$). Pour tout polynôme $P \in \mathbb{R}[X]$ et tout entier $p \geq 1$, il existe des polynômes $A_0, \dots, A_{p-1}$ de degré au plus égal à 1 et un polynôme R_p tels que :

$$\begin{aligned} P(X) &= A_0(X) + A_1(X)P(X) + A_2(X)P^2(X) + \dots \\ &\quad + A_{p-1}(X)P^{p-1}(X) + R_p(X)P^p(X) \end{aligned}$$

Preuve. Pour $n = 1$, il s'agit de la division euclidienne par A . Supposant le résultat acquis pour $n \geq 1$, la division euclidienne de R_p par A , nous donne le résultat pour $n + 1$. $\square$

Le lemme précédent est en fait valable pour tout polynôme A non constant et les A_k tels que $\deg(A_k) < \deg(A)$ sont uniquement déterminés ainsi que R_p .

En utilisant le lemme, on peut écrire pour tous j, k :

$$\frac{\mu_{j,k}(X - \bar{\alpha}_j)^k + \overline{\mu_{j,k}}(X - \alpha_j)^k}{|X - \alpha_j|^{2k}} = \sum_{i=1}^k \frac{u_i X + v_i}{(X^2 + \beta_j X + \gamma_j)^i} + R_{j,k}(X)$$

En définitive, on a le théorème de décomposition en éléments simples pour les fractions rationnelles à coefficients réels.

Théorème 9.23.

Soient $f = \frac{P}{Q}$ une fraction rationnelle réelle (les polynômes P et Q étant premiers entre eux) avec :

$$Q(X) = \prod_{k=1}^r (X - x_k)^{m_k} \prod_{k=1}^s (X^2 + \beta_k X + \gamma_k)^{n_k}$$

où r, s sont deux entiers naturels, $x_1, \dots, x_r$ des réels deux à deux distincts, $(\beta_1, \gamma_1), \dots, (\beta_s, \gamma_s)$ des couples de réels deux à deux distincts tels que $\beta_k^2 - 4\gamma_k < 0$ pour tout k et $m_1, \dots, m_r, n_1, \dots, n_s$ des entiers naturels non nuls. Il existe des réels $(\lambda_{j,k})_{\substack{1 \leq j \leq r \\ 1 \leq k \leq m_j}}, (\mu_{j,k})_{\substack{1 \leq j \leq s \\ 1 \leq k \leq n_j}}, (\nu_{j,k})_{\substack{1 \leq j \leq s \\ 1 \leq k \leq n_j}}$ et une fonction polynomiale réelle R , uniquement déterminés, tels que :

$$f(X) = \sum_{j=1}^r \sum_{k=1}^{m_j} \frac{\lambda_{j,k}}{(X - \alpha_j)^k} + \sum_{j=1}^s \sum_{k=1}^{n_j} \frac{\mu_{j,k}X + \nu_{j,k}}{(X^2 + \beta_j X + \gamma_j)^k} + R(X)$$

Preuve. Nous avons prouvé l'existence d'une telle décomposition et l'unicité se prouve comme dans le cas complexe en raisonnant sur les degrés. $\square$

9.8 Polynômes d'interpolation de Lagrange

Pour tout entier naturel n , on se donne une famille $(x_i)_{0 \leq i \leq n}$ de $n+1$ éléments de $\mathbb{K}$ deux à deux distincts et à cette famille on associe la famille $(L_i)_{0 \leq i \leq n}$ de polynômes définie par :

$$L_i(X) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{X - x_j}{x_i - x_j} \quad (0 \leq i \leq n)$$

On peut remarquer que $L_i(x_j) = \delta_{i,j}$ pour $0 \leq i, j \leq n$.

Théorème 9.24.

Avec les notations qui précèdent, la famille de polynômes $(L_i)_{0 \leq i \leq n}$ est une base de $\mathbb{K}_n[X]$.

Preuve. Comme on a $n+1 = \dim(\mathbb{K}_n[X])$ polynômes, il suffit de montrer que la famille $(L_i)_{0 \leq i \leq n}$ est libre. Si $\sum_{i=1}^n \lambda_i L_i = 0$, l'évaluation en x_j , pour j compris entre 0 et n , nous donne $\lambda_j = 0$. La famille de polynômes $(L_i)_{0 \leq i \leq n}$ est donc libre et c'est une base de $\mathbb{K}_n[X]$. $\square$

Tout polynôme $P \in \mathbb{K}_n[X]$ s'écrit donc, $P = \sum_{i=1}^n \lambda_i L_i$, l'évaluation en x_j nous donnant $\lambda_j = P(x_j)$, pour tout j compris entre 0 et n .

Théorème 9.25.

Avec les notations qui précèdent, pour tous scalaires $y_0, y_1, \dots, y_n$, il existe un unique polynôme $P \in \mathbb{K}_n[X]$ tel que $P(x_j) = y_j$ pour tout j compris entre 0 et n . Ce polynôme est donné par $P = \sum_{i=1}^n y_i L_i$.

Preuve. L'application linéaire $\varphi : \mathbb{K}_n[X] \rightarrow \mathbb{K}^{n+1}$ définie par :

$$\forall P \in \mathbb{K}_n[X], \varphi(P) = (P(x_0), P(x_1), \dots, P(x_n))$$

est surjective. En effet, pour $y = (y_0, \dots, y_n) \in \mathbb{K}^{n+1}$, le polynôme $P = \sum_{i=1}^n y_i L_i$ est tel que $\varphi(P) = y$. Comme les espaces de départ et d'arrivée ont la même dimension, cette application linéaire φ est un isomorphisme, donc pour tout $y = (y_0, \dots, y_n) \in \mathbb{K}^{n+1}$, il existe un unique polynôme P est tel que $\varphi(P) = y$. $\square$

Avec les notations du théorème précédent, on dit que P est le polynôme d'interpolation de Lagrange associé à $x = (x_i)_{0 \leq i \leq n}$ et $y = (y_i)_{0 \leq i \leq n}$.

Pour tout entier p compris entre 0 et n , on a $\sum_{i=1}^n x_i^p L_i(X) = X^p$. En effet, le polynôme X^p est l'unique polynôme $P \in \mathbb{K}_n[X]$ tel que $P(x_j) = x_j^p$ pour tout j compris entre 0 et n , c'est donc le polynôme d'interpolation de Lagrange $\sum_{i=1}^n x_i^p L_i(X)$.

Les polynômes d'interpolation de Lagrange ont des applications intéressantes. Par exemple, les méthode de quadrature de Newton-Cotes sont basées sur l'interpolation de Lagrange. Le point de départ est le résultat suivant.

Théorème 9.26.

Soient $a < b$ deux réels, n un entier naturel non nul et $(x_k)_{0 \leq k \leq n}$ une suite de points deux à deux distincts dans $[a, b]$. Il existe une unique suite de réels $(\mu_k)_{0 \leq k \leq n}$ telle que pour tout polynôme P dans $\mathbb{R}_n[X]$ on a :

$$\int_a^b P(x) dx = \sum_{k=0}^n \mu_k P(x_k) \quad (9.3)$$

Preuve. Soit $(L_k)_{0 \leq k \leq n}$ la base de Lagrange de $\mathbb{R}_n[X]$ associée à la suite $(x_k)_{0 \leq k \leq n}$. Par linéarité, la propriété (9.3) est vérifiée pour tout P dans $\mathbb{R}_n[X]$ si, et seulement si, elle est vérifiée pour tous les L_k , ce qui équivaut à $\int_a^b L_k(x) dx = \mu_k$ pour tout entier k compris entre 0 et n . On a donc ainsi prouvé l'existence et l'unicité des coefficients μ_k . $\square$

9.9 Exercices

Exercice 9.1. Soit B un polynôme non nul de degré m . On désigne par $F = \mathbb{K}[X] \cdot B = \{QB \mid Q \in \mathbb{K}[X]\}$ le sous-espace vectoriel de $\mathbb{K}[X]$ formé des multiples de B . Montrer que pour tout entier naturel n , on a $\mathbb{K}_n[X] = F_n \oplus G_n$, où $F_n = \mathbb{K}_n[X] \cap F$ et $G_n = \mathbb{K}_n[X] \cap \mathbb{K}_{m-1}[X]$ (pour $m = 0$, on a $G_n = \{0\}$). En déduire le théorème de division euclidienne.

Solution. Pour $m = 0$, on a $G_n = \{0\}$, $F = \mathbb{K}[X]$ et $F_n = \mathbb{K}_n[X]$. Pour $m \geq 1$ et $n \leq m - 1$, on a $F_n = \{0\}$ (pour $Q \neq 0$, on a $\deg(QB) \geq \deg(B) = m > n$, donc $QB \notin \mathbb{K}_n[X]$) et $G_n = \mathbb{K}_n[X]$. Supposons que $n \geq m \geq 1$. Dans ce cas, $G_n = \mathbb{K}_{m-1}[X]$ est de dimension m , F_n est de dimension $n + 1 - m$ avec $(B, XB, \dots, X^{n-m}B)$ pour base (si $QB \in F$, on a $\deg(QB) = \deg(Q) + m \leq n$, donc $\deg(Q) \leq n - m$) et $F_n \cap G_n = \{0\}$ (un polynôme dans cette intersection est de degré au plus égal à $m - 1 < n$ et multiple de B , c'est donc le polynôme nul). Il en résulte que $\mathbb{K}_n[X] = F_n \oplus G_n$. Soit $A \in \mathbb{K}[X]$. Si $A = 0$, il s'écrit $A = BQ + R$ avec $Q = B = 0$ et cette écriture est unique si on veut $R \in \mathbb{K}_{m-1}[X]$. Si $A \neq 0$, il est de degré $n \geq 0$, donc dans $\mathbb{K}_n[X] = F_n \oplus G_n$, il s'écrit donc de manière unique $A = BQ + R$ avec $BQ \in F_n$ et $R \in G_n$ (la condition $R \in \mathbb{K}_{m-1}[X]$ impose $BQ = A - R \in \mathbb{K}_n[X]$).

Exercice 9.2. Soient $n \in \mathbb{N}^*$ et $\theta \in \mathbb{R}$. Déterminer le reste dans la division euclidienne de $A_n(X) = (\cos(\theta) + X \sin(\theta))^n$ par $B(X) = X^2 + 1$.

Solution. On a $A_n = BQ_n + a_nX + b_n$ dans $\mathbb{R}[X] \subset \mathbb{C}[X]$ avec $(a_n, b_n) \in \mathbb{R}^2$. En évaluant en $i \in \mathbb{C}$, on a $A_n(i) = e^{in\theta} = a_n i + b_n$, ce qui donne :

$$a_n = \sin(n\theta) \text{ et } b_n = \cos(n\theta)$$

Exercice 9.3. On définit la suite $(H_n)_{n \in \mathbb{N}}$ des polynômes de Hilbert par :

$$H_0(X) = 1, \forall n \in \mathbb{N}^*, H_n(X) = \frac{X(X-1) \cdots (X-n+1)}{n!}$$

et $u : \mathbb{K}[X] \rightarrow \mathbb{K}[X]$ est l'application linéaire définie par :

$$\forall P \in \mathbb{K}[X], u(P)(X) = P(X+1) - P(X)$$

On note $u^0 = \text{Id}$ et pour tout entier $r \geq 1$, $u^r = u \circ u \circ \cdots \circ u$ (r fois).

1. Montrer que, pour tout entier naturel n , $(H_k)_{0 \leq k \leq n}$ est une base de $\mathbb{K}_n[X]$.
2. Montrer que $u(H_0) = 0$ et $u(H_n) = H_{n-1}$ pour tout $n \in \mathbb{N}^*$.
3. Calculer $u^k(H_n)$ pour tous k, n dans $\mathbb{N}$.

4. Calculer $u^k(H_n)(0)$ pour tous k, n dans $\mathbb{N}$.

5. Montrer que pour tout $P \in \mathbb{K}_n[X]$, on a $P = \sum_{k=0}^n u^k(P)(0) H_k$.

Solution.

1. $(H_k)_{0 \leq k \leq n}$ est formée de $n+1$ polynômes échelonnés en degrés dans $\mathbb{K}_n[X]$, c'est donc une base.
2. Comme H_0 est constant, on a $u(H_0) = 0$. Pour $n = 1$, on a $u(H_1) = 1 = H_0$. Pour $n \geq 2$, on a :

$$\begin{aligned} u(H_n)(X) &= H_n(X+1) - H_n(X) = \frac{1}{n!} \left(\prod_{k=0}^{n-1} (X+1-k) - \prod_{k=0}^{n-1} (X-k) \right) \\ &= \frac{1}{n!} \left(\prod_{k=0}^{n-1} (X+1-k) - \prod_{j=1}^n (X-(j-1)) \right) \\ &= \frac{1}{n!} \left(\prod_{k=1}^{n-1} (X+1-k) \right) (X+1 - (X+1-n)) \\ &= \frac{1}{(n-1)!} \prod_{k=1}^{n-1} (X+1-k) = \frac{1}{(n-1)!} \prod_{j=0}^{n-2} (X-j) = H_{n-1}(X) \end{aligned}$$

On a donc $u(H_n) = H_{n-1}$.

3. Par récurrence, on vérifie que $u^k(H_n) = H_{n-k}$ pour $0 \leq k \leq n$ et en conséquence $u^k(H_n) = 0$ pour $k \geq n+1$.
4. Ce qui donne $u^k(H_n)(0) = \delta_{n,k}$ puisque $H_{n-k}(0) = 0$ pour $0 \leq k \leq n-1$ et $H_0 = 1$.
5. Si P est un polynôme non nul de degré $p \geq 1$, il s'écrit alors $P = \sum_{k=0}^p \lambda_k H_k$ et pour $0 \leq j \leq p$, on a :

$$u^j(P)(0) = \sum_{k=0}^p \lambda_k u^j(H_k)(0) = \sum_{k=0}^p \lambda_k \delta_{k,j} = \lambda_j$$

ce qui donne $P = \sum_{j=0}^p u^j(P)(0) H_j$ ($u^j(P)(0) = 0$ pour $j > p$). Pour $P = 0$, cette relation est encore vérifiée.

Exercice 9.4. Soient P un polynôme de degré $n \geq 1$, α un scalaire et m un entier compris entre 1 et n . Déterminer le quotient et le reste dans la division euclidienne de P par $(X - \alpha)^m$.

Solution. On a la formule de Taylor :

$$\begin{aligned} P(X) &= \sum_{k=0}^n \frac{P^{(k)}(\alpha)}{k!} (X - \alpha)^k \\ &= \sum_{k=0}^{m-1} \frac{P^{(k)}(\alpha)}{k!} (X - \alpha)^k + (X - \alpha)^m \sum_{k=m}^n \frac{P^{(k)}(\alpha)}{k!} (X - \alpha)^{k-m} \end{aligned}$$

donc le quotient et le reste dans la division euclidienne de P par $(X - \alpha)^m$ sont donnés par :

$$Q(X) = \sum_{k=m}^n \frac{P^{(k)}(\alpha)}{k!} (X - \alpha)^{k-m} \text{ et } R(X) = \sum_{k=0}^{m-1} \frac{P^{(k)}(\alpha)}{k!} (X - \alpha)^k$$

Exercice 9.5. Pour tout entier $n \geq 1$, on note $P_n(X) = \sum_{k=0}^n \frac{1}{k!} X^k$ dans $\mathbb{K}[X]$. Montrer que les racines de P_n sont nécessairement simples.

Solution. Pour $n = 1$, $P_1(X) = 1 + X$ et $x = -1$ est racine simple de P_1 . Pour $n \geq 2$, on a $P'_n(X) = \sum_{k=1}^n \frac{1}{(k-1)!} X^{k-1} = P_{n-1}(X)$. Il en résulte que si α est une racine d'ordre au moins égal à 2 de P_n , on a alors $P_n(\alpha) = P_{n-1}(\alpha) = 0$ et $\frac{1}{n!} \alpha^n = P_n(\alpha) - P_{n-1}(\alpha) = 0$, donc $\alpha = 0$ et $P_n(\alpha) = 1 \neq 0$, ce qui est absurde. Les éventuelles racines de P_n dans $\mathbb{K}$ sont donc simples.

Exercice 9.6. Soit $P(X) = \sum_{k=0}^n a_k X^k = \prod_{k=1}^n (X - \alpha_k) \in \mathbb{K}[X]$ un polynôme unitaire de degré $n \geq 2$ (donc $a_n = 1$) scindé sur $\mathbb{K}$. On lui associe les fonctions $(S_{n,k})_{1 \leq k \leq n}$ définies par :

$$\forall \alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{K}^n, S_{n,k}(\alpha) = \sum_{i=1}^n \alpha_i^k \quad (1 \leq k \leq n)$$

On note $\sigma_{n,k}$ pour $\sigma_{n,k}(\alpha)$ et S_k pour $S_{n,k}(\alpha)$. Pour tout entier i compris entre 1 et n on désigne par P_i le polynôme de degré $n-1$ défini par $P(X) = (X - \alpha_i) P_i(X)$.

1. Montrer que pour tout entier i compris entre 1 et n et tout entier k compris entre 1 et $n-1$, le coefficient de X^{n-1-k} dans P_i est donné par

$$a_{i,n-1-k} = \alpha_i^k + \sum_{j=1}^k (-1)^j \sigma_j \alpha_i^{k-j}.$$

2. En écrivant le polynôme dérivé de P sous la forme $P' = \sum_{i=1}^n P_i$, montrer

que $(-1)^k (n-k) \sigma_k = \sum_{i=1}^n a_{i,n-1-k}$ pour tout k compris entre 1 et $n-1$.

3. Dédurre de ce qui précède que pour tout entier k compris entre 1 et n on a :

$$S_k - \sigma_1 S_{k-1} + \cdots + (-1)^{k-1} \sigma_{k-1} S_1 + (-1)^k \sigma_k k = 0$$

puis que $S_k + a_{n-1} S_{k-1} + \cdots + a_{n-(k-1)} S_1 + a_{n-k} k = 0$ (formules de Newton).

Solution.

1. On a :

$$\begin{aligned} (X - \alpha_i) P_i(X) &= P(X) = P(X) - P(\alpha_i) = X^n - \alpha_i^n + \sum_{k=1}^{n-1} a_k (X^k - \alpha_i^k) \\ &= X^n - \alpha_i^n + \sum_{k=1}^{n-1} (-1)^{n-k} \sigma_{n-k} (X^k - \alpha_i^k) \end{aligned}$$

donc $P_i(X) = \frac{X^n - \alpha_i^n}{X - \alpha_i} + \sum_{k=1}^{n-1} (-1)^{n-k} \sigma_{n-k} \frac{X^k - \alpha_i^k}{X - \alpha_i}$, chaque polynôme :

$$\frac{X^k - \alpha_i^k}{X - \alpha_i} = X^{k-1} + X^{k-2} \alpha_i + \cdots + X \alpha_i^{k-2} + \alpha_i^{k-1}$$

étant de degré $k-1$. Il en résulte que pour $1 \leq k \leq n-1$, le coefficient de X^{n-1-k} dans P_i est donné par $a_{i,n-1-k} = \alpha_i^k + \sum_{j=1}^k (-1)^j \sigma_j \alpha_i^{k-j}$.

2. On a

$$\begin{aligned} P'(X) &= nX^{n-1} + \sum_{k=1}^{n-1} (-1)^{n-k} k \sigma_{n-k} X^{k-1} \\ &= nX^{n-1} + \sum_{k=0}^{n-2} (-1)^{n-k-1} (k+1) \sigma_{n-k-1} X^k = \sum_{i=1}^n P_i \end{aligned}$$

et en identifiant les coefficients de X^{n-1-k} , pour tout entier k compris entre 1 et $n-1$, on déduit que $(-1)^k (n-k) \sigma_k = \sum_{i=1}^n a_{i,n-1-k}$.

3. L'identité précédente peut s'écrire, pour tout entier k compris entre 1 et $n-1$:

$$(-1)^k (n-k) \sigma_k = S_k + \sum_{j=1}^{k-1} (-1)^j \sigma_j S_{k-j} + (-1)^k n \sigma_k$$

ou encore $S_k - \sigma_1 S_{k-1} + \cdots + (-1)^{k-1} \sigma_{k-1} S_1 + (-1)^k \sigma_k k = 0$. En utilisant les égalités $\sigma_k = (-1)^k a_{n-k}$, on déduit que, pour tout entier k compris entre 1 et $n-1$, on a $S_k + a_{n-1} S_{k-1} + \cdots + a_{n-(k-1)} S_1 + a_{n-k} k = 0$. Pour $k = n$ l'identité s'obtient en écrivant que pour tout entier i compris entre 1 et n on a $0 = P(\alpha_i) = \alpha_i^n + a_{n-1} \alpha_i^{n-1} + \cdots + a_1 \alpha_i + a_0$ et en faisant la somme de toutes ces égalités.

Exercice 9.7. Soient $m \geq n$ deux entiers naturels non nuls et q, r le quotient et le reste dans la division euclidienne dans $\mathbb{Z}$ de m par n .

1. Déterminer le quotient et le reste quand on effectue la division euclidienne de $X^m - 1$ par $X^n - 1$ dans $\mathbb{K}[X]$.
2. Donner une condition nécessaire et suffisante pour que $X^n - 1$ divise $X^m - 1$ dans $\mathbb{K}[X]$.
3. Soit $a \geq 2$ un entier. Montrer que $a^n - 1$ divise $a^m - 1$ dans $\mathbb{Z}$ si, et seulement si, n divise m .
4. Montrer que $(X^n - 1) \wedge (X^m - 1) = X^{n \wedge m} - 1$ dans $\mathbb{K}[X]$.

Solution.

1. En utilisant la division euclidienne $m = qn + r$, dans $\mathbb{Z}$, on a :

$$X^m - 1 = X^{qn} X^r - 1 = (X^{qn} - 1) X^r + X^r - 1$$

Comme $m \geq n \geq 1$, on a $q \geq 1$ et :

$$X^{qn} - 1 = (X^n)^q - 1 = (X^n - 1) \left(1 + X^n + \cdots + X^{(q-1)n} \right)$$

ce qui donne $X^m - 1 = Q(X)(X^n - 1) + R(X)$ avec :

$$Q(X) = X^r \sum_{k=0}^{q-1} X^{kn} \text{ et } R(X) = X^r - 1$$

le polynôme R étant de degré $r < n = \deg(X^n - 1)$, donc R est le reste et Q le quotient dans la division euclidienne de $X^m - 1$ par $X^n - 1$.

2. On en déduit que $X^n - 1$ divise $X^m - 1$ si, et seulement si, $r = 0$, ce qui revient à dire que n divise m .
3. On a $a^n - 1 \geq 1$ puisque $a \geq 2$ et $n \geq 1$. En utilisant les notations de la question précédente, on a $a^m - 1 = Q(a)(a^n - 1) + R(a)$ avec $0 \leq R(a) = a^r - 1 < a^n - 1$ puisque $r < n$. Donc $R(a)$ est le reste dans la division euclidienne de $a^m - 1$ par $a^n - 1$ dans $\mathbb{Z}$ et $a^n - 1$ divise $a^m - 1$ dans $\mathbb{Z}$ si, et seulement si, $R(a) = a^r - 1 = 0$, ce qui équivaut à $r = 0$ (puisque $a \geq 2$) et revient à dire que n divise m .
4. Pour $n = 1$ et tout $m \geq 1$, on a $n \wedge m = 1$ et $(X^m - 1) \wedge (X - 1) = X - 1$ puisque $X - 1$ divise $X^m - 1$. Supposons le résultat acquis pour les entiers compris entre 1 et $n-1$ avec $n \geq 2$. On a alors dans le cas où $1 \leq r \leq n-1$, pour tout $m \geq n$:

$$(X^m - 1) \wedge (X^n - 1) = (X^n - 1) \wedge (X^r - 1) = X^{n \wedge r} - 1 = X^{m \wedge n} - 1$$

Et dans le cas où $r = 0$, n divise m , $X^n - 1$ divise $X^m - 1$ et :

$$(X^m - 1) \wedge (X^n - 1) = X^n - 1 = X^{m \wedge n} - 1$$

Exercice 9.8. Soient $n \in \mathbb{N}^*$ et $P(X) = \sum_{k=0}^m a_k X^k \in \mathbb{K}[X]$ un polynôme de degré $m \geq 1$. Montrer que le reste dans la division euclidienne de P par $X^n - 1$ est $R(X) = \sum_{k=0}^m a_k X^{r_k}$, où r_k est le reste dans la division euclidienne de k par n .

Solution. Pour tout entier k compris entre 0 et m , on a $k = q_k n + r_k$ avec $0 \leq r_k \leq n - 1$ (en particulier $r_0 = 0$). Le polynôme $R(X) = \sum_{k=0}^m a_k X^{r_k}$ est donc de degré au plus égal à $n - 1$ et il s'agit alors de vérifier que le polynôme $P - R$ est divisible par $X^n - 1$. On a $P(X) - R(X) = \sum_{k=1}^m a_k (X^k - X^{r_k})$ avec, pour k compris entre 1 et m , $X^k - X^{r_k} = X^{q_k n + r_k} - X^{r_k} = X^{r_k} ((X^n)^{q_k} - 1)$, le polynôme $(X^n)^{q_k} - 1$ étant divisible par $X^n - 1$. En effet, pour $q_k = 0$, il est nul et pour $q_k \geq 1$, il s'écrit $(X^n)^{q_k} - 1 = (X^n - 1)(1 + X^n + \dots + X^{(q_k-1)n})$. Pour $P(X) = X^m - 1$, on a $r_k = 0$ pour $k = 0, 1, \dots, m-1$ et $r_m = r$ est le reste dans la division euclidienne de m par n . On retrouve donc l'un des résultats de l'exercice précédent.

Exercice 9.9. Soit $P = \lambda \prod_{k=1}^r (X - \lambda_k)^{m_k}$ un polynôme scindé dans $\mathbb{K}[X]$ de degré $n \geq 1$, les λ_k étant deux à deux distincts. Calculer $P \wedge P'$.

Solution. Si $P = \lambda \prod_{k=1}^r (X - \lambda_k)^{m_k}$, les racines λ_k étant deux à deux distinctes, on a alors $\frac{P'}{P} = \sum_{k=1}^r \frac{m_k}{X - \lambda_k}$ (dérivée logarithmique), soit :

$$P'(X) = \sum_{k=1}^r m_k \frac{P(X)}{X - \lambda_k} = \lambda \prod_{k=1}^r (X - \lambda_k)^{m_k-1} \sum_{k=1}^r m_k \prod_{\substack{j=1 \\ j \neq k}}^r (X - \lambda_j) = Q(X) \Delta(X)$$

en notant $\Delta(X) = \prod_{k=1}^r (X - \lambda_k)^{m_k-1}$ et $Q(X) = \lambda \sum_{k=1}^r m_k \prod_{\substack{j=1 \\ j \neq k}}^r (X - \lambda_j)$. Comme

$Q(\lambda_k) = m_k \prod_{\substack{j=1 \\ j \neq k}}^r (\lambda_k - \lambda_j) \neq 0$, le polynôme Q est premier avec P et :

$$P \wedge P' = P \wedge Q\Delta = P \wedge \Delta = \Delta = \prod_{k=1}^r (X - \lambda_k)^{m_k-1}$$

En particulier $\deg(P \wedge P') = \sum_{k=1}^r (m_k - 1) = n - r$. Ce dernier résultat est faux

pour P non scindé. Par exemple le polynôme $P(X) = (X - 1)^2 (X + 1) (X^2 + 1)$ dans $\mathbb{R}[X]$ est de degré $n = 5$ avec $r = 2$ racines distinctes et $P \wedge P' = X - 1$ (on peut calculer ce pgcd dans $\mathbb{C}[X]$ avec le résultat précédent) est de degré $1 \neq n - r = 3$.

Exercice 9.10. Déterminer tous les polynômes non constants P dans $\mathbb{C}[X]$ tels que P' divise P .

Solution. Soit $P = \alpha \prod_{k=1}^r (X - \lambda_k)^{m_k}$ un polynôme non constant dans $\mathbb{C}[X]$. Si P' divise P , on a alors $P \wedge P' = P'$ et $\deg(P') = n - 1 = \deg(P \wedge P') = n - r$, donc $r = 1$ et $P = \alpha (X - \lambda_1)^{m_1}$. La réciproque est évidente.

Exercice 9.11. Soit $P \in \mathbb{K}[X]$ un polynôme de degré $n \geq 1$ et scindé sur $\mathbb{K}$. Montrer que toutes ses racines sont simples si, et seulement si, $P \wedge P' = 1$.

Solution. Si $P \wedge P' = 1$, le théorème de Bézout nous dit qu'il existe deux polynômes U, V tels que $UP + VP' = 1$ et P, P' ne peuvent avoir de racines en commun. Les racines de P sont donc toutes simples. Pour la réciproque, on peut utiliser l'exercice précédent ou procéder comme suit. Soit $P = \lambda \prod_{k=1}^r (X - \lambda_k)$, les racines λ_k étant deux à deux distinctes. Si $\Delta = P \wedge P' \neq 1$, comme Δ divise P , il a aussi des racines simples dans $\{\lambda_1, \dots, \lambda_r\}$ et si λ_j est l'une de ces racines, elle est aussi racine de P' , ce qui veut dire que c'est une racine double de P et contredit l'hypothèse de départ.

Chapitre 10

Réduction des endomorphismes

$\mathbb{K}$ désigne le corps de réels ou des complexes, E est un $\mathbb{K}$ -espace vectoriel de dimension finie $n \geq 1$ et on se donne $u \in \mathcal{L}(E)$.

10.1 Polynômes d'endomorphismes, polynômes annulateurs, polynôme minimal

On note $u^0 = Id$ et on définit les puissances successives de u par la relation de récurrence $u^{k+1} = u \circ u^k$ pour tout $k \in \mathbb{N}$, ce qui nous permet de définir, pour tout polynôme $P(X) = \sum_{k=0}^p a_k X^k \in \mathbb{K}[X]$, l'endomorphisme $P(u)$, par

$P(u) = \sum_{k=0}^p a_k u^k$. On note naturellement $\mathbb{K}[u] = \{P(u) \mid P \in \mathbb{K}[X]\}$ et il est

facile de vérifier que cet ensemble est un sous-espace vectoriel et un sous-anneau commutatif de $\mathcal{L}(E)$. Précisément, $\mathbb{K}[u]$ est une sous-algèbre de $\mathcal{L}(E)$ et pour ce qui est de la commutativité, on a pour tous P, Q dans $\mathbb{K}[X]$:

$$(PQ)(u) = P(u) \circ Q(u) = Q(u) \circ P(u) = (QP)(u)$$

On définit de manière analogue la sous algèbre $\mathbb{K}[A]$ de $\mathcal{M}_n(\mathbb{K})$ engendrée par une matrice $A \in \mathcal{M}_n(\mathbb{K})$.

On vérifie facilement que si $A \in \mathcal{M}_n(\mathbb{K})$ est la matrice de $u \in \mathcal{L}(E)$ dans une base $\mathcal{B}$, alors $P(A)$ est la matrice de $P(u)$ dans $\mathcal{B}$.

L'espace vectoriel $\mathcal{L}(E)$ étant de dimension n^2 (il est isomorphe à $\mathcal{M}_n(\mathbb{K})$ par le choix d'une base), la famille $\{u^k \mid 0 \leq k \leq n^2\}$ est liée, ce qui équivaut à dire qu'il existe un polynôme $P \in \mathbb{K}[X] \setminus \{0\}$ tel que $P(u) = 0$. Il en résulte que l'ensemble $I_u = \{P \in \mathbb{K}[X] \mid P(u) = 0\}$ n'est pas réduit au polynôme nul. Cet ensemble est le noyau du morphisme d'algèbres $\varphi_u : P \mapsto P(u)$, c'est donc un idéal de l'anneau $\mathbb{K}[X]$ (exemple 9.1) et on peut donner la définition suivante justifiée par le théorème 9.7.

Définition 10.1. On appelle idéal annulateur de u l'idéal I_u et polynôme minimal de u le générateur unitaire de cet idéal. On note π_u ce polynôme.

On a donc $I_u = \{P \in \mathbb{K}[X] \mid P(u) = 0\} = \mathbb{K}[X]\pi_u$ et π_u est le polynôme unitaire de plus petit degré annihilant u . Ce polynôme π_u est non constant.

On définit de manière analogue l'idéal annulateur et le polynôme minimal d'une matrice $A \in \mathcal{M}_n(\mathbb{K})$.

On vérifie facilement que si u a pour matrice A dans une base $\mathcal{B}$ de E , alors A et u ont même idéal annulateur et même polynôme minimal (pour $P \in \mathbb{K}[X]$, la matrice de $P(u)$ dans $\mathcal{B}$ est $P(A)$ et $P(u) = 0$ équivaut à $P(A) = 0$, donc $I_u = I_A$ et $\pi_u = \pi_A$). On vérifie aussi que deux matrices semblables ont même idéal annulateur et même polynôme minimal (si $B = Q^{-1}AQ$ et $P \in \mathbb{K}[X]$, on a alors $P(B) = Q^{-1}P(A)Q$ et $P(A) = 0$ équivaut à $P(B) = 0$, donc $I_A = I_B$ et $\pi_A = \pi_B$).

On a $u = 0$ si, et seulement si, $\pi_u(X) = X$. De manière plus générale, les homothéties sont les seuls endomorphismes ayant un polynôme minimal de degré égal à 1, c'est-à-dire que $u = \lambda Id$ si, et seulement si, $\pi_u(X) = X - \lambda$.

Exemples 10.1

1. Un endomorphisme $u \in \mathcal{L}(E)$ est dit nilpotent s'il existe un entier $q \geq 1$ tel que $u^{q-1} \neq 0$ et $u^q = 0$ (q est l'indice de nilpotence de u), ce qui équivaut à dire que $\pi_u(X) = X^q$.
2. Un projecteur est un endomorphisme u de E tel que $u \circ u = u$. Un tel projecteur étant annulé par $X^2 - X$, son polynôme minimal est $\pi_u(X) = X$ si $u = 0$, $\pi_u(X) = X - 1$ si $u = Id$, ou $\pi_u(X) = X^2 - X$ dans les autres cas.

10.2 Le théorème de décomposition des noyaux

Lemme 10.1 Soient p un entier supérieur ou égal à 2, $P_1, \dots, P_p$ des polynômes non nuls dans $\mathbb{K}[X]$ et $Q_1, \dots, Q_p$ les polynômes définis par $Q_k = \prod_{\substack{j=1 \\ j \neq k}}^p P_j$ pour tout

k compris entre 1 et p . Si les polynômes P_k sont deux à deux premiers entre eux dans $\mathbb{K}[X]$, alors les polynômes Q_k sont premiers entre eux dans leur ensemble et pour tout k compris entre 1 et p , les polynômes P_k et Q_k sont premiers entre eux.

Preuve. Si $\Delta = \text{pgcd}(Q_1, \dots, Q_p)$ est non constant, il admet alors un diviseur irréductible R qui divise $Q_1 = \prod_{j=2}^p P_j$, donc l'un des P_j avec $j \neq 1$ et comme R divise aussi Q_j , il divise l'un des P_k avec $k \neq j$, ce qui contredit le fait que P_j et P_k sont premiers entre eux. Si $\Delta_k = \text{pgcd}(P_k, Q_k)$ est non constant, il admet un diviseur irréductible R_k qui divise P_k et le produit Q_k , il divise donc l'un des P_j avec $j \neq k$, ce qui contredit P_j et P_k premiers entre eux pour $k \neq j$. $\square$

Théorème 10.1.

Soient p un entier supérieur ou égal à 2, $P_1, \dots, P_p$ des polynômes non nuls dans $\mathbb{K}[X]$ deux à deux premiers entre eux et $P = \prod_{k=1}^p P_k$. On a

$\ker(P(u)) = \bigoplus_{k=1}^p \ker(P_k(u))$ et pour tout k compris entre 1 et p , le projecteur $\pi_k : \ker(P(u)) \rightarrow \ker(P_k(u))$ est un élément de $\mathbb{K}[u]$.

Preuve. En utilisant les notations du lemme précédent, le théorème de Bézout nous dit qu'il existe des polynômes $R_1, \dots, R_p$ tels que $\sum_{k=1}^p R_k Q_k = 1$ et dans

$\mathcal{L}(E)$, on a $Id = \sum_{k=1}^p R_k(u) \circ Q_k(u)$, soit $x = \sum_{k=1}^p R_k(u) \circ Q_k(u)(x)$ pour tout x dans E . Pour tout $x \in \ker(P(u))$ et tout k compris entre 1 et p , on a par commutativité de $\mathbb{K}[u]$, $P_k(u)(R_k(u) \circ Q_k(u)(x)) = R_k(u) \circ P(u)(x) = 0$, soit $x_k = R_k(u) \circ Q_k(u)(x) \in \ker(P_k(u))$. On a donc $\ker(P(u)) \subset \sum_{k=1}^p \ker(P_k(u))$ et comme $\ker(P_k(u)) \subset \ker(P(u))$ pour tout k compris entre 1 et p , on a l'égalité $\ker(P(u)) = \sum_{k=1}^p \ker(P_k(u))$.

Si $(x_1, \dots, x_p)$ dans $\prod_{k=1}^p \ker(P_k(u))$ est tel que $\sum_{j=1}^p x_j = 0$, pour tout k compris entre 1 et p , on a $0 = Q_k(u) \left(\sum_{j=1}^p x_j \right) = Q_k(u)(x_k)$ et $P_k(u)(x_k) = 0$. Sachant que P_k et Q_k sont premiers entre eux, il existe deux polynômes A et B tels que $AP_k + BQ_k = 1$ et on a $A(u) \circ P_k(u) + B(u) \circ Q_k(u) = Id$ dans $\mathcal{L}(E)$, ce qui donne $x_k = (A(u) \circ P_k(u))(x_k) + (B(u) \circ Q_k(u))(x_k) = 0$. On a donc au final, $\ker(P(u)) = \bigoplus_{k=1}^p \ker(P_k(u))$ et les projecteurs de $\ker(P(u))$ sur $\ker(P_k(u))$ sont les $\pi_k = R_k(u) \circ Q_k(u) \in \mathbb{K}[u]$. $\square$

En notant $\pi_u = \prod_{k=1}^p P_k^{\beta_k}$ la décomposition du polynôme minimal π_u en facteurs irréductibles dans $\mathbb{K}[X]$, on déduit du théorème de décomposition des noyaux que $E = \bigoplus_{k=1}^p \ker(P_k^{\beta_k}(u))$, chaque sous-espace vectoriel $E_k = \ker(P_k^{\beta_k}(u))$ étant stable par u puisque u commute à $P_k(u)$ et non réduit à $\{0\}$. En effet dire que $\ker(P_k^{\beta_k}(u))$ est réduit à $\{0\}$ revient à dire que $P_k^{\beta_k}(u)$ est inversible et de l'égalité $0 = \pi_u(u) = P_k^{\beta_k}(u) \circ Q(u)$, on déduit que $Q(u) = 0$ avec $0 \leq \deg(Q) < \deg(\pi_u)$, ce qui n'est pas possible.

10.3 Valeurs et vecteurs propres

u est un endomorphisme de E .

Définition 10.2. On dit que $\lambda \in \mathbb{K}$ est valeur propre de l'endomorphisme u si on a, $\ker(u - \lambda Id) \neq \{0\}$.

Exemple 10.1 L'unique valeur propre d'une homothétie λId est λ .

Dire que $\lambda \in \mathbb{K}$ est valeur propre de u équivaut à dire qu'il existe $x \in E \setminus \{0\}$ tel que $u(x) = \lambda x$, ce qui revient aussi à dire que $u - \lambda Id$ est non inversible puisque E est de dimension finie. On dit alors que x est un vecteur propre de u associé à la valeur propre λ et que le sous-espace vectoriel $E_\lambda = \ker(u - \lambda Id)$ de E est le sous-espace propre associé à λ .

L'ensemble des valeurs propres de u est appelé le spectre de u et noté $\text{Sp}(u)$.

Un scalaire $\lambda \in \mathbb{K}$ est valeur propre de $u \in \mathcal{L}(E)$ si, et seulement si, $\lambda Id - u$ est non inversible, ce qui équivaut à dire que le polynôme caractéristique de u défini par $P_u(X) = \det(XId - u)$ est nul pour $X = \lambda$. On a donc :

$$\text{Sp}(u) = P_u^{-1}\{0\} = \{\lambda \in \mathbb{K} \mid \det(\lambda Id - u) = 0\}$$

et c'est une partie finie de $\mathbb{K}$ ayant au plus n éléments. Pour $\mathbb{K} = \mathbb{R}$, ce spectre peut être vide.

Théorème 10.2.

Les valeurs propres de u sont les racines de son polynôme minimal.

Preuve. Si $\lambda \in \mathbb{K}$ est une valeur propre de u et $x \in E \setminus \{0\}$ un vecteur propre associé, de l'égalité $0 = \pi_u(u)(x) = \pi_u(\lambda)x$, on déduit que $\pi_u(\lambda) = 0$, c'est-à-dire que λ est racine de π_u . Réciproquement si $\lambda \in \mathbb{K}$ est racine de π_u , on a alors $\pi_u(X) = (X - \lambda)Q(X)$ et avec $\pi_u(u) = (u - \lambda Id) \circ Q(u) = 0$ et le caractère minimal de π_u on déduit que $u - \lambda Id$ est non inversible ce qui équivaut à dire que λ est une valeur propre de u . $\square$

Exemple 10.2 Si u est nilpotent d'ordre $q \geq 1$ (i. e. $u^{q-1} \neq 0$ et $u^q = 0$), on a alors $\pi_u(X) = X^q$ et 0 est l'unique valeur propre de u .

En identifiant une matrice $A \in \mathcal{M}_n(\mathbb{K})$ à l'endomorphisme de $\mathbb{K}^n$ qu'elle définit dans la base canonique de $\mathbb{K}^n$, on peut donner la définition suivante.

Définition 10.3. On dit que λ dans $\mathbb{K}$ est valeur propre de $A \in \mathcal{M}_n(\mathbb{K})$ s'il existe un vecteur non nul x dans $\mathbb{K}^n$ tel que $Ax = \lambda x$. On dit alors que x est un vecteur propre de A associé à la valeur propre λ et que le sous-espace vectoriel de $\mathbb{K}^n$, $E_\lambda = \{x \in \mathbb{K}^n \mid Ax = \lambda x\}$ (que l'on peut noter $\ker(A - \lambda I_n)$) est le sous-espace propre associé à λ .

L'ensemble $\text{Sp}(A)$ des valeurs propres de A est appelé le spectre de A et le polynôme caractéristique de $A \in \mathcal{M}_n(\mathbb{K})$ est défini par $P_A(X) = \det(XId - A)$.

Si u a pour matrice A dans une base de E , alors A et u ont même polynôme caractéristique et mêmes valeurs propres puisque $\det(v) = \det(M)$ pour tout $u \in \mathcal{L}(E)$ de matrice $M \in \mathcal{M}_n(\mathbb{K})$ dans une base de E . Cela se traduit aussi en disant que deux matrices semblables ont même polynôme caractéristique.

Exemples 10.2

1. Le polynôme caractéristique d'une matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{K})$ est :

$$P_A(X) = \begin{vmatrix} X - a & -b \\ -c & X - d \end{vmatrix} = X^2 - \text{Tr}(A)X + \det(A)$$

2. Le polynôme caractéristique de $A = \begin{pmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{pmatrix} \in \mathcal{M}_3(\mathbb{K})$ est :

$$P_A(X) = \begin{vmatrix} X - a_1 & -a_2 & -a_3 \\ -a_4 & X - a_5 & -a_6 \\ -a_7 & -a_8 & X - a_9 \end{vmatrix} = X^3 - aX^2 + bX - c$$

où $a = \text{Tr}(A)$, $c = \det(A)$ et $b = \begin{vmatrix} a_5 & a_6 \\ a_8 & a_9 \end{vmatrix} + \begin{vmatrix} a_1 & a_3 \\ a_7 & a_9 \end{vmatrix} + \begin{vmatrix} a_1 & a_2 \\ a_4 & a_5 \end{vmatrix}$ (ce sont tous les déterminants extraits d'ordre 2 qui contiennent deux termes diagonaux de A).

3. Si $A \in \mathcal{M}_n(\mathbb{K})$ est une matrice triangulaire de termes diagonaux $\lambda_1, \dots, \lambda_n$, son polynôme caractéristique est alors $P_A(X) = \prod_{k=1}^n (X - \lambda_k)$ et les λ_k sont les valeurs propres de A .

Pour $A \in \mathcal{M}_n(\mathbb{K})$, on a $P_A(X) = X^n - \alpha_{n-1}X^{n-1} + \alpha_{n-2}X^{n-2} + \dots + (-1)^n \alpha_0$ avec $\alpha_{n-1} = \text{Tr}(A)$, $\alpha_0 = \det(A)$ et $\alpha_{n-2} = \sum_{1 \leq i < j \leq n} \det(A_{i,j})$, où les $\det(A_{i,j}) =$

$\begin{vmatrix} a_{ii} & a_{ij} \\ a_{ji} & a_{jj} \end{vmatrix}$, pour $1 \leq i < j \leq n$, sont tous les déterminants extraits d'ordre 2 qui contiennent deux termes diagonaux de A . Avec ce résultat, on retrouve le fait que deux matrices semblables ont même polynôme caractéristique. On en déduit également que deux matrices semblables ont même trace.

Théorème 10.3.

Soit $P \in \mathbb{K}[X]$. Pour toute valeur propre $\lambda \in \text{Sp}(u)$, $P(\lambda)$ est valeur propre de $P(u)$. Pour $\mathbb{K} = \mathbb{C}$, on a $\text{Sp}(P(u)) = \{P(\lambda) \mid \lambda \in \text{Sp}(u)\}$.

Preuve. Si $\lambda \in \mathbb{K}$ est une valeur propre de u et $x \in E \setminus \{0\}$ un vecteur propre associé, on a alors pour tout polynôme $P \in \mathbb{K}[X]$, $P(u)x = P(\lambda)x$. En effet, de $u(x) = \lambda x$, on déduit par récurrence sur $k \geq 0$ que $u^k(x) = \lambda^k x$ pour tout $k \in \mathbb{N}$.

et par linéarité, il en résulte que $P(u)x = P(\lambda)x$ pour tout $P \in \mathbb{K}[X]$. Ce qui signifie que x est un vecteur propre de $P(u)$ associé à la valeur propre $P(\lambda)$. Si $P(X) = a_0$ est un polynôme constant, $P(u) = a_0 Id$ a alors a_0 pour unique valeur, l'espace propre associé étant E .

On suppose que $\mathbb{K} = \mathbb{C}$ et que P est non constant ($\text{Sp}(u)$ et $\text{Sp}(P(u))$ sont donc non vides). On sait déjà que $\{P(\lambda) \mid \lambda \in \text{Sp}(u)\} \subset \text{Sp}(P(u))$. Si $\mu \in \mathbb{K}$ est une valeur propre de $P(u)$, en notant $Q(X) = P(X) - \mu$, l'endomorphisme

$Q(u)$ est non injectif et en écrivant que $Q(X) = \alpha \prod_{i=1}^p (X - \lambda_i)^{m_i}$, on déduit qu'il

existe un indice i tel que l'endomorphisme $u - \lambda_i Id$ soit non injectif c'est-à-dire que λ_i est une valeur propre de u et avec $Q(\lambda_i) = 0$ on déduit que $\mu = P(\lambda_i)$. En définitive on a $\text{Sp}(P(u)) = \{P(\lambda) \mid \lambda \in \text{Sp}(u)\}$. $\square$

Pour $\mathbb{K} = \mathbb{R}$, l'inclusion $\{P(\lambda) \mid \lambda \in \text{Sp}(u)\} \subset \text{Sp}(P(u))$ peut être stricte. Par exemple pour $P(X) = X^2$ et -1 valeur propre de $-I_2 = A^2$, où $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ est la matrice de rotation d'angle $\frac{\pi}{2}$, on a $\text{Sp}(u) = \emptyset$, donc l'inclusion est stricte.

Lemme 10.2 *Si F est un sous-espace vectoriel de E stable par u (i. e. $u(F) \subset F$), le polynôme caractéristique de la restriction de u à F divise alors celui de u .*

Preuve. Soit $\mathcal{B}_1$ une base de F complétée en une base $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2$ de E . Dans cette base la matrice de u est $A = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix}$ où A_1 est la matrice dans la base $\mathcal{B}_1$ de la restriction de u à F (F est stable par u) et le polynôme caractéristique de u s'écrit $P_u(X) = \det(A_1 - XI_{n_1}) \det(A_3 - XI_{n_3})$. On en déduit alors que P_u est un multiple du polynôme caractéristique de la restriction de u à F . $\square$

Théorème 10.4.

Si $\lambda \in \mathbb{K}$ est une valeur propre de u de multiplicité α en tant que racine du polynôme caractéristique de u , on a alors $1 \leq \dim(\ker(u - \lambda Id)) \leq \alpha$.

Preuve. Le sous-espace propre $E_\lambda = \ker(u - \lambda Id)$ n'étant pas réduit au vecteur nul, sa dimension est supérieure ou égale à 1. Ce sous-espace vectoriel étant stable par u , le polynôme caractéristique P_λ de la restriction de u à E_λ divise le polynôme caractéristique P_u de u . En remarquant que $P_\lambda(X) = (\lambda - X)^\delta$ où δ est la dimension de E_λ , on déduit que $P_u(X) = (\lambda - X)^\delta Q(X)$ et la racine λ de P_u étant de multiplicité α , on a nécessairement $\delta \leq \alpha$. $\square$

Exemples 10.3

1. Si $R_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ est une matrice de rotation dans $\mathcal{M}_2(\mathbb{R})$, son polynôme caractéristique est :

$$P_\theta(X) = X^2 - 2\cos(\theta)X + 1 = (X - \cos(\theta))^2 + \sin^2(\theta)$$

et ce polynôme n'a pas de racine réelle pour $\theta \in \mathbb{R} \setminus \pi\mathbb{Z}$. On a donc $\text{Sp}(R_\theta) = \emptyset$ dans ce cas ($R_\theta = \pm I_2$ si $\theta \in \pi\mathbb{Z}$).

2. Si $S_\theta = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$ est une matrice de réflexion dans $\mathcal{M}_2(\mathbb{R})$, son polynôme caractéristique est $P_\theta(X) = X^2 - 1$ et $\text{Sp}(S_\theta) = \{-1, 1\}$.

Théorème 10.5.

Si $\text{Sp}(u) = \{\lambda_1, \dots, \lambda_p\}$, alors les sous-espaces propres $E_k = \ker(u - \lambda_k \text{Id})$, pour k compris entre 1 et p , sont en somme directe.

Preuve. Pour $p = 1$, c'est clair. Supposons le résultat acquis pour $p - 1 \geq 1$. Si

$$\sum_{k=1}^p x_k = 0 \text{ avec } x_k \in E_k, \text{ en appliquant } u, \text{ on a } \sum_{k=1}^p u(x_k) = \sum_{k=1}^p \lambda_k x_k = 0 \text{ et :}$$

$$\sum_{k=1}^p \lambda_k x_k - \lambda_p \sum_{k=1}^p x_k = \sum_{k=1}^{p-1} (\lambda_k - \lambda_p) x_k = 0$$

avec $\lambda_k - \lambda_p \neq 0$ pour $1 \leq k \leq p - 1$, ce qui nous donne $x_k = 0$ pour $1 \leq k \leq p - 1$, puis $x_p = 0$. $\square$

10.4 Le théorème de Cayley-Hamilton

Pour tout vecteur non nul x dans E , on note $I_{u,x} = \{P \in \mathbb{K}[X] \mid P(u)(x) = 0\}$. On vérifie facilement que $I_{u,x}$ est un idéal de $\mathbb{K}[X]$. Comme $\pi_u \in I_{u,x}$, cet idéal n'est pas réduit à $\{0\}$. On désigne alors par $\pi_{u,x}$ le générateur unitaire de $I_{u,x}$ et par $p_{u,x} \geq 1$ le degré de $\pi_{u,x}$.

Définition 10.4. Avec les notations qui précèdent, le sous-espace vectoriel $E_{u,x} = \text{Vect}\{u^k(x) \mid k \in \mathbb{N}\}$ est le sous-espace cyclique engendré par x .

Lemme 10.3 $E_{u,x}$ est un sous-espace vectoriel de dimension $p_{u,x}$ de E , stable par u . Le polynôme minimal et le polynôme caractéristique de la restriction de u à $E_{u,x}$ sont respectivement $\pi_{u,x}$ et $(-1)^{p_{u,x}} \pi_{u,x}$.

Preuve. Pour tout $k \in \mathbb{N}$, on a $u(u^k(x)) = u^{k+1}(x) \in E_{u,x}$, donc le sous-espace vectoriel $E_{u,x}$ est stable par u . Comme $\pi_{u,x}$ est de degré minimum dans $I_{u,x} \setminus \{0\}$, le système $\mathcal{B}_{u,x} = \{u^k(x) \mid 0 \leq k \leq p_{u,x} - 1\}$ est libre. En notant

$$\pi_{u,x}(X) = X^{p_{u,x}} - \sum_{k=0}^{p_{u,x}-1} a_k X^k, \text{ de } \pi_{u,x}(u)(x) = 0, \text{ on déduit que } u^{p_{u,x}}(x)$$

est dans $\text{Vect}(\mathcal{B}_{u,x})$ et par récurrence sur $k \geq 0$, on vérifie que $u^{p_{u,x}+k}(x) \in \text{Vect}(\mathcal{B}_{u,x})$ pour tout entier naturel k , ce qui signifie que $\mathcal{B}_{u,x}$ est un système générateur et donc une base de $E_{u,x}$. On a donc $\dim(E_{u,x}) = p_{u,x} \leq n$. Avec

$$\pi_{u,x}(u)(x) = u^{p_{u,x}}(x) - \sum_{k=0}^{p_{u,x}-1} a_k u^k(x) = 0 \text{ et la commutativité de } \mathbb{K}[u], \text{ on dé-$$

duit que $\pi_{u,x}(u)(u^k(x)) = u^k(\pi_{u,x}(u)(x)) = 0$ pour tout entier k , ce qui signifie

que $\pi_{u,x}(u|_{E_{u,x}}) = 0$. Si $Q \in \mathbb{K}_{p_{u,x}-1}[X] \setminus \{0\}$ annule $v_{u,x} = u|_{E_{u,x}}$, on a alors $Q(u)(x) = 0$ et le système $\mathcal{B}_x$ est lié, ce qui contredit la définition de $p_{u,x}$. Donc $\pi_{u,x}$ est le polynôme minimal de $v_{u,x}$. Enfin, en écrivant que la matrice de $v_{u,x}$

dans la base $\mathcal{B}_x$ est $A_{u,x} = \begin{pmatrix} 0 & \cdots & 0 & a_0 \\ 1 & \ddots & \vdots & a_1 \\ \vdots & \ddots & 0 & \vdots \\ 0 & \cdots & 1 & a_{p_x-1} \end{pmatrix}$, on déduit que $(-1)^{p_{u,x}} \pi_{u,x}$

est le polynôme caractéristique de $v_{u,x}$. En effet en notant $P_{A_{u,x}} = P_{(a_0, \dots, a_{p_x-1})}$ ce polynôme caractéristique et en le développant par rapport à la première ligne, on a :

$$P_{(a_0, \dots, a_{p_x-1})}(X) = -X P_{(a_1, \dots, a_{p_x-1})}(X) + (-1)^{p_x+1} a_0$$

et par récurrence, on a $P_{A_x}(X) = (-1)^{p_x} \left(X^{p_x} - \sum_{k=0}^{p_x-1} a_k X^k \right)$. □

Théorème 10.6. Cayley-Hamilton

Si P_u est le polynôme caractéristique de u , on a alors $P_u(u) = 0$.

Preuve. Pour tout $x \in E$ le sous-espace cyclique $E_{u,x}$ étant stable par u , le polynôme caractéristique $\pi_{u,x}$ de $u|_{E_{u,x}}$ divise celui de u . C'est-à-dire qu'il existe un polynôme Q_x tel que $P_u = Q_x \cdot \pi_{u,x}$ et $P_u(u)(x) = Q_x(u) \circ \pi_{u,x}(u)(x) = 0$. On a donc ainsi montré que pour tout $x \in E$, on a $P_u(u)(x) = 0$ (pour $x = 0$ c'est clair) et donc que $P_u(u) = 0$. □

Corollaire 10.1. *Le polynôme minimal π_u divise le polynôme caractéristique P_u . On a donc $\dim(\mathbb{K}[u]) = \deg(\pi_u) \leq n$.*

Preuve. Le polynôme minimal divisant tout polynôme annulateur de u , on déduit du théorème de Cayley-Hamilton que π_u divise le polynôme P_u . □

Du corollaire précédent, on retrouve le fait que les racines de π_u sont valeurs propres de u .

Pour $\mathbb{K} = \mathbb{C}$, le polynôme caractéristique de u s'écrit $P_u(X) = \prod_{k=1}^p (X - \lambda_k)^{\alpha_k}$ avec $\alpha_k \in \mathbb{N}^*$ et les λ_k deux à deux distincts. Le polynôme minimal π_u étant un diviseur de P_u avec les mêmes racines, il s'écrit $\pi_u(X) = \prod_{k=1}^p (X - \lambda_k)^{\beta_k}$ avec $1 \leq \beta_k \leq \alpha_k$.

On en déduit qu'un endomorphisme $u \in \mathcal{L}(E)$ est nilpotent si, et seulement si, $P_u(X) = X^n$.

Dans le cas où l'endomorphisme u est inversible, le théorème de Cayley-Hamilton nous donne un moyen de calculer l'inverse de u , si on connaît son polynôme caractéristique P_u . En effet l'égalité $P_u(u) = 0$ avec $P_u(X) = \sum_{k=0}^n a_k X^k$ donne :

$$u^{-1} = -\frac{1}{a_0} \sum_{k=1}^n a_k u^{k-1} = -\frac{1}{\det(u)} \sum_{k=1}^n a_k u^{k-1}$$

On constate que u^{-1} est un polynôme en u .

Le théorème de Cayley-Hamilton permet également de calculer u^p pour tout entier p supérieur ou égal à n en fonction de $Id, u, \dots, u^{n-1}$. En effet pour $p = n$, de $P_u(u) = \sum_{k=0}^n a_k u^k = 0$ avec $a_n = 1$, on déduit que $u^n = -\sum_{k=0}^{n-1} a_k u^k$ et pour $p > n$ la division euclidienne de X^p par P_u , $X^p = QP_u + R$ avec $\deg(R) < n$, donne $u^p = R(u)$.

10.5 Diagonalisation

On sait que si $\text{Sp}(u) = \{\lambda_1, \dots, \lambda_p\}$, on a alors $\bigoplus_{k=1}^p \ker(u - \lambda_k Id) \subset E$ (théorème 10.5). On s'intéresse ici au cas d'égalité.

Si $\bigoplus_{k=1}^p \ker(u - \lambda_k Id) = E$, en désignant pour tout k compris entre 1 et p par $\mathcal{B}_k$ une base de $\ker(u - \lambda_k Id)$, la réunion $\bigcup_{k=1}^p \mathcal{B}_k$ est une base de E dans laquelle la matrice de u est diagonale.

Définition 10.5. On dit que u est diagonalisable, s'il existe une base de E dans laquelle la matrice de u est diagonale.

Si u diagonalisable a pour matrice A dans une base $\mathcal{B} = (e_i)_{1 \leq i \leq n}$, il existe alors une base $\mathcal{B}' = (e'_i)_{1 \leq i \leq n}$ dans laquelle la matrice D de u est diagonale. Les matrices A et D sont alors semblables, ce qui signifie qu'il existe une matrice $P \in GL_n(\mathbb{K})$ telle que $D = P^{-1}AP$. Ce qui nous amène à poser la définition suivante.

Définition 10.6. On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est diagonalisable, si elle est semblable à une matrice diagonale.

Une condition suffisante de diagonalisation est donnée par le résultat suivant.

Théorème 10.7.

Si u a n valeurs propres distinctes dans $\mathbb{K}$, il est alors diagonalisable.

Preuve. Si $\text{Sp}(u) = \{\lambda_1, \dots, \lambda_n\}$, alors toutes les valeurs propres de u sont simples et chaque sous-espace propre est de dimension 1 (théorème 10.4). Comme ces sous-espaces propres sont en somme directe, on a $\dim \left(\bigoplus_{k=1}^n \ker(u - \lambda_k Id) \right) = n$, donc $E = \bigoplus_{k=1}^n \ker(u - \lambda_k Id)$ et u est diagonalisable. □

Des conditions nécessaires et suffisantes pour qu'un endomorphisme soit diagonalisable sont données par le résultat suivant.

Théorème 10.8.

Les conditions suivantes sont équivalentes.

1. *l'endomorphisme u est diagonalisable ;*
2. *si $\text{Sp}(u) = \{\lambda_1, \dots, \lambda_p\}$, on a alors $E = \bigoplus_{k=1}^p \ker(u - \lambda_k \text{Id})$;*
3. *si $\text{Sp}(u) = \{\lambda_1, \dots, \lambda_p\}$, on a alors $\sum_{k=1}^p \dim(\ker(u - \lambda_k \text{Id})) = n$;*
4. *le polynôme caractéristique de u est scindé sur $\mathbb{K}$ de racines deux à deux distinctes $\lambda_1, \dots, \lambda_p$ dans $\mathbb{K}$, chaque λ_k ($1 \leq k \leq p$) étant de multiplicité $\alpha_k = \dim(\ker(u - \lambda_k \text{Id}))$;*
5. *il existe un polynôme annulateur de u scindé à racines simples dans $\mathbb{K}$;*
6. *le polynôme minimal π_u est scindé à racines simples dans $\mathbb{K}$.*

Preuve. (i) $\Leftrightarrow$ (ii) Si u est diagonalisable il existe alors une base de E :

$$\mathcal{B} = \{e_{11}, \dots, e_{1,\alpha_1}, \dots, e_{p1}, \dots, e_{p,\alpha_p}\}$$

formée de vecteurs propres avec :

$$u(e_{k,j}) = \lambda_k e_{k,j} \quad (1 \leq k \leq p, 1 \leq j \leq \alpha_k)$$

où les $\lambda_1, \dots, \lambda_p$ sont les valeurs propres distinctes deux à deux de u . On a alors $\text{Vect}\{e_{k1}, \dots, e_{k,\alpha_k}\} = \ker(u - \lambda_k \text{Id})$ pour tout k compris entre 1 et p et :

$$E = \bigoplus_{k=1}^p \text{Vect}\{e_{k1}, \dots, e_{k,\alpha_k}\} = \bigoplus_{k=1}^p \ker(u - \lambda_k \text{Id})$$

La réciproque est évidente.

(ii) $\Leftrightarrow$ (iii) Cette équivalence résulte du fait que les espaces propres sont en somme directe.

$$(iii) \Rightarrow (iv) \text{ Si } \sum_{k=1}^p \dim(\ker(u - \lambda_k \text{Id})) = n, \text{ on a alors } E = \bigoplus_{k=1}^p \ker(u - \lambda_k \text{Id})$$

et le polynôme caractéristique de u s'écrit $\chi_u(X) = \prod_{k=1}^p (X - \lambda_k)^{\alpha_k}$.

(iv) $\Rightarrow$ (v) Comme $\dim(\ker(u - \lambda_k \text{Id}))$ est égal à la multiplicité de λ_k comme racine du polynôme caractéristique, on déduit que $\sum_{k=1}^p \dim(\ker(u - \lambda_k \text{Id})) = n$ et

donc que $E = \bigoplus_{k=1}^p \ker(u - \lambda_k \text{Id})$. On a alors $P(u) = 0$ avec $P(X) = \prod_{k=1}^p (X - \lambda_k)$

scindé sur $\mathbb{K}$ à racines simples.

(v) $\Rightarrow$ (vi) Le polynôme minimal étant un diviseur de tout polynôme annulateur, il est également scindé sur $\mathbb{K}$ à racines simples si (v) est vérifié.

(vi) $\Rightarrow$ (i) Si le polynôme minimal de u s'écrit $\pi_u(X) = \prod_{k=1}^p (X - \lambda_k)$, les λ_k étant deux à deux distincts, le théorème de décomposition des noyaux (théorème 10.1) nous dit alors que $E = \bigoplus_{k=1}^p \ker(u - \lambda_k Id)$, ce qui signifie que u est diagonalisable. $\square$

Un endomorphisme u est diagonalisable à racines simples si, et seulement si, son polynôme minimal ou caractéristique est de degré n et scindé à racines simples. Sur $\mathbb{C}$, il suffit de vérifier que $P_u \wedge P'_u = 1$, où P'_u est le polynôme dérivé du polynôme caractéristique P_u .

10.6 Endomorphismes trigonalisables

Définition 10.7. On dit que u est trigonalisable, s'il existe une base de E dans laquelle la matrice de u est triangulaire.

Si u a pour matrice A dans une base $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ de E et si u est trigonalisable, il existe alors une base $\mathcal{B}' = (e'_i)_{1 \leq i \leq n}$ dans laquelle la matrice T de u est triangulaire. On sait alors que les matrices A et T sont semblables c'est-à-dire qu'il existe une matrice d'ordre n à coefficients dans $\mathbb{K}$ inversible P telle que $T = P^{-1}AP$. Ce qui nous amène à poser la définition suivante.

Définition 10.8. On dit qu'une matrice $A \in \mathcal{M}_n(\mathbb{K})$ est trigonalisable, si elle est semblable à une matrice triangulaire.

Deux matrices semblables ayant même déterminant on en déduit qu'elles ont même polynôme caractéristique. En conséquence si la matrice A est semblable à une matrice triangulaire T , les termes diagonaux de T sont alors les valeurs propres de A .

Exemple 10.3 Une matrice à coefficients réels ayant des valeurs propres complexes non réelles n'est pas trigonalisable dans $\mathcal{M}_n(\mathbb{R})$. Par exemple la matrice $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ de valeurs propres complexes i et $-i$ n'est pas trigonalisable dans $M_2(\mathbb{R})$.

Lemme 10.4 Si $n \geq 2$ et si le polynôme caractéristique P_u de u est scindé sur $\mathbb{K}$, il existe alors un hyperplan de E stable par u .

Preuve. Comme P_u est scindé sur $\mathbb{K}$, on a $\text{Sp}(u) \neq \emptyset$ et pour $\lambda \in \text{Sp}(u)$, $\ker(u - \lambda Id) \neq \{0\}$, donc $\text{Im}(u - \lambda Id)$ est un sous-espace vectoriel de dimension inférieure ou égale à $n - 1$. Il existe donc un hyperplan H de E qui contient $\text{Im}(u - \lambda Id)$. On a donc $u(x) - \lambda x \in H$ pour tout $x \in E$ et :

$$\forall x \in H, u(x) = (u(x) - \lambda x) + \lambda x \in H$$

ce qui signifie que l'hyperplan H est stable par u . $\square$

Théorème 10.9.

L'endomorphisme u est trigonalisable sur $\mathbb{K}$ si, et seulement si, son polynôme caractéristique est scindé sur $\mathbb{K}$.

Preuve. Il est clair que si u est trigonalisable, son polynôme caractéristique est alors scindé sur $\mathbb{K}$ (on utilise la matrice de u dans une base de trigonalisation). Pour la réciproque, on raisonne par récurrence sur la dimension $n \geq 1$ de E . Pour $n = 1$ le résultat est évident. Supposons le acquis pour les espaces vectoriels de dimension $n - 1 \geq 1$ et soit $u \in \mathcal{L}(E)$ avec E de dimension $n \geq 2$ tel que P_u soit scindé sur $\mathbb{K}$. Le lemme précédent nous dit qu'il existe un hyperplan H de E qui est stable par u . Le polynôme caractéristique de la restriction de u à H qui divise χ_u (lemme 10.2) est également scindé sur $\mathbb{K}$, donc la restriction v de u à H est trigonalisable, c'est-à-dire qu'il existe une base $\mathcal{B}_1 = \mathcal{B} = (e_i)_{1 \leq i \leq n-1}$ de H dans laquelle la matrice de v est triangulaire supérieure. Pour tout vecteur $e_n \in E - H$ le système $\mathcal{B} = \mathcal{B} = (e_i)_{1 \leq i \leq n}$ est une base de E et la matrice de u dans cette base est triangulaire supérieure. $\square$

Corollaire 10.2. *Sur $\mathbb{C}$, tout endomorphisme $u \in \mathcal{L}(E)$ est trigonalisable.*

Preuve. Sur $\mathbb{C}$, P_u est scindé. $\square$

Corollaire 10.3. *Sur $\mathbb{C}$, toute matrice $A \in \mathcal{M}_n(\mathbb{C})$ est trigonalisable, c'est-à-dire qu'il existe une matrice $P \in GL_n(\mathbb{C})$ telle que $P^{-1}AP$ soit triangulaire.*

Corollaire 10.4. *Si $u \in \mathcal{L}(E)$ est trigonalisable et si F est un sous-espace vectoriel de E stable par u , alors la restriction de u à F est aussi trigonalisable.*

Preuve. Le polynôme caractéristique de la restriction de u à F qui divise P_u (lemme 10.2) est également scindé sur $\mathbb{C}$. $\square$

Corollaire 10.5. *Si $u \in \mathcal{L}(E)$ [resp. $A \in \mathcal{M}_n(\mathbb{K})$] est trigonalisable, alors la trace de u [resp. de A] est égale à la somme des valeurs propres de u [resp. de A] et le déterminant de u [resp. de A] est égal au produit des valeurs propres de u [resp. de A].*

Preuve. On utilise la matrice de u dans une base de trigonalisation. $\square$

Dans le cas où u [resp. de A] n'est pas trigonalisable, il est faux, a priori, de dire que la trace est égale à la somme des valeurs propres et le déterminant est égal au produit des valeurs propres. Par exemple pour $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 \cos(\theta) & -2 \sin(\theta) \\ 0 & 2 \sin(\theta) & 2 \cos(\theta) \end{pmatrix}$ dans $\mathcal{M}_3(\mathbb{R})$ avec $\theta \in]0, \pi[$, on a une seule valeur propre réelle, $\lambda_1 = 1$ et $\text{Tr}(A) = 1 + 4 \cos(\theta) \neq \lambda_1$, $\det(A) = 4 \neq \lambda_1$.

Théorème 10.10.

Les assertions suivantes sont équivalentes :

1. *l'endomorphisme u est trigonalisable ;*
2. *il existe un polynôme annulateur de u qui est scindé sur $\mathbb{K}$;*
3. *le polynôme minimal π_u est scindé sur $\mathbb{K}$.*

Preuve. Si u est trigonalisable de valeurs propres distinctes $\lambda_1, \dots, \lambda_p$, son polynôme caractéristique est $P_u(X) = (-1)^n \prod_{k=1}^p (X - \lambda_k)^{\alpha_k}$ (utiliser la matrice de u dans une base de trigonalisation) et ce dernier annule u . Le polynôme minimal étant un diviseur de tout polynôme annulateur de u , il est également scindé sur $\mathbb{K}$ si **2.** est vérifié.

Pour l'implication **3.** $\Rightarrow$ **1.** on procède par récurrence sur la dimension $n \geq 1$ de E . Pour $n = 1$ le résultat est évident, puisque dans ce cas tout endomorphisme est une homothétie. Supposons le acquis pour les espaces vectoriels de dimension $n - 1 \geq 1$ et soit $u \in \mathcal{L}(E)$ avec E de dimension $n \geq 2$ tel que π_u soit scindé sur $\mathbb{K}$. Les valeurs propres de u étant les racines de π_u , on peut considérer l'une d'elles $\lambda \in \mathbb{K}$ et l'espace propre associé $E_\lambda = \ker(u - \lambda Id) \neq \{0\}$. Le théorème du rang nous dit que le sous-espace vectoriel $\text{Im}(u - \lambda Id)$ est de dimension inférieure ou égale à $n - 1$ et il est contenu dans un hyperplan H . On a alors $u(x) - \lambda x \in H$ pour tout $x \in E$ et $u(x) = (u(x) - \lambda x) + \lambda x \in H$ pour tout $x \in H$, donc H est stable par u . Le polynôme minimal de la restriction de u à H qui divise π_u est également scindé sur $\mathbb{K}$, donc la restriction v de u à H est trigonalisable, c'est-à-dire qu'il existe une base $\mathcal{B}_1 = \mathcal{B} = (e_i)_{1 \leq i \leq n-1}$ de H dans laquelle la matrice de v est triangulaire supérieure. Pour tout vecteur $e_n \in E - H$ le système $\mathcal{B} = \mathcal{B} \cup (e_n)_{1 \leq i \leq n}$ est une base de E et la matrice de u dans cette base est triangulaire supérieure. $\square$

Corollaire 10.6. *Dans le cas où π_u est scindé sur $\mathbb{K}$, le déterminant de u est le produit des valeurs propres et la trace de u est la somme des valeurs propres, ces dernières étant comptées avec leurs multiplicités.*

10.7 Exercices

Exercice 10.1. *Calculer, pour $n \geq 3$, les valeurs propres de la matrice*

$$A = \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 & 1 \\ 1 & 1 & \cdots & 1 & 1 \end{pmatrix} \in \mathcal{M}_n(\mathbb{K}).$$

Solution. Le polynôme caractéristique de A est :

$$P_A(X) = \begin{vmatrix} 1-X & 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & -X & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & -X & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & \ddots & 0 & 1 \\ 1 & 0 & 0 & 0 & \cdots & -X & 1 \\ 1 & 1 & 1 & 1 & \cdots & 1 & 1-X \end{vmatrix}$$

En retranchant la colonne n à la colonne 1 et la colonne $n-1$ aux colonnes $k=2, \dots, n-2$, on obtient :

$$\begin{aligned} P_A(X) &= \begin{vmatrix} -X & 0 & 0 & 0 & \cdots & 1 & 1 \\ 0 & -X & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & -X & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ddots & 0 & 1 \\ 0 & X & X & X & \cdots & -X & 1 \\ X & 0 & 0 & 0 & \cdots & 1 & 1-X \end{vmatrix} \\ &= X^{n-2} \begin{vmatrix} -1 & 0 & 0 & 0 & \cdots & 1 & 1 \\ 0 & -1 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & -1 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ddots & 0 & 1 \\ 0 & 1 & 1 & 1 & \cdots & -X & 1 \\ 1 & 0 & 0 & 0 & \cdots & 1 & 1-X \end{vmatrix} \end{aligned}$$

et en ajoutant la ligne 1 à la ligne n , cela donne :

$$\begin{aligned} P_A(X) &= X^{n-2} \begin{vmatrix} -1 & 0 & 0 & 0 & \cdots & 1 & 1 \\ 0 & -1 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & -1 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ddots & 0 & 1 \\ 0 & 1 & 1 & 1 & \cdots & -X & 1 \\ 0 & 0 & 0 & 0 & \cdots & 2 & 2-X \end{vmatrix} \\ &= -X^{n-2} \begin{vmatrix} -1 & 0 & 0 & \cdots & 0 & 1 \\ 0 & -1 & 0 & \cdots & 0 & 1 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & 0 & 1 \\ 1 & 1 & 1 & \cdots & -X & 1 \\ 0 & 0 & 0 & \cdots & 2 & 2-X \end{vmatrix} \end{aligned}$$

En ajoutant la ligne k à la ligne $n-2$, pour $k = 1, \dots, n-3$, on obtient :

$$\begin{aligned} P_A(X) &= -X^{n-2}(-1)^{n-3} \begin{vmatrix} -X & n-2 \\ 2 & 2-X \end{vmatrix} = (-1)^n X^{n-2} (X(X-2) - 2(n-2)) \\ &= (-1)^n X^{n-2} ((X-1)^2 - (2n-3)) \end{aligned}$$

Les valeurs propres de A sont donc $\lambda_1 = 0$ d'ordre $n-2$, $\lambda_2 = 1 - \sqrt{2n-3}$ et $\lambda_3 = 1 + \sqrt{2n-3}$ qui sont simples. L'espace propre associé à la valeur propre 0 est l'espace de dimension $n-2$ d'équations :

$$\begin{cases} x_2 + \dots + x_{n-1} = 0 \\ x_1 + x_n = 0 \end{cases}$$

Un vecteur dans ce noyau s'écrit donc :

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-2} \\ -x_2 - \dots - x_{n-2} \\ -x_1 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ 0 \\ -1 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ -1 \\ 0 \end{pmatrix} + \dots + x_{n-2} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ -1 \\ 0 \end{pmatrix}$$

et en notant $(e_i)_{1 \leq i \leq n}$ la base canonique de $\mathbb{K}^n$, on déduit qu'une base de ce noyau est $(e_1 - e_n, e_2 - e_{n-1}, \dots, e_{n-2} - e_{n-1})$. Cette matrice est donc diagonalisable.

Exercice 10.2. On suppose que $\mathbb{K} = \mathbb{R}$ et que u^2 a une valeur propre $\mu > 0$. Montrer que $\sqrt{\mu}$ ou $-\sqrt{\mu}$ est valeur propre de u .

Solution. L'endomorphisme $u^2 - \mu Id = (u - \sqrt{\mu} Id) \circ (u + \sqrt{\mu} Id)$ est non injectif, donc l'un des deux endomorphismes $u - \sqrt{\mu} Id$ ou $u + \sqrt{\mu} Id$ est non injectif, ce qui signifie que $\sqrt{\mu}$ ou $-\sqrt{\mu}$ est valeur propre de u .

Exercice 10.3. Soit $A = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix} \in \mathcal{M}_n(\mathbb{C})$. Déterminer les valeurs propres de A et les espaces propres associés.

Solution. En désignant par $(e_i)_{1 \leq i \leq n}$ la base canonique de $\mathbb{C}^n$, on a :

$$Ae_k = e_{k+1} \quad (1 \leq k \leq n-1) \quad \text{et} \quad Ae_n = e_1$$

Il en résulte que $A^n e_k = e_k$ pour tout k compris entre 1 et n et $A^{n-1} e_1 = e_n \neq e_1$. Donc $A^n = I_n$ et $A^{n-1} \neq I_n$, c'est-à-dire que A est d'ordre n (comme σ). On

en déduit que les valeurs propres de A sont des racines n -èmes de l'unité. Soient $\omega = e^{\frac{2i\pi}{n}}$ et pour k compris entre 1 et n , $\lambda_k = \omega^k$ et f_k le vecteur :

$$f_k = \begin{pmatrix} \omega^{-k} \\ \omega^{-2k} \\ \vdots \\ \omega^{-(n-1)k} \\ \omega^{-nk} \end{pmatrix} = \sum_{j=1}^n \omega^{-jk} e_j$$

On a alors :

$$\begin{aligned} Af_k &= \sum_{j=1}^n \omega^{-jk} A e_j = \sum_{j=1}^{n-1} \omega^{-jk} e_{j+1} + \omega^{-nk} e_1 \\ &= \sum_{i=2}^n \omega^{-(i-1)k} e_i + e_1 = \sum_{i=1}^n \omega^{-(i-1)k} e_i = \omega^k \sum_{i=1}^n \omega^{-ik} e_i = \lambda_k f_k \end{aligned}$$

Les valeurs propres de A sont donc les n racines n -èmes de l'unité et les espaces propres sont des droites. La matrice A est donc diagonalisable.

Exercice 10.4. Soient a, b dans $\mathbb{K}$ et $A(a, b) = ((a_{ij}))_{1 \leq i, j \leq n}$ dans $\mathcal{M}_n(\mathbb{K})$ avec $n \geq 3$ définie par :

$$\forall i \in \{1, 2, \dots, n\}, \begin{cases} a_{ii} = b \\ a_{ij} = a \text{ si } j \in \{1, 2, \dots, n\} - \{i\} \end{cases}$$

1. Calculer $\Delta(a, b) = \det(A(a, b))$.
2. Calculer le polynôme caractéristique et les valeurs propres avec leur multiplicité de la matrice $A(a, b)$.

Solution.

1. La matrice $A(a, b)$ est de la forme $A(a, b) = \begin{pmatrix} b & a & a & \cdots & a \\ a & b & a & \cdots & a \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a & \cdots & a & b & a \\ a & \cdots & a & a & b \end{pmatrix}$. En

ajoutant les lignes 2 à n à la première ligne on a :

$$\Delta(a, b) = \det(A(a, b)) = (b + (n-1)a) \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ a & b & a & \cdots & a \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a & \cdots & a & b & a \\ a & \cdots & a & a & b \end{vmatrix}$$

Puis en retranchant la première colonne aux colonnes 2 à n on obtient :

$$\Delta(a, b) = (b + (n-1)a)(b-a)^{n-1}$$

2. Le polynôme caractéristique de $A(a, b)$ est donné par :

$$P_{(a,b)}(X) = \Delta(a, b - X) = (-1)^n (X - (b + (n-1)a))(X - (b-a))^{n-1}$$

Les valeurs propres de $A(a, b)$ sont donc $\lambda_1 = b + (n-1)a$ et $\lambda_2 = b - a$. Pour $a = 0$, la matrice $A(0, b)$ est celle d'une homothétie et $\lambda_1 = \lambda_2 = b$ est valeur propre d'ordre n de cette matrice. Pour $a \neq 0$, λ_1 est valeur propre simple et λ_2 est valeur propre d'ordre $n-1$. On peut aussi remarquer que $A(a, b) + (a-b)I_n = aA(1, 1)$ avec $A(1, 1)$ de rang 1. On déduit donc que 0 est valeur propre d'ordre $n-1$ de $A(1, 1)$ et que $\text{Tr}(A(1, 1)) = n$ est valeur propre simple (le polynôme caractéristique de $A(1, 1)$ est $(-1)^n X^{n-1}(X - \lambda)$). On retrouve alors facilement les valeurs propres de $A(a, b)$.

Exercice 10.5. Soit u un endomorphisme de E de rang égal à 1. Montrer que u est diagonalisable si, et seulement si, sa trace est non nulle.

Solution. Comme u est de rang égal à 1, son noyau $\ker(u)$ est de dimension $n-1$ et 0 est valeur propre d'espace propre associé $E_0 = \ker(u)$. On a alors $E = E_0 \oplus \mathbb{R}a$ où a est un vecteur non nul et la matrice de u dans une base

adaptée à cette décomposition est de la forme $A = \begin{pmatrix} 0 & \cdots & 0 & \lambda_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$, donc

$\lambda_n = \text{Tr}(A) = \text{Tr}(u)$ est valeur propre de u . Si $\text{Tr}(u) \neq 0$, ce scalaire est une valeur propre non nulle de u et son espace propre associé E_1 est de dimension au moins égale à 1 et comme $E_0 \oplus E_1$ est contenu dans E avec $\dim(E_0) = n-1$, on a nécessairement $\dim(E_1) = 1$ et l'égalité $E = E_0 \oplus E_1$, ce qui signifie que u est diagonalisable. Si $\text{Tr}(u) = 0$, l'endomorphisme admet alors 0 comme valeur propre d'ordre n et il est diagonalisable si, et seulement si, c'est l'endomorphisme nul, ce qui n'est pas possible puisque $\text{rg}(u) = 1$. En définitive, u est diagonalisable si, et seulement si, sa trace est non nulle.

Exercice 10.6. Montrer que si u est diagonalisable et F un sous-espace vectoriel de E stable par u , alors la restriction de u à F est un endomorphisme de F diagonalisable.

Solution. Si F est un sous-espace vectoriel de E stable par u , la restriction v de u à F est un endomorphisme de F . Comme u est diagonalisable, il annule par un polynôme scindé à racines simples et ce polynôme annule v qui est alors diagonalisable.

Exercice 10.7. Soit u un endomorphisme de E diagonalisable avec $\text{Sp}(u) = \{\lambda_1, \dots, \lambda_p\}$. Montrer que pour $1 \leq k \leq p$ la projection de E sur

le sous-espace propre $\ker(u - \lambda_k Id)$ est donnée par $p_k = \alpha_k \prod_{\substack{j=1 \\ j \neq k}}^p (u - \lambda_j Id)$,
 où $\alpha_k = \frac{1}{\prod_{\substack{j=1 \\ j \neq k}}^p (\lambda_k - \lambda_j)}$ (utiliser la décomposition en éléments simples de $\frac{1}{\pi_u}$).

Solution. Comme u est diagonalisable, son polynôme minimal s'écrit $\pi_u(X) = \prod_{k=1}^p (X - \lambda_k)$. La décomposition $\frac{1}{\pi_u(X)} = \sum_{k=1}^p \frac{\alpha_k}{X - \lambda_k}$, où $\alpha_k = \frac{1}{\prod_{\substack{j=1 \\ j \neq k}}^p (\lambda_k - \lambda_j)}$

permet d'obtenir l'égalité de Bézout $\sum_{k=1}^p \alpha_k \prod_{\substack{j=1 \\ j \neq k}}^p (X - \lambda_j) = 1$. De cette égalité, on déduit que :

$$\forall x \in E, x = \sum_{k=1}^p \alpha_k \prod_{\substack{j=1 \\ j \neq k}}^p (u - \lambda_j Id) x$$

Considérant que pour $1 \leq k \leq p$ et tout $x \in E$, on a :

$$(u - \lambda_k Id) \prod_{\substack{j=1 \\ j \neq k}}^p (u - \lambda_j Id) x = \pi_u(u) x = 0$$

on déduit que $\alpha_k \prod_{\substack{j=1 \\ j \neq k}}^p (u - \lambda_j Id) x \in \ker(u - \lambda_k Id)$, c'est-à-dire que les projecteurs p_k s'écrivent :

$$p_k = \alpha_k \prod_{\substack{j=1 \\ j \neq k}}^p (u - \lambda_j Id) \text{ avec } \alpha_k = \frac{1}{\prod_{\substack{j=1 \\ j \neq k}}^p (\lambda_k - \lambda_j)}$$

En particulier ces projecteurs sont des polynômes en u .

Chapitre 11

Formes bilinéaires et quadratiques réelles ou complexes

On notera $\mathbb{K}$ le corps de réels ou des complexes, en précisant quand cela sera nécessaire s'il s'agit de $\mathbb{R}$ ou $\mathbb{C}$. Par scalaire on entend réel ou complexe.

On désigne pour ce chapitre par E un espace vectoriel réel ou complexe de dimension finie ou non et non réduit à $\{0\}$.

11.1 Formes linéaires

On rappelle la définition suivante déjà donnée au paragraphe 6.4.

Définition 11.1. Une forme linéaire sur E est une application linéaire de E dans $\mathbb{K}$.

Exemples 11.1

1. Si E est un espace vectoriel de dimension n et $\mathcal{B} = (e_j)_{1 \leq j \leq n}$ une base de E , alors la j -ième projection, $p_j : x = \sum_{i=1}^n x_i e_i \mapsto x_j$, où j est un entier compris entre 1 et n , est une forme linéaire sur E .
2. Si E est un espace vectoriel de dimension n , $\mathcal{B} = (e_j)_{1 \leq j \leq n}$ une base de E et $\alpha_1, \dots, \alpha_n$ des scalaires, alors l'application :

$$\ell : x = \sum_{i=1}^n x_i e_i \mapsto \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n$$

est une forme linéaire sur E .

En fait toutes les formes linéaires sur E de dimension n sont de la forme précédente. En effet, tout vecteur x de E s'écrit $x = \sum_{j=1}^n x_j e_j$ et pour tout forme linéaire

ℓ sur E , on a $\ell(x) = \ell\left(\sum_{j=1}^n x_j e_j\right) = \sum_{j=1}^n x_j \ell(e_j) = \sum_{j=1}^n \alpha_j x_j$, en notant $\alpha_j = \ell(e_j)$ pour tout entier j compris entre 1 et n .

La matrice ligne $L = (\alpha_1, \dots, \alpha_n) = (\ell(e_1), \dots, \ell(e_n))$ est tout simplement la matrice de ℓ dans la base $\mathcal{B} = (e_j)_{1 \leq j \leq n}$ de E et on a :

$$\forall x \in E, \ell(x) = L \cdot x = (\alpha_1, \alpha_2, \dots, \alpha_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{j=1}^n \alpha_j x_j$$

Ce résultat peut aussi s'exprimer sous la forme :

$$\forall x \in E, \ell(x) = \sum_{j=1}^n \alpha_j p_j(x) = \left(\sum_{j=1}^n \alpha_j p_j \right) (x)$$

où p_j désigne, pour j compris entre 1 et n , la projection $x \mapsto x_j$. On peut donc écrire, une base $\mathcal{B}$ de E étant donnée, toute forme linéaire ℓ sur E sous la forme $\ell = \sum_{j=1}^n \alpha_j p_j$, où les $\alpha_j \in \mathbb{K}$ sont uniquement déterminés par $\alpha_j = \ell(e_j)$ pour tout entier j compris entre 1 et n . Nous avons donc montré le résultat suivant.

Théorème 11.1.

Si E est un espace vectoriel de dimension n et $\mathcal{B} = (e_j)_{1 \leq j \leq n}$ une base de E , alors l'ensemble de toutes les formes linéaires sur E est un espace vectoriel de dimension n de base $(p_1, \dots, p_n)$.

On rappelle que si on dispose d'une base $\mathcal{B}$ d'un espace vectoriel E dire qu'une famille $(v_1, \dots, v_p)$ d'éléments de E est libre (ou que ces éléments sont linéairement indépendants) équivaut à dire que les vecteurs colonnes $X_1, \dots, X_p$ formés des composantes de ces vecteurs dans la base $\mathcal{B}$ sont linéairement indépendants dans $\mathbb{K}^n$. On peut donc parler de formes linéaires linéairement indépendantes. On rappelle également que pour montrer que le système $(X_1, \dots, X_p)$ est libre dans $\mathbb{K}^n$, il suffit d'extraire de la matrice $(X_1, \dots, X_p)$ un déterminant d'ordre p non nul (ce qui impose bien sur que $p \leq n$).

Dire que le système $(X_1, \dots, X_p)$ est libre dans $\mathbb{K}^n$ équivaut aussi à dire que la matrice $(X_1, \dots, X_p)$ est de rang p . Comme une matrice et sa transposée ont même

rang, il revient au même de calculer le rang de la matrice transposée $\begin{pmatrix} {}^t X_1 \\ \vdots \\ {}^t X_p \end{pmatrix}$.

On retiendra que des formes linéaires $\ell_1, \dots, \ell_p$ définies sur E , de base $(e_j)_{1 \leq j \leq n}$, par :

$$\forall x \in E, \ell_i(x) = \sum_{j=1}^n \alpha_{i,j} x_j \quad (1 \leq i \leq p)$$

sont linéairement indépendantes si, et seulement si la matrice :

$$A = \begin{pmatrix} L_1 \\ \vdots \\ L_p \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & & \ddots & \vdots \\ \alpha_{p1} & \alpha_{p2} & \cdots & \alpha_{pn} \end{pmatrix}$$

est de rang p (L_i est la matrice de ℓ_i dans la base $\mathcal{B}$ de E), ce qui revient à dire qu'on peut en extraire un déterminant d'ordre p non nul.

La somme de deux formes linéaires sur E est une forme linéaire, mais en général le produit de deux formes linéaires sur E n'est pas une forme linéaire.

Définition 11.2. On appelle *hyperplan vectoriel* de E , le noyau d'une forme linéaire non nulle sur E .

De manière un peu plus générale, un hyperplan affine de E est un sous-ensemble de E de la forme $H = \{x \in E \mid \ell(x) = \lambda\}$, où ℓ est une forme linéaire non nulle et λ un réel. Sur E , de base $\mathcal{B} = (e_j)_{1 \leq j \leq n}$, un hyperplan est donc l'ensemble des vecteurs $x = \sum_{j=1}^n x_j e_j$ tels que $\sum_{j=1}^n \alpha_j x_j = 0$, où les scalaires α_j ne sont pas tous nuls. De plus une forme linéaire non nulle ℓ étant surjective (exercice 6.5), le théorème du rang nous dit que, pour E de dimension n , on a $\dim(\ker(\ell)) = n - 1$. Réciproquement si H est un sous-espace de dimension $n - 1$ dans E de dimension n , il admet une base $(e_i)_{1 \leq i \leq n-1}$ qui peut se compléter en une base $(e_i)_{1 \leq i \leq n}$ de E et H est le noyau de la n -ième projection $p_n : x = \sum_{j=1}^n x_j e_j \mapsto x_n$. Nous avons donc montré le résultat suivant.

Théorème 11.2.

Sur un espace vectoriel E de dimension n un hyperplan est un sous-espace de E de dimension $n - 1$.

Les supplémentaires d'un hyperplan dans E de dimension finie sont donc des droites. En fait ce résultat est général.

Théorème 11.3.

Si H est un hyperplan d'un espace vectoriel E , il existe alors une droite D telle que $E = H \oplus D$.

Preuve. On a $H = \ker(\ell)$ où ℓ est une forme linéaire non nulle sur E . Il existe donc un vecteur non nul a dans E tel que $\ell(a) \neq 0$. En désignant par $D = \mathbb{K}a$ la droite dirigée par a , on a alors $E = H \oplus D$. En effet, si $x \in H \cap D$, il existe un scalaire λ tel que $x = \lambda a$ et $\ell(x) = \lambda \ell(a) = 0$ nous donne $\lambda = 0$. On a donc $H \cap D = \{0\}$. De plus pour tout vecteur $x \in E$, le vecteur $y = x - \frac{\ell(x)}{\ell(a)}a$ est

dans $H = \ker(\ell)$ et avec $x = y + \frac{\ell(x)}{\ell(a)}a$, on déduit que $x \in H + D$. On a donc $E = H + D$ et $E = H \oplus D$. $\square$

Réciproquement un sous-espace vectoriel H de E supplémentaire d'une droite D est le noyau de la forme linéaire ℓ qui associe à tout vecteur x de E sa projection sur D , c'est donc un hyperplan. On a donc le résultat suivant.

Théorème 11.4.

Un hyperplan de E est un sous-espace de E supplémentaire d'une droite.

11.2 Formes bilinéaires

Définition 11.3. Une forme bilinéaire sur E est une application :

$$\begin{aligned} \varphi : E \times E &\rightarrow \mathbb{K} \\ (x, y) &\mapsto \varphi(x, y) \end{aligned}$$

telle que pour tout x dans E l'application $y \mapsto \varphi(x, y)$ est linéaire et pour tout y dans E l'application $x \mapsto \varphi(x, y)$ est linéaire.

Définition 11.4. On dit qu'une forme bilinéaire φ sur E est symétrique si $\varphi(y, x) = \varphi(x, y)$ pour tous x, y dans E .

Définition 11.5. On dit qu'une forme bilinéaire φ sur E est anti-symétrique (ou alternée) si $\varphi(y, x) = -\varphi(x, y)$ pour tous x, y dans E .

Une application symétrique φ de E^2 dans $\mathbb{K}$ est bilinéaire si, et seulement si, l'une des deux applications $y \mapsto \varphi(x, y)$ (pour tout x dans E) ou $x \mapsto \varphi(x, y)$ (pour tout y dans E) est linéaire.

Exemples 11.2

1. Si ℓ_1 et ℓ_2 sont deux formes linéaires sur E , alors l'application $(x, y) \mapsto \ell_1(x) \ell_2(y)$ est une forme bilinéaire sur E .
2. Si E est l'espace $C^0([a, b], \mathbb{R})$ des fonctions continues de $[a, b]$ dans $\mathbb{R}$, alors l'application $\varphi : (f, g) \mapsto \int_a^b f(t) g(t) dt$ est une forme bilinéaire.
3. Si $\ell_1, \dots, \ell_p$ sont des formes linéaires sur E et $(\alpha_{ij})_{1 \leq i, j \leq p}$ une famille de scalaires, alors l'application $(x, y) \mapsto \sum_{1 \leq i, j \leq p} \alpha_{ij} \ell_i(x) \ell_j(y)$ est une forme bilinéaire.

Nous verrons un peu plus loin que, sur un espace de dimension n , toutes les formes bilinéaires sont de la forme précédente.

On vérifie facilement que l'ensemble $\text{Bil}(E)$ de toutes les formes bilinéaires sur E est un espace vectoriel.

11.3 Expression matricielle des formes bilinéaires (en dimension finie)

Comme pour les applications linéaires, les matrices nous serviront à décrire une forme bilinéaire dans le cas des espaces de dimension finie.

Pour ce paragraphe E est un espace vectoriel de dimension n et on désigne par $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une base de E . Tout vecteur $x \in E$ s'écrit de manière unique sous

la forme $x = \sum_{j=1}^n x_j e_j$. On associe à un tel x le vecteur colonne $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ de $\mathbb{K}^n$.

Plaçons nous tout d'abord sur $E = \mathbb{R}^2$ (ou $\mathbb{C}^2$) muni de sa base canonique (e_1, e_2) . Si φ est une forme bilinéaire sur E , on a alors pour tous vecteurs $x = x_1 e_1 + x_2 e_2$ et $y = y_1 e_1 + y_2 e_2$ dans E :

$$\begin{aligned} \varphi(x, y) &= \varphi(x_1 e_1 + x_2 e_2, y) = x_1 \varphi(e_1, y) + x_2 \varphi(e_2, y) \\ &= x_1 \varphi(e_1, y_1 e_1 + y_2 e_2) + x_2 \varphi(e_2, y_1 e_1 + y_2 e_2) \\ &= x_1 (y_1 \varphi(e_1, e_1) + y_2 \varphi(e_1, e_2)) + x_2 (y_1 \varphi(e_2, e_1) + y_2 \varphi(e_2, e_2)) \end{aligned}$$

En désignant par A la matrice $A = \begin{pmatrix} \varphi(e_1, e_1) & \varphi(e_1, e_2) \\ \varphi(e_2, e_1) & \varphi(e_2, e_2) \end{pmatrix}$, on remarque que :

$$AY = \begin{pmatrix} \varphi(e_1, e_1) & \varphi(e_1, e_2) \\ \varphi(e_2, e_1) & \varphi(e_2, e_2) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} y_1 \varphi(e_1, e_1) + y_2 \varphi(e_1, e_2) \\ y_1 \varphi(e_2, e_1) + y_2 \varphi(e_2, e_2) \end{pmatrix}$$

et :

$${}^t X (AY) = (x_1, x_2) \begin{pmatrix} y_1 \varphi(e_1, e_1) + y_2 \varphi(e_1, e_2) \\ y_1 \varphi(e_2, e_1) + y_2 \varphi(e_2, e_2) \end{pmatrix} = \varphi(x, y)$$

Le produit des matrices étant associatif, cela s'écrit $\varphi(x, y) = {}^t X A Y$. Le cas d'une forme bilinéaire sur un espace de dimension n se traite de manière analogue.

Définition 11.6. La matrice d'une forme bilinéaire $\varphi \in \text{Bil}(E)$ dans une base $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ de E est la matrice $A = ((\varphi(e_i, e_j)))_{1 \leq i, j \leq n}$.

Théorème 11.5.

Soit φ une forme bilinéaire sur E et A la matrice de φ dans la base $\mathcal{B}$. Pour tous vecteurs x, y dans E , on a $\varphi(x, y) = {}^t X A Y$.

Preuve. En utilisant la bilinéarité de φ , on a :

$$\begin{aligned}\varphi(x, y) &= \varphi\left(\sum_{i=1}^n x_i e_i, y\right) = \sum_{i=1}^n x_i \varphi(e_i, y) \\ &= \sum_{i=1}^n x_i \varphi\left(e_i, \sum_{j=1}^n y_j e_j\right) = \sum_{i=1}^n x_i \sum_{j=1}^n y_j \varphi(e_i, e_j)\end{aligned}$$

$$\text{et avec } AY = \left(\sum_{j=1}^n y_j \varphi(e_i, e_j)\right)_{1 \leq i \leq n}, \quad {}^t XAY = {}^t X(AY) = \sum_{i=1}^n x_i \sum_{j=1}^n y_j \varphi(e_i, e_j),$$

on a le résultat annoncé. $\square$

On retiendra que l'expression d'une forme bilinéaire φ dans une base est :

$$\varphi(x, y) = {}^t XAY = \sum_{i=1}^n x_i \sum_{j=1}^n a_{ij} y_j = \sum_{1 \leq i, j \leq n} a_{ij} x_i y_j$$

où on a posé $a_{ij} = \varphi(e_i, e_j)$ pour i, j compris entre 1 et n . Réciproquement une telle fonction sur E^2 définit bien une forme bilinéaire sur E . Ce résultat peut aussi s'exprimer comme suit.

Théorème 11.6.

Une application φ de $E \times E$ dans $\mathbb{K}$ est une forme bilinéaire sur E si, et seulement si, et seulement si, il existe une matrice $A = (a_{ij})_{1 \leq i, j \leq n}$ dans $\mathcal{M}_n(\mathbb{K})$ et des formes linéaires $\ell_1, \dots, \ell_n$ linéairement indépendantes telles que :

$$\forall (x, y) \in E \times E, \quad \varphi(x, y) = \sum_{1 \leq i, j \leq n} a_{ij} \ell_i(x) \ell_j(y).$$

Preuve. Si φ est bilinéaire, on a alors dans une base $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ de E , pour tous x, y dans E :

$$\varphi(x, y) = \sum_{1 \leq i, j \leq n} a_{ij} x_i y_j = \sum_{1 \leq i, j \leq n} a_{ij} \ell_i(x) \ell_j(y)$$

où $(\ell_i)_{1 \leq i \leq n}$ est la base duale de $\mathcal{B}$. La réciproque est claire. $\square$

L'application qui associe à une forme bilinéaire φ sur un espace vectoriel E de dimension n sa matrice dans une base donnée de E réalise un isomorphisme de $\text{Bil}(E)$ sur l'espace $\mathcal{M}_n(\mathbb{K})$ des matrices carrées d'ordre n . Cet espace étant de dimension n^2 , on en déduit que $\dim_{\mathbb{K}}(\text{Bil}(E)) = n^2$.

Théorème 11.7.

Une forme bilinéaire φ sur E est symétrique [resp. alternée] si, et seulement si, sa matrice A dans une quelconque base $\mathcal{B}$ de E est symétrique [resp. alternée].

Preuve. Si φ est symétrique [resp. alternée], on a alors en particulier $\varphi(e_i, e_j) = \varphi(e_j, e_i)$ [resp. $\varphi(e_i, e_j) = -\varphi(e_j, e_i)$] pour tous i, j compris entre 1 et n , ce qui signifie que la matrice A de φ dans $\mathcal{B}$ est symétrique [resp. alternée]. Réciproquement si cette matrice est symétrique [resp. alternée], on a alors pour tous x, y dans E :

$$\varphi(y, x) = {}^tYAX = {}^t({}^tYAX) = {}^tX {}^tAY = {}^tXAY = \varphi(x, y)$$

$$\text{resp. } \varphi(y, x) = {}^tYAX = {}^t({}^tYAX) = {}^tX {}^tAY = -{}^tXAY = -\varphi(x, y)$$

(le produit matriciel $T = {}^tYAX$ étant un scalaire, on a bien ${}^tT = T$). $\square$

Théorème 11.8.

Soient $\mathcal{B}_1$ et $\mathcal{B}_2$ deux bases de E et P la matrice de passage de $\mathcal{B}_1$ à $\mathcal{B}_2$. Si A_1 et A_2 sont les matrices d'une forme bilinéaire φ sur E dans les bases $\mathcal{B}_1$ et $\mathcal{B}_2$ respectivement, on a alors $A_2 = {}^tPA_1P$.

Preuve. Pour $x \in E$, on note respectivement X_1 et X_2 les vecteurs colonnes formés des composantes de x dans les bases $\mathcal{B}_1$ et $\mathcal{B}_2$ respectivement. Pour tous vecteurs x, y dans E , on a alors :

$$\varphi(x, y) = {}^tX_1A_1Y_1 = {}^t(PX_2)A_1(PY_2) = {}^tX_2({}^tPA_1P)Y_2$$

ce qui signifie exactement que $A_2 = {}^tPA_1P$ du fait de l'unicité de la matrice de φ dans la base $\mathcal{B}_2$. $\square$

Définition 11.7. Le discriminant dans une base $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ de E d'une forme bilinéaire φ est le déterminant de la matrice $A = ((\varphi(e_i, e_j)))_{1 \leq i, j \leq n}$ de φ dans cette base. On le note $\Delta_{\mathcal{B}}(\varphi)$.

En utilisant le théorème 11.8, on déduit que si $\mathcal{B}_1$ et $\mathcal{B}_2$ sont deux bases de E et P la matrice de passage de $\mathcal{B}_1$ à $\mathcal{B}_2$, on a alors pour toute forme bilinéaire φ sur E , $\Delta_{\mathcal{B}_2}(\varphi) = (\det(P))^2 \Delta_{\mathcal{B}_1}(\varphi)$.

11.4 Formes quadratiques

Définition 11.8. On appelle forme quadratique sur E une application q définie de E dans $\mathbb{K}$ par :

$$\forall x \in E, q(x) = \varphi(x, x)$$

où φ est une forme bilinéaire.

Il est facile de vérifier que l'ensemble $Q(E)$ des formes quadratiques sur E est un espace vectoriel.

A priori, il n'y a pas unicité des formes bilinéaires associées à une forme quadratique. Par exemple sur $\mathbb{R}^2$, les formes bilinéaires φ et ψ définies par :

$$\begin{cases} \varphi(x, y) = x_1 y_1 + x_2 y_2 \\ \psi(x, y) = x_1 y_1 + x_1 y_2 - x_2 y_1 + x_2 y_2 \end{cases}$$

définissent la même forme quadratique :

$$q(x) = \varphi(x, x) = x_1^2 + x_2^2 = \psi(x, x) = x_1^2 + x_1 x_2 - x_2 x_1 + x_2^2$$

L'unicité de φ est assurée par le résultat suivant.

Théorème 11.9.

Si q est une forme quadratique sur E , il existe alors une unique forme bilinéaire symétrique φ telle que $q(x) = \varphi(x, x)$ pour tout $x \in E$.

Preuve. La forme quadratique q est définie par $q(x) = \varphi_0(x, x)$ pour tout $x \in E$, où φ_0 est une forme bilinéaire sur E . L'application φ définie sur $E \times E$ par :

$$\varphi(x, y) = \frac{\varphi_0(x, y) + \varphi_0(y, x)}{2}$$

est bilinéaire et symétrique avec $\varphi(x, x) = q(x)$ pour tout $x \in E$, ce qui prouve l'existence de φ . Comme φ est bilinéaire et symétrique, on a pour x, y dans E :

$$\begin{aligned} q(x + y) &= \varphi(x + y, x + y) = \varphi(x, x) + 2\varphi(x, y) + \varphi(y, y) \\ &= q(x) + 2\varphi(x, y) + q(y) \end{aligned}$$

de sorte que $\varphi(x, y) = \frac{q(x + y) - q(x) - q(y)}{2}$, ce qui prouve l'unicité de φ . $\square$

Définition 11.9. Avec les notations du théorème qui précède, on dit que φ est la forme polaire de la forme quadratique q .

On retiendra l'expression de cette forme polaire :

$$\forall (x, y) \in E^2, \varphi(x, y) = \frac{q(x + y) - q(x) - q(y)}{2}$$

En écrivant que :

$$\begin{cases} q(x + y) = q(x) + 2\varphi(x, y) + q(y) \\ q(x - y) = q(x) - 2\varphi(x, y) + q(y) \end{cases}$$

on déduit que cette forme polaire est aussi définie par :

$$\forall (x, y) \in E^2, \varphi(x, y) = \frac{q(x + y) - q(x - y)}{4}$$

On notera aussi que pour tout scalaire λ et tout vecteur x , on a :

$$q(\lambda x) = \varphi(\lambda x, \lambda x) = \lambda^2 \varphi(x, x) = \lambda^2 q(x)$$

ce qui se traduit en disant que q est une fonction homogène de degré 2.

L'application qui associe à une forme quadratique q sa forme polaire φ réalise un isomorphisme d'espaces vectoriels de $Q(E)$ sur l'espace $Bil_s(E)$ des formes bilinéaires symétriques sur E . Pour E de dimension n , $Q(E)$ est de dimension $\frac{n(n+1)}{2}$. De cet isomorphisme, on déduit aussi que deux formes bilinéaires symétriques φ_1 et φ_2 sur E sont égales si, et seulement si, $\varphi_1(x, x) = \varphi_2(x, x)$ pour tout $x \in E$.

Dans le cas des espaces vectoriels de dimension finie on peut utiliser les matrices pour définir les formes quadratiques.

Définition 11.10. Soit E un espace vectoriel de dimension n et $\mathcal{B}$ une base de E . Si q est une forme quadratique sur E de forme polaire φ , on dit alors que la matrice de φ dans la base $\mathcal{B}$ est la matrice de q dans cette base.

En reprenant les notations du paragraphe 11.3, une forme quadratique est définie sur E de base $\mathcal{B}$ par $q(x) = \varphi(x, x) = {}^t X A X = \sum_{1 \leq i, j \leq n} a_{ij} x_i x_j$ avec $a_{ij} = a_{ji}$, ce qui peut s'écrire :

$$q(x) = \sum_{i=1}^n a_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j$$

Réciproquement une fonction q ainsi définie est une forme quadratique sur E de matrice $A = ((a_{ij}))_{1 \leq i, j \leq n}$ dans la base $\mathcal{B}$.

Le choix d'une base de E permet donc de réaliser un isomorphisme d'espaces vectoriels de $Q(E)$ sur l'espace des polynômes homogènes de degré 2 à n variables.

11.5 Théorème de réduction de Gauss

11.5.1 Cas des espaces de dimension 2

On désigne par E un $\mathbb{K}$ -espace vectoriel de dimension 2, $\mathcal{B} = (e_1, e_2)$ une base de E et pour tout vecteur v de E , on note x, y les coordonnées de v dans cette base, soit $v = xe_1 + ye_2$. Dans cette base, une forme quadratique q s'écrit sous la forme $q(v) = ax^2 + 2bxy + cy^2$. La matrice de cette forme quadratique dans la base $\mathcal{B}$ est donc $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$. On suppose que $q \neq 0$, soit que $(a, b, c) \neq (0, 0, 0)$.

- Si $a \neq 0$, on a $q(v) = a \left(x + \frac{b}{a} y \right)^2 + \frac{\delta}{a} y^2$, où $\delta = ac - b^2$ est le déterminant de A . Il y a alors deux possibilités :
 - soit $\delta = 0$ et $q(v) = a \left(x + \frac{b}{a} y \right)^2 = a \ell_1^2(v)$, où $\ell_1 : v \mapsto x + \frac{b}{a} y$ est une forme linéaire non nulle

- soit $\delta \neq 0$ et $q(v) = a \left(x + \frac{b}{a}y\right)^2 + \frac{\delta}{a}y^2 = a\ell_1^2(v) + \frac{\delta}{a}\ell_2^2(v)$, où $\ell_1 : v \mapsto x + \frac{b}{a}y$ et $\ell_2 : v \mapsto y$ sont deux formes linéaires indépendantes puisque $\begin{vmatrix} 1 & 0 \\ \frac{b}{a} & 1 \end{vmatrix} = 1 \neq 0$.
- Si $a = 0$ et $c \neq 0$, on a $q(v) = 2bxy + cy^2 = c \left(y + \frac{b}{c}x\right)^2 + \frac{\delta}{c}x^2$, où $\delta = -b^2$ est encore le déterminant de A et il y a deux possibilités :
 - soit $b = 0$ et $q(v) = cy^2 = c\ell_1^2(v)$, où $\ell_1 : v \mapsto y$ est une forme linéaire non nulle ;
 - soit $b \neq 0$ et $q(v) = c \left(y + \frac{b}{c}x\right)^2 + \frac{\delta}{c}x^2 = a\ell_1^2(v) + \frac{\delta}{a}\ell_2^2(v)$, où $\ell_1 : v \mapsto y + \frac{b}{c}x$ et $\ell_2 : v \mapsto x$ sont deux formes linéaires indépendantes.
- Si $a = 0$ et $c = 0$, on a alors $b \neq 0$ et :

$$q(x, y) = 2bxy = \frac{b}{2} \left((x+y)^2 - (x-y)^2 \right) = \frac{b}{2}\ell_1^2(v) - \frac{b}{2}\ell_2^2(vy)$$

où $\ell_1 : v \mapsto x + y$ et $\ell_2 : v \mapsto x - y$ sont deux formes linéaires indépendantes puisque $\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2 \neq 0$.

On a donc montré le résultat suivant.

Théorème 11.10.

Toute forme quadratique non nulle q sur $\mathbb{K}$ -espace vectoriel E de dimension 2 peut s'écrire sous la forme $q = \lambda_1 \ell_1^2$ où λ_1 est un scalaire non nul et ℓ_1 une forme linéaire non nulle ou $q = \lambda_1 \ell_1^2 + \lambda_2 \ell_2^2$ où λ_1, λ_2 sont deux scalaires non nuls et ℓ_1, ℓ_2 deux formes linéaires indépendantes.

Ce résultat se généralise dans le cas des espaces de dimension n comme on le verra au paragraphe suivant. Quand on a trouvé une telle décomposition, on dit qu'on a réduit la forme quadratique, sous-entendu sous forme de combinaison linéaire de carrés de formes linéaires indépendantes.

Si q est une forme quadratique sur $E = \mathbb{R}^2$, on notera $q(x, y)$ pour $q(v)$, où $v = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$.

On peut remarquer qu'une telle décomposition n'est pas unique. Par exemple, pour q_2 , on peut aussi écrire $q_2(x, y) = \left(\frac{1}{2}x + \frac{1}{2}y\right)^2 - \left(\frac{1}{2}x - \frac{1}{2}y\right)^2$.

11.5.2 Cas des espaces de dimension $n \geq 1$

Commençons par un exemple. En s'inspirant de la méthode exposée au paragraphe précédent, nous allons réduire la forme quadratique q définie sur $\mathbb{R}^3$ par

$q(x) = x_1^2 + x_2^2 + x_3^2 + 2x_1x_3 + 2x_2x_3$. On regroupe les termes contenant x_1 pour l'écrire comme le début d'un carré, soit $x_1^2 + 2x_1x_3 = (x_1 + x_3)^2 - x_3^2$, ce qui donne $q(x) = (x_1 + x_3)^2 + x_2^2 + 2x_2x_3$. On utilise ensuite la méthode développée pour le cas $n = 2$ à la forme q' définie sur $\mathbb{R}^2$ par $q'(x_2, x_3) = x_2^2 + 2x_2x_3$, soit $q'(x_2, x_3) = (x_2 + x_3)^2 - x_3^2$, ce qui donne :

$$q(x) = (x_1 + x_3)^2 + (x_2 + x_3)^2 - x_3^2 = \ell_1^2(x) + \ell_2^2(x) - \ell_3^2(x)$$

les formes ℓ_1, ℓ_2 et ℓ_3 étant indépendantes puisque $\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{vmatrix} = 1 \neq 0$.

La démonstration du théorème qui suit s'inspire de cette méthode.

Si E est un $\mathbb{K}$ -espace vectoriel de dimension $n \geq 1$, on notera $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une base de E et pour tout vecteur x de E , $x_1, \dots, x_n$ désignent les coordonnées de x dans cette base, soit $x = \sum_{i=1}^n x_i e_i$. On associe toujours à ce vecteur x de E le vecteur colonne $X = (x_i)_{1 \leq i \leq n}$ dans $\mathbb{K}^n$.

Théorème 11.11.

Pour toute forme quadratique non nulle q sur E , il existe un entier p compris entre 1 et n , des scalaires non nuls $\lambda_1, \dots, \lambda_p$ et des formes linéaires $\ell_1, \dots, \ell_p$ indépendantes dans E^* tels que :

$$\forall x \in E, q(x) = \sum_{j=1}^p \lambda_j \ell_j^2(x)$$

Preuve. Pour $n = 1$, il n'y a rien à montrer et pour $n = 2$ c'est fait. On suppose le résultat acquis au rang $n - 1$ et on se donne une forme quadratique non nulle q définie dans une base $\mathcal{B}$ d'un espace vectoriel E de dimension $n \geq 3$ par :

$$q(x) = \sum_{i=1}^n a_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j$$

Supposons tout d'abord que cette expression contient au moins un terme carré, c'est-à-dire qu'il existe un indice i compris entre 1 et n tel que $a_{ii} \neq 0$. Quitte à effectuer une permutation sur les vecteurs de base, on peut supposer que $a_{11} \neq 0$. En regroupant les termes contenant x_1 , on écrit que :

$$\begin{aligned} a_{11} x_1^2 + 2 \sum_{j=2}^n a_{1j} x_1 x_j &= a_{11} \left(x_1^2 + 2x_1 \sum_{j=2}^n \frac{a_{1j}}{a_{11}} x_j \right) \\ &= a_{11} \left(\left(x_1 + \sum_{j=2}^n \frac{a_{1j}}{a_{11}} x_j \right)^2 - \left(\sum_{j=2}^n \frac{a_{1j}}{a_{11}} x_j \right)^2 \right) \end{aligned}$$

et :

$$q(x) = a_{11} \left(x_1 + \sum_{j=2}^n \frac{a_{1j}}{a_{11}} x_j \right)^2 + q'(x') = a_{11} \ell_1^2(x) + q'(x')$$

où $\ell_1(x) = x_1 + \sum_{j=2}^n \frac{a_{1j}}{a_{11}} x_j$, q' est une forme quadratique définie sur le sous-espace

vectorel H de E engendré par $e_2, \dots, e_n$ et $x' = \sum_{i=2}^n x_i e_i$ si $x = \sum_{i=1}^n x_i e_i$.

Si $q' = 0$, on a alors $q = a_{11} \ell_1^2$ avec a_{11} et ℓ_1 non nuls.

Si $q' \neq 0$, l'hypothèse de récurrence nous dit qu'il existe un entier p compris entre 2 et n , des scalaires non nuls $\lambda_2, \dots, \lambda_p$ et des formes linéaires indépendantes $\ell_2, \dots, \ell_p$ définies sur H tels que :

$$\forall x' \in H, q'(x') = \sum_{j=2}^p \lambda_j \ell_j^2(x')$$

et en prolongeant les formes linéaires $\ell_2, \dots, \ell_p$ à E (en posant $\ell_j(x) = \ell_j(x')$),

on a $q(x) = a_{11} \ell_1^2(x) + \sum_{j=2}^p \lambda_j \ell_j^2(x)$, ce qui donne une décomposition de q comme combinaison linéaire de carrés de formes linéaires.

Il reste à vérifier que les formes $\ell_1, \ell_2, \dots, \ell_p$ sont linéairement indépendantes dans E^* .

L'égalité $\sum_{j=1}^p \lambda_j \ell_j$ équivaut à dire que $\sum_{j=1}^p \lambda_j \ell_j(x) = 0$ pour tout $x \in E$. Prenant $x = e_1$, on a $\ell_1(x) = 1$ et $\ell_j(x) = 0$ pour j compris entre 2 et p , ce qui donne $\lambda_1 = 0$ et $\sum_{j=2}^p \lambda_j \ell_j(x') = 0$ pour tout $x' \in H$, ce qui équivaut à $\sum_{j=2}^p \lambda_j \ell_j = 0$ et la nullité de tous les λ_j puisque le système $(\ell_2, \dots, \ell_p)$ est libre dans H^* . On a donc le résultat annoncé.

Il reste enfin à traiter le cas où q est sans facteurs carrés, c'est-à-dire le cas où tous les coefficients a_{ii} sont nuls. Comme q est non nulle, il existe deux indices $i < j$ tels que $a_{ij} \neq 0$. Quitte à effectuer une permutation sur les vecteurs de base, on peut supposer que $a_{12} \neq 0$. On regroupe alors dans l'expression de q tous les termes contenant x_1 et x_2 que l'on fait apparaître comme fragment d'un produit de deux formes linéaires, soit :

$$\begin{aligned} Q &= a_{12} x_1 x_2 + x_1 \sum_{j=3}^n a_{1j} x_j + x_2 \sum_{j=3}^n a_{2j} x_j \\ &= \left(a_{12} x_1 + \sum_{j=3}^n a_{2j} x_j \right) \left(x_2 + \sum_{j=3}^n \frac{a_{1j}}{a_{12}} x_j \right) - \left(\sum_{j=3}^n a_{2j} x_j \right) \left(\sum_{j=3}^n \frac{a_{1j}}{a_{12}} x_j \right) \end{aligned}$$

ce qui donne :

$$q(x) = 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j = 2Q + 2 \sum_{3 \leq i < j \leq n} a_{ij} x_i x_j = 2L_1(x) L_2(x) + q'(x')$$

où $L_1(x) = a_{12}x_1 + \sum_{j=3}^n a_{2j}x_j$, $L_2(x) = x_2 + \sum_{j=3}^n \frac{a_{1j}}{a_{12}}x_j$ et q' est une forme quadratique définie sur le sous-espace vectoriel H de E engendré par $e_3, \dots, e_n$ et $x' = \sum_{i=3}^n x_i e_i$ si $x = \sum_{i=1}^n x_i e_i$. En écrivant que :

$$2L_1(x)L_2(x) = \frac{1}{2}(L_1(x) + L_2(x))^2 - \frac{1}{2}(L_1(x) - L_2(x))^2 = \frac{1}{2}\ell_1^2(x) - \frac{1}{2}\ell_2^2(x)$$

on a $q(x) = \frac{1}{2}\ell_1^2(x) - \frac{1}{2}\ell_2^2(x) + q'(x')$.

Si $q' = 0$, on a alors $q = \frac{1}{2}\ell_1^2 - \frac{1}{2}\ell_2^2$, les formes linéaires ℓ_1 et ℓ_2 étant indépendantes puisque la matrice $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \begin{pmatrix} a_{12} & 1 & \alpha_{13} & \dots & \alpha_{1n} \\ a_{12} & -1 & \alpha_{23} & \dots & \alpha_{2n} \end{pmatrix}$ est de rang 2 car le déterminant extrait $\begin{vmatrix} a_{12} & 1 \\ a_{12} & -1 \end{vmatrix} = -2a_{12}$ est non nul.

Si $q' \neq 0$, l'hypothèse de récurrence nous dit qu'il existe un entier p compris entre 3 et n , des scalaires non nuls $\lambda_3, \dots, \lambda_p$ et des formes linéaires indépendantes $\ell_3, \dots, \ell_p$ définies sur H tels que :

$$\forall x' \in H, q'(x') = \sum_{j=3}^p \lambda_j \ell_j^2(x')$$

et en prolongeant les formes linéaires $\ell_3, \dots, \ell_n$ à E , on a :

$$q(x) = \frac{1}{2}\ell_1^2(x) - \frac{1}{2}\ell_2^2(x) + \sum_{j=3}^p \lambda_j \ell_j^2(x)$$

ce qui donne une décomposition de q comme combinaison linéaire de carrés de formes linéaires.

Il reste à vérifier que les formes $\ell_1, \ell_2, \dots, \ell_p$ sont linéairement indépendantes dans E^* .

L'égalité $\sum_{j=1}^p \lambda_j \ell_j$ équivaut à dire que $\sum_{j=1}^p \lambda_j \ell_j(x) = 0$ pour tout $x \in E$. Prenant $x = e_1$ et $x = e_2$, on obtient $\lambda_1 a_{12} + \lambda_2 a_{21} = 0$ et $\lambda_1 - \lambda_2 = 0$, ce qui équivaut à $\lambda_1 = \lambda_2 = 0$ puisque $a_{21} \neq 0$ et $\sum_{j=3}^p \lambda_j \ell_j(x') = 0$ pour tout $x' \in H$, ce qui équivaut à $\sum_{j=3}^p \lambda_j \ell_j = 0$ et la nullité de tous les λ_j puisque le système $(\ell_3, \dots, \ell_p)$ est libre dans H^* . On a donc le résultat annoncé. $\square$

On peut remarquer que cette démonstration est constructive, c'est-à-dire qu'elle fournit un algorithme permettant d'obtenir une réduction en combinaison linéaire de carrés.

Une telle décomposition est appelée réduction de Gauss, ou plus simplement réduction, de la forme quadratique q .

Le théorème 11.11 nous fournit aussi une expression intéressante de la forme polaire de la forme quadratique q comme nous allons le voir avec le théorème qui suit.

Pour la suite de ce paragraphe, q désigne une forme quadratique non nulle sur E et $q = \sum_{j=1}^p \lambda_j \ell_j^2$ une réduction de Gauss de cette forme quadratique où p est un entier compris entre 1 et n , $\lambda_1, \dots, \lambda_p$ sont des scalaires non nuls et $\ell_1, \dots, \ell_p$ des formes linéaires indépendantes. On notera φ la forme polaire de q .

Théorème 11.12.

Avec les notations qui précèdent, la forme polaire φ de q est alors définie par :

$$\forall (x, y) \in E \times E, \varphi(x, y) = \sum_{j=1}^p \lambda_j \ell_j(x) \ell_j(y)$$

Preuve. Il est clair que φ est une forme bilinéaire symétrique sur E et pour tout $x \in E$, on a $\varphi(x, x) = \sum_{j=1}^p \lambda_j \ell_j^2(x) = q(x)$, ce qui signifie que φ est la forme polaire de q . □

Les formes linéaires $\ell_1, \dots, \ell_p$ étant linéairement indépendantes dans l'espace vectoriel $E^* = \mathcal{L}(E, \mathbb{K})$ des formes linéaires sur E , elles peuvent se compléter en une base de cet espace (qui on le sait est de dimension n), c'est-à-dire qu'il existe des formes linéaires $\ell_{p+1}, \dots, \ell_n$ (dans le cas où $p \leq n-1$) telles que $(\ell_1, \dots, \ell_n)$ soit une base de E^* . La réduction de Gauss du théorème 11.11 peut alors s'écrire :

$$\forall x \in E, q(x) = \sum_{j=1}^n \lambda_j \ell_j^2(x)$$

où on a posé $\lambda_{p+1} = \dots = \lambda_n = 0$ dans le cas où $p \leq n-1$.

Théorème 11.13.

Étant donnée une base $(\ell_i)_{1 \leq i \leq n}$ de l'espace vectoriel E^* des formes linéaires sur E , il existe une base $(f_i)_{1 \leq i \leq n}$ de E telle que :

$$\ell_i(f_j) = \delta_{ij} \quad (1 \leq i, j \leq n)$$

où les δ_{ij} sont les symboles de Kronecker définis par :

$$\delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

Preuve. En notant, pour $1 \leq i \leq n$ et tout $x \in E$, $\ell_i(x) = \alpha_{i1}x_1 + \dots + \alpha_{in}x_n$, l'expression de ℓ_i dans la base $\mathcal{B} = (e_i)_{1 \leq i \leq n}$, la matrice $Q = ((\alpha_{ij}))_{1 \leq i, j \leq n}$ est inversible puisque les ℓ_i forment une base de E^* (la ligne i de Q est la matrice de ℓ_i dans la base $\mathcal{B}$). En notant $F_1, \dots, F_n$ les colonnes de la matrice Q^{-1} , l'égalité $QQ^{-1} = I_n$ s'écrit $Q(F_1, \dots, F_n) = (QF_1, \dots, QF_n) = (E_1, \dots, E_n)$, où

$(E_i)_{1 \leq i \leq n}$ est la base canonique de $\mathbb{K}^n$. On a donc $QF_j = E_j$ pour tout entier j compris entre 1 et n . En remarquant que pour tout $x \in E$, on a :

$$QX = \begin{pmatrix} \alpha_{11}x_1 + \cdots + \alpha_{1n}x_n \\ \vdots \\ \alpha_{n1}x_1 + \cdots + \alpha_{nn}x_n \end{pmatrix} = \begin{pmatrix} \ell_1(x) \\ \vdots \\ \ell_n(x) \end{pmatrix}$$

et en désignant par f_j le vecteur de E de composantes F_j dans la base $\mathcal{B}$, les égalités $QF_j = E_j$ se traduisent par :

$$\begin{pmatrix} \ell_1(f_j) \\ \vdots \\ \ell_n(f_j) \end{pmatrix} = E_j = \begin{pmatrix} \delta_{1j} \\ \vdots \\ \delta_{nj} \end{pmatrix}$$

et on a bien $\ell_i(f_j) = \delta_{ij}$ pour tous i, j compris entre 1 et n . $\square$

Dans la situation du théorème précédent, on dit que $(\ell_i)_{1 \leq i \leq n}$ est la base duale de $(f_i)_{1 \leq i \leq n}$ ou que $(f_i)_{1 \leq i \leq n}$ est la base anté-duale de $(\ell_i)_{1 \leq i \leq n}$. Le théorème de réduction de Gauss peut alors se traduire matriciellement comme suit.

Théorème 11.14.

Avec les notations qui précèdent, il existe une base $(f_i)_{1 \leq i \leq n}$ de E dans laquelle la matrice de q est diagonale de la forme :

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

(les p premiers λ_i sont non nuls et les suivants sont nuls).

Preuve. Partant de la réduction de Gauss $q = \sum_{j=1}^p \lambda_j \ell_j^2$ avec $1 \leq p \leq n$ et les λ_i

tous non nuls, on complète $(\ell_1, \dots, \ell_p)$ en une base $(\ell_i)_{1 \leq i \leq n}$ de E^* et on construit une base $(f_i)_{1 \leq i \leq n}$ de E telle que $\ell_i(f_j) = \delta_{ij}$ pour tous i, j compris entre 1 et n . En posant $\lambda_{p+1} = \cdots = \lambda_n = 0$ dans le cas où $p \leq n-1$, la forme polaire φ de q est définie par $\varphi(x, y) = \sum_{k=1}^n \lambda_k \ell_k(x) \ell_k(y)$ et pour i, j compris entre 1 et n , on a :

$$\varphi(f_i, f_j) = \sum_{k=1}^n \lambda_k \ell_k(f_i) \ell_k(f_j) = \lambda_i \ell_i(f_j) = \begin{cases} \lambda_i & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

ce qui donne le résultat annoncé. $\square$

Une telle base $(f_i)_{1 \leq i \leq n}$ est dite orthogonale pour la forme quadratique q (cette définition sera précisée un peu plus loin).

Comme une matrice symétrique définit une unique forme quadratique dans la base canonique de $E = \mathbb{K}^n$, on déduit de tout ce qui précède, en utilisant la formule de changement de base pour les formes quadratiques, le corollaire qui suit.

Corollaire 11.1. Si A est une matrice symétrique d'ordre n à coefficients dans $\mathbb{K}$, il existe alors une matrice inversible P telle que la matrice ${}^tPAP$ soit diagonale.

Preuve. En gardant toujours les mêmes notations, la matrice $P = Q^{-1}$ de colonnes $f_1, \dots, f_n$ est la matrice de passage de la base canonique de $\mathbb{K}^n$ à la base $(f_i)_{1 \leq i \leq n}$ et la matrice D de q dans cette base est diagonale et s'écrit $D = {}^tPAP$. $\square$

11.6 Orthogonalité, noyau et rang

Pour ce paragraphe, φ est une forme bilinéaire symétrique sur un espace vectoriel E et q la forme quadratique associée.

Définition 11.11. On dit que deux vecteurs x, y de E sont orthogonaux relativement à φ si $\varphi(x, y) = 0$.

Exemple 11.1 Sur $\mathbb{R}^2$ ou $\mathbb{R}^3$ le produit scalaire usuel :

$$(x, y) \mapsto x \cdot y = x_1y_1 + x_2y_2 \text{ ou } (x, y) \mapsto x \cdot y = x_1y_1 + x_2y_2 + x_3y_3$$

définit une forme bilinéaire symétrique et la définition de l'orthogonalité correspond bien à celle étudiée au Lycée.

Définition 11.12. Si X est une partie non vide E , l'orthogonal de X relativement à φ est le sous-ensemble de E formé des vecteurs orthogonaux à tous les vecteurs de X .

L'orthogonal d'une partie non vide X de E est notée $X^\perp$ et on a :

$$X^\perp = \{y \in E \mid \forall x \in X, \varphi(x, y) = 0\}$$

Exemple 11.2 Pour $X = \{0\}$, on a $X^\perp = E$.

Les propriétés suivantes se déduisent immédiatement de la définition.

Théorème 11.15.

Soient X, Y deux parties non vide de E .

1. $X^\perp$ est un sous-espace vectoriel de E .
2. $X \subset (X^\perp)^\perp$.
3. Si $X \subset Y$, alors $Y^\perp \subset X^\perp$.

Comme, pour toute partie non vide X de E , $X^\perp$ est un sous-espace vectoriel de E , l'inclusion $X \subset (X^\perp)^\perp$ sera stricte pour X non sous-espace vectoriel.

Pour le produit scalaire usuel sur $E = \mathbb{R}^2$, on a $E^\perp = \{0\}$. En effet si $y \in E^\perp$, il est en particulier orthogonal à lui même, donc $y \cdot y = y_1^2 + y_2^2 = 0$ et $y_1 = y_2 = 0$, soit $y = 0$. Mais de manière général un vecteur peut être orthogonal à lui même sans être nécessairement nul. Considérons par exemple la forme bilinéaire symétrique φ définie sur $\mathbb{R}^2$ par $\varphi(x, y) = x_1 y_1 - x_2 y_2$. Un vecteur x est orthogonal à lui même si, et seulement si, $x_1^2 - x_2^2 = 0$, ce qui équivaut à $x_2 = \pm x_1$.

Définition 11.13. On dit qu'un vecteur x de E est *isotrope relativement à φ* s'il est orthogonal à lui même.

Définition 11.14. L'ensemble des vecteurs isotropes de E , relativement à φ , est le *cône isotrope* de φ .

Le cône isotrope de φ est donc le sous-ensemble de E :

$$C_\varphi = \{x \in E \mid q(x) = \varphi(x, x) = 0\}$$

On dit aussi que C_φ est le cône isotrope de la forme quadratique q et on le note alors C_q ou $q^{-1}\{0\}$.

Définition 11.15. Le noyau de φ est l'orthogonal de E .

En notant $\ker(\varphi)$ le noyau de φ , on a :

$$\ker(\varphi) = E^\perp = \{y \in E \mid \forall x \in E, \varphi(x, y) = 0\}$$

et ce noyau est un sous-espace vectoriel de E . On dit aussi que $\ker(\varphi)$ est le noyau de la forme quadratique q et on le note alors $\ker(q)$.

Lemme 11.1 *Le noyau de φ est contenu dans son cône isotrope.*

Preuve. Si $x \in \ker(\varphi)$, il est orthogonal à tout vecteur de E et en particulier à lui même, ce qui signifie qu'il est dans le cône isotrope de φ . $\square$

Dans le cas où E est de dimension finie, en désignant par A la matrice de φ dans une base $\mathcal{B}$ et u l'endomorphisme de E ayant A pour matrice dans cette base, le noyau de φ est égal au noyau de u .

Théorème 11.16.

Soient E un espace vectoriel de dimension n , $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une base de E , A la matrice de la forme bilinéaire φ dans la base $\mathcal{B}$ et u l'endomorphisme de E de matrice A dans la base $\mathcal{B}$. On a alors $\ker(\varphi) = \ker(u)$.

Preuve. Un vecteur x est dans le noyau de φ si, et seulement si, il est orthogonal à tout vecteur de E , ce qui équivaut à dire du fait de la linéarité à droite de φ que x est orthogonal à chacun des vecteurs de la base $\mathcal{B}$, soit :

$$(x \in \ker(\varphi)) \Leftrightarrow (\forall i \in \{1, 2, \dots, n\}, \varphi(x, e_i) = 0)$$

ce qui revient à dire les coordonnées $x_1, \dots, x_n$ de x dans la base $\mathcal{B}$ sont solutions du système linéaire de n équations à n inconnues :

$$\varphi \left(\sum_{j=1}^n x_j e_j, e_i \right) = \sum_{j=1}^n x_j \varphi(e_j, e_i) = \sum_{j=1}^n \varphi(e_i, e_j) x_j = 0 \quad (1 \leq i \leq n)$$

Ce système s'écrit $AX = 0$ où $A = ((\varphi(e_i, e_j)))_{1 \leq i, j \leq n}$ où A est la matrice de φ dans $\mathcal{B}$ et X le vecteur colonne formé des composantes de x dans cette base. Ce système est encore équivalent à $u(x) = 0$, où u l'endomorphisme de E de matrice A dans $\mathcal{B}$, ce qui revient à dire que $x \in \ker(u)$. $\square$

On retiendra qu'en dimension finie, le noyau de φ se calcule en résolvant le système $AX = 0$, en utilisant les notations du théorème précédent. Ce résultat peut aussi se montrer comme suit. Dire que $x \in \ker(\varphi)$ équivaut à dire que $\varphi(y, x) = 0$ pour tout $y \in E$, soit à ${}^t Y A X = 0$ pour tout $Y \in \mathbb{R}^n$ et prenant $Y = A X$, on a ${}^t (A X) A X = 0$. Mais pour $Z \in \mathbb{R}^n$, on a ${}^t Z Z = \sum_{i=1}^n z_i^2$ et ${}^t Z Z = 0$ équivaut à $Z = 0$. Donc $A X = 0$ pour $x \in \ker(\varphi)$. La réciproque est évidente.

Définition 11.16. On dit que la forme bilinéaire symétrique φ (ou de manière équivalente la forme quadratique q) est non dégénérée si son noyau est réduit à $\{0\}$.

Du théorème précédent, on déduit qu'en dimension finie, une forme bilinéaire symétrique est non dégénérée si, et seulement si, sa matrice dans une quelconque base de E est inversible, ce qui équivaut à dire que son déterminant est non nul.

Comme pour les applications linéaires, on peut définir le rang d'une forme quadratique à partir de la dimension de son noyau.

Définition 11.17. Si E est de dimension finie égale à n , le rang de φ (ou de q) est l'entier $\text{rg}(q) = n - \dim(\ker(q))$.

Du théorème précédent, on déduit qu'en dimension finie le rang d'une forme quadratique est égal à celui de sa matrice dans une quelconque base.

En dimension finie la réduction de Gauss d'une forme quadratique nous permet d'obtenir son rang et son noyau.

Pour la suite de ce paragraphe, q désigne une forme quadratique non nulle sur un espace vectoriel E de dimension n et $q = \sum_{j=1}^p \lambda_j \ell_j^2$ la réduction de Gauss de cette forme quadratique où p est un entier compris entre 1 et n , $\lambda_1, \dots, \lambda_p$ sont des scalaires non nuls et $\ell_1, \dots, \ell_p$ des formes linéaires indépendantes.

On a vu que la forme polaire de q est définie par $\varphi(x, y) = \sum_{j=1}^p \lambda_j \ell_j(x) \ell_j(y)$.

Théorème 11.17.

Avec les notations qui précèdent, on a $\text{rg}(q) = p$ et :

$$\ker(q) = \{x \in E \mid \ell_1(x) = \cdots = \ell_p(x) = 0\}$$

Preuve. À la réduction de Gauss $q = \sum_{j=1}^p \lambda_j \ell_j^2$ est associée une base $(f_i)_{1 \leq i \leq n}$ de E dans laquelle la matrice de q est diagonale de la forme :

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

où les p premiers λ_i sont non nuls et les suivants nuls (théorème 11.14). Il en résulte que $\text{rg}(q) = \text{rg}(D) = p$ et $\ker(q)$ est de dimension $n - p$. Comme les formes linéaires $\ell_1, \ell_2, \dots, \ell_p$ sont linéairement indépendantes, l'espace vectoriel $F = \{x \in E \mid \ell_1(x) = \cdots = \ell_p(x) = 0\}$ est de dimension $n - p$. De plus pour tout $x \in F$ et $y \in E$, on a $\varphi(x, y) = \sum_{j=1}^p \lambda_j \ell_j(x) \ell_j(y) = 0$, ce qui signifie que F est contenu dans le noyau de q . Ces espaces étant de même dimension, on a l'égalité $F = \ker(q)$. $\square$

Le résultat précédent nous permet de simplifier la recherche d'une base q -orthogonale $(f_i)_{1 \leq i \leq n}$ de E en se passant de compléter le système libre $(\ell_i)_{1 \leq i \leq n}$ en une base du dual de E .

Dans le cas où $p = n$, la forme q est non dégénérée et une telle base q -orthogonale se calcule en résolvant les n systèmes linéaires :

$$\ell_i(f_j) = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \quad (1 \leq i, j \leq n)$$

ce qui revient à inverser la matrice $Q = ((\alpha_{ij}))_{1 \leq i, j \leq n}$, où les α_{ij} sont définis par :

$$\ell_i(x) = \alpha_{i1}x_1 + \cdots + \alpha_{in}x_n$$

(les ℓ_i étant exprimés dans une base canonique donnée de E).

Dans le cas où $1 \leq p \leq n-1$, on détermine tout d'abord une base $(f_{p+1}, \dots, f_n)$ du noyau de q en résolvant le système linéaire de p équations à n inconnues :

$$\ell_i(x) = 0 \quad (1 \leq i \leq p)$$

Ces vecteurs sont deux à deux orthogonaux puisque orthogonaux à tout vecteur de E . Il suffit ensuite de résoudre les p systèmes linéaires :

$$\ell_i(f_j) = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases} \quad (1 \leq i, j \leq p)$$

ce qui fournit une famille q -orthogonale $(f_1, \dots, f_p)$ formée de vecteurs non nuls. Pour j fixé entre 1 et p , le système linéaire $\ell_i(f_j) = \delta_{ij}$ où i varie de 1 à p a des solutions puisque la matrice de ce système est de rang p et deux solutions de ce système diffèrent d'un élément du noyau de q . La famille $(f_i)_{1 \leq i \leq n}$ est alors une base q -orthogonale de E (exercice : vérifier qu'on a bien une base). Dans la pratique, on résout d'abord le système :

$$\begin{cases} \ell_1(x) = b_1 \\ \vdots \\ \ell_p(x) = b_p \end{cases}$$

où $b = (b_1, \dots, b_p)$ est un élément quelconque de $\mathbb{K}^p$. La valeur $b = 0$ nous donne une base du noyau de q , puis les valeurs successives $b = (1, 0, \dots, 0)$, $b = (0, 1, 0, \dots, 0)$, $\dots$, $b = (0, \dots, 0, 1)$ nous permettent de déterminer des vecteurs $f_1, \dots, f_p$.

11.7 Signature d'une forme quadratique réelle en dimension finie

Pour ce paragraphe, q est une forme quadratique non nulle a priori sur un espace vectoriel réel E de dimension finie égale à $n \geq 1$ et on note φ sa forme polaire.

Théorème 11.18.

Il existe un unique couple (s, t) d'entiers naturels tel que pour toute base $(e_i)_{1 \leq i \leq n}$ de E qui est orthogonale relativement à q , le nombre de vecteurs e_i tels que $q(e_i) > 0$ est égal à s et le nombre de vecteurs e_i tels que $q(e_i) < 0$ est égal à t . De plus, on a $s + t = \text{rg}(q)$.

Preuve. Soient $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ et $\mathcal{B}' = (e'_i)_{1 \leq i \leq n}$ deux bases q -orthogonales de E telles que :

$$\begin{cases} q(e_i) > 0 \quad (1 \leq i \leq s), \\ q(e'_i) > 0 \quad (1 \leq i \leq s') \\ q(e_i) < 0 \quad (s+1 \leq i \leq s+t), \quad q(e'_i) < 0 \quad (s'+1 \leq i \leq s'+t') \\ q(e_i) = 0 \quad (s+t+1 \leq i \leq n), \quad q(e'_i) = 0 \quad (s'+t'+1 \leq i \leq n) \end{cases}$$

où s, t, s', t' sont des entiers compris entre 0 et n avec la convention que la condition correspondante sur le signe de $q(e_i)$ ou $q(e'_i)$ n'a pas lieu quand l'encadrement de l'indice i n'a pas de sens. Considérant les matrices de q dans chacune de ces bases, on voit que nécessairement on a $s + t = s' + t' = \text{rg}(q)$. On désigne par F le sous-espace vectoriel de E engendré par $\{e_1, \dots, e_s\}$ ($F = \{0\}$ pour $s = 0$) et par G' celui engendré par $\{e'_{s'+1}, \dots, e'_n\}$ ($G' = \{0\}$ pour $s' = n$). En supposant que $s \geq 1$, on a alors :

$$\forall x \in F \setminus \{0\}, \quad q(x) = \sum_{i=1}^s \lambda_i x_i^2 > 0$$

et :

$$\forall x \in G', q(x) = \sum_{i=s'+1}^n \lambda'_i x_i^2 \leq 0$$

et en conséquence $F \cap G' = \{0\}$. Ce dernier résultat étant encore valable pour $s = 0$. On en déduit alors que :

$$\dim(F \oplus G') = \dim(F) + \dim(G') = s + n - s' \leq n$$

et $s \leq s'$. En permutant les rôles joués par s et s' , on montre de même que $s' \leq s$. On a donc $s = s'$ et $t = t'$ puisque $s + t = s' + t' = \text{rg}(q)$. $\square$

Définition 11.18. Le couple (s, t) d'entiers naturels défini par le théorème précédent est appelé signature de q et on le note $\text{sgn}(q)$.

Une forme quadratique q est donc de signature (s, t) si, et seulement si, elle admet une réduction de Gauss de la forme $q = \sum_{j=1}^s \lambda_j \ell_j^2 - \sum_{j=s+1}^{s+t} \lambda_j \ell_j^2$, où les λ_j sont tous strictement positifs (pour $s = 0$ la première somme n'existe pas et $s = n$ c'est la deuxième qui n'existe pas). En définissant les formes linéaires L_j par $L_j(x) = \ell_j(\sqrt{\lambda_j}x)$, on a la décomposition $q = \sum_{j=1}^s L_j^2 - \sum_{j=s+1}^{s+t} L_j^2$ et à cette décomposition est associée une base q -orthogonale de E dans laquelle la matrice de q est $D = \begin{pmatrix} I_s & 0 & 0 \\ 0 & -I_t & 0 \\ 0 & 0 & 0 \end{pmatrix}$, où I_r est la matrice identité d'ordre r . Les blocs diagonaux I_s , $-I_t$ ou 0 n'existent pas si $s = 0$, $s = n$ ou $s + t = n$.

Définition 11.19. On dit que la forme bilinéaire symétrique φ (ou de manière équivalente la forme quadratique q) est positive [resp. définie positive] si $q(x) \geq 0$ [resp. $q(x) > 0$] pour tout x dans E [resp. dans $E \setminus \{0\}$].

Une forme quadratique non nulle est donc positive [resp. définie positive] si, et seulement si, sa signature est $(s, 0)$ [resp. $(n, 0)$] où s est compris entre 1 et n .

On définit de manière analogue les formes quadratiques négative [resp. définie négative] et une forme quadratique non nulle est négative [resp. définie négative] si, et seulement si, sa signature est $(0, t)$ [resp. $(0, n)$] avec t compris entre 1 et n .

À partir d'une réduction de Gauss, $q = \sum_{j=1}^p \lambda_j \ell_j^2$, on déduit que q est positive [resp. définie positive] si, et seulement si, tous les λ_j sont strictement positifs [resp. $p = n$ et tous les λ_j sont strictement positifs]. En effet, la condition suffisante est évidente et pour la condition nécessaire, en supposant $\lambda_1 < 0$ (on peut toujours s'y ramener) et en désignant par $(e_i)_{1 \leq i \leq n}$ une base q -orthogonale de E déduite de cette réduction de Gauss, on a $q(e_1) = \lambda_1 < 0$ et q n'est pas positive.

Une définition équivalente de la signature qu'une forme quadratique est donnée par le théorème qui suit. La démonstration de ce théorème nécessite les lemmes suivants.

Lemme 11.2 *Soit F un sous-espace vectoriel de E . La restriction de q à F est non dégénérée si, et seulement si, $F \cap F^\perp = \{0\}$.*

Preuve. Dire que la restriction de q à F est non dégénérée équivaut à dire que :

$$\{x \in F \mid \forall y \in F, \varphi(x, y) = 0\} = \{0\}$$

et ce ensemble est justement $F \cap F^\perp$ (c'est aussi le noyau la restriction de q à F).
□

Lemme 11.3 *Soit F un sous-espace vectoriel de E . Si la restriction de q à F est non dégénérée on a alors $E = F \oplus F^\perp$.*

Preuve. Laissée au lecteur. □

Théorème 11.19.

En désignant par $\mathcal{P}$ [resp. $\mathcal{N}$] l'ensemble de tous les sous-espaces vectoriels F de E tels que la restriction de q à F soit définie positive [resp. définie négative] ($\mathcal{P}$ ou $\mathcal{N}$ peut être vide), la signature (s, t) de q est donnée par :

$$s = \begin{cases} 0 & \text{si } \mathcal{P} = \emptyset \\ \max_{F \in \mathcal{P}} \dim(F) & \text{si } \mathcal{P} \neq \emptyset \end{cases} \quad \text{et } t = \begin{cases} 0 & \text{si } \mathcal{N} = \emptyset \\ \max_{F \in \mathcal{N}} \dim(F) & \text{si } \mathcal{N} \neq \emptyset \end{cases}$$

Preuve. Notons :

$$s' = \begin{cases} 0 & \text{si } \mathcal{P} = \emptyset \\ \max_{F \in \mathcal{P}} \dim(F) & \text{si } \mathcal{P} \neq \emptyset \end{cases} \quad \text{et } t' = \begin{cases} 0 & \text{si } \mathcal{N} = \emptyset \\ \max_{F \in \mathcal{N}} \dim(F) & \text{si } \mathcal{N} \neq \emptyset \end{cases}$$

Par définition de la signature, on a $s \leq s'$ et $t \leq t'$. Si $\mathcal{P} = \emptyset$, on a alors $s = s' = 0$. Si $\mathcal{P} \neq \emptyset$, on peut trouver $F \in \mathcal{P}$ tel que $\dim(F) = s'$ et on a nécessairement $\dim(F) \leq s$. En effet si $\dim(F) > s$, on désigne par $(e_1, \dots, e_{s'})$ une base q -orthogonale de F et on peut compléter cette base en une base $(e_1, \dots, e_n)$ de E qui est aussi q -orthogonale puisque la restriction de q à F est non dégénérée (elle est définie positive) et $E = F \oplus F^\perp$. Comme $F \in \mathcal{P}$ est de dimension maximale, la restriction de q à $F^\perp$ est négative et la signature de q est (s', t') avec $s' > s$, ce qui n'est pas possible. On a donc $s' \leq s$ et $s = s'$. On montre de manière analogue que $t = t'$. □

Si q est définie positive, on a donc une réduction de Gauss $q = \sum_{j=1}^n \lambda_j \ell_j^2$ où tous les λ_j sont strictement positifs et la matrice de q dans une base q -orthogonale adaptée à cette réduction est diagonale de termes diagonaux $\lambda_1, \dots, \lambda_n$. En notant D cette matrice, on a $\det(D) = \prod_{k=1}^n \lambda_k > 0$. La matrice de q dans une autre base de $\mathbb{R}^n$ s'écrivant $A = {}^t P D P$ avec P inversible, on a $\det(A) = (\det(P))^2 \det(D) > 0$.

L'utilisation des mineurs principaux de la matrice de q dans une quelconque base de $\mathbb{R}^n$ nous permet de savoir si une forme quadratique est définie positive ou non. On rappelle que si $A = ((a_{ij}))_{1 \leq i, j \leq n}$ est une matrice carrée d'ordre n , les mineurs principaux de A sont les déterminants des matrices extraites $A_k = ((a_{ij}))_{1 \leq i, j \leq k}$ où k est un entier compris entre 1 et n .

Théorème 11.20.

Soit q une forme quadratique non nulle sur un espace vectoriel réel E de dimension n de matrice $A = ((a_{ij}))_{1 \leq i, j \leq n}$ dans une base $(e_i)_{1 \leq i \leq n}$. La forme q est définie positive si, et seulement si, tous les mineurs principaux de A sont strictement positifs.

Preuve. Supposons q définie positive sur E . Pour k compris entre 1 et n , la matrice $A_k = ((a_{ij}))_{1 \leq i, j \leq k}$ est la matrice de la forme quadratique q_k égale à la restriction de q au sous-espace vectoriel E_k de E engendré par les vecteurs $e_1, \dots, e_k$. Cette forme q_k étant définie positive comme q , il en résulte que $\det(A_k) > 0$. Pour la réciproque, on raisonne par récurrence sur la dimension $n \geq 1$ de E . Pour $n = 1$, le résultat est évident puisque $E = \mathbb{R}e_1$ est une droite vectoriel et q s'écrit $q(x) = q(x_1 e_1) = \lambda x_1^2$ avec $\lambda = q(e_1) = \det(A)$. Supposons le résultat acquis pour tous les espaces de dimension au plus égal à n et soit q une forme quadratique sur un espace E de dimension $n+1$. On se donne une base $(e_i)_{1 \leq i \leq n+1}$ de E et on suppose que tous les mineurs principaux de la matrice $A = ((a_{ij}))_{1 \leq i, j \leq n+1}$ de q dans cette base sont strictement positifs. En désignant par H le sous-espace vectoriel de E engendré par les vecteurs $e_1, \dots, e_n$, la matrice extraite $A_n = ((a_{ij}))_{1 \leq i, j \leq n}$ est la matrice de la forme quadratique q_n égale à la restriction de q à H . Tous les mineurs principaux de A_n étant strictement positifs, cette forme q_1 est définie positive sur H . La restriction de q à H étant définie positive et q non dégénérée ($\det(A) \neq 0$), la signature de q ne peut être que $(n, 1)$ ou $(n+1, 0)$ (par définition de la signature). Si cette signature est $(n, 1)$, cela signifie qu'on a une décomposition de Gauss de la forme $q = \sum_{j=1}^n \lambda_j \ell_j^2 - \lambda_{n+1} \ell_n^2$ où tous les λ_j sont strictement positifs et la matrice de q dans une base q -orthogonale adaptée à cette réduction est diagonale de termes diagonaux $\lambda_1, \dots, \lambda_n, -\lambda_{n+1}$. En notant D cette matrice, on a $\det(D) = -\lambda_{n+1} \prod_{k=1}^n \lambda_k < 0$, ce qui contredit $\det(D) = (\det(P))^2 \det(A) > 0$. La signature de q est donc $(n+1, 0)$ et q est définie positive. $\square$

11.8 Quadriques dans $\mathbb{R}^n$ ou $\mathbb{C}^n$

Pour $n \geq 2$, on munit l'espace vectoriel $\mathbb{K}^n$ de sa base canonique et les coordonnées d'un vecteur X de $\mathbb{K}^n$ sont notées $x_1, \dots, x_n$.

Pour $n = 2$ [resp. $n = 3$] et $\mathbb{K} = \mathbb{R}$ les coordonnées seront notées x, y [resp. x, y, z].

Définition 11.20. On appelle quadrique dans $\mathbb{K}^n$ toute partie $\mathcal{C}$ de $\mathbb{K}^n$ définie par :

$$\mathcal{C} = \left\{ X \in \mathbb{K}^n \mid \sum_{i=1}^n a_{ii}x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij}x_i x_j + \sum_{i=1}^n b_i x_i + c = 0 \right\}$$

où c , les a_{ij} et les b_i sont des scalaires avec $(a_{11}, \dots, a_{nn}) \neq 0$.

On dit aussi que $\mathcal{C}$ est la courbe d'équation :

$$P(X) = \sum_{i=1}^n a_{ii}x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij}x_i x_j + \sum_{i=1}^n b_i x_i + c = 0$$

dans la base canonique. On notera aussi $\mathcal{C} = P^{-1}\{0\}$ où P est une fonction polynomiale de degré 2 sur $\mathbb{K}^n$. Une telle courbe peut être vide comme le montre l'exemple de $P(X) = x^2 + y^2 + 1$ dans $\mathbb{R}^2$.

Exemples 11.3 Pour $P(X) = x^2 + y^2$ dans $\mathbb{R}^2$, $\mathcal{C}$ est réduit à $\{0\}$. Pour $P(X) = x^2 + y^2 - 1$ dans $\mathbb{R}^2$, $\mathcal{C}$ est le cercle de centre 0 et de rayon 1. Pour $P(X) = x^2$ dans $\mathbb{K}^2$, $\mathcal{C}$ est la droite d'équation $x = 0$. Pour $P(X) = xy$ dans $\mathbb{K}^2$, $\mathcal{C}$ est la réunion des droites d'équations respectives $x = 0$ et $y = 0$.

On peut remarquer que le polynôme P s'écrit $P = q + \ell + c$, où q est une forme quadratique, ℓ une forme linéaire et c une constante.

En désignant par A et L les matrices de q et ℓ dans la base canonique de $\mathbb{K}^n$, on a $P(X) = {}^t X A X + L X + c$.

Pour tout X_0 dans $\mathbb{K}^n$, on a pour tout X dans $\mathbb{K}^n$, en désignant par φ la forme polaire de q :

$$\begin{aligned} P(X + X_0) &= q(X + X_0) + \ell(X + X_0) + c \\ &= q(X) + 2\varphi(X, X_0) + \ell(X) + \ell(X_0) + q(X_0) + c \end{aligned}$$

l'application $X \mapsto \varphi(X, X_0)$ étant une forme linéaire sur $\mathbb{K}^n$.

Définition 11.21. On dit que la quadrique $\mathcal{C} = P^{-1}\{0\}$ est à centre s'il existe un unique élément X_0 dans $\mathbb{K}^n$ tel que $P(X + X_0) = q(X) + d$ pour tout X dans $\mathbb{K}^n$, où d est une constante.

Théorème 11.21.

La quadrique $\mathcal{C} = P^{-1}\{0\}$ est à centre si, et seulement si, la forme quadratique q est non dégénérée.

Preuve. Dire que $\mathcal{C}$ est à centre revient à dire qu'il existe un unique X_0 dans $\mathbb{K}^n$ tel que pour tout X dans $\mathbb{K}^n$, on ait $2\varphi(X, X_0) + \ell(X) = 0$, soit $2 {}^t X_0 A X + L X = 0$, c'est-à-dire :

$$\forall X \in \mathbb{K}^n, (2 {}^t X_0 A + L) X = 0$$

ce qui équivaut à $2 {}^tX_0A + L = 0$ ou à $2 {}^tAX_0 = 2AX_0 = - {}^tL$ (la matrice A de q est symétrique) X_0 étant unique. En définitive, $\mathcal{C}$ est à centre si, et seulement si, l'équation $2AX_0 = - {}^tL$ a une unique solution, ce qui équivaut à dire que A est inversible. En effet, pour A inversible, la solution est $X_0 = -\frac{1}{2}A^{-1} {}^tL$ et pour A non inversible, l'ensemble des solutions de ce système est soit vide soit infini puisque pour toute solution X_0 , l'ensemble $X_0 + \ker(A)$, avec $\dim(\ker(A)) \geq 1$, nous donne une infinité de solutions. Et A inversible signifie que q non dégénérée. $\square$

Si la quadrique $\mathcal{C}$ est à centre de centre X_0 , en effectuant le changement de variable $X' = X - X_0$ (on ramène l'origine en X_0), on a :

$$\begin{aligned} P(X) &= P((X - X_0) + X_0) = q(X - X_0) + d = q(X') + d \\ &= \sum_{i=1}^n a_{ii} (x'_i)^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x'_i x'_j + d \end{aligned}$$

Sachant que pour tout $X' \in \mathbb{R}^n$ on a $q(X') = q(-X')$, on déduit que :

$$\begin{aligned} (X = X_0 + X' \in \mathcal{C}) &\Leftrightarrow (P(X) = P(X' + X_0) = q(X') + d = 0) \\ &\Leftrightarrow (q(-X') + d = P(-X' + X_0) = 0) \Leftrightarrow (X_0 - X' \in \mathcal{C}) \end{aligned}$$

ce qui se traduit en disant que le centre X_0 est un centre de symétrie pour $\mathcal{C}$.

Le système linéaire permettant de déterminer le centre, quand il est unique, est donné par :

$$2 \sum_{j=1}^n a_{ij} x_j + b_i = 0 \quad (1 \leq i \leq n)$$

Dans $\mathbb{R}^n$, on a pour $1 \leq k \leq n$:

$$\begin{aligned} \frac{\partial}{\partial x_k} P(X) &= \frac{\partial}{\partial x_k} \left(\sum_{1 \leq i, j \leq n} a_{ij} x_i x_j + \sum_{i=1}^n b_i x_i + c \right) \\ &= \frac{\partial}{\partial x_k} \left(\sum_{i=1}^n x_i \sum_{j=1}^n a_{ij} x_j + \sum_{i=1}^n b_i x_i + c \right) = \sum_{j=1}^n a_{kj} x_j + \sum_{i=1}^n x_i a_{ik} + b_k \\ &= \sum_{j=1}^n a_{kj} x_j + \sum_{i=1}^n a_{ki} x_i + b_k = 2 \sum_{j=1}^n a_{kj} x_j + b_k \end{aligned}$$

et notre système linéaire s'écrit simplement :

$$\frac{\partial}{\partial x_k} P(X) = 0 \quad (1 \leq k \leq n)$$

Exemples 11.4

1. Dans $\mathbb{R}^2$ la quadrique d'équation $y - x^2 = 0$ n'est pas à centre puisque la forme quadratique $q : (x, y) \mapsto -x^2$ est dégénérée. Cette quadrique du plan $\mathbb{R}^2$ est une parabole.

2. Soit dans $\mathbb{R}^2$ la quadrique d'équation $x^2 + y^2 - 2xy - 8(x + y) + 16 = 0$. La forme quadratique q de matrice $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ est dégénérée ($\det(A) = 0$) et donc n'est pas à centre. En effectuant le changement de variable $x' = x - y$, $y' = x + y$, cette équation s'écrit $(x')^2 - 8y' + 16 = 0$, soit $y' = \frac{1}{8}(x')^2 - 16$ et $\mathcal{C}$ est une parabole.

En définitive, si $\mathcal{C}$ est une quadrique à centre, en plaçant l'origine au centre, cette conique à une équation de la forme $q(X) = \alpha$, où q est une forme quadratique non dégénérée et α une constante.

Le théorème de réduction de Gauss nous permet d'écrire q comme combinaison linéaire de n carrés de formes linéaires, ce qui revient à dire qu'il existe une base de $\mathbb{K}^n$ dans laquelle l'expression de q est $q(X) = \sum_{i=1}^n \lambda_i y_i^2$, les scalaires λ_i étant non

nuls et dans cette base (orthogonale pour q), une équation de $\mathcal{C}$ est $\sum_{i=1}^n \lambda_i y_i^2 = \alpha$.

Dans le cas des quadratiques à centre réelles, en désignant par (s, t) la signature de q avec $s + t = n$, il existe une base de $\mathbb{R}^n$ dans laquelle une équation de $\mathcal{C}$ est $\sum_{i=1}^s y_i^2 - \sum_{i=s+1}^n y_i^2 = \alpha$ avec la convention que $\sum_{i=1}^s = 0$ pour $s = 0$ et $\sum_{i=s+1}^n = 0$ pour $t = 0$.

En particulier dans le plan $\mathbb{R}^2$, on a les possibilités suivantes en désignant par x, y les coordonnées de X dans une base q -orthogonale, l'origine étant ramenée au centre de la quadrique :

- $x^2 + y^2 = \pm\alpha = \beta$ pour q de signature $(2, 0)$ ou $(0, 2)$ et $\mathcal{C}$ est vide pour $\beta < 0$, réduite à $\{(0, 0)\}$ pour $\beta = 0$ ou une ellipse pour $\beta > 0$;
- $x^2 - y^2 = \alpha$ pour q de signature $(1, 1)$ et $\mathcal{C}$ est une hyperbole.

Exemple 11.3 Considérons dans $\mathbb{R}^2$ la quadrique d'équation :

$$x^2 + y^2 + 4xy + 4(x + y) - 8 = 0$$

La forme quadratique q est non dégénérée puisque sa matrice dans la base canonique est $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ de déterminant $\det(A) = -3 \neq 0$. Son centre est solution du système linéaire :

$$\begin{cases} \frac{\partial P}{\partial x} P(X) = 2x + 4y + 4 = 0 \\ \frac{\partial P}{\partial y} P(X) = 2y + 4x + 4 = 0 \end{cases}$$

soit $x + 2y = -2$ et $2x + y = -2$, ce qui donne $X_0 = -\frac{2}{3}(1, 1)$. Le changement de variables $X' = X - X_0$, soit $x' = x + \frac{2}{3}$, $y' = y + \frac{2}{3}$ donne :

$$\left(x' - \frac{2}{3}\right)^2 + \left(y' - \frac{2}{3}\right)^2 + 4\left(x' - \frac{2}{3}\right)\left(y' - \frac{2}{3}\right) + 4\left(x' - \frac{2}{3} + y' - \frac{2}{3}\right) - 8 = 0$$

soit $(x')^2 + (y')^2 + 4x'y' = \frac{32}{3}$ comme prévu. La réduction de Gauss nous donne $(x' + 2y')^2 - 3(y')^2 = \frac{32}{3}$ et $\mathcal{C}$ est une hyperbole.

En restant dans $\mathbb{R}^2$, une quadrique $\mathcal{C}$ a une équation de la forme :

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0$$

Si la forme quadratique q est dégénérée et non nulle, elle est de rang 1, ce qui équivaut à dire que la matrice $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ est de rang 1 et il existe un réel non nul λ tel que $\begin{pmatrix} b \\ c \end{pmatrix} = \lambda \begin{pmatrix} a \\ b \end{pmatrix}$ et l'équation de $\mathcal{C}$ devient :

$$a(x^2 + 2\lambda xy + \lambda^2 y^2) + dx + ey + f = 0$$

soit $a(x + \lambda y)^2 + dx + ey + f = 0$. Pour $a = 0$, on a l'équation $dx + ey + f = 0$ qui définit une droite si $(e, d) \neq (0, 0)$, l'ensemble vide si $(e, d) = (0, 0)$ et $f \neq 0$ ou $\mathbb{R}^2$ tout entier si $(e, d) = (0, 0)$ et $f = 0$. Pour $a \neq 0$, on distingue deux cas de figure. Soit $e = \lambda d$ et notre équation devient $(x + \lambda y)^2 + \frac{d}{a}(x + \lambda y) + \frac{f}{a} = 0$, soit :

$$\left(x + \lambda y + \frac{d}{2a}\right)^2 = \frac{d^2 - 4af}{4a^2}$$

ce qui définit la réunion de deux droites si $d^2 - 4af > 0$ (les droites d'équations $x + \lambda y = \frac{\sqrt{d^2 - 4af} - d}{2a}$ et $x + \lambda y = -\frac{\sqrt{d^2 - 4af} + d}{2a}$), une droite si $d^2 - 4af = 0$

(la droite d'équation $x + \lambda y = -\frac{d}{2a}$) ou l'ensemble vide si $d^2 - 4af < 0$. Soit $e \neq \lambda d$ et le changement de variable $x' = x + \lambda y$, $y' = dx + ey$ nous donne l'équation $a(x')^2 + y' + f = 0$, ce qui définit une parabole (le changement de variable est validé par $\begin{vmatrix} 1 & \lambda \\ d & e \end{vmatrix} = e - \lambda d \neq 0$). En définitive, dans une base adaptée, une quadrique de $\mathbb{R}^2$ a une équation de l'une des formes suivantes : $x^2 + y^2 = \beta$; $x^2 - y^2 = \alpha$; $dx + ey + f = 0$; $x^2 = \alpha$; $ax^2 + y + f = 0$.

11.9 Exercices

Exercice 11.1. Soient ℓ_1 et ℓ_2 deux formes linéaires sur E . Montrer que l'application $\ell_1 \ell_2$ est une forme linéaire sur E si, et seulement si, l'une de ces deux formes est l'application nulle.

Solution. Il est clair que si l'une de ces deux formes est l'application nulle, alors $\ell_1 \ell_2$ est une forme linéaire sur E . Réciproquement supposons que $\ell_1 \ell_2$ soit linéaire.

On a alors pour tout scalaire λ et tous vecteurs x, y dans E :

$$\begin{aligned} \ell_1(x) \ell_2(x) + \lambda \ell_1(y) \ell_2(y) &= (\ell_1 \ell_2)(x) + \lambda (\ell_1 \ell_2)(y) \\ &= (\ell_1 \ell_2)(x + \lambda y) = \ell_1(x + \lambda y) \ell_2(x + \lambda y) \\ &= (\ell_1(x) + \lambda \ell_1(y)) (\ell_2(x) + \lambda \ell_2(y)) \\ &= \ell_1(x) \ell_2(x) + \lambda (\ell_1(x) \ell_2(y) + \ell_1(y) \ell_2(x)) + \lambda^2 \ell_1(y) \ell_2(y) \end{aligned}$$

et le polynôme $\ell_1(y) \ell_2(y) \lambda^2 + (\ell_1(x) \ell_2(y) + \ell_1(y) \ell_2(x) - \ell_1(y) \ell_2(y)) \lambda$ est identiquement nul, ce qui équivaut à :

$$\ell_1(y) \ell_2(y) = 0 \text{ et } \ell_1(x) \ell_2(y) + \ell_1(y) \ell_2(x) - \ell_1(y) \ell_2(y) = 0$$

ou encore à $\ell_1(y) \ell_2(y) = 0$ et $\ell_1(x) \ell_2(y) + \ell_1(y) \ell_2(x) = 0$ pour tous x, y dans E . Si $\ell_1 \neq 0$, il existe alors $y \in E$ tel que $\ell_1(y) \neq 0$, donc $\ell_2(y) = 0$ et $\ell_1(y) \ell_2(x) = 0$ pour tout $x \in E$, ce qui équivaut à $\ell_2 = 0$.

Exercice 11.2. Montrer que les formes linéaires $(\ell_j)_{1 \leq j \leq 3}$ définies sur $\mathbb{K}^5$ par :

$$\begin{cases} \ell_1(x) = x_1 + x_2 + x_3 + x_4 + x_5 \\ \ell_2(x) = 3x_1 - 2x_3 + 2x_4 + x_5 \\ \ell_3(x) = 3x_2 + x_3 + 3x_4 \end{cases}$$

sont linéairement indépendantes.

Solution. Il s'agit de vérifier que la matrice :

$$A = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 3 & 0 & -2 & 2 & 1 \\ 0 & 3 & 1 & 3 & 0 \end{pmatrix}$$

est de rang 3, ce qui résulte de $\begin{vmatrix} 1 & 1 & 1 \\ 3 & 0 & -2 \\ 0 & 3 & 1 \end{vmatrix} = 12 \neq 0$.

Exercice 11.3. Montrer que toute forme bilinéaire φ sur E s'écrit de manière unique comme somme d'une forme bilinéaire symétrique et d'une forme bilinéaire alternée.

Solution. Soit φ une forme bilinéaire sur E . Les applications φ_1 et φ_2 définies sur E^2 par :

$$\begin{cases} \varphi_1(x, y) = \frac{1}{2} (\varphi(x, y) + \varphi(y, x)) \\ \varphi_2(x, y) = \frac{1}{2} (\varphi(x, y) - \varphi(y, x)) \end{cases}$$

sont bilinéaires, la forme φ_1 étant symétrique et φ_2 étant alternée. Et on a bien $\varphi = \varphi_1 + \varphi_2$. Réciproquement si $\varphi = \varphi_1 + \varphi_2$ avec φ_1 bilinéaire symétrique et φ_2

bilinéaire alternée, on a alors :

$$\begin{cases} \varphi(x, y) = \varphi_1(x, y) + \varphi_2(x, y) \\ \varphi(y, x) = \varphi_1(y, x) + \varphi_2(y, x) = \varphi_1(x, y) - \varphi_2(x, y) \end{cases}$$

et $\varphi(x, y) + \varphi(y, x) = 2\varphi_1(x, y)$, $\varphi(x, y) - \varphi(y, x) = \varphi_2(x, y)$, ce qui prouve l'unicité de φ_1 et φ_2 .

Exercice 11.4. Montrer que chacune des applications φ qui suivent est bilinéaire et calculer sa matrice dans la base canonique de $\mathbb{R}^n$.

1. $n = 2$, $\varphi(x, y) = x_1y_1 - 2x_2y_2$.
2. $n = 3$, $\varphi(x, y) = x_1y_1 - x_1y_2 - x_2y_2 - 2x_2y_3 - x_3y_1 - 2x_3y_3$.
3. $n = 3$, $\varphi(x, y) = 2x_1y_1 - 3x_2y_2 - x_3y_3$.

Solution. La bilinéarité de chacune de ces applications est évidente. Les matrices respectives dans les bases canoniques sont :

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & -2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & -1 & 0 \\ 0 & -1 & -2 \\ -1 & 0 & -2 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

La première et la troisième sont symétriques, mais pas la deuxième.

Exercice 11.5. Soit $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}$. Déterminer la forme bilinéaire φ sur $\mathbb{R}^3$ de matrice A dans la base canonique.

Solution. On a :

$$\begin{aligned} \varphi(x, y) &= {}^t X A Y = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \\ &= x_1(y_1 + 2y_2 + 3y_3) + x_2(2y_1 + 3y_2 + 4y_3) + x_3(3y_1 + 4y_2 + 5y_3) \end{aligned}$$

Exercice 11.6. Soient E, F deux espaces vectoriels, u une application linéaire de E dans F et φ une forme bilinéaire sur F .

1. Montrer que l'application $\psi : (x, y) \in E^2 \mapsto \varphi(u(x), u(y))$ est bilinéaire.
2. En supposant E et F de dimension finie et en désignant par $\mathcal{B}_1$ une base de E , $\mathcal{B}_2$ une base de F , A la matrice de u dans les bases $\mathcal{B}_1$ et $\mathcal{B}_2$ et par B la matrice de φ dans la base $\mathcal{B}_2$, déterminer la matrice de ψ dans la base $\mathcal{B}_1$.

3. On suppose ici que E est de dimension $n \geq 1$ et que φ est une forme bilinéaire sur E . On appelle matrice de Gram d'une famille $(x_i)_{1 \leq i \leq n}$ de vecteurs de E , la matrice $G(x_1, \dots, x_n) = ((\varphi(x_i, x_j)))_{1 \leq i, j \leq n}$ et le déterminant de cette matrice, noté $g(x_1, \dots, x_n)$, est appelé déterminant de Gram de la famille $(x_i)_{1 \leq i \leq n}$.

(a) En désignant par $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une base de E , montrer que :

$$g(x_1, \dots, x_n) = (\det_{\mathcal{B}}(x_1, \dots, x_n))^2 \Delta_{\mathcal{B}}(\varphi)$$

(b) Montrer que pour tout endomorphisme u de E , on a :

$$g(u(x_1), \dots, u(x_n)) = (\det(u))^2 g(x_1, \dots, x_n)$$

Solution.

1. Pour x [resp. y] fixé dans E , l'application $y \mapsto \varphi(u(x), u(y))$ [resp. $x \mapsto \varphi(u(x), u(y))$] est linéaire comme composée de deux applications linéaires. L'application ψ est donc bilinéaire sur E .
2. On note, pour tout vecteur x de E , X le vecteur colonne formé des composantes de x dans la base $\mathcal{B}_1$. Pour x, y dans E , on a :

$$\varphi(u(x), u(y)) = {}^t(AX)B(AY) = {}^tX({}^tABA)Y$$

et en conséquence ${}^tABA$ est la matrice de ψ dans la base $\mathcal{B}_1$.

3.

- (a) Soient $u \in \mathcal{L}(E)$ défini par $u(e_i) = x_i$ pour tout i compris entre 1 et n et ψ la forme bilinéaire définie sur E^2 par $\psi(x, y) = \varphi(u(x), u(y))$. On a :

$$\begin{aligned} \Delta_{\mathcal{B}}(\psi) &= \det \left(((\psi(e_i, e_j)))_{1 \leq i, j \leq n} \right) = \det \left(((\varphi(u(e_i), u(e_j))))_{1 \leq i, j \leq n} \right) \\ &= \det \left(((\varphi(x_i, x_j)))_{1 \leq i, j \leq n} \right) = g(x_1, \dots, x_n) \end{aligned}$$

et avec la question précédente, on a aussi :

$$\Delta_{\mathcal{B}}(\psi) = \det({}^tABA) = (\det(A))^2 \det(B) = (\det_{\mathcal{B}}(x_1, \dots, x_n))^2 \Delta_{\mathcal{B}}(\varphi)$$

- (b) On a $g(u(x_1), \dots, u(x_n)) = (\det_{\mathcal{B}}(u(x_1), \dots, u(x_n)))^2 \Delta_{\mathcal{B}}(\varphi)$ avec :

$$\det_{\mathcal{B}}(u(x_1), \dots, u(x_n)) = \det(u) \det_{\mathcal{B}}(x_1, \dots, x_n)$$

ce qui donne :

$$\begin{aligned} g(u(x_1), \dots, u(x_n)) &= (\det(u))^2 (\det_{\mathcal{B}}(x_1, \dots, x_n))^2 \Delta_{\mathcal{B}}(\varphi) \\ &= (\det(u))^2 g(x_1, \dots, x_n) \end{aligned}$$

Exercice 11.7. Déterminer la matrice et la forme polaire de la forme quadratique q définie dans la base canonique de $\mathbb{K}^3$ par :

$$q(x) = x_1^2 + 3x_2^2 + 5x_3^2 + 4x_1x_2 + 6x_1x_3 + 2x_2x_3$$

Solution. La matrice de q dans la base canonique est $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 5 \end{pmatrix}$ et sa forme polaire est définie par :

$$\begin{aligned} \varphi(x, y) &= {}^tXAY = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 5 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \\ &= x_1y_1 + 3x_2y_2 + 5x_3y_3 + 2(x_1y_2 + x_2y_1) + 3(x_1y_3 + y_1x_3) + x_2y_3 + x_3y_2 \end{aligned}$$

Exercice 11.8. Soient ℓ_1, ℓ_2 deux formes linéaires indépendantes sur E .

1. Montrer que l'application q définie sur E par $q(x) = \ell_1(x)\ell_2(x)$ est une forme quadratique et déterminer sa forme polaire.
2. Montrer que q peut s'écrire comme différence de deux carrés de formes linéaires indépendantes.
3. On suppose que $E = \mathbb{K}^n$. Donner la matrice de q dans la base canonique de E .

Solution.

1. L'application φ définie sur E^2 par $\varphi(x, y) = \frac{1}{2}\ell_1(x)\ell_2(y) + \frac{1}{2}\ell_1(y)\ell_2(x)$ est bilinéaire symétrique et $q(x) = \varphi(x, x)$ pour tout $x \in E$. Donc q est une forme quadratique de forme polaire φ .
2. On a $q = \frac{1}{4}((\ell_1 + \ell_2)^2 - (\ell_1 - \ell_2)^2)$, les formes linéaires $\ell'_1 = \ell_1 + \ell_2$ et $\ell'_2 = \ell_1 - \ell_2$ étant indépendantes puisque ℓ_1, ℓ_2 le sont. En effet si $\alpha\ell'_1 + \beta\ell'_2 = 0$, on a alors $(\alpha + \beta)\ell_1 + (\alpha - \beta)\ell_2 = 0$, donc $\alpha + \beta = \alpha - \beta = 0$ et $\alpha = \beta = 0$.
3. Notons $\ell_1(x) = \sum_{j=1}^n \alpha_j x_j$ et $\ell_2(x) = \sum_{j=1}^n \beta_j x_j$ dans la base canonique de $E = \mathbb{K}^n$.

On a alors :

$$\begin{aligned} q(x) &= \left(\sum_{j=1}^n \alpha_j x_j \right) \left(\sum_{j=1}^n \beta_j x_j \right) \\ &= \sum_{1 \leq i, j \leq n} \alpha_i \beta_j x_i x_j = \sum_{i=1}^n \alpha_i \beta_i x_i^2 + \sum_{1 \leq i < j \leq n} (\alpha_i \beta_j + \alpha_j \beta_i) x_i x_j \end{aligned}$$

ce qui signifie que la matrice de q est :

$$A = \frac{1}{2}((\alpha_i \beta_j + \alpha_j \beta_i))_{1 \leq i, j \leq n} = \frac{1}{2}({}^tL_1L_2 + {}^tL_2L_1)$$

où $L_1 = \begin{pmatrix} \alpha_1 & \cdots & \alpha_n \end{pmatrix}$ et $L_2 = \begin{pmatrix} \beta_1 & \cdots & \beta_n \end{pmatrix}$ sont les matrices de ℓ_1, ℓ_2 dans la base canonique de $\mathbb{K}^n$. En remarquant que ${}^t(\alpha) = (\alpha)$ pour α réel ou complexe, on peut aussi écrire que :

$$\begin{aligned} \varphi(x, y) &= \frac{1}{2} (\ell_1(x) \ell_2(y) + \ell_1(y) \ell_2(x)) = \frac{1}{2} ((L_1 X)(L_2 Y) + (L_1 Y)(L_2 X)) \\ &= \frac{1}{2} ({}^t(L_1 X)(L_2 Y) + {}^t(L_2 X)(L_1 Y)) \\ &= \frac{1}{2} (({}^t X {}^t L_1)(L_2 Y) + ({}^t X {}^t L_2)(L_1 Y)) \\ &= \frac{1}{2} ({}^t X ({}^t L_1 L_2 + {}^t L_2 L_1) Y) \end{aligned}$$

et la matrice A de φ , ou de q , est $A = \frac{1}{2} ({}^t L_1 L_2 + {}^t L_2 L_1)$. Ou encore revenir à la définition de la matrice $A = ((a_{ij}))_{1 \leq i, j \leq n}$ de q dans la base canonique $(e_i)_{1 \leq i \leq n}$:

$$a_{ij} = \varphi(e_i, e_j) = \frac{1}{2} (\ell_1(e_i) \ell_2(e_j) + \ell_1(e_j) \ell_2(e_i)) = \frac{1}{2} (\alpha_i \beta_j + \alpha_j \beta_i)$$

Exercice 11.9. Soit L une matrice ligne à n colonnes. Montrer que la matrice $A = {}^t L L$ est une matrice carrée symétrique et que la forme quadratique q de matrice A dans la base canonique de $\mathbb{K}^n$ est le carré d'une forme linéaire.

Solution. Si $L = \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{pmatrix}$, on a alors :

$$A = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{pmatrix} = ((\alpha_i \alpha_j))_{1 \leq i, j \leq n}$$

ce qui définit bien une matrice symétrique d'ordre n . La forme quadratique q de matrice A est alors définie par :

$$q(x) = {}^t X A X = {}^t X ({}^t L L) X = {}^t (L X) (L X) = (L X)^2$$

avec $LX = \ell(x)$ où ℓ est la forme linéaire de matrice L dans la base canonique de $\mathbb{K}^n$. On a donc $q = \ell^2$.

Exercice 11.10. Soient p un entier naturel non nul, $\ell_1, \dots, \ell_p$ des formes linéaires sur E et $\lambda_1, \dots, \lambda_p$ des scalaires. Montrer que l'application q définie sur E par $q(x) = \sum_{j=1}^p \alpha_j \ell_j^2(x)$ est une forme quadratique et déterminer sa forme polaire.

Solution. L'application φ définie sur E^2 par $\varphi(x, y) = \sum_{j=1}^p \alpha_j \ell_j(x) \ell_j(y)$ est bilinéaire symétrique et $q(x) = \varphi(x, x)$ pour tout $x \in E$.

Exercice 11.11. Montrer que si q est une forme quadratique sur $\mathbb{R}^n$, alors sa forme polaire φ est donnée par :

$$\varphi(x, y) = \frac{1}{2} \sum_{j=1}^n \frac{\partial q}{\partial x_j}(x) y_j = \frac{1}{2} \sum_{i=1}^n \frac{\partial q}{\partial y_i}(y) x_i$$

Solution. En notant $A = ((a_{ij}))_{1 \leq i, j \leq n}$ la matrice de q dans la base canonique, on a :

$$q(x) = \sum_{i=1}^n a_{ii} x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j = \sum_{1 \leq i, j \leq n} a_{ij} x_i x_j$$

et la forme polaire de q est définie par $\varphi(x, y) = \sum_{1 \leq i, j \leq n} a_{ij} x_i y_j$. Pour tout entier k compris entre 1 et n , on a alors :

$$\begin{aligned} \frac{\partial q}{\partial x_k}(x) &= \frac{\partial}{\partial x_k} \left(\sum_{i=1}^n x_i \sum_{j=1}^n a_{ij} x_j \right) = \sum_{i=1}^n \frac{\partial}{\partial x_k} \left(x_i \sum_{j=1}^n a_{ij} x_j \right) \\ &= \frac{\partial}{\partial x_k} \left(x_k \sum_{j=1}^n a_{kj} x_j \right) + \sum_{\substack{i=1 \\ i \neq k}}^n \frac{\partial}{\partial x_k} \left(x_i \sum_{j=1}^n a_{ij} x_j \right) \\ &= \sum_{j=1}^n a_{kj} x_j + x_k a_{kk} + \sum_{\substack{i=1 \\ i \neq k}}^n x_i a_{ik} = \sum_{j=1}^n a_{kj} x_j + \sum_{i=1}^n a_{ik} x_i \\ &= \sum_{j=1}^n a_{jk} x_j + \sum_{i=1}^n a_{ki} x_i = 2 \sum_{i=1}^n a_{ik} x_i \end{aligned}$$

(les égalités $a_{kj} = a_{jk}$ sont justifiées par la symétrie de la matrice A). On en déduit alors que :

$$\varphi(x, y) = \sum_{1 \leq i, j \leq n} a_{ij} x_i y_j = \sum_{j=1}^n \left(\sum_{i=1}^n a_{ij} x_i \right) y_j = \frac{1}{2} \sum_{j=1}^n \frac{\partial q}{\partial x_j}(x) y_j$$

Par symétrie, on a la deuxième formule. Par exemple, la forme polaire de la forme quadratique q définie dans la base canonique de $\mathbb{R}^3$ par :

$$q(x) = x_1^2 + 3x_2^2 + 5x_3^2 + 4x_1x_2 + 6x_1x_3 + 2x_2x_3.$$

est donnée par :

$$\begin{aligned} \varphi(x, y) &= \frac{1}{2} ((2x_1 + 4x_2 + 6x_3) y_1 + (4x_1 + 6x_2 + 2x_3) y_2 + (10x_3 + 6x_1 + 2x_2) y_3) \\ &= (x_1 + 2x_2 + 3x_3) y_1 + (2x_1 + 3x_2 + x_3) y_2 + (5x_3 + 3x_1 + x_2) y_3 \end{aligned}$$

qui est bien le résultat obtenu à l'exercice 11.7.

Exercice 11.12. Réduire la forme quadratique définie sur $\mathbb{R}^3$ par :

$$q(x) = x_1x_2 + 2x_1x_3 + 2x_1x_4 + x_2x_3 + 4x_2x_4 + 2x_3x_4$$

Solution. On a :

$$\begin{aligned} q(x) &= x_1x_2 + 2x_1x_3 + 2x_1x_4 + x_2x_3 + 4x_2x_4 + 2x_3x_4 \\ &= (x_1 + x_3 + 4x_4)(x_2 + 2x_3 + 2x_4) - 2x_3^2 - 8x_4^2 - 8x_3x_4 \\ &= (x_1 + x_3 + 4x_4)(x_2 + 2x_3 + 2x_4) - 2(x_3 + 2x_4)^2 \\ &= \frac{1}{4}(x_1 + x_2 + 3x_3 + 6x_4)^2 - \frac{1}{4}(x_1 - x_2 - x_3 + 2x_4)^2 - 2(x_3 + 2x_4)^2 \end{aligned}$$

Exercice 11.13. Donner la matrice de la forme quadratique définie dans la base canonique de $\mathbb{R}^n$ par $q(x) = \sum_{i=1}^n x_i^2 + \sum_{1 \leq i < j \leq n} x_i x_j$. Réduire q dans les cas $n = 2$, $n = 3$ et $n = 4$.

Solution. On a $A = \frac{1}{2} \begin{pmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \ddots & 1 \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 2 \end{pmatrix}$. Pour $n = 2$, on a :

$$q(x) = x_1^2 + x_2^2 + x_1x_2 = \left(x_1 + \frac{1}{2}x_2\right)^2 + \frac{3}{4}x_2^2$$

Pour $n = 3$, on a :

$$\begin{aligned} q(x) &= x_1^2 + x_2^2 + x_3^2 + x_1x_2 + x_1x_3 + x_2x_3 \\ &= \left(x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3\right)^2 + \frac{3}{4}\left(x_2^2 + x_3^2 + \frac{2}{3}x_2x_3\right) \\ &= \left(x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3\right)^2 + \frac{3}{4}\left(x_2 + \frac{1}{3}x_3\right)^2 + \frac{2}{3}x_3^2 \end{aligned}$$

Pour $n = 4$, on a :

$$\begin{aligned} q(x) &= x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4 \\ &= \left(x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3\right)^2 + \frac{3}{4}\left(x_2^2 + x_3^2 + \frac{2}{3}x_2x_3\right) \\ &= \left(x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 + \frac{1}{2}x_4\right)^2 + \frac{3}{4}\left(x_2 + \frac{1}{3}x_3 + \frac{1}{3}x_4\right)^2 \\ &\quad + \frac{2}{3}\left(x_3 + \frac{1}{4}x_4\right)^2 + \frac{5}{8}x_4^2 \end{aligned}$$

Exercice 11.14. On considère à nouveau la forme quadratique q définie sur $\mathbb{R}^n$ par $q(x) = \sum_{i=1}^n x_i^2 + \sum_{1 \leq i < j \leq n} x_i x_j$ avec $n \geq 3$.

1. Écrire q sous la forme $q(x) = \ell_1^2(x) + q_1(x')$, où ℓ_1 est une forme linéaire sur $\mathbb{R}^n$, q_1 une forme quadratique sur $\mathbb{R}^{n-1}$ et $x' = (x_2, \dots, x_n)$.
2. Écrire q_1 sous la forme $q_1(x) = \frac{3}{4}\ell_2^2(x) + q_2(x'')$, où ℓ_2 est une forme linéaire sur $\mathbb{R}^{n-1}$, q_2 une forme quadratique sur $\mathbb{R}^{n-2}$ en notant $x'' = (x_3, \dots, x_n)$.
3. Montrer que pour tout p compris entre 1 et $n-1$, on peut écrire q sous la forme :

$$q(x) = \ell_1^2(x) + \frac{3}{4}\ell_2^2(x) + \dots + \frac{p+1}{2p}\ell_p^2(x) + \frac{p+2}{2p+2}q_{p+1}(x)$$

$$\text{où on a noté } \ell_1(x) = x_1 + \frac{1}{2}\sum_{j=2}^n x_j, \ell_2(x) = x_2 + \frac{1}{3}\sum_{j=3}^n x_j, \dots, \ell_p(x) =$$

$$x_p + \frac{1}{p+1}\sum_{j=p+1}^n x_j \text{ et } q_{p+1}(x) = \sum_{i=p+1}^n x_i^2 + \frac{2}{p+2}\sum_{p+1 \leq i < j \leq n} x_i x_j.$$

4. Réduire q .

Solution.

1. On a :

$$\begin{aligned} q(x) &= x_1^2 + x_1 \sum_{j=2}^n x_j + \sum_{i=2}^n x_i^2 + \sum_{2 \leq i < j \leq n} x_i x_j \\ &= \left(x_1 + \frac{1}{2} \sum_{j=2}^n x_j \right)^2 - \frac{1}{4} \left(\sum_{j=2}^n x_j \right)^2 + \sum_{i=2}^n x_i^2 + \sum_{2 \leq i < j \leq n} x_i x_j \\ &= \left(x_1 + \frac{1}{2} \sum_{j=2}^n x_j \right)^2 + \frac{3}{4} \sum_{i=2}^n x_i^2 + \frac{1}{2} \sum_{2 \leq i < j \leq n} x_i x_j = \ell_1^2(x) + q_1(x') \end{aligned}$$

$$\text{avec } \ell_1(x) = x_1 + \frac{1}{2} \sum_{j=2}^n x_j \text{ et } q_1(x') = \frac{3}{4} \sum_{i=2}^n x_i^2 + \frac{1}{2} \sum_{2 \leq i < j \leq n} x_i x_j.$$

2. On a :

$$\begin{aligned}
 q_1(x') &= \frac{3}{4}x_2^2 + \frac{1}{2}x_2 \sum_{j=3}^n x_j + \frac{3}{4} \sum_{i=3}^n x_i^2 + \frac{1}{2} \sum_{3 \leq i < j \leq n} x_i x_j \\
 &= \frac{3}{4} \left(x_2 + \frac{1}{3} \sum_{j=3}^n x_j \right)^2 - \frac{1}{12} \left(\sum_{j=3}^n x_j \right)^2 + \frac{3}{4} \sum_{i=3}^n x_i^2 + \frac{1}{2} \sum_{3 \leq i < j \leq n} x_i x_j \\
 &= \frac{3}{4} \left(x_2 + \frac{1}{3} \sum_{j=3}^n x_j \right)^2 + \frac{2}{3} \sum_{i=3}^n x_i^2 + \frac{1}{3} \sum_{3 \leq i < j \leq n} x_i x_j = \frac{3}{4} \ell_2^2(x) + q_2(x'')
 \end{aligned}$$

avec $\ell_2(x) = x_2 + \frac{1}{3} \sum_{j=3}^n x_j$ et $q_2(x'') = \frac{2}{3} \sum_{i=3}^n x_i^2 + \frac{1}{3} \sum_{3 \leq i < j \leq n} x_i x_j$.

3. Le résultat est vrai pour $p = 1$. Supposons le acquis pour p compris entre 1 et $n - 2$. On a alors :

$$\begin{aligned}
 q_{p+1}(x) &= \sum_{i=p+1}^n x_i^2 + \frac{2}{p+2} \sum_{p+1 \leq i < j \leq n} x_i x_j \\
 &= x_{p+1}^2 + \frac{2}{p+2} x_{p+1} \sum_{j=p+2}^n x_j + \sum_{i=p+2}^n x_i^2 + \frac{2}{p+2} \sum_{p+2 \leq i < j \leq n} x_i x_j \\
 &= \left(x_{p+1} + \frac{1}{p+2} \sum_{j=p+2}^n x_j \right)^2 - \frac{1}{(p+2)^2} \left(\sum_{j=p+2}^n x_j \right)^2 + \sum_{i=p+2}^n x_i^2 \\
 &\quad + \frac{2}{p+2} \sum_{p+2 \leq i < j \leq n} x_i x_j = \ell_{p+1}^2(x) + Q_{p+2}(x)
 \end{aligned}$$

avec $\ell_{p+1}(x) = x_{p+1} + \frac{1}{p+2} \sum_{j=p+2}^n x_j$ et :

$$\begin{aligned}
 Q_{p+2}(x) &= \left(1 - \frac{1}{(p+2)^2} \right) \sum_{i=p+2}^n x_i^2 + \frac{2}{p+2} \left(1 - \frac{1}{p+2} \right) \sum_{p+2 \leq i < j \leq n} x_i x_j \\
 &= \frac{(p+1)(p+3)}{(p+2)^2} \sum_{i=p+2}^n x_i^2 + \frac{2(p+1)}{(p+2)^2} \sum_{p+2 \leq i < j \leq n} x_i x_j \\
 &= \frac{(p+1)(p+3)}{(p+2)^2} \left(\sum_{i=p+2}^n x_i^2 + \frac{2}{p+3} \sum_{p+2 \leq i < j \leq n} x_i x_j \right) = \frac{(p+1)(p+3)}{(p+2)^2} q_{p+2}(x)
 \end{aligned}$$

Ce qui donne :

$$\begin{aligned}
 q(x) &= \ell_1^2(x) + \frac{3}{4}\ell_2^2(x) + \cdots + \frac{p+1}{2p}\ell_p^2(x) + \frac{p+2}{2p+2}\ell_{p+1}^2(x) \\
 &\quad + \frac{p+2}{2p+2}Q_{p+2}(x) \\
 &= \ell_1^2(x) + \frac{3}{4}\ell_2^2(x) + \cdots + \frac{p+1}{2p}\ell_p^2(x) + \frac{p+2}{2p+2}\ell_{p+1}^2(x) \\
 &\quad + \frac{p+2}{2p+2} \frac{(p+1)(p+3)}{(p+2)^2} q_{p+2}(x) \\
 &= \ell_1^2(x) + \frac{3}{4}\ell_2^2(x) + \cdots + \frac{p+1}{2p}\ell_p^2(x) + \frac{p+2}{2p+2}\ell_{p+1}^2(x) + \frac{p+3}{2p+4}q_{p+2}(x)
 \end{aligned}$$

soit le résultat au rang $p+1$.

4. Faisant $p = n-1$, on a :

$$q(x) = \ell_1^2(x) + \frac{3}{4}\ell_2^2(x) + \cdots + \frac{n}{2(n-1)}\ell_{n-1}^2(x) + \frac{n+1}{2n}q_n(x)$$

avec $q_n(x) = x_n^2 = \ell_n^2(x)$, soit $q(x) = \sum_{p=1}^n \frac{p+1}{2p}\ell_p^2(x)$ avec :

$$\ell_p(x) = x_p + \frac{1}{p+1} \sum_{j=p+1}^n x_j$$

pour p compris entre 1 et n (pour $p = n$, la somme $\sum_{j=p+1}^n$ est nulle).

Exercice 11.15. Soit $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}$.

1. Déterminer la forme quadratique q sur $\mathbb{R}^3$ (ou $\mathbb{C}^3$) ayant pour matrice A dans la base canonique.
2. Déterminer 2 formes linéaires indépendantes ℓ_1, ℓ_2 telles que $q = \ell_1^2 - \ell_2^2$.
3. Déterminer une forme linéaire ℓ_3 (aussi simple que possible) telle que (ℓ_1, ℓ_2, ℓ_3) soit une base $\mathcal{L}(\mathbb{R}^3, \mathbb{R})$.
4. Déterminer une base (f_1, f_2, f_3) de $\mathbb{R}^3$ telle que $\ell_i(f_j) = \delta_{ij}$ pour tous i, j compris entre 1 et 3.
5. Déterminer une base de $\mathbb{R}^3$ orthogonale pour q et une matrice inversible P telle que $D = {}^t P A P$ soit diagonale.

Solution. On note x, y, z les coordonnées d'un vecteur v de $\mathbb{R}^3$ et $q(x, y, z)$ pour $q(v)$.

1. La forme q est définie par $q(x, y, z) = x^2 + 3y^2 + 5z^2 + 2(2xy + 3xz + 4yz)$.
2. On a la réduction de Gauss $q = \ell_1^2 - \ell_2^2$ avec $\ell_1(x, y, z) = x + 2y + 3z$ et $\ell_2(x, y, z) = y + 2z$.
3. On peut prendre ℓ_3 définie par $\ell_3(x, y, z) = z(\ell_1, \ell_2, \ell_3)$ est bien une base

$$\mathcal{L}(\mathbb{R}^3, \mathbb{R}) \text{ puisque } \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{vmatrix} = 1 \neq 0.$$

4. Il s'agit d'inverser la matrice $Q = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$. Pour ce faire on résout le système :

$$\begin{cases} x + 2y + 3z = x' \\ y + 2z = y' \\ z = z' \end{cases}$$

ce qui donne $x = x' - 2y' + z'$, $y = y' - 2z'$, $z' = z$ et f_1, f_2, f_3 sont les colonnes

$$\text{de } P = Q^{-1} = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

5. La base (f_1, f_2, f_3) est alors orthogonale pour q , ce qui signifie que la matrice de q dans cette base est diagonale. Précisément, on a $D = {}^tPAP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, ce qui peut aussi se vérifier par le calcul :

$${}^tPAP = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix} \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Exercice 11.16. Soit q la forme quadratique définie sur $\mathbb{R}^3$ par :

$$q(x) = x^2 + (1+a)y^2 + (1+a+a^2)z^2 + 2xy - 2ayz$$

1. Donner la matrice A de q dans la base canonique de $\mathbb{R}^3$.
2. Calculer le déterminant de A .
3. Pour quelles valeurs de A la forme q est-elle non dégénérée ?
4. Réduire q et donner son rang et sa signature en fonction de a .
5. Déterminer une base orthogonale pour q .
6. En déduire une matrice inversible P telle que $D = {}^tPAP$ soit diagonale.

Solution.

1. La matrice de q dans la base canonique de $\mathbb{R}^3$ est :

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1+a & -a \\ 0 & -a & 1+a+a^2 \end{pmatrix}$$

2. On a $\det(A) = a(1 + a^2)$.
3. La forme q est dégénérée si, et seulement si, $a = 0$.
4. On a $q(x) = (x + y)^2 + a(y - z)^2 + (1 + a^2)z^2$. Pour $a = 0$, q est de rang 2 et de signature $(2, 0)$. Pour $a \neq 0$, q est de rang 3 et de signature $(3, 0)$ pour $a > 0$ et $(2, 1)$ pour $a < 0$.
5. Dans tous les cas, il s'agit de résoudre le système :

$$\begin{cases} x + y = \alpha \\ y - z = \beta \\ z = \gamma \end{cases}$$

Ce système a pour solution $x = \alpha - \beta - \gamma$, $y = \beta + \gamma$, $z = \gamma$, ce qui donne pour base q -orthogonale :

$$f_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, f_2 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, f_3 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

6. La matrice de q dans la base (f_1, f_2, f_3) est $D = {}^tPAP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & 1 + a^2 \end{pmatrix}$,

$$\text{où } P = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Exercice 11.17. Déterminer le noyau et le cône isotrope de la forme bilinéaire symétrique φ définie sur $\mathbb{R}^3$ par $\varphi(x, y) = x_1y_1 + x_2y_2 - x_3y_3$.

Solution. Dire que y est dans le noyau de φ signifie que $\varphi(x, y) = 0$ pour tout vecteur x de $\mathbb{R}^3$, ce qui équivaut à $\varphi(e_i, y) = 0$ pour chacun des vecteurs de base canonique e_1, e_2, e_3 . Le noyau de φ est donc l'ensemble des solutions du système linéaire :

$$\begin{cases} \varphi(e_1, y) = y_1 = 0 \\ \varphi(e_2, y) = y_2 = 0 \\ \varphi(e_3, y) = -y_3 = 0 \end{cases}$$

soit $\ker(\varphi) = \{0\}$. Le cône isotrope de φ est formé des vecteurs x tels que $x_1^2 + x_2^2 - x_3^2 = 0$, et on reconnaît là l'équation d'un cône de $\mathbb{R}^3$.

Exercice 11.18. Soient F, G deux sous-espaces vectoriels de E . Montrer que $(F + G)^\perp = F^\perp \cap G^\perp$ et $(F \cap G)^\perp \supset F^\perp + G^\perp$.

Solution. Si $x \in (F + G)^\perp$, on a alors $\varphi(x, y + z) = 0$ pour tous vecteurs $y \in F$ et $z \in G$ et en particulier :

$$\begin{cases} \forall y \in F, \varphi(x, y) = \varphi(x, y + 0) = 0 \\ \forall z \in G, \varphi(x, z) = \varphi(x, 0 + z) = 0 \end{cases}$$

ce qui nous dit que $x \in F^\perp \cap G^\perp$. Réciproquement si $x \in F^\perp \cap G^\perp$, on a alors $\varphi(x, y) = \varphi(x, z) = 0$ pour tous vecteurs $y \in F$ et $z \in G$ et conséquence $\varphi(x, y+z) = 0$ pour ces vecteurs y, z , ce qui nous dit que $x \in (F+G)^\perp$. Si $x = u + v \in F^\perp + G^\perp$ avec $u \in F^\perp$ et $v \in G^\perp$, on a alors pour tout $y \in F \cap G$:

$$\varphi(x, y) = \varphi(u, y) + \varphi(v, y) = 0$$

et $x \in (F \cap G)^\perp$. L'égalité $(F \cap G)^\perp = F^\perp + G^\perp$ n'est pas assurée en général. Par exemple pour F, G supplémentaires dans E , on a $(F \cap G)^\perp = \{0\}^\perp = E$ n'est en général pas égal à $F^\perp + G^\perp$.

Exercice 11.19. On note q la forme quadratique définie dans la base canonique $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ de $\mathbb{R}^n$ par $q(x) = \sum_{i=1}^n x_i^2 + \sum_{1 \leq i < j \leq n} x_i x_j$.

1. Déterminer la matrice de q dans la base $\mathcal{B}$, son noyau et son rang.
2. Pour $n = 2$ et pour $n = 3$, effectuer la décomposition en carrés de Gauss de q , en déduire une base q -orthogonale de $\mathbb{R}^2$ et écrire la matrice de q dans cette base.
3. On suppose que $n \geq 4$ et on note $f_1 = e_1$.
 - (a) Déterminer l'orthogonal H relativement à q de e_1 .
 - (b) Pour tout j compris entre 2 et n , on note $f_j = e_1 + \dots + e_{j-1} - j e_j$. Montrer que $(f_j)_{2 \leq j \leq n}$ est une base de H .
 - (c) Calculer $A f_j$ pour tout j compris entre 2 et n .
 - (d) Montrer que $\mathcal{B}' = (f_j)_{1 \leq j \leq n}$ est une base q -orthogonale de $\mathbb{R}^n$.
 - (e) Écrire la matrice de q dans la base $\mathcal{B}'$. En déduire une décomposition de q comme combinaison linéaire de carrés de formes linéaires indépendantes.

Solution.

$$1. A = \begin{pmatrix} 1 & \frac{1}{2} & \cdots & \frac{1}{2} \\ \frac{1}{2} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{1}{2} \\ \frac{1}{2} & \cdots & \frac{1}{2} & 1 \end{pmatrix}. \text{ Pour } 1 \leq j \leq n, \text{ on a :}$$

$$(x \in \ker(q)) \Leftrightarrow (Ax = 0) \Leftrightarrow (x_1 + \dots + x_{j-1} + 2x_j + x_{j+1} + \dots + x_n = 0)$$

En ajoutant toutes ces équations on obtient $\sum_{j=1}^n x_j = 0$ qui retranchée à l'équation j donne $x_j = 0$. On a donc $\ker(q) = \{0\}$ et $\text{rang}(q) = n$.

2. Pour $n = 2$, on a $q(x) = x_1^2 + x_2^2 + x_1x_2 = \left(x_1 + \frac{1}{2}x_2\right)^2 + \frac{3}{4}x_2^2$. En résolvant le système :

$$\begin{cases} x_1 + \frac{1}{2}x_2 = a \\ x_2 = b \end{cases}$$

pour $(a, b) = (1, 0)$ et $(a, b) = (0, 1)$, on obtient la base q -orthogonale : $f_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $f_2 = \begin{pmatrix} -\frac{1}{2} \\ 1 \end{pmatrix}$. La matrice de q dans cette base est $D = \begin{pmatrix} 1 & 0 \\ 0 & \frac{3}{4} \end{pmatrix}$.

Pour $n = 3$, on a :

$$\begin{aligned} q(x) &= x_1^2 + x_2^2 + x_3^2 + x_1x_2 + x_1x_3 + x_2x_3 \\ &= \left(x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3\right)^2 + \frac{3}{4}\left(x_2^2 + x_3^2 + \frac{2}{3}x_2x_3\right) \\ &= \left(x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3\right)^2 + \frac{3}{4}\left(x_2 + \frac{1}{3}x_3\right)^2 + \frac{2}{3}x_3^2. \end{aligned}$$

En résolvant le système :

$$\begin{cases} x_1 + \frac{1}{2}x_2 + \frac{1}{2}x_3 = a \\ x_2 + \frac{1}{3}x_3 = b \\ x_3 = c \end{cases}$$

pour $(a, b, c) = (1, 0, 0)$, $(0, 1, 0)$ et $(0, 0, 1)$ on obtient la base q -orthogonale :

$$f_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, f_2 = \begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \end{pmatrix}, f_3 = \begin{pmatrix} -\frac{1}{3} \\ -\frac{1}{3} \\ 1 \end{pmatrix}$$

La matrice de q dans cette base est $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{4} & 0 \\ 0 & 0 & \frac{2}{3} \end{pmatrix}$.

3.

$$(a) \quad x \in \{e_1\}^\perp \Leftrightarrow \varphi(x, e_1) = 0 \Leftrightarrow {}^t x A e_1 = 0 \Leftrightarrow (x_1, \dots, x_n) \begin{pmatrix} 1 \\ \frac{1}{2} \\ \vdots \\ \frac{1}{2} \end{pmatrix} = 0. \text{ Une}$$

équation de H est donc : $2x_1 + x_2 + \dots + x_n = 0$.

- (b) Les coordonnées de f_j dans $\mathcal{B}$ sont données par $x_1 = \dots = x_{j-1} = 1$, $x_j = -j$, $x_{j+1} = \dots = x_n = 0$ et $2x_1 + x_2 + \dots + x_n = 2 + (j-2) - j = 0$. Les vecteurs f_j sont bien dans H et ils sont libres, donc forment une base.

$$(c) \quad A f_j = \frac{1}{2} \begin{pmatrix} 2 & 1 & \cdots & 1 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \\ -j \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -j-1 \\ -1 \\ \vdots \\ -1 \end{pmatrix}.$$

(d) Pour $2 \leq i < j$, on a :

$$\varphi(f_i, f_j) = \frac{1}{2} (1, \dots, 1, -i, 0, \dots, 0) \begin{pmatrix} 0 \\ \vdots \\ 0 \\ -j-1 \\ -1 \\ \vdots \\ -1 \end{pmatrix} = 0$$

et on sait déjà que f_1 est q -orthogonal aux f_j pour $j \geq 2$.

(e) On a $q(f_j) = \frac{j(j+1)}{2}$ et la matrice de q dans $\mathcal{B}'$ est :

$$D = \frac{1}{2} \begin{pmatrix} 2 & 0 & \cdots & \cdots & 0 \\ 0 & 6 & \ddots & & \vdots \\ \vdots & \ddots & 12 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & n(n+1) \end{pmatrix}$$

L'expression de q dans $\mathcal{B}'$ est $q(x) = \frac{1}{2} \sum_{j=1}^n j(j+1) x_j'^2 = \frac{1}{2} \sum_{j=1}^n j(j+1) \ell_j^2(x)$

avec $X' = P^{-1}X$ où $P = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & -2 & 1 & & \vdots \\ \vdots & 0 & -3 & \ddots & \vdots \\ \vdots & & \vdots & \ddots & 1 \\ 0 & 0 & 0 & 0 & -n \end{pmatrix}$ est la matrice de

passage de $\mathcal{B}$ à $\mathcal{B}'$. Pour $n = 5$, on a $P^{-1} = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & -\frac{1}{2} & -\frac{1}{6} & -\frac{1}{6} & -\frac{1}{6} \\ 0 & 0 & -\frac{1}{3} & -\frac{1}{12} & -\frac{1}{12} \\ 0 & 0 & 0 & -\frac{1}{4} & -\frac{1}{20} \\ 0 & 0 & 0 & 0 & -\frac{1}{5} \end{pmatrix}$

et pour $n \geq 4$, la ligne 1 de P^{-1} est $(1 \quad \frac{1}{2} \quad \frac{1}{2} \quad \cdots \quad \frac{1}{2})$ et la ligne $j \geq 2$ est $(0, \dots, 0, -\frac{1}{j}, -\frac{1}{j(j+1)}, \dots, -\frac{1}{j(j+1)})$. On a donc :

$$\begin{cases} \ell_1(x) = x_1 + \frac{1}{2}x_2 + \cdots + \frac{1}{2}x_n \\ \ell_j(x) = \frac{1}{j}x_j + \frac{1}{j(j+1)}x_{j+1} + \cdots + \frac{1}{j(j+1)}x_n \\ \ell_n(x) = \frac{1}{n}x_n \end{cases}$$

ou encore :

$$q(x) = \frac{1}{2} \sum_{j=1}^{n-1} \frac{1}{j(j+1)} ((j+1)x_j + x_{j+1} + \cdots + x_n)^2 + \frac{n+1}{2n} x_n^2$$

Chapitre 12

Espaces préhilbertiens

On désigne par E un espace vectoriel réel non réduit à $\{0\}$.

12.1 Produit scalaire

Définition 12.1. On dit qu'une forme bilinéaire symétrique φ sur E est :

- positive si $\varphi(x, x) \geq 0$ pour tout x dans E ;
- définie si pour x dans E l'égalité $\varphi(x, x) = 0$ équivaut à $x = 0$.

Définition 12.2. On appelle produit scalaire sur E toute forme bilinéaire symétrique définie positive. Un espace préhilbertien est un espace vectoriel réel muni d'un produit scalaire. Un espace préhilbertien de dimension finie est dit euclidien.

Dans le cas où E est un espace euclidien, on peut aussi dire qu'un produit scalaire sur E est la forme polaire d'une forme quadratique de signature $(n, 0)$. On notera $(x, y) \mapsto \langle x | y \rangle$ quand il n'y a pas d'ambiguïté un tel produit scalaire et pour $y = x$, on note $\|x\| = \sqrt{\langle x | x \rangle}$. L'application $x \mapsto \|x\|^2 = \langle x | x \rangle$ est tout simplement la forme quadratique associée à $\langle \cdot | \cdot \rangle$.

Les trois égalités qui suivent, expressions de la forme polaire d'une forme quadratique, sont utiles en pratique.

Théorème 12.1.

Pour tous x, y dans E on a :

$$\langle x | y \rangle = \frac{\|x + y\|^2 - \|x\|^2 - \|y\|^2}{2} = \frac{\|x + y\|^2 - \|x - y\|^2}{4}$$

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

La deuxième identité est l'égalité du parallélogramme. Elle est caractéristique des produits scalaires dans le sens où une norme est déduite d'un produit scalaire si, et seulement si, elle vérifie l'identité du parallélogramme.

Exemple 12.1 L'espace vectoriel $\mathbb{R}^n$ étant muni de sa base canonique $(e_i)_{1 \leq i \leq n}$, l'application $(x, y) \mapsto \langle x | y \rangle = \sum_{k=1}^n x_k y_k$ définit un produit scalaire sur $\mathbb{R}^n$. On dit que c'est le produit scalaire euclidien canonique de $\mathbb{R}^n$.

Dans tout ce qui suit E désigne un espace préhilbertien.

Théorème 12.2. Inégalité de Cauchy-Schwarz

Pour tous x, y dans E on a $|\langle x | y \rangle| \leq \|x\| \|y\|$, l'égalité étant réalisée si, et seulement si, x et y sont liés.

Preuve. Si $x = 0$, on a alors l'égalité pour tout $y \in E$. Si $x \neq 0$ et $y = \lambda x$ avec $\lambda \in \mathbb{R}$, on a encore l'égalité. On suppose donc que x est non nul et y non lié à x . La fonction polynomiale P défini par :

$$P(t) = \|y + tx\|^2 = \|x\|^2 t^2 + 2 \langle x | y \rangle t + \|y\|^2$$

est alors à valeurs strictement positives, le coefficient de t^2 étant non nul, il en résulte que son discriminant est strictement négatif, soit $\langle x | y \rangle^2 - \|x\|^2 \|y\|^2 < 0$, ce qui équivaut à $|\langle x | y \rangle| < \|x\| \|y\|$. $\square$

Sur $\mathbb{R}^n$ muni du produit scalaire canonique, l'inégalité de Cauchy-Schwarz prend la forme suivante :

$$\left| \sum_{k=1}^n x_k y_k \right|^2 \leq \left(\sum_{k=1}^n x_k^2 \right) \left(\sum_{k=1}^n y_k^2 \right)$$

Dans l'espace $C^0([a, b], \mathbb{R})$ des fonctions continues de $[a, b]$ dans $\mathbb{R}$ muni du produit scalaire $(f, g) \mapsto \int_a^b f(t) g(t) dt$, l'inégalité de Cauchy-Schwarz s'écrit :

$$\left(\int_a^b f(t) g(t) dt \right)^2 \leq \int_a^b f^2(t) dt \int_a^b g^2(t) dt$$

On peut déduire de cette inégalité quelques inégalités intéressantes sur les réels.

Une conséquence importante de l'inégalité de Cauchy-Schwarz est l'inégalité triangulaire de Minkowski.

Théorème 12.3. Inégalité de Minkowski

Pour tous x, y dans E on a $\|x + y\| \leq \|x\| + \|y\|$, l'égalité étant réalisée si, et seulement si, $x = 0$ ou $x \neq 0$ et $y = \lambda x$ avec $\lambda \geq 0$ (on dit que x et y sont positivement liés).

Preuve. Si $x = 0$, on a alors l'égalité pour tout $y \in E$. Si $x \neq 0$ et $y = \lambda x$ avec $\lambda \in \mathbb{R}$, on a alors $\|x + y\| = |1 + \lambda| \|x\| \leq (1 + |\lambda|) \|x\| = \|x\| + \|\lambda x\| = \|x\| + \|y\|$, l'égalité

étant réalisée pour $\lambda \geq 0$. Pour $\lambda < 0$, l'inégalité est stricte puisque dans ce cas $|1 + \lambda| < 1 + |\lambda| = 1 - \lambda$. On suppose que x est non nul et y non lié à x . On a $\|x + y\|^2 = \|x\|^2 + 2 \langle x | y \rangle + \|y\|^2$ et avec l'inégalité de Cauchy-Schwarz :

$$\|x + y\|^2 < \|x\|^2 + 2 \|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$$

ce qui équivaut à $\|x + y\| < \|x\| + \|y\|$. □

L'inégalité de Minkowski ajoutée aux propriétés de positivité ($\|x\| > 0$ pour tout $x \neq 0$) et d'homogénéité ($\|\lambda x\| = |\lambda| \|x\|$ pour tout réel λ et tout vecteur x) se traduit en disant que l'application $x \mapsto \|x\| = \sqrt{\langle x | x \rangle}$ définit une norme sur E . De cette inégalité, on déduit l'inégalité suivante :

$$\forall (x, y) \in E^2, \quad \left| \|x\| - \|y\| \right| \leq \|x + y\|$$

Par récurrence, on montre facilement que pour tous vecteurs $x_1, \dots, x_p$, on a :

$$\|x_1 + \dots + x_p\| \leq \|x_1\| + \dots + \|x_p\|.$$

Sur $\mathbb{R}^n$ muni du produit scalaire canonique, l'inégalité de Minkowski prend la forme suivante :

$$\sqrt{\sum_{k=1}^n (x_k + y_k)^2} \leq \sqrt{\sum_{k=1}^n x_k^2} + \sqrt{\sum_{k=1}^n y_k^2}$$

Dans l'espace $\mathcal{C}^0([a, b], \mathbb{R})$ des fonctions continues de $[a, b]$ dans $\mathbb{R}$ muni du produit scalaire $(f, g) \mapsto \int_a^b f(t) g(t) dt$, l'inégalité de Minkowski s'écrit :

$$\sqrt{\int_a^b (f(t) + g(t))^2 dt} \leq \sqrt{\int_a^b f^2(t) dt} + \sqrt{\int_a^b g^2(t) dt}$$

12.2 Orthogonalité

Définition 12.3. On dit que deux vecteurs x et y appartenant à E sont orthogonaux si $\langle x | y \rangle = 0$.

Théorème 12.4. Pythagore

Les vecteurs x et y sont orthogonaux dans E si, et seulement si, on a :

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2$$

On montre facilement par récurrence sur $p \geq 2$, que si $x_1, \dots, x_p$ sont deux à deux orthogonaux, on a alors $\left\| \sum_{k=1}^p x_k \right\|^2 = \sum_{k=1}^p \|x_k\|^2$.

Définition 12.4. On appelle famille orthogonale dans E toute famille $(e_i)_{i \in I}$ de vecteurs de E telle que $\langle e_i | e_j \rangle = 0$ pour tous $i \neq j$ dans I . Si de plus $\|e_i\| = 1$ pour tout $i \in I$, on dit alors que cette famille est orthonormée ou orthonormale.

Définition 12.5. L'orthogonal d'une partie non vide X de E est l'ensemble $X^\perp = \{y \in E \mid \forall x \in X, \langle x | y \rangle = 0\}$.

Il est facile de vérifier que $X^\perp$ est un sous-espace vectoriel de E .

Théorème 12.5.

Une famille orthogonale de vecteurs non nuls de E est libre.

Preuve. Si $(e_i)_{i \in I}$ est une telle famille et si $\sum_{j \in J} \lambda_j e_j = 0$ où J est une partie finie

de I on a alors $0 = \left\langle \sum_{j \in J} \lambda_j e_j \mid e_k \right\rangle = \lambda_k \|e_k\|^2$ pour tout $k \in J$ avec $\|e_k\| \neq 0$ et nécessairement $\lambda_k = 0$. $\square$

Exemple 12.2 La famille $\{\cos(nt), \sin(mt) \mid (n, m) \in \mathbb{N} \times \mathbb{N}^*\}$ est orthogonale dans l'espace vectoriel $\mathcal{F}$ des fonctions continues et 2π -périodiques sur $\mathbb{R}$ muni du produit scalaire $(f, g) \mapsto \langle f \mid g \rangle = \int_{-\pi}^{\pi} f(t) g(t) dt$. Pour $n \neq m$ dans $\mathbb{N}$, on a :

$$\int_{-\pi}^{\pi} \cos(nt) \cos(mt) dt = \frac{1}{2} \int_{-\pi}^{\pi} (\cos((n+m)t) + \cos((n-m)t)) dt = 0$$

pour $n \neq m$ dans $\mathbb{N}^*$, on a :

$$\int_{-\pi}^{\pi} \sin(nt) \sin(mt) dt = \frac{1}{2} \int_{-\pi}^{\pi} (\cos((n-m)t) - \cos((n+m)t)) dt = 0$$

et pour $(n, m) \in \mathbb{N} \times \mathbb{N}^*$, on a :

$$\int_{-\pi}^{\pi} \cos(nt) \sin(mt) dt = \frac{1}{2} \int_{-\pi}^{\pi} (\sin((n+m)t) - \sin((n-m)t)) dt = 0$$

Pour $n = 0$, on a $\int_{-\pi}^{\pi} dt = 2\pi$ et pour $n \geq 1$:

$$\begin{cases} \int_{-\pi}^{\pi} \cos^2(nt) dt = \frac{1}{2} \int_{-\pi}^{\pi} (\cos(2nt) + 1) dt = \pi \\ \int_{-\pi}^{\pi} \sin^2(nt) dt = \frac{1}{2} \int_{-\pi}^{\pi} (1 - \cos(2nt)) dt = \pi \end{cases}$$

De l'exemple précédent, on déduit que la famille de fonctions :

$$\left\{ \frac{1}{\sqrt{2\pi}} \right\} \cup \left\{ \frac{1}{\sqrt{\pi}} \cos(nt), \frac{1}{\sqrt{\pi}} \sin(mt) \mid (n, m) \in \mathbb{N}^* \times \mathbb{N}^* \right\}$$

est orthonormée dans $\mathcal{F}$. Pour tout $n \in \mathbb{N}^*$, la famille :

$$\mathcal{T}_n = \left\{ \frac{1}{\sqrt{2\pi}} \right\} \cup \left\{ \frac{\cos(jt)}{\sqrt{\pi}}, \frac{\sin(kt)}{\sqrt{\pi}} \mid 1 \leq j, k \leq n \right\}$$

est une base orthonormée de l'espace $\mathcal{P}_n$ des polynômes trigonométriques de degré inférieur ou égal à n .

12.3 Procédé d'orthogonalisation de Gram-Schmidt

Théorème 12.6. Orthonormalisation de Gram-Schmidt

Pour toute famille libre $(x_i)_{1 \leq i \leq p}$ dans E , il existe une unique famille orthonormée $(e_i)_{1 \leq i \leq p}$ dans E telle que :

$$\forall k \in \{1, 2, \dots, p\}, \begin{cases} \text{Vect}\{e_1, \dots, e_k\} = \text{Vect}\{x_1, \dots, x_k\} \\ \langle x_k \mid e_k \rangle > 0 \end{cases}$$

Preuve. Pour $p = 1$, on a nécessairement $e_1 = \lambda_1 x_1$ avec $\lambda_1 \in \mathbb{R}^*$ et $1 = \|e_1\|^2 = \lambda_1^2 \|x_1\|^2$, donc $\lambda_1^2 = \frac{1}{\|x_1\|^2}$ ce qui donne deux solutions pour λ_1 . La condition supplémentaire $\langle x_1 \mid e_1 \rangle > 0$ entraîne $\lambda_1 > 0$ et on obtient ainsi l'unique solution $e_1 = \frac{1}{\|x_1\|} x_1$. Supposons $p \geq 2$ et construite la famille orthonormée $(e_i)_{1 \leq i \leq p-1}$ vérifiant les conditions :

$$\forall k \in \{1, 2, \dots, p-1\}, \begin{cases} \text{Vect}\{e_1, \dots, e_k\} = \text{Vect}\{x_1, \dots, x_k\} \\ \langle x_k \mid e_k \rangle > 0 \end{cases}$$

Si $(e'_1, \dots, e'_{p-1}, e_p)$ est une solution à notre problème on a alors nécessairement $e'_k = e_k$ pour tout k compris entre 1 et $p-1$ (unicité pour le cas $p-1$). La condition $\text{Vect}\{e_1, \dots, e_p\} = \text{Vect}\{x_1, \dots, x_p\}$ entraîne $e_p = \sum_{j=1}^{p-1} \lambda_j e_j + \lambda_p x_p$.

Avec les conditions d'orthogonalité $\langle e_p \mid e_j \rangle = 0$ pour $1 \leq j \leq p-1$, on déduit que $\lambda_j + \lambda_p \langle x_p \mid e_j \rangle = 0$ et :

$$e_p = \lambda_p \left(x_p - \sum_{j=1}^{p-1} \langle x_p \mid e_j \rangle e_j \right) = \lambda_p y_p$$

Du fait que $x_p \notin \text{Vect} \{x_1, \dots, x_{p-1}\} = \text{Vect} \{e_1, \dots, e_{p-1}\}$ on déduit que $y_p \neq 0$ et la condition $\|e_p\| = 1$ donne $|\lambda_p| = \frac{1}{\|y_p\|}$. La condition supplémentaire :

$$0 < \langle x_p | e_p \rangle = \left\langle \frac{1}{\lambda_p} \left(e_p - \sum_{j=1}^{p-1} \lambda_j e_j \right) | e_p \right\rangle = \frac{1}{\lambda_p}$$

entraîne $\lambda_p > 0$. Ce qui donne en définitive une unique solution pour e_p . $\square$

La construction d'une famille orthonormée $(e_i)_{1 \leq i \leq p}$ peut se faire en utilisant l'algorithme suivant :

$$\begin{cases} y_1 = x_1, & e_1 = \frac{1}{\|y_1\|} y_1 \\ y_k = x_k - \sum_{j=1}^{k-1} \langle x_k | e_j \rangle e_j, & e_k = \frac{1}{\|f_k\|} f_k, \quad (k = 2, \dots, p) \end{cases}$$

Le calcul de $\|y_k\|$ peut être simplifié en écrivant que :

$$\begin{aligned} \|y_k\|^2 &= \left\langle y_k | x_k - \sum_{j=1}^{k-1} \langle x_k | e_j \rangle e_j \right\rangle \\ &= \langle y_k | x_k \rangle = \left\langle x_k - \sum_{j=1}^{k-1} \langle x_k | e_j \rangle e_j | x_k \right\rangle = \|x_k\|^2 - \sum_{j=1}^{k-1} \langle x_k | e_j \rangle^2 \end{aligned}$$

(y_k est orthogonal à e_j pour $1 \leq j \leq k-1$). Les $\langle x_k | e_j \rangle$ étant déjà calculés (pour obtenir y_k), il suffit donc de calculer $\|x_k\|^2$. En fait le calcul de $\langle y_k | x_k \rangle$ est souvent plus rapide.

Corollaire 12.1. *Si F est un sous-espace vectoriel de dimension finie ou infinie dénombrable de E , alors il existe une base orthonormée pour F .*

Preuve. On raisonne par récurrence en utilisant le théorème de Gram-Schmidt. $\square$

Si E est un espace euclidien de dimension finie et $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ une base orthonormée de E , alors tout vecteur $x \in E$ s'écrit $x = \sum_{k=1}^n \langle x | e_k \rangle e_k = \sum_{k=1}^n x_k e_k$ et on a pour tous vecteurs x, y dans E , en notant X la matrice de x dans la base $\mathcal{B}$:

$$\langle x | y \rangle = \sum_{k=1}^n \langle x | e_k \rangle \langle y | e_k \rangle = \sum_{k=1}^n x_k y_k = {}^t X Y$$

$$\text{et } \|x\|^2 = \sum_{k=1}^n \langle x | e_k \rangle^2 = \sum_{k=1}^n x_k^2.$$

Ces égalités sont des cas particuliers des égalités de Parseval valables de manière plus générale dans les espaces de Hilbert.

Théorème 12.7.

Si E est un espace euclidien de dimension $n \geq 1$, $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ et $\mathcal{B}' = (e'_i)_{1 \leq i \leq n}$ sont deux bases orthonormées de E , alors la matrice de passage P de $\mathcal{B}$ à $\mathcal{B}'$ est telle que $P^{-1} = {}^tP$. En particulier, on a $\det(P) = \pm 1$.

Preuve. Les colonnes de la matrice P sont formées des vecteurs colonnes $E'_1, \dots, E'_n$, où E'_j est la matrice de e'_j dans la base $\mathcal{B}$ et on a :

$$\begin{aligned} {}^tPP &= \begin{pmatrix} {}^tE'_1 \\ \vdots \\ {}^tE'_n \end{pmatrix} (E'_1, \dots, E'_n) = (({}^tE'_i E'_j))_{1 \leq i, j \leq n} \\ &= ((\langle e'_i | e'_j \rangle))_{1 \leq i, j \leq n} = ((\delta_{ij}))_{1 \leq i, j \leq n} = I_n \end{aligned}$$

ce équivaut à dire que $P^{-1} = {}^tP$. On a alors :

$$1 = \det(I_n) = \det({}^tPP) = \det({}^tP) \det(P) = \det^2(P)$$

et $\det(P) = \pm 1$. □

Définition 12.6. On appelle matrice orthogonale toute matrice réelle d'ordre n inversible telle que $P^{-1} = {}^tP$.

La matrice de passage d'une base orthonormée $\mathcal{B}$ de E à une autre base orthonormée $\mathcal{B}'$ est donc une matrice orthogonale et réciproquement une telle matrice est la matrice de passage d'une base orthonormée de E à une autre.

12.4 Projection orthogonale sur un sous-espace de dimension finie

$(E, \langle \cdot | \cdot \rangle)$ est toujours un espace préhilbertien de dimension finie ou non.

Théorème 12.8. Projection orthogonale

Soit F un sous-espace vectoriel de dimension finie de E non réduit à $\{0\}$. Pour tout vecteur $x \in E$, il existe un unique vecteur y dans F tel que $\|x - y\| = d(x, F) = \inf_{z \in F} \|x - z\|$. Ce vecteur est également l'unique vecteur appartenant à F tel que $x - y \in F^\perp$. Son expression dans une base orthonormée $(e_i)_{1 \leq i \leq n}$ de F est donnée par $y = \sum_{k=1}^n \langle x | e_k \rangle e_k$ et on a :

$$\|x - y\|^2 = \|x\|^2 - \|y\|^2 = \|x\|^2 - \sum_{k=1}^n \langle x | e_k \rangle^2 \quad (12.1)$$

Preuve. Soit $(e_i)_{1 \leq i \leq n}$ une base orthonormée de F (le théorème de Gram-Schmidt nous assure l'existence d'une telle base). Pour x dans E , on définit le vecteur $y \in F$ par $y = \sum_{k=1}^n \langle x | e_k \rangle e_k$. On a alors $\langle x - y | e_j \rangle = 0$ pour tout $j \in \{1, \dots, n\}$, c'est-à-dire que $x - y \in F^\perp$. Le théorème de Pythagore donne alors, pour tout $z \in F$:

$$\|x - z\|^2 = \|(x - y) + (y - z)\|^2 = \|x - y\|^2 + \|y - z\|^2 \geq \|x - y\|^2$$

et on a bien $\|x - y\| = d(x, F)$. S'il existe un autre vecteur $u \in F$ tel que $\|x - u\| = d(x, F) = \delta$, de $\delta^2 = \|x - u\|^2 = \|x - y\|^2 + \|y - u\|^2 = \delta^2 + \|y - u\|^2$, on déduit alors que $\|y - u\| = 0$ et $y = u$. On sait déjà que le vecteur $y \in F$ est tel que $x - y \in F^\perp$. Supposons qu'il existe un autre vecteur $u \in F$ tel que $x - u \in F^\perp$, pour tout $z \in F$, on a alors :

$$\|x - z\|^2 = \|(x - u) + (u - z)\|^2 = \|x - u\|^2 + \|u - z\|^2 \geq \|x - u\|^2$$

donc $\|x - u\| = d(x, F)$ et $u = y$ d'après ce qui précède. La dernière égalité se déduit de $\|x\|^2 = \|(x - y) + y\|^2 = \|x - y\|^2 + \|y\|^2$. $\square$

Si x est un vecteur de E , alors le vecteur y de F qui lui est associé dans le théorème précédent est la meilleure approximation de x dans F . En considérant la caractérisation géométrique $x - y \in F^\perp$, on dit aussi que y est la projection orthogonale de x sur F . On note $y = p_F(x)$. On a donc :

$$(y = p_F(x)) \Leftrightarrow (y \in F \text{ et } x - y \in F^\perp) \Leftrightarrow (y \in F \text{ et } \|x - y\| = d(x, F))$$

et dans une base orthonormée de F , une expression de p_F est :

$$\forall x \in E, p_F(x) = \sum_{k=1}^n \langle x | e_k \rangle e_k$$

On dit que l'application p_F est la projection orthogonale de E sur F .

Si $F = \{0\}$, on peut définir p_F et c'est l'application nulle. On suppose donc, a priori, F non réduit à $\{0\}$. Dans le cas où E est de dimension finie et $F = E$, p_F est l'application identité.

$p_F(x) = x$ équivaut à dire que $x \in F$ et $p_F(x) = 0$ équivaut à dire que $x \in F^\perp$.

Exemple 12.3 Si $D = \mathbb{R}a$ est une droite vectorielle, une base orthonormée de D est $\left(\frac{1}{\|a\|}a\right)$ et pour tout $x \in E$, on a $p_D(x) = \frac{\langle x | a \rangle}{\|a\|^2}a$.

De l'inégalité (12.1), on déduit que pour tout vecteur $x \in E$, on a :

$$\|p_F(x)\|^2 = \sum_{k=1}^n \langle x | e_k \rangle^2 \leq \|x\|^2$$

Cette inégalité est l'inégalité de Bessel.

Si $(e_i)_{1 \leq i \leq n}$ est une base (non nécessairement orthonormée) de F , alors la projection orthogonale d'un vecteur x de E sur F est le vecteur $y = \sum_{j=1}^n y_j e_j$, où les composantes y_j , pour j compris entre 1 et n , sont solutions du système linéaire :

$$\langle x - y \mid e_i \rangle = 0 \quad (1 \leq i \leq n)$$

soit :

$$\sum_{j=1}^n \langle e_i \mid e_j \rangle y_j = \langle x \mid e_i \rangle \quad (1 \leq i \leq n)$$

Ce système est appelé système d'équations normales.

Sur l'espace vectoriel $\mathcal{F}$ des fonctions continues et 2π -périodiques muni du produit scalaire :

$$(f, g) \mapsto \langle f \mid g \rangle = \int_{-\pi}^{\pi} f(x) g(x) dx$$

la meilleure approximation, pour la norme déduite de ce produit scalaire, d'une fonction $f \in \mathcal{F}$ par un polynôme trigonométrique de degré inférieur ou égal à n est donnée par :

$$\begin{aligned} S_n(f) &= \left\langle f \mid \frac{c_0}{\sqrt{2\pi}} \right\rangle \frac{c_0}{\sqrt{2\pi}} + \sum_{k=1}^n \left\langle f \mid \frac{c_k}{\sqrt{\pi}} \right\rangle \frac{c_k}{\sqrt{\pi}} + \sum_{k=1}^n \left\langle f \mid \frac{s_k}{\sqrt{\pi}} \right\rangle \frac{s_k}{\sqrt{\pi}} \\ &= \frac{1}{2\pi} \langle f \mid c_0 \rangle c_0 + \frac{1}{\pi} \sum_{k=1}^n \langle f \mid c_k \rangle c_k + \frac{1}{\pi} \sum_{k=1}^n \langle f \mid s_k \rangle s_k \end{aligned}$$

où $c_k : x \mapsto \cos(kx)$ pour $k \geq 0$ et $s_k : x \mapsto \sin(kx)$ pour $k \geq 1$. Soit :

$$S_n(f)(x) = \frac{a_0(f)}{2} + \sum_{k=1}^n a_k(f) \cos(kx) + \sum_{k=1}^n b_k(f) \sin(kx)$$

où les $a_k(f)$ et $b_k(f)$ sont les coefficients de Fourier trigonométriques de f définis par :

$$a_k(f) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(kt) dt \quad \text{et} \quad b_k(f) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin(kt) dt$$

L'opérateur S_n de projection orthogonale de $\mathcal{F}$ sur l'espace $\mathcal{P}_n$ des polynômes trigonométriques de degré inférieur ou égal à n est l'opérateur de Fourier.

La série $a_0(f) + \sum (a_n(f) \cos(nx) + b_n(f) \sin(nx))$ est la série de Fourier de f .

L'inégalité de Bessel s'écrit :

$$\left\langle f \mid \frac{c_0}{\sqrt{2\pi}} \right\rangle^2 + \sum_{k=1}^n \left\langle f \mid \frac{c_k}{\sqrt{\pi}} \right\rangle^2 + \sum_{k=1}^n \left\langle f \mid \frac{s_k}{\sqrt{\pi}} \right\rangle^2 \leq \|f\|^2$$

ou encore :

$$\frac{a_0^2(f)}{2} + \sum_{k=1}^n (a_k^2(f) + b_k^2(f)) \leq \frac{1}{\pi} \int_{-\pi}^{\pi} f^2(t) dt$$

Il en résulte que la série numérique $a_0^2(f) + \frac{1}{2} \sum (a_n^2(f) + b_n^2(f))$ converge avec :

$$\frac{a_0^2(f)}{2} + \sum_{n=1}^{+\infty} (a_n^2(f) + b_n^2(f)) \leq \frac{1}{\pi} \int_{-\pi}^{\pi} f^2(t) dt$$

(théorème de Bessel). On peut monter qu'on a fait l'égalité (théorème de Parseval).

De l'inégalité de Bessel, on déduit que $\lim_{n \rightarrow +\infty} a_n(f) = \lim_{n \rightarrow +\infty} b_n(f) = 0$ (théorème de Riemann-Lebesgue).

Exemple 12.4 Si $f \in \mathcal{F}$ est la fonction 2π -périodique, paire valant $x(\pi - x)$ sur $[0, \pi]$, on a alors $b_n(f) = 0$ pour tout $n \geq 1$ puisque f est paire et :

$$a_0(f) = \frac{2}{\pi} \int_0^{\pi} t(\pi - t) dt = \frac{\pi^2}{3}$$

Pour $n \geq 1$, on a :

$$\begin{aligned} a_n(f) &= \frac{2}{\pi} \int_0^{\pi} t(\pi - t) \cos(nt) dt = -\frac{2}{n^2} (1 + (-1)^n) \\ &= \begin{cases} 0 & \text{si } n = 2p + 1 \\ -\frac{1}{p^2} & \text{si } n = 2p \end{cases} \end{aligned}$$

L'identité de Parseval nous donne $\frac{\pi^4}{18} + \sum_{p=1}^{+\infty} \frac{1}{p^4} = \frac{2}{\pi} \int_0^{\pi} t^2(\pi - t)^2 dt = \frac{\pi^4}{15}$, soit

$$\sum_{p=1}^{+\infty} \frac{1}{p^4} = \frac{\pi^4}{90}.$$

Du théorème de projection orthogonale, on déduit le résultat suivant valable en dimension finie.

Corollaire 12.2. Pour tout sous-espace vectoriel F de dimension finie de E on a $E = F \oplus F^\perp$ et $(F^\perp)^\perp = F$.

Preuve. Pour tout $x \in F \cap F^\perp$, on a $\|x\|^2 = \langle x | x \rangle = 0$ et $x = 0$. Donc $F \cap F^\perp = \{0\}$. Soit $x \in E$ et $y \in F$ sa projection orthogonale dans F . On a $x - y \in F^\perp$ et $x = y + (x - y) \in F + F^\perp$. D'où l'égalité $E = F \oplus F^\perp$. Il en résulte que $\dim(F^\perp) = \dim(E) - \dim(F)$. On a donc $\dim((F^\perp)^\perp) = \dim(F)$ et avec l'inclusion $F \subset (F^\perp)^\perp$, on déduit qu'on a l'égalité. $\square$

Pour F de dimension infinie, on a toujours $F \cap F^\perp = \{0\}$ mais pas nécessairement $E = F \oplus F^\perp$, ni même $(F^\perp)^\perp = F$. On considère par exemple l'espace vectoriel $E = C^0([0, 1], \mathbb{R})$ muni du produit scalaire $\langle f | g \rangle = \int_0^1 f(t)g(t) dt$. Pour $F = \mathbb{R}[x]$, du théorème de Weierstrass on déduit que $F^\perp = \{0\}$ et pourtant on a $E \neq F \oplus F^\perp$ et $(F^\perp)^\perp = E \neq F$.

Avec le théorème qui suit, on donne les principales propriétés des projections orthogonales.

Théorème 12.9.

Soit F un sous-espace vectoriel de dimension finie de E .

1. *Pour $x \in E$, on a $x \in F$ si, et seulement si, $p_F(x) = x$.*
2. *$p_F \circ p_F = p_F$.*
3. *La projection orthogonale p_F de E sur F est une application linéaire surjective de E sur F .*
4. *Le noyau de p_F est $F^\perp$.*
5. *Pour tous x, y dans E , on a $\langle p_F(x) | y \rangle = \langle x | p_F(y) \rangle = \langle p_F(x) | p_F(y) \rangle$ (on dit que p_F est auto-adjoint).*
6. *Pour E de dimension finie, on a $p_F + p_{F^\perp} = Id$.*
7. *Pour E de dimension finie, on a $p_F \circ p_{F^\perp} = p_{F^\perp} \circ p_F = 0$.*

Preuve.

1. Si $x = p_F(x)$, on a alors $x \in F$. Réciproquement si $x \in F$, avec $x - x = 0 \in F^\perp$, on déduit que $p_F(x) = x$.
2. Résulte de $p_F(x) = x$ pour tout $x \in F$.
3. Si $(e_i)_{1 \leq i \leq n}$ est une base orthonormée de F , on a alors $p_F(x) = \sum_{k=1}^n \langle x | e_k \rangle e_k$ pour tout $x \in E$ et p_F est linéaire puisque chaque application $x \mapsto \langle x | e_k \rangle e_k$ est linéaire. L'égalité $p_F(x) = x$ pour tout $x \in F$ nous dit en particulier que p_F est surjective de E sur F .
4. Si $x \in \ker(p_F)$, on a $p_F(x) = 0$ et $x = x - p_F(x) \in F^\perp$. Réciproquement si $x = x - 0 \in F^\perp$, on a $p_F(x) = 0$ puisque $0 \in F$.
5. Pour x, y dans E , on a $p_F(x) \in F$ et $y - p_F(y) \in F^\perp$, donc :

$$\langle p_F(x) | y \rangle = \langle p_F(x) | y - p_F(y) + p_F(y) \rangle = \langle p_F(x) | p_F(y) \rangle$$

l'expression $\langle p_F(x) | p_F(y) \rangle$ étant symétrique en x, y . Il en résulte que l'on a $\langle p_F(x) | y \rangle = \langle x | p_F(y) \rangle$. En utilisant une base orthonormée $(e_i)_{1 \leq i \leq n}$ de F

on peut aussi écrire que $p_F(x) = \sum_{k=1}^n \langle x | e_k \rangle e_k$, $p_F(y) = \sum_{k=1}^n \langle y | e_k \rangle e_k$ et :

$$\langle p_F(x) | y \rangle = \sum_{k=1}^n \langle x | e_k \rangle \langle y | e_k \rangle = \langle x | p_F(y) \rangle = \langle p_F(x) | p_F(y) \rangle$$

6. Dans $E = F \oplus F^\perp$, on a les deux écritures :

$$x = (x - p_F(x)) + p_F(x) = (x - p_{F^\perp}(x)) + p_{F^\perp}(x)$$

avec $(p_F(x), x - p_F(x))$ et $(x - p_{F^\perp}(x), p_{F^\perp}(x))$ dans $F \times F^\perp$, ce qui entraîne $x - p_F(x) = p_{F^\perp}(x)$ du fait de l'unicité de l'écriture dans une somme directe. On a donc bien $p_F(x) + p_{F^\perp}(x) = x$ pour tout $x \in E$.

7. On en déduit que $p_F(x - p_{F^\perp}(x)) = p_F(p_F(x)) = p_F(x)$ et $p_F(p_{F^\perp}(x)) = 0$. L'égalité $p_{F^\perp} \circ p_F = 0$ se montre de manière analogue. $\square$

On rappelle que, sur un espace vectoriel E , un projecteur est une application linéaire p de E dans E telle que $p \circ p = p$. Il est facile de vérifier que si p est un projecteur de E , alors $\ker(p)$ et $\text{Im}(p)$ sont en somme directe et pour tout $x = y + z$ avec $(y, z) \in \ker(p) \times \text{Im}(p)$, on a $p(x) = y$. En effet, si $x \in \ker(p) \cap \text{Im}(p)$, on a $x = p(y)$ et $0 = p(x) = p \circ p(y) = p(y) = x$, donc $\ker(p) \cap \text{Im}(p) = \{0\}$ et tout $x \in E$ s'écrit $x = x - p(x) + p(x)$ avec $x - p(x) \in \ker(p)$ et $p(x) \in \text{Im}(p)$, donc $E = \ker(p) + \text{Im}(p)$. On a donc bien $E = \ker(p) \oplus \text{Im}(p)$ et pour tout $x = y + z \in E$ avec $(y, z) \in \ker(p) \times \text{Im}(p)$, on a $p(x) = p(y) + p(z) = p(z) = z$. On dit que p est le projecteur sur $F = \text{Im}(p)$ parallèlement à $\ker(p)$. Réciproquement si $E = F \oplus G$, l'application qui associe à $x = y + z$, où $(y, z) \in F \times G$, le vecteur y est un projecteur sur F parallèlement à G .

Les projecteurs orthogonaux sont des cas particuliers de projecteurs. Ce sont en fait les projecteurs de E caractérisés par la propriété **5.** du théorème 12.9 ou par $\|p(x)\| \leq \|x\|$ pour tout $x \in E$.

Théorème 12.10.

Soit p un projecteur d'un espace euclidien E . Les propriétés suivantes sont équivalentes :

1. *p est un projecteur orthogonal ;*
2. *pour tous x, y dans E , on a $\langle p(x) | y \rangle = \langle x | p(y) \rangle$;*
3. *pour tout $x \in E$, on a $\|p(x)\| \leq \|x\|$.*

Preuve. Si $p = p_F$ est un projecteur orthogonal, on sait déjà qu'il est auto-adjoint, c'est-à-dire que **1.** implique **2.**

Si p est un projecteur qui vérifie **2.** on a, en utilisant l'inégalité de Cauchy-Schwarz, pour tout $x \in E$:

$$\|p(x)\|^2 = \langle p(x) | p(x) \rangle = \langle x | p(p(x)) \rangle = \langle x | p(x) \rangle \leq \|x\| \|p(x)\|$$

et $\|p(x)\| \leq \|x\|$ pour $x \neq 0$, l'égalité étant réalisée pour $x = 0$.

Supposons que p soit un projecteur vérifiant **3.** On a $E = \ker(p) \oplus \text{Im}(p)$ et si p est un projecteur orthogonal sur F , on a nécessairement $\ker(p) = F^\perp$ et $\text{Im}(p) = F = (\ker(p))^\perp$. Réciproquement si $\text{Im}(p) = (\ker(p))^\perp$, on a alors pour tout $x \in E$, $p(x) \in F = \text{Im}(p)$ et $x - p(x) \in \ker(p) = F^\perp$, ce qui signifie que $p(x)$ est le projeté orthogonal de x sur F . Il s'agit donc de montrer que $\text{Im}(p) = (\ker(p))^\perp$. Pour $x \in \ker(p)$ et $y \in \text{Im}(p)$, en notant $z = y - \lambda x$ où λ est un réel, on a $p(z) = p(y) = y$ et :

$$\|y\|^2 = \|p(z)\|^2 \leq \|z\|^2 = \|y\|^2 - 2\lambda \langle x | y \rangle + \lambda^2 \|x\|^2$$

soit $\lambda (\lambda \|x\|^2 - 2 \langle x | y \rangle) \geq 0$, ce qui entraîne $\lambda \|x\|^2 - 2 \langle x | y \rangle \geq 0$ pour $\lambda > 0$ et $\lambda \|x\|^2 - 2 \langle x | y \rangle \leq 0$ pour $\lambda < 0$. Faisant tendre λ vers 0 par valeurs positives et négatives respectivement, on obtient $\langle x | y \rangle \leq 0$ et $\langle x | y \rangle \geq 0$, soit $\langle x | y \rangle = 0$. Le projecteur p est donc un projecteur orthogonal. $\square$

12.5 Les endomorphismes symétriques réels

Théorème 12.11.

Les valeurs propres d'une matrice symétrique réelle A sont toutes réelles.

Preuve. Soient λ une valeur propre complexe de A et $x \in \mathbb{C}^n$ un vecteur propre non nul associé. On a alors $Ax = \lambda x$, qui par conjugaison complexe donne $A\bar{x} = \bar{\lambda}\bar{x}$ puisque A est réelle et ${}^t x A \bar{x} = \bar{\lambda} {}^t x \bar{x} = \bar{\lambda} \sum_{k=1}^n |x_k|^2$. Mais on a aussi, du fait que A est symétrique :

$${}^t x A \bar{x} = {}^t (Ax) \bar{x} = {}^t (\lambda x) \bar{x} = \lambda {}^t x \bar{x} = \lambda \sum_{k=1}^n |x_k|^2$$

et en conséquence $\bar{\lambda} = \lambda$ puisque $\sum_{k=1}^n |x_k|^2 \neq 0$. □

Pour ce paragraphe, $(E, \langle \cdot | \cdot \rangle)$ est un espace réel euclidien de dimension $n \geq 1$. On rappelle qu'un endomorphisme $u \in \mathcal{L}(E)$ est dit symétrique si :

$$\forall (x, y) \in E \times E, \langle u(x) | y \rangle = \langle x | u(y) \rangle$$

Un endomorphisme $u \in \mathcal{L}(E)$ est symétrique si, et seulement si, pour toute base orthonormée $\mathcal{B}$ de E la matrice A de u dans $\mathcal{B}$ est symétrique, c'est-à-dire que ${}^t A = A$.

On note $\mathcal{S}(E)$ l'ensemble des endomorphismes symétriques de E et $\mathcal{S}_n(\mathbb{R})$ l'ensemble des matrices symétriques réelles.

Lemme 12.1 *On suppose que $n = 2$. Un endomorphisme symétrique réel u a deux valeurs propres réelles distinctes ou confondues et se diagonalise dans une base orthonormée.*

Preuve. Soit $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathcal{S}_2(\mathbb{R})$ la matrice de u dans une base orthonormée. Le polynôme caractéristique de A est :

$$P_A(\lambda) = \lambda^2 - \text{Tr}(A)\lambda + \det(A) = \lambda^2 - (a+c)\lambda + (ac-b^2)$$

et son discriminant est $\delta = (a+c)^2 - 4(ac-b^2) = (a-c)^2 + 4b^2 \geq 0$. On a donc une valeur propre réelle double pour $a = c$ et $b = 0$ et deux valeurs propres réelles distinctes pour $a \neq c$ ou $b \neq 0$. Pour $a = c$ et $b = 0$, u est l'homothétie de rapport a et toute base orthonormée est une base de diagonalisation pour u . Sinon, l'endomorphisme u a deux valeurs propres réelles distinctes $\lambda_1 = \frac{a+c-\sqrt{\delta}}{2}$ et $\lambda_2 = \frac{a+c+\sqrt{\delta}}{2}$. Il est donc diagonalisable et désignant par (e_1, e_2) une base de vecteurs propres associée formée de vecteurs unitaires, avec :

$$\lambda_1 \langle e_1 | e_2 \rangle = \langle u(e_1) | e_2 \rangle = \langle e_1 | u(e_2) \rangle = \lambda_2 \langle e_1 | e_2 \rangle$$

et $\lambda_1 \neq \lambda_2$, on déduit que $\langle e_1 | e_2 \rangle = 0$, c'est-à-dire que (e_1, e_2) est une base orthonormée de vecteurs propres. $\square$

Lemme 12.2 Soit $u \in \mathcal{S}(E)$. Si F est un sous-espace vectoriel de E stable par u , alors $F^\perp$ est aussi stable par u .

Preuve. Pour tout $x \in F^\perp$ et tout $y \in F$, on a $u(y) \in F$ et $\langle u(x) | y \rangle = \langle x | u(y) \rangle = 0$, ce qui prouve que $u(x) \in F^\perp$. $\square$

Lemme 12.3 Pour tout endomorphisme u de E il existe un sous-espace vectoriel P de E de dimension 1 ou 2 stable par u .

Preuve. Si u a une valeur propre réelle λ , alors pour tout vecteur propre associé $x \in E \setminus \{0\}$, la droite $D = \mathbb{R}x$ est stable par u . Sinon $n \geq 2$ et le polynôme minimal π_u se décompose dans $\mathbb{R}[X]$ en produit de facteurs irréductibles de degré 2 (les valeurs propres de u sont les racines de π_u). Ce polynôme π_u s'écrit donc $\pi_u(X) = (X^2 + bX + c)Q(X)$ avec $b^2 - 4c < 0$ et $Q(u) \neq 0$ (π_u est le polynôme unitaire de plus petit degré annulant u). De l'égalité $0 = \pi_u(u) = (u^2 + bu + cId) \circ Q(u)$, on déduit alors que l'endomorphisme $u^2 + bu + cId$ n'est pas injectif, c'est-à-dire que son noyau n'est pas réduit à $\{0\}$. Pour tout vecteur x non nul dans ce noyau on vérifie alors que $P = \text{Vect}\{x, u(x)\}$ est un sous-espace vectoriel de dimension 2 stable par u . En effet, P est de dimension 2 puisque x n'est pas vecteur propre de u et avec $u^2(x) + bu(x) + cx = 0$ on déduit que $u^2(x)$ est dans P , ce qui entraîne que P est stable par u . $\square$

Théorème 12.12. Théorème spectral

Un endomorphisme symétrique réel $u \in \mathcal{S}(E)$ a n valeurs propres réelles distinctes ou confondues et se diagonalise dans une base orthonormée.

Preuve. C'est vrai pour $n = 1$ et $n = 2$. Supposant le résultat acquis jusqu'au rang $n - 1 \geq 1$ et soit $u \in \mathcal{S}(E)$ avec $\dim(E) = n$. En désignant par P un sous-espace vectoriel de E de dimension 1 ou 2 stable par u , on a $E = P \oplus P^\perp$, l'espace $P^\perp$ étant également stable par u . Comme les restrictions de u à P et à $P^\perp$ sont des endomorphismes symétriques, elles ont toutes leurs valeurs propres réelles et il existe une base orthonormée de diagonalisation $\mathcal{B}_1$ de $u|_P$ et une autre $\mathcal{B}_2$ de $u|_{P^\perp}$. La réunion $\mathcal{B}_1 \cup \mathcal{B}_2$ est alors une base orthonormée de diagonalisation de u et u a toutes ses valeurs propres réelles. $\square$

Corollaire 12.3. Toute matrice symétrique réelle $A \in \mathcal{S}_n(\mathbb{R})$ se diagonalise dans une base orthonormée, c'est-à-dire qu'il existe une matrice orthogonale $P \in \mathcal{O}_n(\mathbb{R})$ et une matrice diagonale D telles que ${}^tPAP$ soit diagonale.

12.6 Espaces vectoriels hermitiens

Pour ce paragraphe, E est un espace vectoriel sur $\mathbb{C}$ de dimension $n \geq 1$.

Définition 12.7. On appelle *produit scalaire hermitien* sur E toute application

$$\begin{aligned} \varphi : E \times E &\rightarrow \mathbb{C} \\ (x, y) &\mapsto \varphi(x, y) \end{aligned}$$

qui vérifie les propriétés suivantes :

- pour tout y dans E l'application $x \mapsto \varphi(x, y)$ est linéaire ;
- $\varphi(y, x) = \overline{\varphi(x, y)}$ pour tous x, y dans E ;
- $\varphi(x, x) \geq 0$ pour tout x dans E ;
- pour x dans E , $\varphi(x, x) = 0$ équivaut à $x = 0$.

On note en général $(x, y) \mapsto \langle x | y \rangle$ un tel produit scalaire.

Définition 12.8. Un *espace hermitien* est un $\mathbb{C}$ -espace vectoriel de dimension finie muni d'un produit scalaire hermitien.

Dans la suite de ce paragraphe E désigne un espace hermitien et on note pour tout x dans E , $\|x\| = \sqrt{\langle x | x \rangle}$.

Théorème 12.13. Inégalité de Cauchy-Schwarz

Pour tous x, y dans E on a $|\langle x | y \rangle| \leq \|x\| \|y\|$, l'égalité étant réalisée si, et seulement si, x et y sont liés.

Preuve. Pour x, y fixés dans E on pose $\langle x | y \rangle = \rho e^{i\theta}$ avec $\rho = |\langle x | y \rangle| \in \mathbb{R}^+$ et $\theta \in [0, 2\pi[$. On désigne par P la fonction polynomiale définie par :

$$\forall t \in \mathbb{R}, P(t) = \|x + te^{i\theta}y\|^2 = \|y\|^2 t^2 + 2t|\langle x | y \rangle| + \|x\|^2$$

On a $P(t) \geq 0$ pour tout réel t et nécessairement le discriminant de ce polynôme de degré 2 est toujours négatif ou nul, soit $|\langle x | y \rangle|^2 - \|x\|^2 \|y\|^2 \leq 0$, ce qui équivaut à $|\langle x | y \rangle| \leq \|x\| \|y\|$. Si x et y sont liés, il est alors clair que l'égalité est réalisée. Réciproquement si l'égalité est réalisée, le polynôme P admet alors une racine réelle λ et $x + \lambda e^{i\theta}y = 0$, c'est-à-dire que x et y sont liés. $\square$

Théorème 12.14. Inégalité de Minkowski

Pour tous x, y dans E on a $\|x + y\| \leq \|x\| + \|y\|$, l'égalité étant réalisée si, et seulement si, $x = \lambda y$ avec $\lambda \geq 0$ (on dit que x et y sont positivement liés).

Preuve. On a $\|x + y\|^2 = \|x\|^2 + 2 \operatorname{Re} \langle x | y \rangle + \|y\|^2$ et avec l'inégalité de Cauchy-Schwarz :

$$\|x + y\|^2 \leq \|x\|^2 + 2 \|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2$$

ce qui équivaut à $\|x + y\| \leq \|x\| + \|y\|$. L'égalité est réalisée si, et seulement si, $\operatorname{Re} \langle x | y \rangle = |\langle x | y \rangle| = \|x\| \|y\|$, ce qui entraîne $|\langle x | y \rangle| = \|x\| \|y\|$ et il existe un nombre complexe λ tel que $x = \lambda y$. Pour $x = 0$, l'égalité est réalisée et pour $x \neq 0$, on a $\operatorname{Re} \langle x | y \rangle = \operatorname{Re} (\lambda \|y\|^2) = |\lambda| \|y\|^2$ avec $\|y\| \neq 0$, donc $\operatorname{Re}(\lambda) = |\lambda|$ et λ est un réel positif. $\square$

Comme dans le cas euclidien, on dit que deux vecteurs x et y dans E sont orthogonaux si $\langle x | y \rangle = 0$.

Le théorème de Pythagore est encore vrai sous la forme : si les vecteurs x et y sont orthogonaux dans E , on a alors $\|x + y\|^2 = \|x\|^2 + \|y\|^2$. La réciproque étant fautive dans le cas hermitien (l'égalité $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ entraîne seulement $\operatorname{Re} \langle x | y \rangle = 0$).

On définit, comme dans le cas euclidien, les notions de famille orthogonale et orthonormale. Une famille orthogonale de vecteurs non nuls de E est libre.

Définition 12.9. L'orthogonal d'une partie non vide X de E est l'ensemble $X^\perp = \{y \in E \mid \forall x \in X, \langle x | y \rangle = 0\}$.

Il est facile de vérifier que $X^\perp$ est un sous-espace vectoriel de E .

Théorème 12.15. procédé d'orthonormalisation de Gram-Schmidt

Pour toute famille libre $(x_k)_{1 \leq k \leq p}$ dans E , il existe une unique famille orthonormée $(e_k)_{1 \leq k \leq p}$ dans E telle que :

$$\forall k \in \{1, 2, \dots, p\}, \begin{cases} \operatorname{Vect}\{e_1, \dots, e_k\} = \operatorname{Vect}\{x_1, \dots, x_k\} \\ \langle x_k | e_k \rangle > 0 \end{cases}$$

Preuve. La démonstration est analogue à celle du théorème ??.

En notant $\mathcal{E} = (e_k)_{1 \leq k \leq n}$ la base canonique de $\mathbb{C}^n$ muni de sa structure hermitienne canonique avec le produit scalaire $(x, y) \mapsto \langle x | y \rangle = \sum_{k=1}^n x_k \overline{y_k}$, tout système orthonormé $\mathcal{F} = (f_k)_{1 \leq k \leq n}$ est une base de $\mathbb{C}^n$ (ce système est libre formé de n vecteurs) et la matrice de passage U de la base canonique $\mathcal{E}$ à la base $\mathcal{F}$ est telle que $U^*U = I_n$, où $U^* = ((u_{ij}^*))_{1 \leq i, j \leq n}$ désigne la matrice adjointe de la matrice complexe $U = ((u_{ij}))_{1 \leq i, j \leq n}$, définie par $u_{ij}^* = \overline{u_{ji}}$ pour $1 \leq i, j \leq n$.

En effet les colonnes de U sont formées des composantes des vecteurs f_j dans la base $\mathcal{E}$ et le coefficient d'indice (i, j) de U^*U est $\langle f_j | f_i \rangle = \delta_{ij}$.

Définition 12.10. On dit qu'une matrice carrée A à coefficients complexes d'ordre n est unitaire si $U^*U = I_n$. On note $\mathcal{U}_n(\mathbb{C})$ l'ensemble des matrices unitaires.

On peut également dire qu'une matrice unitaire est la matrice de passage de la base canonique de $\mathbb{C}^n$ à une base orthonormée où $\mathbb{C}^n$ est muni de sa structure hermitienne canonique.

12.7 Réduction des matrices normales

On note $(e_k)_{1 \leq k \leq n}$ la base canonique de $\mathbb{C}^n$ et le produit scalaire hermitien canonique de $\mathbb{C}^n$ est noté :

$$(x, y) = \left(\sum_{i=1}^n x_i e_i, \sum_{i=1}^n y_i e_i \right) \mapsto \langle x | y \rangle = \sum_{i=1}^n x_i \bar{y}_i$$

La norme associée est la norme hermitienne notée $x \mapsto \|x\|$.

Définition 12.11. Une matrice complexe A est dite hermitienne si $A^* = A$. On dit qu'elle est hermitienne positive, si on a de plus $\langle Ax | x \rangle \geq 0$ pour tout $x \in E$. On dit qu'elle est hermitienne définie positive, si on a de plus $\langle Ax | x \rangle > 0$ pour tout $x \in E \setminus \{0\}$.

Définition 12.12. Une matrice complexe A est dite normale si $A^* A = A A^*$.

Les matrices hermitiennes et unitaires sont des cas particuliers de matrices normales.

Lemme 12.4 Les valeurs propres d'une matrice hermitienne [resp. hermitienne positive] sont réelles [resp. réelles et positives]. Les valeurs propres d'une matrice unitaire sont de module égal à 1.

Preuve. Soient $A \in \mathcal{M}_n(\mathbb{C})$ hermitienne, $\lambda \in \mathbb{C}$ une valeur propre de A et $x \in \mathbb{C}^n$ un vecteur propre non nul associé. On a :

$$\lambda \|x\|^2 = \langle Ax | x \rangle = \langle x | A^* x \rangle = \langle x | Ax \rangle = \bar{\lambda} \|x\|^2$$

donc $\lambda = \bar{\lambda}$ et $\lambda \in \mathbb{R}$. Si de plus A est positive, on a alors $\lambda \|x\|^2 = \langle Ax | x \rangle \geq 0$ avec $x \neq 0$, ce qui entraîne $\lambda \geq 0$.

Soient $A \in \mathcal{M}_n(\mathbb{C})$ unitaire, $\lambda \in \mathbb{C}$ une valeur propre de A et $x \in \mathbb{C}^n \setminus \{0\}$ un vecteur propre associé. On a :

$$|\lambda|^2 \|x\|^2 = \|Ax\|^2 = \langle Ax | Ax \rangle = \langle x | A^* Ax \rangle = \|x\|^2$$

et nécessairement $|\lambda| = 1$. □

Théorème 12.16.

Si A une matrice complexe normale, elle se diagonalise alors dans une base orthonormée, c'est-à-dire qu'il existe une matrice unitaire U et une matrice diagonale D telles que $A = UDU^*$, les coefficients diagonaux de D étant les valeurs propres dans $\mathbb{C}$ de la matrice A .

Preuve. On montre tout d'abord que pour toute valeur propre $\lambda \in \mathbb{C}$ de A , le sous-espace propre associé $E_\lambda = \ker(A - \lambda I_n)$ est stable par A^* et que son orthogonal $E_\lambda^\perp$ est stable par A et par A^* , ce qui permet de raisonner par récurrence sur l'ordre n de la matrice A .

Pour tout $x \in E_\lambda$ on a, puisque A est normale, $AA^*x = A^*Ax = \lambda A^*x$ et $A^*x \in E_\lambda$, c'est-à-dire que E_λ est stable par A^* . Pour $x \in E_\lambda^\perp$ et $y \in E_\lambda$, on a $A^*y \in E_\lambda$ et $\langle Ax | y \rangle = \langle x | A^*y \rangle = 0$, c'est-à-dire que $Ax \in E_\lambda^\perp$ et $E_\lambda^\perp$ est stable par A . Puis en écrivant que :

$$\langle A^*x | y \rangle = \langle x | Ay \rangle = \langle x | \lambda y \rangle = \bar{\lambda} \langle x | y \rangle = 0$$

on déduit que $A^*x \in E_\lambda^\perp$ et $E_\lambda^\perp$ est stable par A^* . On raisonne ensuite par récurrence sur la dimension n de l'espace vectoriel hermitien E . Pour $n = 1$ le résultat est évident. Supposons le acquis pour toute matrice normale d'ordre $p \in \{1, \dots, n-1\}$. Soient λ une valeur propre de A et E_λ l'espace propre associé. Si $E_\lambda = E$, la matrice A est alors une homothétie et toute base orthonormée de E convient. Si $E_\lambda \neq E$, $E_\lambda^\perp$ est alors de dimension $p \in \{1, \dots, n-1\}$ et la restriction B de A à $E_\lambda^\perp$ est un endomorphisme normal de $E_\lambda^\perp$ (une matrice est identifiée à l'application linéaire qu'elle définit dans la base canonique de E). Il existe donc une base orthonormée de $E_\lambda^\perp$ formée de vecteurs propres de B , donc de A . En complétant cette base par une base orthonormée de E_λ on obtient une base orthonormée de E formée de vecteurs propres de A . $\square$

Si λ est une valeur propre de A , la stabilité de $E_\lambda^\perp$ par A^* est en fait vraie pour toute matrice A .

Du fait qu'une matrice hermitienne ou unitaire est normale on déduit les résultats suivants.

Corollaire 12.4. *Une matrice hermitienne [resp. unitaire] a ses valeurs propres réelles [resp. de module 1] et se diagonalise dans une base orthonormée.*

12.8 Exercices

Exercice 12.1. *L'espace vectoriel $\mathbb{R}^n$ est toujours muni de sa base canonique $(e_i)_{1 \leq i \leq n}$. Soit $\omega \in \mathbb{R}^n$. À quelle condition sur ω l'application*

$\varphi : (x, y) \mapsto \sum_{k=1}^n \omega_k x_k y_k$ *définit-elle un produit scalaire sur l'espace vectoriel $\mathbb{R}^n$?*

Solution. L'application φ définit une forme bilinéaire symétrique sur $\mathbb{R}^n$ pour tout $\omega \in \mathbb{R}^n$. Si φ est un produit scalaire, on a alors $\omega_j = \varphi(e_j, e_j) > 0$ pour tout j compris entre 1 et n . Réciproquement si tous les ω_j sont strictement positifs, on a $\varphi(x, x) = \sum_{i=1}^n \omega_i x_i^2 \geq 0$ pour tout $x \in \mathbb{R}^n$ et $\varphi(x, x) = 0$ équivaut à $\omega_i x_i^2 = 0$ pour tout i , ce qui équivaut à $x_i = 0$ pour tout i , soit à $x = 0$. En conclusion, φ définit un produit scalaire sur $\mathbb{R}^n$ si, et seulement si, tous les ω_i sont strictement positifs.

Exercice 12.2. Soient n un entier naturel non nul, $x_0, \dots, x_n$ des réels deux à deux distincts et $\omega \in \mathbb{R}^{n+1}$. À quelle condition sur ω l'application $\varphi : (P, Q) \mapsto \sum_{i=0}^n \omega_i P(x_i) Q(x_i)$ définit-elle un produit scalaire sur l'espace vectoriel $\mathbb{R}_n[X]$?

Solution. L'application φ définit une forme bilinéaire symétrique sur $\mathbb{R}_n[X]$ pour tout $\omega \in \mathbb{R}^n$. Si φ est un produit scalaire, en désignant par $(L_i)_{0 \leq i \leq n}$ la base de Lagrange de $\mathbb{R}_n[X]$ définie par $L_i(X) = \prod_{\substack{k=0 \\ k \neq i}}^n \frac{X - x_k}{x_i - x_k}$ pour $0 \leq i \leq n$ (L_i est le polynôme de degré n qui vaut 1 en x_i et 0 en x_k pour $k \neq i$), on a alors $\omega_j = \varphi(L_j, L_j) > 0$ pour tout j compris entre 1 et n . Réciproquement si tous les ω_j sont strictement positifs, on a $\varphi(P, P) = \sum_{i=1}^n \omega_i P^2(x_i) \geq 0$ pour tout $x \in \mathbb{R}^n$ et $\varphi(P, P) = 0$ équivaut à $\omega_i P^2(x_i) = 0$ pour tout i , ce qui équivaut à $P(x_i) = 0$ pour tout i compris entre 0 et n soit à $P = 0$ (P est un polynôme dans $\mathbb{R}_n[X]$ qui a $n+1$ racines distinctes, c'est donc le polynôme nul). En conclusion, φ définit un produit scalaire sur $\mathbb{R}_n[X]$ si, et seulement si, tous les ω_i sont strictement positifs.

Exercice 12.3. n étant un entier naturel non nul, on note $\mathcal{P}_n$ l'ensemble des polynômes trigonométriques de degré inférieur ou égal à n , c'est-à-dire l'ensemble des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ la forme :

$$P : x \mapsto P(x) = a_0 + \sum_{k=1}^n (a_k \cos(kx) + b_k \sin(kx))$$

1. Montrer que $\mathcal{P}_n$ est un espace vectoriel et préciser sa dimension.
2. Montrer que si $P \in \mathcal{P}_n$ s'annule en $2n+1$ points deux à deux distincts dans $[-\pi, \pi[$, alors $P = 0$ (utiliser les expressions complexes des fonctions $\cos$ et $\sin$).

3. Montrer que si $x_0, \dots, x_{2n}$ sont des réels deux à deux distincts dans $[-\pi, \pi[$, alors l'application $\varphi : (P, Q) \mapsto \sum_{i=0}^{2n} P(x_i) Q(x_i)$ définit un produit scalaire sur l'espace vectoriel $\mathcal{P}_n$.

Solution.

1. Il est clair que $\mathcal{P}_n$ est un sous-espace vectoriel de l'espace vectoriel des applications de $\mathbb{R}$ dans $\mathbb{R}$. En notant respectivement c_k et s_k les fonctions $x \mapsto \cos(kx)$ pour $k \geq 0$ et $x \mapsto \sin(kx)$ pour $k \geq 1$, $\mathcal{P}_n$ est engendré par la famille $\mathcal{B}_n = \{c_k \mid 0 \leq k \leq n\} \cup \{s_k \mid 1 \leq k \leq n\}$, c'est donc un espace vectoriel de dimension au plus égale à $2n + 1$. Montrons que cette famille de fonctions est libre. Pour ce faire, on procède par récurrence sur $n \geq 1$ (comme avec l'exercice 7.2). Pour $n = 1$, si $a_0 + a_1 \cos(x) + b_1 \sin(x) = 0$, en évaluant cette fonction en $0, \frac{\pi}{2}$ et π successivement, on aboutit au système linéaire :

$$\begin{cases} a_0 + a_1 = 0 \\ a_0 + b_1 = 0 \\ a_0 - a_1 = 0 \end{cases}$$

qui équivaut à $a_0 = b_0 = b_1 = 0$. La famille $\{c_0, c_1, s_1\}$ est donc libre. Supposons le résultat acquis au rang $n - 1 \geq 1$. Si $P = a_0 + \sum_{k=1}^n (a_k c_k + b_k s_k) = 0$, en dérivant deux fois, on a $P'' = -\sum_{k=1}^n k^2 (a_k c_k + b_k s_k) = 0$. Il en résulte que :

$$n^2 P + P'' = n^2 a_0 + \sum_{k=1}^{n-1} (n^2 - k^2) (a_k c_k + b_k s_k) = 0$$

et l'hypothèse de récurrence nous dit que $n^2 a_0 = 0$, $(n^2 - k^2) a_k = 0$ et $(n^2 - k^2) b_k$ pour tout k compris entre 1 et $n - 1$, ce qui équivaut à dire que $a_0 = 0$ et $a_k = b_k = 0$ pour tout k compris entre 1 et $n - 1$ puisque $n^2 - k^2 \neq 0$. Il reste alors $a_n c_n + b_n s_n = 0$, ce qui implique $a_n = 0$, en évaluant en $x = 0$ et $b_n = 0$. La famille $\mathcal{B}_n$ est donc libre. On verra un peu plus loin que cette famille est orthogonale, formée de fonctions non nulles, et en conséquence libre (exercice 12.2).

2. Posant $z = e^{ix}$ pour tout réel x , on a :

$$\begin{aligned} P(x) &= a_0 + \sum_{k=1}^n \left(a_k \frac{z^k + \bar{z}^k}{2} + b_k \frac{z^k - \bar{z}^k}{2i} \right) \\ &= a_0 + \frac{1}{2} \sum_{k=1}^n \left(a_k \left(z^k + \frac{1}{z^k} \right) - i b_k \left(z^k - \frac{1}{z^k} \right) \right) \\ &= a_0 + \frac{1}{2} \sum_{k=1}^n \left((a_k - i b_k) z^k + (a_k + i b_k) \frac{1}{z^k} \right) \end{aligned}$$

ou encore :

$$z^n P(x) = a_0 z^{2n} + \frac{1}{2} \sum_{k=1}^n ((a_k - ib_k) z^{n+k} + (a_k + ib_k) z^{n-k}) = Q(z)$$

Il en résulte que si P s'annule en $2n+1$ points deux à deux distincts, $x_0, \dots, x_{2n}$, dans $[-\pi, \pi[$, alors le polynôme complexe $Q \in \mathbb{C}_{2n}[z]$ s'annule en $2n+1$ points distincts du cercle unité, $e^{ix_0}, \dots, e^{ix_{2n}}$, ce qui revient à dire que c'est le polynôme nul et $P = 0$.

3. On vérifie facilement que φ est une forme bilinéaire symétrique et positive. L'égalité $\varphi(P, P) = 0$ entraîne que $P \in \mathcal{P}_n$ s'annule en $2n+1$ points deux à deux distincts dans $[-\pi, \pi[$ et en conséquence $P = 0$.

Exercice 12.4. Montrer que, pour toute fonction $\alpha \in \mathcal{C}^0([a, b], \mathbb{R}_+^*)$, l'application $\varphi : (f, g) \mapsto \int_a^b f(t) g(t) \alpha(t) dt$ définit un produit scalaire sur l'espace vectoriel $E = \mathcal{C}^0([a, b], \mathbb{R})$.

Solution. Avec la structure de corps commutatif de $\mathbb{R}$ et la linéarité et positivité de l'intégrale, on déduit que φ est une forme bilinéaire symétrique positive sur E . Sachant que l'intégrale sur $[a, b]$ d'une fonction continue et à valeurs positives est nulle si, et seulement si, cette fonction est nulle, on déduit que φ est une forme définie. Donc φ est un produit scalaire sur E .

Exercice 12.5. Montrer que pour tout $n \in \mathbb{N}^*$, on a :

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

Solution. L'inégalité de Cauchy-Schwarz nous donne $\sum_{k=1}^n k\sqrt{k} \leq \sqrt{\sum_{k=1}^n k^2} \sqrt{\sum_{k=1}^n k}$ avec $\sum_{k=1}^n k = \frac{n(n+1)}{2}$ et $\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$, ce qui donne :

$$\sum_{k=1}^n k\sqrt{k} \leq \sqrt{\frac{n^2(n+1)^2(2n+1)}{12}} = \frac{n(n+1)}{2\sqrt{3}} \sqrt{2n+1}$$

Exercice 12.6. On se donne un entier $n \geq 1$ et des réels $x_1, \dots, x_n$ strictement positifs. Montrer que :

$$\left(\sum_{k=1}^n x_k \right) \left(\sum_{k=1}^n \frac{1}{x_k} \right) \geq n^2 \text{ et } \sum_{k=1}^n \frac{1}{k^2} \geq \frac{6n}{(n+1)(2n+1)}$$

Solution.

1. L'inégalité de Cauchy-Schwarz nous donne $n = \sum_{k=1}^n \sqrt{x_k} \frac{1}{\sqrt{x_k}} \leq \sqrt{\sum_{k=1}^n x_k} \sqrt{\sum_{k=1}^n \frac{1}{x_k}}$, encore équivalent à l'inégalité proposée. L'égalité est réalisée si, et seulement si, il existe un réel λ tel que $\sqrt{x_k} = \lambda \frac{1}{\sqrt{x_k}}$ pour tout k compris entre 1 et n , ce qui équivaut à $x_k = \lambda$ pour tout k compris entre 1 et n , où λ est un réel strictement positif.
2. Prenant $x_k = k^2$ pour $1 \leq k \leq n$, on en déduit que $\left(\sum_{k=1}^n k^2 \right) \left(\sum_{k=1}^n \frac{1}{k^2} \right) \geq n^2$ et avec $\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$, on en déduit que :

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

Exercice 12.7.

1. Montrer que pour tous réels a, b et λ , on a :

$$(2\lambda - 1)a^2 - 2\lambda ab = \lambda(a - b)^2 - \lambda b^2 + (\lambda - 1)a^2$$

2. Soit q la forme quadratique définie sur $\mathbb{R}^n$ par :

$$q(x) = \sum_{k=1}^n (2k - 1)x_k^2 - 2 \sum_{k=1}^{n-1} kx_k x_{k+1}$$

Effectuer une réduction de q en combinaison linéaire de carrés de formes linéaires indépendantes. Préciser le rang le noyau et la signature de q .

3. On note $(x, y) = (x_1, \dots, x_n, y_1, \dots, y_n)$ un vecteur de $H = \mathbb{R}^{2n}$ et

Q la forme quadratique définie sur H par $Q(x, y) = \sum_{k=1}^n (y_k^2 - 2x_k y_k)$.

Effectuer une réduction de Q en combinaison linéaire de carrés de formes linéaires indépendantes. Préciser le rang le noyau et la signature de Q .

4. Pour $n \geq 1$ et $x = (x_1, \dots, x_n)$ dans $\mathbb{R}^n$, on définit $y = (y_1, \dots, y_n)$ par

$$y_k = \frac{1}{k} \sum_{j=1}^k x_j.$$

- (a) Montrer que $x_1 = y_1$ et $x_k = ky_k - (k-1)y_{k-1}$ pour tout k compris entre 2 et n .
- (b) Montrer que $Q(x, y) = -q(y)$.
- (c) En déduire $\sum_{k=1}^n y_k^2 \leq \sum_{k=1}^n 2x_k y_k$, puis montrer que $\sum_{k=1}^n y_k^2 \leq 4 \sum_{k=1}^n x_k^2$.
- (d) En déduire que si $(x_n)_{n \geq 1}$ est une suite de réels telle que la série $\sum x_n^2$ soit convergente et si $(y_n)_{n \geq 1}$ est la suite des moyennes de Cesàro définie par $y_n = \frac{1}{n} \sum_{j=1}^n x_j$ pour tout $n \geq 1$, alors la série $\sum y_n^2$ est convergente et $\sum_{n=1}^{+\infty} y_n^2 \leq 4 \sum_{n=1}^{+\infty} x_n^2$.

Solution.

1. On a :

$$\begin{aligned} (2\lambda - 1)a^2 - 2\lambda ab &= \lambda(a^2 - 2ab) + (\lambda - 1)a^2 = \lambda((a-b)^2 - b^2) + (\lambda - 1)a^2 \\ &= \lambda(a-b)^2 - \lambda b^2 + (\lambda - 1)a^2 \end{aligned}$$

2. En utilisant le résultat précédent, on a :

$$\begin{aligned} q(x) &= (2n-1)x_n^2 + \sum_{k=1}^{n-1} ((2k-1)x_k^2 - 2kx_k x_{k+1}) \\ &= (2n-1)x_n^2 + \sum_{k=1}^{n-1} k(x_k - x_{k+1})^2 - \sum_{k=1}^{n-1} kx_{k+1}^2 + \sum_{k=1}^{n-1} (k-1)x_k^2 \\ &= (2n-1)x_n^2 + \sum_{k=1}^{n-1} k(x_k - x_{k+1})^2 + \sum_{k=1}^{n-2} kx_{k+1}^2 - \sum_{k=1}^{n-1} kx_{k+1}^2 \\ &= (2n-1)x_n^2 + \sum_{k=1}^{n-1} k(x_k - x_{k+1})^2 - (n-1)x_n^2 = \sum_{k=1}^{n-1} k(x_k - x_{k+1})^2 + nx_n^2 \end{aligned}$$

soit $q(x) = \sum_{k=1}^n k\ell_k^2(x)$, où les formes linéaires ℓ_k sont définies par :

$$\begin{cases} \ell_k(x) = x_k - x_{k+1} & (1 \leq k \leq n-1) \\ \ell_n(x) = x_n \end{cases}$$

Il est facile de vérifier que ces formes linéaires sont indépendantes. On en déduit que $\text{rg}(q) = n$, $\ker(q) = \{0\}$ et $\text{sgn}(q) = (n, 0)$.

3. On a $Q(x, y) = \sum_{k=1}^n (y_k - x_k)^2 - \sum_{k=1}^n x_k^2$, soit :

$$Q(x, y) = \sum_{k=1}^n L_k^2(x, y) - \sum_{k=n+1}^{2n} L_k^2(x, y)$$

où les formes linéaires L_k sont définies par :

$$\begin{cases} L_k(x, y) = y_k - x_k & (1 \leq k \leq n) \\ L_k(x, y) = x_k & (k+1 \leq k \leq 2n) \end{cases}$$

Il est facile de vérifier que ces formes linéaires sont indépendantes. On en déduit que $\text{rg}(Q) = 2n$, $\ker(Q) = \{0\}$ et $\text{sgn}(Q) = (n, n)$.

4.

(a) On a $x_1 = y_1$ et pour $k \geq 2$, de $ky_k = \sum_{j=1}^k x_j$, on déduit que l'on a

$$x_k = ky_k - (k-1)y_{k-1}.$$

(b) En posant $y_{-1} = 0$, on a :

$$\begin{aligned} Q(x, y) &= \sum_{k=1}^n (y_k - 2x_k) y_k = \sum_{k=1}^n (y_k - 2(ky_k - (k-1)y_{k-1})) y_k \\ &= \sum_{k=1}^n ((1-2k)y_k + 2(k-1)y_{k-1}) y_k \\ &= \sum_{k=1}^n (1-2k)y_k^2 + 2 \sum_{k=1}^n (k-1)y_{k-1}y_k \end{aligned}$$

$$\text{soit } Q(x, y) = \sum_{k=1}^n (1-2k)y_k^2 + 2 \sum_{k=1}^{n-1} ky_k y_{k+1} = -q(y).$$

(c) Comme q est positive, on a $Q(x, y) = \sum_{k=1}^n (y_k^2 - 2x_k y_k) = -q(y) \leq 0$, soit

$\sum_{k=1}^n y_k^2 \leq \sum_{k=1}^n 2x_k y_k$. En utilisant l'inégalité de Cauchy-Schwarz dans $\mathbb{R}^n$, on a :

$$\sum_{k=1}^n y_k^2 \leq 2 \sum_{k=1}^n x_k y_k \leq 2 \sqrt{\sum_{k=1}^n x_k^2} \sqrt{\sum_{k=1}^n y_k^2}$$

$$\text{et } \sum_{k=1}^n y_k^2 \leq 4 \sum_{k=1}^n x_k^2.$$

(d) Résulte de ce qui précède.

Exercice 12.8. Soit $f \in \mathcal{C}^0([a, b], \mathbb{R})$. Montrer que $\left(\int_a^b f(t) dt\right)^2 \leq (b-a) \int_a^b f^2(t) dt$. Dans quel cas a-t-on égalité ?

Solution. L'inégalité de Cauchy-Schwarz nous donne :

$$\left(\int_a^b f(t) \cdot 1 dt\right)^2 \leq \int_a^b f^2(t) dt \int_a^b 1 dt = (b-a) \int_a^b f^2(t) dt$$

l'égalité étant réalisée si, et seulement si, la fonction f est constante.

Exercice 12.9. Soit $f \in \mathcal{C}^0([a, b], \mathbb{R}_+^*)$. Montrer que $\left(\int_a^b \frac{1}{f(t)} dt\right) \left(\int_a^b f(t) dt\right) \geq (b-a)^2$. Dans quel cas a-t-on égalité ?

Solution. L'inégalité de Cauchy-Schwarz nous donne :

$$(b-a)^2 = \left(\int_a^b \sqrt{f}(t) \cdot \frac{1}{\sqrt{f}(t)} dt\right)^2 \leq \int_a^b f(t) dt \int_a^b \frac{1}{f(t)} dt$$

l'égalité étant réalisée si, et seulement si, il existe un réel λ tel que $\sqrt{f} = \lambda \frac{1}{\sqrt{f}}$, ce qui équivaut à dire que f est constante.

Exercice 12.10. Soit $f \in \mathcal{C}^1([a, b], \mathbb{R})$ telle que $f(a) = 0$. Montrer que $\int_a^b |f(t)|^2 dt \leq \frac{(b-a)^2}{2} \int_a^b |f'(t)|^2 dt$.

Solution. Pour tout $t \in]a, b]$, on a :

$$|f(t)|^2 = \left(\int_a^t f'(x) dx\right)^2 \leq \int_a^t 1 dx \int_a^t |f'(x)|^2 dx \leq (t-a) \int_a^t |f'(t)|^2 dt$$

et en intégrant :

$$\int_a^b |f(t)|^2 dt \leq \int_a^b |f'(t)|^2 dt \int_a^b (t-a) dt = \frac{(b-a)^2}{2} \int_a^b |f'(t)|^2 dt$$

Pour $f \in \mathcal{C}^1([a, b], \mathbb{R})$ sans hypothèse sur $f(a)$, la fonction g définie par $g(x) = f(x) - f(a)$ vérifie les hypothèses de l'exercice et avec $g(a) = 0$, $g' = f'$, on obtient l'inégalité :

$$\int_a^b |f(t) - f(a)|^2 dt \leq \frac{(b-a)^2}{2} \int_a^b |f'(t)|^2 dt$$

Exercice 12.11. Étant donnée une famille $(x_i)_{0 \leq i \leq n}$ de $n+1$ réels deux à deux distincts, on munit $\mathbb{R}_n[X]$ du produit scalaire :

$$(P, Q) \mapsto \langle P | Q \rangle = \sum_{i=0}^n P(x_i) Q(x_i)$$

Montrer que la famille $(L_i)_{0 \leq i \leq n}$ des polynômes de Lagrange définie par
$$L_i(x) = \prod_{\substack{k=0 \\ k \neq i}}^n \frac{x - x_k}{x_i - x_k}$$
 pour $0 \leq i \leq n$ est une base orthonormée de $\mathbb{R}_n[X]$.

Solution. Pour i, j compris entre 1 et n , on a :

$$\langle L_i | L_j \rangle = \sum_{k=0}^n L_i(x_k) L_j(x_k) = L_i(x_j) = \delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

La famille $(L_i)_{0 \leq i \leq n}$ est orthonormée, donc libre et comme elle est formée de $n+1$ polynômes, c'est une base de $\mathbb{R}_n[X]$.

Exercice 12.12. Soient $(E, \langle \cdot | \cdot \rangle)$ un espace euclidien et u un automorphisme de E .

1. Montrer que l'application $\varphi : (x, y) \mapsto \langle u(x) | u(y) \rangle$ définit un produit scalaire sur E .
2. Dans le cas où E est de dimension finie, donner la matrice de φ dans une base orthonormée de E en fonction de celle de u .

Solution.

1. De la linéarité de u , on déduit que φ est bilinéaire symétrique. Pour $x \in E$, on a $\varphi(x, x) = \|u(x)\|^2 \geq 0$ et $\varphi(x, x) = 0$ équivaut à $x \in \ker(u)$, soit à $x = 0$ puisque u est bijectif. Donc φ est un produit scalaire sur E .
2. Soit A la matrice de u dans une base orthonormée de E . Pour $(x, y) \in E^2$, on a $\varphi(x, y) = \langle u(x) | u(y) \rangle = {}^t(AX)(AY) = {}^tX({}^tAA)Y$ et la matrice de φ dans $\mathcal{B}$ est ${}^tAA$.

Exercice 12.13. Montrer que l'application :

$$(P, Q) \mapsto \langle P | Q \rangle = \int_0^2 (2-t) P(t) Q(t) dt$$

définit un produit scalaire sur $\mathbb{R}_2[X]$. Donner une base orthonormée.

Solution. La fonction $t \mapsto 2-t$ étant à valeurs strictement positives sur $]0, 2[$, il est facile de vérifier que $\langle \cdot | \cdot \rangle$ est un produit scalaire. En utilisant l'algorithme de

Gram-Schmidt, on définit la base orthonormée $(P_i)_{0 \leq i \leq 2}$ par :

$$\begin{cases} Q_0 = 1, \|Q_0\|^2 = \int_0^2 (2-t) dt = 2, P_0 = \frac{1}{\sqrt{2}} \\ Q_1 = x - \langle x | P_0 \rangle P_0 = x - \frac{2}{3}, \|Q_1\|^2 = \langle Q_1 | x \rangle = \frac{4}{9}, P_1 = \frac{3}{2}x - 1 \\ Q_2 = x^2 - \langle x^2 | P_0 \rangle P_0 - \langle x^2 | P_1 \rangle P_1 = x^2 - \frac{8}{5}x + \frac{2}{5} \\ \|Q_2\|^2 = \langle Q_2 | x^2 \rangle = \frac{8}{75}, P_2 = \frac{\sqrt{6}}{4} (5x^2 - 8x + 2) \end{cases}$$

Une base orthonormée de $\mathbb{R}_2[X]$ est donc $\left(\frac{1}{\sqrt{2}}, \frac{3}{2}x - 1, \frac{\sqrt{6}}{4} (5x^2 - 8x + 2) \right)$.

Exercice 12.14. Montrer que l'application $(P, Q) \mapsto \int_{-1}^1 P(t) Q(t) dt$ définit un produit scalaire sur $\mathbb{R}_2[X]$. Donner la matrice dans la base canonique et déterminer une base orthonormée.

Solution. On sait déjà que $\langle \cdot | \cdot \rangle$ est un produit scalaire. En utilisant l'algorithme de Gram-Schmidt, on définit la base orthonormée $(P_i)_{0 \leq i \leq 2}$ par :

$$\begin{cases} Q_0 = 1, \|Q_0\|^2 = \int_{-1}^1 dt = 2, P_0 = \frac{1}{\sqrt{2}} \\ Q_1 = x - \langle x | P_0 \rangle P_0 = x, \\ \|Q_1\|^2 = \langle Q_1 | x \rangle = \frac{2}{3}, P_1 = \frac{\sqrt{3}}{\sqrt{2}}x, \\ Q_2 = x^2 - \langle x^2 | P_0 \rangle P_0 - \langle x^2 | P_1 \rangle P_1 = x^2 - \frac{1}{3}, \\ \|Q_2\|^2 = \langle Q_2 | x^2 \rangle = \frac{8}{45}, P_2 = \frac{3}{4}\sqrt{10} \left(x^2 - \frac{1}{3} \right) \end{cases}$$

Une base orthonormée de $\mathbb{R}_2[X]$ est donc $\left(\frac{1}{\sqrt{2}}, \frac{\sqrt{3}}{\sqrt{2}}x, \frac{3}{4}\sqrt{10} \left(x^2 - \frac{1}{3} \right) \right)$.

Exercice 12.15. Pour tout entier $n \in \mathbb{N}$, on note $\pi_{2n}(X) = (X^2 - 1)^n$ et $R_n = \pi_{2n}^{(n)}$. On munit $\mathbb{R}[X]$ du produit scalaire défini par :

$$\forall (P, Q) \in E^2, \langle P | Q \rangle = \int_{-1}^1 P(x) Q(x) dx$$

1. Montrer que R_n est un polynôme de degré n de la parité de n .
2. Calculer, pour $n \geq 1$, les coefficients de X^n et X^{n-1} dans R_n .

3. Montrer que, pour $n \geq 1$, pour tout entier k compris entre 1 et n et tout $P \in \mathbb{R}[X]$, on a :

$$\int_{-1}^1 \pi_{2n}^{(k)}(t) P(t) dt = - \int_{-1}^1 \pi_{2n}^{(k-1)}(t) P'(t) dt$$

4. Montrer que, pour $n \geq 1$ et tout polynôme $P \in \mathbb{R}_{n-1}[X]$, on a $\langle R_n | P \rangle = 0$.
5. En déduire que la famille $(R_n)_{n \in \mathbb{N}}$ est orthogonale dans $\mathbb{R}[X]$.
6. Calculer $\|R_n\|$, pour tout entier n positif ou nul. Les polynômes $P_n = \frac{1}{\|R_n\|} R_n$ sont les polynômes de Legendre normalisés.

Solution.

1. Pour $n = 0$ on a $R_0 = \pi_0 = 1$. Pour $n \geq 1$, $\pi_{2n}(X) = \sum_{k=0}^n (-1)^{n-k} \binom{n}{k} X^{2k}$ est de degré $2n$ et sa dérivée d'ordre n :

$$R_n(X) = \sum_{\frac{n}{2} \leq k \leq n} (-1)^{n-k} \binom{n}{k} \frac{(2k)!}{(2k-n)!} X^{2k-n}$$

est un polynôme de degré n . Le polynôme π_{2n} est pair donc sa dérivée d'ordre n , R_n est de la parité de n .

2. Le coefficient dominant de R_n est $\beta_n^{(n)} = \frac{(2n)!}{n!}$ et le coefficient de X^{n-1} est nul du fait que R_n est de la parité de n .
3. Une intégration par parties donne, pour tout entier k compris entre 1 et n et tout $P \in \mathbb{R}[X]$:

$$\int_{-1}^1 \pi_{2n}^{(k)}(t) P(t) dt = \left[\pi_{2n}^{(k-1)}(t) P(t) \right]_{-1}^1 - \int_{-1}^1 \pi_{2n}^{(k-1)}(t) P'(t) dt$$

Et utilisant le fait que -1 et 1 sont racines d'ordre n du polynôme $\pi_{2n}(X) = (X-1)^n (X+1)^n$, on a $\pi_{2n}^{(k-1)}(\pm 1) = 0$, de sorte que :

$$\int_{-1}^1 \pi_{2n}^{(k)}(t) P(t) dt = - \int_{-1}^1 \pi_{2n}^{(k-1)}(t) P'(t) dt$$

4. En effectuant n intégrations par parties, on obtient :

$$\int_{-1}^1 \pi_{2n}^{(n)}(t) P(t) dt = (-1)^n \int_{-1}^1 \pi_{2n}(t) P^{(n)}(t) dt$$

Pour $P \in \mathbb{R}_{n-1}[X]$, on a $P^{(n)} = 0$ et :

$$\langle R_n | P \rangle = \left\langle \pi_{2n}^{(n)} | P \right\rangle = \left\langle \pi_{2n} | P^{(n)} \right\rangle = 0$$

5. Chaque polynôme R_k étant de degré k , on déduit de la question précédente que $\langle R_n | R_m \rangle = 0$ pour $0 \leq n < m$ et par symétrie $\langle R_n | R_m \rangle = 0$ pour $n \neq m$ dans $\mathbb{N}$. La famille $\{R_n | n \in \mathbb{N}\}$ est donc orthogonal dans $\mathbb{R}[X]$.
6. En utilisant 4. on a :

$$\begin{aligned}\|R_n\|^2 &= \int_{-1}^1 \pi_{2n}^{(n)}(t) R_n(t) dt = (-1)^n \int_{-1}^1 \pi_{2n}(t) R_n^{(n)}(t) dt \\ &= (-1)^n \beta_n^{(n)} n! \int_{-1}^1 \pi_{2n}(t) dt = (2n)! (-1)^n I_n\end{aligned}$$

où $I_n = \int_{-1}^1 \pi_{2n}(t) dt = \int_{-1}^1 (t^2 - 1)^n dt$. Pour $n \geq 1$, on a :

$$I_n = \int_{-1}^1 (t^2 - 1)^{n-1} (t^2 - 1) dt = \int_{-1}^1 (t^2 - 1)^{n-1} t \cdot t dt - I_{n-1}$$

et une intégration par parties donne :

$$\int_{-1}^1 (t^2 - 1)^{n-1} t \cdot t dt = \left[t \frac{1}{2n} (t^2 - 1)^n \right] - \int_{-1}^1 \frac{1}{2n} (t^2 - 1)^n dt = -\frac{1}{2n} I_n$$

soit la relation de récurrence $I_n = -\frac{1}{2n} I_n - I_{n-1}$, soit $I_n = -\frac{2n}{2n+1} I_{n-1}$. Il en résulte que :

$$I_n = (-1)^n \frac{(2n)(2(n-1)) \cdots 2}{(2n+1)(2n-1) \cdots 1} I_0 = (-1)^n \frac{2^{2n} (n!)^2}{(2n+1)!} 2$$

$$\text{et } \|R_n\|^2 = (2n)! \frac{2^{2n} (n!)^2}{(2n+1)!} 2 = \frac{2}{2n+1} 2^{2n} (n!)^2, \text{ soit } \|R_n\| = 2^n n! \sqrt{\frac{2}{2n+1}}.$$

Exercice 12.16. On munit l'espace vectoriel $\mathbb{R}[X]$ du produit scalaire :

$$(P, Q) \mapsto \langle P | Q \rangle = \int_0^{+\infty} P(t) Q(t) e^{-t} dt$$

1. Justifier la convergence des intégrales $\langle P | Q \rangle$ pour tous P, Q dans $\mathbb{R}[X]$ et le fait qu'on a bien un produit scalaire.
2. Construire une base orthonormée de $\mathbb{R}_3[X]$.
3. Soit $P = 1 + X + X^3$. Déterminer $Q \in \mathbb{R}_2[X]$ tel que $\|P - Q\|$ soit minimal.

Solution.

1. On vérifie par récurrence que $\int_0^{+\infty} t^k e^{-t} dt = k!$ et de ce résultat on déduit que l'application $\langle \cdot | \cdot \rangle$ est bien définie sur $\mathbb{R}[X]$. On vérifie ensuite facilement que c'est un produit scalaire.

2. En utilisant le procédé de Gram-Schmidt sur la base $(1, x, x^2, x^3)$ de $\mathbb{R}_3[x]$, on a :

$$\left\{ \begin{array}{l} Q_0 = 1, \|Q_0\|^2 = 1, P_0 = \frac{Q_0}{\|Q_0\|} = 1 \\ P_1 = x - \langle x | P_0 \rangle P_0 = x - 1, \|Q_1\|^2 = 1, P_1 = \frac{Q_1}{\|Q_1\|} = x - 1 \\ Q_2 = x^2 - \langle x^2 | P_0 \rangle P_0 - \langle x^2 | P_1 \rangle P_1 = x^2 - 4x + 2 \\ \|Q_2\|^2 = 4, P_2 = \frac{Q_2}{\|Q_2\|} = \frac{1}{2}(x^2 - 4x + 2) \\ Q_3 = x^3 - \langle x^3 | P_0 \rangle P_0 - \langle x^3 | P_1 \rangle P_1 - \langle x^3 | P_2 \rangle P_2 \\ = x^3 - 9x^2 + 18x - 6, \\ \|Q_3\|^2 = 36, P_3 = \frac{Q_3}{\|Q_3\|} = \frac{1}{6}(x^3 - 9x^2 + 18x - 6) \end{array} \right.$$

3. Le polynôme Q est la projection orthogonale de P sur $F = \mathbb{R}_2[X]$ donnée par $Q = \sum_{k=0}^2 \langle P | P_k \rangle P_k$. Le calcul des $\langle P | P_k \rangle$ peut être évité en remarquant que dans la base orthonormée (P_0, P_1, P_2, P_3) de $E = \mathbb{R}_3[x]$, on a :

$$P = \sum_{k=0}^3 \langle P | P_k \rangle P_k = Q + \langle P | P_3 \rangle P_3$$

le coefficient $\langle P | P_3 \rangle$ s'obtenant en identifiant les coefficients de X^3 dans cette égalité (P est de degré 2 au plus), soit $\langle P | P_3 \rangle = 6$. On a donc :

$$Q = P - \langle P | P_3 \rangle P_3 = 9x^2 - 17x + 7$$

et $d(P, \mathbb{R}_2[X]) = \|P - Q\| = |\langle P | P_3 \rangle| = 6$.

Exercice 12.17. Calculer $\inf_{(a,b) \in \mathbb{R}^2} \int_{-1}^1 (x^2 - ax - b)^2 dx$.

Solution. En munissant $E = C^0([-1, 1])$ du produit scalaire $\langle f, g \rangle = \int_{-1}^1 f(x) g(x) dx$, on a :

$$M = \inf_{(a,b) \in \mathbb{R}^2} \int_{-1}^1 (x^2 - ax - b)^2 dx = \inf_{Q \in \mathbb{R}_1[x]} \|f - Q\|^2$$

où $f(x) = x^2$. En utilisant le théorème de projection orthogonale, on a $M = \|f - P\|^2 = \|f\|^2 - \|P\|^2$, où P est la projection orthogonale de f sur $\mathbb{R}_1[X]$, soit $P = \langle f, P_0 \rangle P_0 + \langle f, P_1 \rangle P_1$ où (P_0, P_1) est une base orthonormée de $\mathbb{R}_1[x]$. Le procédé de Gram-Schmidt donne $P_0(x) = \frac{1}{\sqrt{2}}$, $P_1(x) = \frac{\sqrt{3}}{\sqrt{2}}X$ et on a $\langle f, P_0 \rangle = \frac{2}{3} \frac{1}{\sqrt{2}}$, $\langle f, P_1 \rangle = 0$, donc $P(X) = \frac{1}{3}$ et $M = \frac{2}{5} - \frac{2}{9} = \frac{8}{45}$.

Exercice 12.18. Calculer $\inf_{(a,b) \in \mathbb{R}^2} \int_0^1 x^2 (\ln(x) - ax - b)^2 dx$.

Solution. On munit $E = \mathcal{C}([0, 1])$ du produit scalaire $\langle f | g \rangle = \int_0^1 f(x) g(x) dx$ et on note f la fonction définie sur $[0, 1]$ par :

$$f(x) = \begin{cases} x \ln(x) & \text{si } x \in]0, 1] \\ 0 & \text{si } x = 0. \end{cases}$$

Avec $\lim_{x \rightarrow 0} x \ln(x) = 0$, on déduit que $f \in E$. Avec ces notations il s'agit donc de calculer $\delta^2 = d(f, F)^2 = \inf_{(a,b) \in \mathbb{R}^2} \|f - ax^2 - bx\|^2$, où $F = \text{Vect}\{X, X^2\}$. On sait que si (P_1, P_2) est une base orthonormée de F , alors :

$$\delta^2 = \|f - \langle f | P_1 \rangle P_1 - \langle f | P_2 \rangle P_2\|^2 = \|f\|^2 - \langle f | P_1 \rangle^2 - \langle f | P_2 \rangle^2$$

Une telle base orthonormée s'obtient avec le procédé de Gram-Schmidt :

$$\begin{cases} P_1 = \sqrt{3}X \\ P_2 = \sqrt{5}(4X^2 - 3X) \end{cases}$$

Puis avec :

$$\begin{cases} \forall n \in \mathbb{N}, \langle f | x^n \rangle = \int_0^1 x^{n+1} \ln(x) dx = -\frac{1}{(n+2)^2} \\ \|f\|^2 = \int_0^1 x^2 \ln^2(x) dx = \frac{2}{27} \end{cases}$$

on obtient $\langle f | P_1 \rangle = -\frac{\sqrt{3}}{9}$, $\langle f | P_2 \rangle = \frac{\sqrt{5}}{12}$ et $\delta^2 = \frac{1}{2^4 3^3} = \frac{1}{432}$ et la projection orthogonale de f sur F est $P = \langle f | P \rangle P_1 + \langle f | P_2 \rangle P_2 = \frac{5}{3}X^2 - \frac{19}{12}X$.

Exercice 12.19. On suppose que E est euclidien et on se donne une base orthonormée $\mathcal{B}$ de E .

1. Déterminer la matrice dans $\mathcal{B}$ de la projection orthogonale sur la droite $D = \mathbb{R}a$ engendrée par un vecteur non nul a .
2. Déterminer la matrice de la projection orthogonale sur un hyperplan H de E dans $\mathcal{B}$.

Solution.

1. Par définition de p_D , on a, pour tout $x \in E$, $p_D(x) = \frac{\langle x | a \rangle}{\|a\|^2} a$. En écrivant

que $a = \sum_{i=1}^n a_i e_i$, on a $p_D(e_j) = \frac{\langle e_j | a \rangle}{\|a\|^2} a = \sum_{i=1}^n \frac{a_i a_j}{\|a\|^2} e_i$ pour tout j compris

entre 1 et n et la matrice A de p_D dans $\mathcal{B}$ est $A = \left(\left(\frac{a_i a_j}{\|a\|^2} \right) \right)_{1 \leq i, j \leq n}$, ce

qui peut aussi s'écrire $A = \frac{1}{\|a\|^2} C {}^t C$, où C est le vecteur colonne formé des composantes de a dans $\mathcal{B}$.

2. On a $H = \{a\}^\perp = (\mathbb{R}a)^\perp$ avec $a \neq 0$ et pour tout $x \in E$, $p_H(x) = x - \frac{\langle x | a \rangle}{\|a\|^2} a$.

La matrice de p_H dans $\mathcal{B}$ est donc :

$$B = I_n - A = I_n - \frac{1}{\|a\|^2} C {}^t C = \left(\left(\delta_{ij} - \frac{a_i a_j}{\|a\|^2} \right) \right)_{1 \leq i, j \leq n}$$

Exercice 12.20. Diagonaliser la matrice $A = \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 & 1 \\ 1 & 1 & \cdots & 1 & 1 \end{pmatrix}$ dans $\mathcal{M}_n(\mathbb{R})$ pour $n \geq 3$.

Solution. Avec l'exercice 10.1, on a calculé les valeurs propres de A . Nous allons retrouver ces résultats en exploitant le fait que la matrice symétrique réelle est diagonalisable. Comme A est de rang 2, son noyau est de dimension $n - 2$ et $\lambda_1 = 0$ est valeur propre d'ordre $n - 2$ puisque A est diagonalisable. Désignant par λ_2 et λ_3 ses deux autres valeurs propres, on a $\text{Tr}(A) = 2 = \lambda_2 + \lambda_3$. De plus avec

$$A^2 = \begin{pmatrix} n & 2 & \cdots & 2 & n \\ 2 & 2 & \cdots & 2 & 2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 2 & 2 & \cdots & 2 & 2 \\ n & 2 & \cdots & 2 & n \end{pmatrix}, \text{ on déduit que } \text{Tr}(A^2) = 4(n - 1) = \lambda_2^2 + \lambda_3^2. \text{ On a}$$

donc :

$$\begin{cases} \lambda_2 + \lambda_3 = 2 \\ \lambda_2 \lambda_3 = \frac{(\lambda_2 + \lambda_3)^2 - (\lambda_2^2 + \lambda_3^2)}{2} = 2(2 - n) \end{cases}$$

et λ_2, λ_3 sont les racines de $X^2 - 2X + 2(2 - n) = (X - 1)^2 - (2n - 3)$, ce qui donne $\lambda_2 = 1 - \sqrt{2n - 3}$ et $\lambda_3 = 1 + \sqrt{2n - 3}$. La réduction de A s'en déduit.

Bibliographie

- [1] J. M. ARNAUDIES, J. LELONG-FERRAND — *Cours de Mathématiques. Tomes 1*. Dunod (1974).
- [2] J. M. DE KONINCK, A. MERCIER. *Introduction à la théorie des nombres*. Modulo. (1994).
- [3] C. DESCHAMPS, A. WARUSFEL. *Mathématiques tout en un. Série E. Ramis. Volumes 1 et 2*. Dunod. (1999).
- [4] X. GOURDON. *Les Maths en tête. Algèbre*. Ellipses. (2004).
- [5] F. LIRET, D. MARTINAIS. *Cours de mathématiques. Algèbre et géométrie 1-ère et 2-ième année*. Dunod. (1999).
- [6] E. RAMIS, C. DESCHAMPS, J. ODOUX. *Cours de Mathématiques Spéciales. Volume 1*. Masson. (1974).
- [7] J. TRIGNAN — *La géométrie des nombres complexes*. Bréal (1983).

Index

- échelonnés (polynômes), 214
- égalité du parallélogramme, 306
- élément inversible, 33
- élément neutre, 31
- élément régulier, 32
- équation diophantienne, 82, 83
- équivalence, 3
- étagés (polynômes), 214

- abélien, 34
- adjointe (matrice), 320
- algorithme d'Euclide, 80
- anneau, 57
- anneau intègre, 61
- anti-symétrie, 16
- application linéaire, 146
- argument d'un nombre complexe, 121
- assertion, 1
- associés (polynômes), 216
- automorphisme, 41, 149
- axiome, 1

- Bézout, 75, 77, 218, 220
- base, 169, 171
- base de numération, 73
- base incomplète, 175
- bijektivité, 14
- binôme de Newton, 22, 59

- cône isotrope, 279
- Cauchy (produit de), 211
- Cauchy-Schwarz, 114, 306
- Cauchy-Schwarz (inégalité de), 319
- Cayley-Hamilton, 252
- changement de base, 185
- classe d'équivalence, 17
- combinaison linéaire, 141
- complémentaire, 9
- composée (de deux applications), 13
- composée (de deux polynômes), 214
- congruence, 72
- conjonction, 3
- conjugué d'un nombre complexe, 113
- connecteurs logiques, 2
- contraposée, 4
- corps, 62

- décomposition des noyaux, 247
- décomposition en éléments simples, 233
- décomposition en facteurs irréductibles, 228
- décomposition en facteurs premiers, 87
- démonstration par l'absurde, 4
- déterminant, 156, 195, 197, 203, 204
- degré (d'un polynôme), 212
- diagonalisable, 253
- différence symétrique, 9
- dimension, 173
- dimension finie, 173
- discriminant, 118
- dividende, 72
- diviseur, 72
- division euclidienne, 72, 215
- droite vectorielle, 142

- endomorphisme, 41, 146
- endomorphisme symétrique, 317
- ensemble quotient, 17
- Eratosthène, 86
- espace euclidien, 305
- espace préhilbertien, 305
- espace vectoriel, 140
- Euclide, 85, 87, 228
- Euclide (algorithme), 219
- Euler, 94, 124

- famille génératrice, 169, 171
- famille liée, 169, 171

- famille libre, 169, 171
- famille orthogonale, 308
- Fermat, 90, 95
- fonction polynomiale, 222
- forme bilinéaire, 266
- forme bilinéaire alternée, 266
- forme bilinéaire non dégénérée, 280
- forme bilinéaire symétrique, 266
- forme linéaire, 148, 263
- forme polaire, 270
- forme quadratique, 269

- Gauss, 78, 221
- Gram-Schmidt, 320
- groupe, 33
- groupe cyclique, 37
- groupe monogène, 37

- hermitien (espace), 319
- hermitienne (matrice), 321
- hermitienne définie positive (matrice), 321
- hermitienne positive (matrice), 321
- Hilbert (polynômes de), 238
- hyperplan vectoriel, 145, 265

- idéal, 217
- idéal annulateur, 246
- image, 42, 147
- image (d'une application), 12
- image réciproque, 12
- implication, 3
- inégalité de Bessel, 312
- inclusion, 6
- indicatrice d'Euler, 94
- injectivité, 14
- interpolation de Lagrange, 237
- intersection, 9
- isomorphisme, 149
- isomorphisme de groupes, 41

- Kronecker (symboles de), 211

- Lagrange, 40
- Leibniz, 226
- loi associative, 31
- loi commutative, 31
- loi de composition interne, 31

- matrice, 151
- matrice (d'une application linéaire), 182
- matrice élémentaire, 194
- matrice de dilatation, 193
- matrice de passage, 185
- matrice de transvection, 193
- matrice symétrique, 158
- matrice triangulaire, 191
- Minkowski, 306
- Minkowski (inégalité de), 319
- module d'un nombre complexe, 113
- Moivre, 124
- morphisme d'anneaux, 62
- morphisme de corps, 64
- morphisme de groupes, 40
- multiplicité, 223

- négation, 2
- Newton (formules de), 241
- nombre de Fermat, 107
- nombre de Mersenne, 104
- nombre premier, 85
- normale (matrice), 321
- noyau, 42, 62, 147, 279

- opérations élémentaires, 192
- ordre d'un élément, 43
- ordre d'un groupe, 40
- ordre total, 17
- orthogonal, 278
- orthogonalisation de Gram-Schmidt, 309
- orthogonalité, 278, 307
- orthogonaux (vecteurs), 320

- pôle d'une fraction rationnelle, 232
- partition, 11
- Peano, 2
- permutation d'un ensemble, 15
- pgcd, 218
- plan vectoriel, 143
- plus grand commun diviseur, 75
- plus petit commun multiple, 79
- polynôme, 212
- polynôme caractéristique, 248
- polynôme dérivé, 225
- polynôme minimal, 246
- polynômes irréductible, 227
- postulat, 1

- premiers entre eux, 76, 220
- principal (idéal), 217
- produit cartésien, 11
- produit scalaire, 305
- produit scalaire hermitien, 319
- projection orthogonale, 311
- Pythagore, 307

- quadrique, 286
- quadrique à centre, 286
- quantificateurs, 6
- quotient, 72

- récurrence faible, 8
- récurrence forte, 8
- réduction de Gauss, 271, 275
- réflexivité, 16
- réunion, 9
- racine (d'un polynôme), 222
- racine n -ième d'un nombre complexe, 125
- racine n -ième de l'unité, 126
- rang, 179, 280
- relation d'équivalence, 17
- relation d'ordre, 16
- reste, 72

- scindé (polynôme), 224
- signature, 283
- singleton, 5
- somme directe, 160, 161
- sous-anneau, 61
- sous-corps, 63
- sous-espace cyclique, 251
- sous-espace propre, 248
- sous-espace vectoriel, 142
- sous-groupe, 35
- sous-groupe engendré, 37
- spectre, 248, 249
- supplémentaire, 161
- support (d'un polynôme), 212
- surjectivité, 14
- symétrie, 17

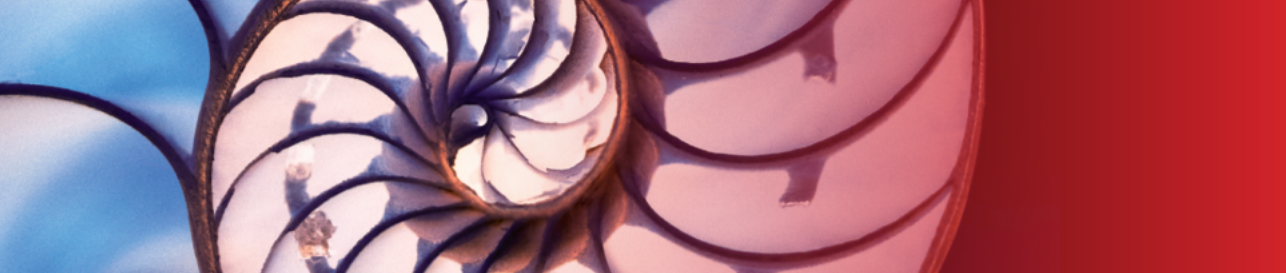
- table de vérité, 3
- tautologie, 3
- Taylor, 226
- théorème chinois, 84, 96
- théorème du rang, 180
- théorème spectral, 318
- trace, 159
- transitivité, 3, 16
- transposée, 157
- trigonalisable, 255

- unitaire (matrice), 320

- valeur propre, 248
- valuation (d'un polynôme), 212
- Vandermonde, 208
- vecteur isotrope, 279
- vecteur propre, 248

- Wilson, 91, 95

- Zermelo-Fränkel, 2



Algèbre et géométrie pour la Licence

Parfaitement adapté à la diversité des parcours scientifiques universitaires, ce manuel couvre l'ensemble des programmes d'algèbre et de géométrie enseignés en première et en deuxième années de licence mathématiques.

Il ne s'agit pas d'un manuel de « méthodes » où l'on sacrifie la notion de rigueur qui est l'essence même des mathématiques. Les notions étudiées ici le sont de façon rigoureuse en démontrant tous les résultats énoncés. Les 12 chapitres sont consacrés à l'étude de quelques notions de logique et de théorie des ensembles, des structures de groupe, d'anneaux et de corps, en se concentrant sur l'anneau des entiers relatifs, le corps des nombres complexes, l'anneau des polynômes à coefficients réels ou complexes, les principales notions d'algèbre linéaire et bilinéaire avec la réduction des endomorphismes et des formes quadratiques ainsi qu'à quelques notions d'arithmétique. Bibliographie sélective et index viennent compléter l'ensemble.

1. Éléments de logique et de théorie des ensembles
 2. Structure de groupe
 3. Structures d'anneau et de corps
 4. Division euclidienne dans $\mathbb{Z}$
 5. Le corps $\mathbb{C}$ des nombres complexes
 6. Espaces vectoriels réels ou complexes
 7. Espaces vectoriels réels ou complexes de dimension finie
 8. Opérations élémentaires et déterminants
 9. Polynômes à coefficients réels ou complexes
 10. Réduction des endomorphismes
 11. Formes bilinéaires et quadratiques réelles ou complexes
 12. Espaces préhilbertiens
- Bibliographie – Index

LES PLUS

- Cours rédigé avec démonstration systématique des résultats énoncés
- Chaque théorème est suivi d'une série d'applications
- Tous les exercices sont intégralement corrigés

Docteur en mathématiques, professeur agrégé à l'université Grenoble-Alpes, **Marie-Cécile Darracq** enseigne les mathématiques en Licence. Membre du jury du Capes externe de 2006 à 2009, puis de l'agrégation interne depuis 2010, elle est directrice des études du Département Sciences Drôme-Ardèche de l'université Grenoble-Alpes.

Agrégé de mathématiques, **Jean-Étienne Rombaldi** a enseigné à l'université Grenoble-Alpes, institut Fourier. Membre du jury du CAPES externe et de l'agrégation interne de mathématiques pendant plusieurs années, il a été responsable de la préparation à l'agrégation interne de l'université de Grenoble et préparateur à l'agrégation interne et externe de cette même université ainsi que pour le CNED.

ISBN : 978-2-8073-3221-8



deboeck
SUPÉRIEUR

www.deboecksuperieur.com