

OBJECTIF
LICENCE

3^e année

Sous la direction de GUY AULIAC
GUY AULIAC - JEAN DELCOURT - RÉMY GOBLOT

MATHÉMATIQUES

Algèbre et géométrie

50% COURS
+ 50% EXOS
= 100%
EFFICACE

Compléments
sur le web

EdiScience

MATHÉMATIQUES

ALGÈBRE ET GÉOMÉTRIE

Consulter nos catalogues sur le web



EdiScience

Titre

Auteur

Collection

ISBN

Index thématique

Je trouve mon libraire

L'achète en ligne

contact@ediscience.net

A la "UNE"

+ de références scientifiques...

Rechercher

EdiScience
101 rue de la République
93000 DUNOD

EdiScience.net

sans (trop d')effort

STS - IUT/Classes préparatoires/Universités/Crises/Écoles/Formation continue/Auto-formation

Mathématiques Physique Chimie SVT Electronique Informatique Eco-gestion

EdiScience, une marque d'ouvrages universitaires qui a bercé des générations d'étudiants !

EdiScience

Les livres EdiScience se positionnent résolument du côté de l'étudiant et de l'apprenant en formation continue : l'exemple, l'exo, la méthode y sont privilégiés. Les ouvrages y sont sélectionnés pour répondre à l'attente spécifique des concours et des examens scientifiques des premiers cycles en classes préparatoires et en universités.

7 séries permettent ainsi de parfaire sa formation : les cours, les exos et problèmes corrigés, les annales, les méthodes, les aide-mémoires, les Schaum's et les mini-Schaum's. Chacune d'entre elle correspond à un type d'apprentissage précis, à un moment privilégié de l'année.

Cours

Exos

Annales

Méthodes

Aide-mémoire



Ces ouvrages couvrent tous les sujets incontournables des sciences, avec des contenus d'une grande qualité, alliant un apprentissage par l'entraînement idéal pour l'auto-formation et une présentation résolument conviviale.

Les "grands" Schaum's

Les "mini" Schaum's

MATHÉMATIQUES

Algèbre et géométrie

50 % Cours + 50 % exos

Sous la direction de

Guy Auliac

Professeur agrégé à l'université de Marne-la-Vallée

Jean Delcourt

Professeur agrégé à l'université de Cergy-Pontoise

Rémy Goblot

Professeur à l'université de Lille



Couverture : Claude Lieber

Le pictogramme qui figure ci-contre mérite une explication. Son objet est d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit, particulièrement dans le domaine de l'édition technique et universitaire, le développement massif du photocopillage.

Le Code de la propriété intellectuelle du 1^{er} juillet 1992 interdit en effet expressément la photocopie à usage collectif sans autorisation des ayants droit. Or, cette pratique s'est généralisée dans les établissements

d'enseignement supérieur, provoquant une baisse brutale des achats de livres et de revues, au point que la possibilité même pour

les auteurs de créer des œuvres nouvelles et de les faire éditer correctement est aujourd'hui menacée. Nous rappelons donc que toute reproduction, partielle ou totale, de la présente publication est interdite sans autorisation de l'auteur, de son éditeur ou du Centre français d'exploitation du

droit de copie (CFC, 20, rue des Grands-Augustins, 75006 Paris).



© Dunod, Paris, 2005

ISBN 2 10 048335 8

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, 2° et 3° a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).

Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

Table des matières

AVANT-PROPOS	IX
CHAPITRE 1 • THÉORIE DES ENSEMBLES	
1.1 Ensembles, applications	1
1.2 Cardinal d'un ensemble	5
1.3 Ensembles ordonnés, bon ordre	8
1.4 Relation d'équivalence, ensembles quotients	11
1.5 Rappels d'arithmétique élémentaire	13
Exercices	14
CHAPITRE 2 • ANNEAUX ET CORPS	
2.1 Anneaux, sous-anneaux, idéaux	25
2.2 Divisibilité	36
2.3 Anneaux principaux, euclidiens et factoriels	39
Exercices	46
CHAPITRE 3 • POLYNÔMES	
3.1 L'anneau des polynômes $A[X]$	59
3.2 Polynômes irréductibles	62
3.3 Polynômes à plusieurs indéterminées	68
Exercices	72

CHAPITRE 4 • ALGÈBRE LINÉAIRE

4.1	Diagonalisation et trigonalisation	85
4.2	Décomposition de Dunford et réduction de Jordan	94
4.3	Réduction de Frobenius	102
	Exercices	108

CHAPITRE 5 • GROUPES

5.1	Sous-groupes, morphismes	119
5.2	Groupe quotient et groupe produit	123
5.3	Groupes commutatifs finis	126
5.4	Actions de groupes	130
5.5	Théorèmes de Sylow	134
	Exercices	136

CHAPITRE 6 • ALGÈBRE BILINÉAIRE

6.1	Formes bilinéaires	147
6.2	Formes bilinéaires symétriques et antisymétriques	150
6.3	Formes quadratiques	155
6.4	Espaces vectoriels euclidiens et hermitiens	159
6.5	Adjoint	164
	Exercices	169

CHAPITRE 7 • GROUPES CLASSIQUES

7.1	Les groupes linéaires et spécial linéaires	179
7.2	Le groupe orthogonal	183
7.3	La dimension deux	187
7.4	Décomposition des transformations orthogonales	193
	Exercices	194

CHAPITRE 8 • ESPACES AFFINES EUCLIDIENS

8.1	Notions affines	203
8.2	Géométrie euclidienne	222
	Exercices	233

CHAPITRE 9 • CALCULS BARYCENTRIQUES

9.1	Espace vectoriel des points pondérés	243
9.2	Application en géométrie plane	248
9.3	Application en géométrie du triangle	254
9.4	Fonction de Leibnitz	255
	Exercices	256

CHAPITRE 10 • TÉTRAÈDRES ET PARALLÉLÉPIPÈDES

10.1	Produit mixte, produit vectoriel	267
10.2	Applications à des configurations	272
	Exercices	277

CHAPITRE 11 • GÉOMÉTRIE DES CERCLES

11.1	Positions relatives de cercles et de droites	283
11.2	Puissance d'un point	287
11.3	Propriété angulaire du cercle	292
11.4	Faisceaux de cercles	294
11.5	L'espace des cercles du plan	296
11.6	Projection stéréographique	304
	Exercices	308

CHAPITRE 12 • CONIQUES

12.1	Coniques dans un plan affine	315
12.2	Coniques dans un plan affine euclidien	325
	Exercices	338

CHAPITRE 13 • NOMBRES COMPLEXES ET GÉOMÉTRIE

13.1	Le corps $\mathbb{C}$ comme plan géométrique	349
13.2	Utilisation de $\mathbb{C}$ en géométrie affine plane	351
13.3	La géométrie des cercles et $\mathbb{C}$	355
13.4	Groupe circulaire	362
	Exercices	367

INDEX	382
--------------	-----

RÉFÉRENCES BIBLIOGRAPHIQUES	386
------------------------------------	-----

Avant-propos

La présente série est destinée aux étudiants de troisième année de Licence qui suivent un parcours de mathématiques. Elle est composée de trois volumes, *Intégration et probabilités*, *Algèbre et géométrie*, *Topologie et analyse*, et elle couvre les notions généralement enseignées sur ces thèmes à ce niveau d'études.

C'est en troisième année de licence que se constituent les bases à partir desquelles un étudiant pourra, soit aborder un master de mathématiques appliquées ou de mathématiques pures, soit préparer le CAPES de mathématiques. De nombreuses notions nouvelles sont abordées et il est indispensable que l'étudiant les fasse siennes, se les approprie.

Cette appropriation nécessite dans un premier temps une lecture attentive et une bonne compréhension des résultats du cours, ainsi que des démonstrations qui les justifient, des motivations et heuristiques qui les sous-tendent. L'acquisition de nouvelles notions mathématiques ne saurait par ailleurs être complète sans une réelle manipulation de ces nouveaux concepts de la part de l'étudiant. C'est pourquoi cette série a le parti pris de proposer, outre un cours complet, une quantité importante d'exercices corrigés. Ces exercices vont d'une application immédiate du cours à un approfondissement de certains résultats. Ils sont un élément fondamental d'assimilation et d'appropriation du contenu du cours. Rappelons à ce propos que chercher un exercice est en soit très formateur et que c'est justement cette recherche qui fait progresser. En d'autres termes, il n'est guère souhaitable de se précipiter sur la solution à la première difficulté...

Dans ce tome consacré à l'algèbre et à la géométrie, on s'est efforcé de présenter toutes les connaissances et tous les outils qui constituent le socle de l'algèbre et de la géométrie.

Parlons d'abord de la forme : il s'agit pour nous, tout en évitant les généralités excessives, d'utiliser les notions abstraites comme par exemple les quotients, qui permettent d'arriver plus vite à l'essentiel. Les démonstrations sont détaillées, les exemples nombreux. Quelques exercices sont insérés dans le fil du cours, ils l'illustrent ou le prolongent. D'autres exercices parfois un peu plus longs sont placés en fin de chapitre. Tous les exercices sont corrigés. De courts problèmes permettent d'approfondir, d'ouvrir d'autres portes. Les corrigés des problèmes sont disponibles sur le site de Dunod : www.dunod.com.

On trouve une bonne partie de ce que doit connaître un étudiant en mathématiques, qu'il se destine à l'enseignement, à la recherche ou aux applications. Les chapitres d'algèbre concernent beaucoup l'algèbre générale (anneaux, groupes) mais aussi des compléments d'algèbre linéaire et bilinéaire. Pour les chapitres de géométrie, on a choisi une approche et un éclairage qui utilisent toutes les ressources de la première partie du livre ; c'est ainsi, par exemple, que la théorie des formes quadratiques est constamment sollicitée pour une étude précise des faisceaux de cercles, des coniques.

Nous souhaitons enfin rappeler que cette série a vu le jour grâce à notre ami Guy Auliac. Il en a conçu le projet et il est co-auteur des volumes d'*Intégration et probabilités*, ainsi que d'*Analyse*. Malgré sa maladie, il a travaillé à ces ouvrages avec l'enthousiasme, l'énergie et la compétence qui le caractérisaient. Il nous a maintenant quitté et nous lui dédions ces livres.

Chapitre 1

Théorie des ensembles : rappels et compléments

Nous n'avons pas l'intention de présenter ici une théorie mathématique rigoureuse et complète. Il faudrait pour cela des prérequis de logique, un appareillage complexe, choisir entre différentes axiomatiques... Nous nous bornerons à une partie de la « théorie naïve des ensembles », selon l'expression de Paul Halmös (voir [18]).

1.1 ENSEMBLES, APPLICATIONS

1.1.1. Ensembles

Acceptons la notion intuitive d'ensemble : un ensemble E est un objet mathématique ; si x est un objet mathématique, la relation d'appartenance $x \in E$ est soit vraie, soit fausse, et les x pour lesquels elle est vraie sont appelés les **éléments** de E . Deux ensembles sont égaux s'ils ont exactement les mêmes éléments.

Un ensemble est défini en **extension** si on en donne la liste des éléments, liste mise entre des accolades. Cas particulier : $\emptyset = \{\}$ est l'ensemble qui n'a aucun élément, on l'appelle **ensemble vide** ; $\{\emptyset\}$ est un ensemble qui a un élément : $\emptyset \in \{\emptyset\}$.

Il est défini en **compréhension** lorsqu'on définit ses éléments par une propriété, exprimée sous forme d'une proposition mathématique. C'est dans ce second cas qu'il pourra être utile de se poser la question : mon ensemble peut-il être vide ? Admettons

qu'on connaisse l'ensemble des nombres entiers naturels $\mathbb{N}$. On peut alors définir les ensembles :

- $P = \{n \in \mathbb{N} \mid \exists k \in \mathbb{N}, n = 2k\}$
- $\mathcal{A} = \{n \in \mathbb{N} \mid n \mid 60\}$
- $\mathcal{Q} = \{1, 2, 3, 4, 5, 6, 10, 12, 15, 20, 30, 60\}$
- $\mathcal{P} = \{p \in \mathbb{N} \mid (\forall n \in \mathbb{N}, n \mid p \Rightarrow n = 1 \text{ ou } n = p) \text{ et } (p \neq 1)\}$
- $\mathcal{I} = \{a \in \mathbb{N} \mid \exists b \in \mathbb{N}, a^2 - 1973b^2 = -1\}$

Il est facile de voir que l'ensemble P est formé des nombres pairs, l'ensemble $\mathcal{P}$ est formé des nombres premiers, les deux ensembles $\mathcal{A}$ et $\mathcal{Q}$ sont égaux, et quant à l'ensemble $\mathcal{I}$, il n'est pas du tout immédiat de décider s'il est ou non vide¹. Mais étant donné un nombre a , on peut décider rapidement s'il est, ou non, dans l'ensemble $\mathcal{I}$.

On appelle **sous-ensemble** d'un ensemble E , un ensemble F tel que : $\forall x \in F, x \in E$, et on écrit $F \subset E$. L'ensemble vide et E lui-même sont des sous-ensembles de E .

Remarques

- Un ensemble est défini en compréhension de la façon suivante

$$\mathcal{A} = \{x \in E \mid p(x)\}$$

et ses éléments sont a priori choisis dans un ensemble E , p étant une propriété qui a un sens pour les éléments de E . Si on ne fait pas cette restriction, on pourrait écrire :

$$\mathcal{A} = \{x \mid x \notin x\}$$

et la proposition $\mathcal{A} \in \mathcal{A}$ risque de donner des maux de têtes : est-elle vraie, mais alors elle est fausse ?...

- Autre remarque : un « vrai » logicien ne fait pas la différence entre des êtres mathématiques qui seraient des ensembles, d'autres qui n'auraient vocation qu'à être des éléments. Dans la « vraie » théorie des ensembles, tout est ensemble.

1.1.2. Union et intersection de deux ensembles, produit cartésien

Si A et B sont deux ensembles, on définit leur union $A \cup B$ et leur intersection $A \cap B$ par :

$$A \cup B = \{x \mid x \in A \text{ ou } x \in B\} \quad A \cap B = \{x \mid x \in A \text{ et } x \in B\}$$

Ce sont deux ensembles (et cela nécessite un axiome pour l'union, en vertu de la remarque précédente, alors que l'intersection peut être définie comme $\{x \in A \mid x \in B\}$). Les opérations ainsi définies ont des propriétés bien connues

1. Il est non vide : essayer avec $a = 88526$ et $b = 1993$.

que nous ne détaillerons pas. Signalons également la définition de la différence ensembliste

$$A \setminus B = \{x \in A \mid x \notin B\}$$

et rappelons que la proposition $A \subset B$ est une abréviation de $A \cap B = A$. Si $A \subset B$, l'ensemble $B \setminus A$ s'appelle le complémentaire de A dans B .

On va maintenant définir le couple (a, b) : c'est l'ensemble

$$(a, b) = \{a, \{a, b\}\}$$

Cela permet d'obtenir l'équivalence :

$$(a, b) = (a', b') \iff (a = a' \text{ et } b = b')$$

Ne pas confondre le couple (a, b) avec l'ensemble (la paire) $\{a, b\}$; avec notre définition, le couple (a, a) désigne l'ensemble $\{a, \{a\}\}$. Cette définition des couples peut paraître inutilement abstraite, et elle masque la « symétrie » qu'il y a entre (a, b) et (b, a) .

Le **produit cartésien**² de deux ensembles est alors l'ensemble des couples :

$$A \times B = \{(a, b) \mid a \in A \text{ et } b \in B\}$$

Enfin rappelons également que $\mathcal{P}(E)$ est l'ensemble de tous les sous-ensembles de E . Il contient en particulier l'ensemble vide et E lui-même.

1.1.3. Relations, Applications

Des définitions :

- Une relation binaire $\mathcal{R}$ est un sous-ensemble de $A \times B$; on écrit $a \mathcal{R} b$ plutôt que $(a, b) \in \mathcal{R}$.
- Une application de A dans B est une relation f qui vérifie :

$$\forall a \in A, \quad \exists b \in B, \quad a f b$$

et

$$\forall a \in A, \forall (b, b') \in B \times B, \quad (a f b \text{ et } a f b') \Rightarrow (b = b')$$

Cela signifie qu'il y a toujours un b tel que $a f b$ et qu'il y a unicité de b . On écrit :

$$\begin{array}{ccc} f : & A & \longrightarrow B \\ & a & \longmapsto b = f(a) \end{array}$$

On peut composer des relations, $\mathcal{R} \subset A \times B$ et $\mathcal{S} \subset B \times C$ en posant pour x dans A et z dans C :

$$(x, z) \in \mathcal{R} \bullet \mathcal{S} \iff \exists y \in B \quad (x, y) \in \mathcal{R} \text{ et } (y, z) \in \mathcal{S}$$

2. En l'honneur de René Descartes, qui a utilisé les couples de coordonnées pour repérer des points.

Dans le cas où $\mathcal{R}$ et $\mathcal{S}$ sont des applications, $\mathcal{R} \bullet \mathcal{S}$ est une application et :

$$z = \mathcal{R} \bullet \mathcal{S}(x) \iff \exists y \in B \quad y = \mathcal{R}(x) \text{ et } z = \mathcal{S}(y)$$

et donc $\mathcal{R} \bullet \mathcal{S}$ sera une application telle que $\mathcal{R} \bullet \mathcal{S}(x) = \mathcal{S}(\mathcal{R}(x))$. On préfère en ce cas noter $z = \mathcal{R} \bullet \mathcal{S}(x) = \mathcal{S} \circ \mathcal{R}(x)$, c'est ce qu'on appelle « la loi rond ».

L'ensemble des applications d'un ensemble A dans un ensemble B est noté B^A (chercher une justification de cette notation...). On peut alors définir les applications **injectives**, **surjectives** et **bijectives**.

1.1.4. Familles, produit

On appelle **famille** indexée par un ensemble I , une application de I dans un ensemble $\mathcal{A}$. On note a_i l'image de $i \in I$ et $(a_i)_{i \in I}$ la famille. Il est possible bien sûr que les a_i soient eux-mêmes des ensembles. Si $(A_i)_{i \in I}$ est une famille d'ensembles, il existe un ensemble qui est la **réunion** des ensembles A_i ; on le note

$$A = \bigcup_{i \in I} A_i$$

et il est caractérisé par :

$$a \in A \iff \exists i \in I, a \in A_i$$

Si I est de la forme $I = \{i_1, i_2\}$, on retrouve la réunion traditionnelle de deux ensembles. Si I est vide, la réunion est vide. Et si I est non vide, on peut définir l'intersection de la famille :

$$B = \bigcap_{i \in I} A_i$$

caractérisée par :

$$a \in B \iff \forall i \in I, a \in A_i$$

Il y a un peu de subtilité dans ces définitions : l'intersection d'une famille vide $I = \emptyset$ ne peut être définie sans contradiction, (c'est lié à l'impossibilité d'accepter l'existence de l'ensemble de tous les ensembles); bien sûr, l'intersection peut être vide, par exemple quand l'un des A_i est vide, mais pas seulement dans ce cas...

De même qu'on a défini le produit cartésien de deux ensembles, définissons le produit d'une famille par :

$$\prod_{i \in I} A_i = \{f : I \longrightarrow \bigcup_{i \in I} A_i \mid f(i) \in A_i\}$$

Se donner un élément de ce produit, c'est finalement se donner une famille, indexée par I , de la forme $(a_i)_{i \in I}$ où $a_i \in A_i$ pour tout $i \in I$, et si I a deux éléments, on retrouve (moyennant une identification : laquelle ?), le produit cartésien habituel.

Exercice 1.1. Examiner le rôle que joue $\emptyset$ vis-à-vis des opérations d'union, intersection, produit cartésien. Déterminer

$$A^\emptyset, \emptyset^A, \emptyset^\emptyset$$

Exercice 1.2. Déterminer $\mathcal{P}(\emptyset)$, $\mathcal{P}(\mathcal{P}(\emptyset))$, etc.

1.2 CARDINAL D'UN ENSEMBLE

On ne peut se contenter d'une définition circulaire comme « le cardinal d'un ensemble est son nombre d'éléments ». Depuis Cantor, on procède en gros de la façon suivante : on dit que deux ensembles sont **équipotents** s'il existe une bijection de l'un vers l'autre, et on conviendra que le cardinal d'un ensemble est la « classe » de tous les ensembles qui sont en bijection avec lui.

Attention, on ne peut parler d'« ensemble de tous les ensembles » sans contradiction. C'est pour cela qu'on a employé le terme un peu vague de classe.

1.2.1. Ensembles finis, infinis, dénombrables

Commençons par admettre qu'on a défini l'ensemble des nombres entiers $\mathbb{N}$: cet ensemble contient un élément noté 0 et tout élément n a un successeur, noté $n + 1$; deux éléments qui ont même successeur sont égaux, et 0 n'est pas un successeur. De plus, $\mathbb{N}$ a la propriété de récurrence, que l'on peut énoncer ainsi : pour tout $F \subset \mathbb{N}$, on a

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

On peut alors définir dans $\mathbb{N}$ l'addition, la multiplication et la relation d'ordre habituelle, que l'on appelle **ordre naturel** et que l'on note $\leq$. Nous ne détaillerons pas ces constructions. En suivant la démarche proposée dans le cas général, on va définir les ensembles finis, infinis, dénombrables :

- On dit qu'un ensemble A est **fini** s'il existe un entier n tel que A soit en bijection avec $\{1, \dots, n\}$ ³.
- Un ensemble est **infini** s'il n'est pas fini, et **dénombrable** s'il est en bijection avec $\mathbb{N}$.

3. Notation qui représente bien sûr l'ensemble $\{k \in \mathbb{N} \mid 1 \leq k \leq n\}$

Dans le cas d'un ensemble fini E , l'entier n est appelé **cardinal** de E et est noté $\text{Card}(E)$.

Remarque : Cette définition a un sens, car on peut démontrer (par récurrence) que si $n \neq p$, l'ensemble $\{1, \dots, n\}$ et l'ensemble $\{1, \dots, p\}$ ne sont pas en bijection. On dit également que le cardinal de l'ensemble vide est 0.

On peut vérifier (voir par exemple la section suivante) que $\mathbb{Q}$ est dénombrable, mais un théorème célèbre attribué à Cantor affirme que $\mathbb{R}$ n'est pas dénombrable.

Il est alors possible d'énoncer le théorème suivant :

Théorème 1.1. *Si E et F sont deux ensembles finis de même cardinal, toute injection (resp. surjection) de l'un dans l'autre, est bijective.*

Avant de faire la démonstration, remarquons l'analogie avec le théorème qui concerne les applications linéaires entre espaces vectoriels de même dimension finie.

Démonstration. Il convient d'abord de montrer que si F' est un sous-ensemble de F , alors F' est fini, de cardinal inférieur à celui de F . De plus, l'inclusion est stricte si et seulement si $\text{Card}(F') < \text{Card}(F)$. Ce lemme se démontre par récurrence sur le cardinal de F : si $\text{Card}(F) = 1$, il est facile de déterminer tous les sous-ensembles de F et de vérifier le lemme. Supposons la proposition vraie pour tous les ensembles de cardinal n , et prenons F de cardinal $n + 1$, une bijection permet de supposer que $F = \{1, \dots, n + 1\}$. Soit F' un sous-ensemble strict de F . On peut supposer, quitte à composer par une bijection, que F' ne contient pas $n + 1$, ainsi F' est un sous-ensemble de $\{1, \dots, n\}$. Si cette inclusion est stricte, l'hypothèse de récurrence s'applique, F' est fini de cardinal strictement inférieur à n donc à $n + 1$. Sinon, F' coïncide avec $\{1, \dots, n\}$, il est de cardinal n , strictement inférieur à $n + 1$.

Passons à la démonstration du théorème, et soit f injective de E dans F . Alors $f(E)$ est un sous-ensemble de F , qui est en bijection avec E , donc de même cardinal que E donc que F . On en déduit que $f(E) = F$ et que f est surjective. Le lecteur terminera la démonstration. $\square$

1.2.2. Théorème de Cantor-Bernstein

Il n'est pas toujours facile de définir une bijection entre deux ensembles. Le théorème suivant, qui n'est pas complètement banal, permet de prouver que deux ensembles ont même cardinal.

Théorème 1.2. Théorème de Cantor-Bernstein. *Soit A et B deux ensembles. On suppose qu'il existe une application f **injective** de A vers B , et une application g **injective** de B vers A . Alors A et B ont même cardinal.*

Démonstration. Commençons par observer que tout élément de A a au maximum un antécédent par g . Et tout élément de B a au maximum un antécédent par f . Partant

d'un élément a de A , on peut construire une suite, éventuellement finie des antécédents successifs, par g puis par f , etc.

- Si cette suite est infinie, on dit que a est dans A_∞ .
- Si elle est finie, mais s'arrête en un élément de B , on dit que a est dans A_1 .
- Sinon, elle est finie et s'arrête en A , ou elle est vide, (cas où a n'a pas d'antécédent par g), on dit que a est dans A_0 .

On a ainsi défini une partition de A et on va en déduire une bijection ϕ de A dans B . Si $a \in A_0$ ou $a \in A_\infty$, on pose $\phi(a) = f(a)$. Si $a \in A_1$, il a un antécédent par g , et on pose $\phi(a) = g^{-1}(a)$, notation un peu abusive pour désigner cet antécédent unique. On peut définir de même dans B trois ensembles B_∞ , B_1 et B_0 .

- L'image de A_∞ par ϕ est alors incluse dans B_∞ , puisque si a a une infinité d'antécédents successifs, $f(a)$ aussi, et tout élément de B_∞ est l'image par ϕ d'un élément qui a une infinité d'antécédents successifs : ϕ restreint à A_∞ est donc bijective sur B_∞ .
- L'image de A_0 est formée de $f(a)$, où a a un nombre pair d'antécédents successifs ; les $f(a)$ ont donc un nombre impair d'antécédents successifs, donc $f(a)$ est dans B_1 , et tout élément de B_1 est l'image par ϕ d'un élément de A_0 , pour les mêmes raisons ; ϕ est bijective de A_0 sur B_1 .
- De même, ϕ réalise une bijection de A_1 sur B_0 .

En conclusion, ϕ est bien bijective de A sur B . □

1.2.3. Le théorème de Cantor

Il s'agit de montrer que E et $\mathcal{P}(E)$ n'ont pas le même cardinal.

Théorème 1.3. Théorème de Cantor. *Il n'existe pas de surjection d'un ensemble E dans l'ensemble de ses sous-ensembles $\mathcal{P}(E)$*

Démonstration. Supposons que $f : E \rightarrow \mathcal{P}(E)$ soit surjective. L'image d'un élément de E est un sous-ensemble de E , et on peut donc considérer

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

Mais comme f est surjective, il existe $y \in E$ tel que $F = f(y)$. Si $y \in F$, c'est donc que $y \notin f(y)$ soit $y \notin F$, contradiction. L'alternative $y \notin F$ conduit à la même impasse, car y doit vérifier $y \in f(y) = F$, encore une contradiction⁴. □

Il existe bien sûr une application injective de E dans $\mathcal{P}(E)$, et on peut donc dire que le cardinal de $\mathcal{P}(E)$ est strictement plus grand que le cardinal de E .

4. On pourra réfléchir à la phrase : le barbier de cette ville rase tous les hommes qui ne se rasent pas eux-mêmes, et ceux-là seulement.

Exercice 1.3. Montrer que $\mathcal{P}(E)$ est en bijection avec $\{0, 1\}^E$. En déduire le cardinal de $\mathcal{P}(E)$ lorsque E est fini.

Exercice 1.4. Soit f l'application de $\mathbb{N}$ dans lui-même définie par $f(n) = 2n$, et g l'application de $\mathbb{N}$ dans lui-même définie par $g(n) = 3n$. Décrire la bijection ϕ donnée par le théorème de Cantor-Bernstein. Si on échange les rôles de f et de g , que dire des bijections correspondantes ?

Exercice 1.5. En utilisant le théorème de Cantor-Bernstein, montrer que $\mathbb{N}$ est en bijection avec $\mathbb{N} \times \mathbb{N}$, puis que $\mathbb{N}$ est en bijection avec $\mathbb{Q}$.

1.3 ENSEMBLES ORDONNÉS, BON ORDRE

1.3.1. Relations d'ordre

Parmi les relations, on va privilégier maintenant des relations binaires définies dans $E \times E$, et pour commencer, consacrons-nous aux **relations d'ordre**.

Définition 1.4. Une relation binaire, notée $\preceq$ dans E est une relation d'ordre si elle a les trois propriétés :

- | | | |
|--|---|----------------|
| (1) $\forall x \in E,$ | $x \preceq x$ | (réflexivité) |
| (2) $\forall (x, y) \in E \times E,$ | $(x \preceq y \text{ et } y \preceq x) \Rightarrow (x = y)$ | (antisymétrie) |
| (3) $\forall (x, y, z) \in E \times E \times E,$ | $(x \preceq y \text{ et } y \preceq z) \Rightarrow (x \preceq z)$ | (transitivité) |

Avant de donner des exemples, donnons quelques définitions supplémentaires.

- Une relation d'ordre dans E permet parfois d'ordonner les éléments comme les points d'une droite. On dit qu'un ordre est **total** si

$$\forall (x, y) \in E \times E, \quad (x \preceq y \quad \text{ou} \quad y \preceq x)$$

c'est-à-dire si deux éléments quelconques sont toujours comparables. Une relation d'ordre qui n'est pas totale est dite **partielle**.

- Un élément M de E est un **majorant** de $A \subset E$ si

$$\forall a \in A, \quad a \preceq M$$

- α est le **plus grand élément** de A lorsque

$$\alpha \in A \quad \text{et} \quad \forall a \in A, \quad a \preceq \alpha$$

Et bien sûr, on définit de même minorant et plus petit élément. Deux notions un peu

plus difficiles :

- s est **borne supérieure** de A si c'est le plus petit élément de l'ensemble des majorants de A .
- un élément μ de A est **maximal** si

$$\forall a \in A, \quad (\mu \preceq a) \Rightarrow (a = \mu)$$

Des exemples de relations d'ordre, de recherche de majorants, borne supérieures... sont donnés en exercice.

1.3.2. Bon ordre

Autre vocabulaire, une relation d'ordre dans E est un **bon ordre** si on a la propriété : *Toute partie non vide de E admet un plus petit élément.* C'est le cas de $\mathbb{N}$, qui sert un peu de modèle, et de ses sous-ensembles. Une première remarque

Proposition 1.5. *Un bon ordre est un ordre total.*

Démonstration. Il suffit de dire que toute partie de la forme $\{a, b\}$ admet un plus petit élément. $\square$

En revanche, l'exemple de $\mathbb{Z}$ avec l'ordre naturel montre qu'un ordre total n'est pas forcément un bon ordre. C'est également le cas de $\mathbb{R}^+$, avec l'ordre naturel : les intervalles ouverts à gauche, par exemple, n'ont pas de plus petit élément.

Passons maintenant à des considérations un peu plus délicates. Un théorème affirme :

Théorème 1.6. Théorème de Zermelo.

Tout ensemble peut être muni d'une relation de bon ordre.

Ce théorème est un peu surprenant, si on pense à des ensembles « grands » comme $\mathbb{R}$, pour lequel l'ordre naturel n'est certes pas un bon ordre. En fait, ce théorème résulte d'un axiome que nous n'avons pas encore énoncé, et qui est nommé l'**axiome du choix**.

Axiome 1.7. Axiome du choix

Tout produit $\prod_{i \in I} A_i$ d'ensembles non vides est non vide

Cet axiome est ainsi nommé car il signifie qu'il existe f dans ce produit, c'est-à-dire une application de I dans la réunion de A_i , telle que $f(i) \in A_i$ pour tout i ; on dit que c'est une **fonction de choix**. Lorsque la famille est infinie, l'existence d'une telle fonction n'est pas évidente : il s'agit de « choisir » d'un seul coup une infinité d'éléments, sans forcément avoir un algorithme.

Terminons avec un troisième théorème, encore équivalent aux deux énoncés précédents :

Théorème 1.8. Théorème de Zorn

Soit E un ensemble tel que toute partie $\mathcal{P}$ de $\mathcal{P}(E)$ qui est totalement ordonnée admet un majorant : on dit parfois que E est **inductif**.

Alors E admet (au moins) un élément maximal.

On peut déduire les théorèmes de l'axiome du choix, mais l'axiome est aussi impliqué par chacun des théorèmes. On va l'admettre, mais donnons néanmoins un exemple de démonstration : le théorème de Zorn implique le théorème de Zermelo.

Démonstration. Soit X un ensemble. On considère l'ensemble des parties de X qui peuvent être munies d'un bon ordre ; si A est une telle partie, on notera $(A, \preceq)$ le couple formé de cette partie et d'un bon ordre. Soit $\mathcal{X}$ l'ensemble de tous les couples possibles. Cet ensemble est non vide (il contient l'ensemble vide et l'ordre vide, et, si X est non vide, il contient les singletons), et il est muni d'un ordre partiel, le prolongement : on dira que A et B deux parties bien ordonnées de X vérifient $A \prec B$ si $A \subset B$ et si l'ordre de B prolonge celui de A , les éléments de $B \setminus A$ étant tous plus grands que ceux de A . Soit alors $\mathcal{P}$ une partie totalement ordonnée de $\mathcal{X}$. La réunion des éléments de $\mathcal{P}$ est bien ordonnée et c'est donc un majorant des éléments de $\mathcal{P}$. On en déduit que $\mathcal{X}$ contient un élément maximal M . Mais cet élément maximal est X lui-même, muni d'un bon ordre. En effet, si $x \in X \setminus M$, l'ensemble $\{x\} \cup M$ peut être muni d'un bon ordre (en prenant x supérieur à tous les autres éléments de M), contredisant ainsi la maximalité de M . $\square$

Reconnaissons qu'on n'utilise pas très souvent ces énoncés. Donnons un des exemples pour lesquels le théorème de Zorn est incontournable.

Théorème 1.9. Tout espace vectoriel sur un corps K admet une base.

Démonstration. E est un espace vectoriel et $\mathcal{V}$ l'ensemble des systèmes libres. On peut ordonner $\mathcal{V}$ par l'inclusion, et si $\mathcal{C}$ est un sous-ensemble totalement ordonné de $\mathcal{V}$, il admet un majorant, qui est la réunion des ensembles de $\mathcal{C}$: cette réunion U est un système libre (car si un sous-ensemble fini de U donnait une relation de liaison, ce sous-ensemble serait inclus dans un des éléments de $\mathcal{C}$, contradiction) et c'est un majorant de $\mathcal{C}$. On en déduit que l'ensemble des systèmes libres admet un élément maximal : un tel élément $\mathcal{B}$ est une base ; si en effet un vecteur x n'était pas dans $\text{vect}(\mathcal{B})$, l'ensemble $\mathcal{B} \cup \{x\}$ serait libre, contradiction avec la maximalité de $\mathcal{B}$. $\square$

Exercice 1.6. Justifier : l'unicité du plus grand élément, de la borne supérieure, (sous réserve d'existence). Montrer que si A admet un plus grand élément, c'est également sa borne supérieure. Donner un exemple où la borne supérieure de A existe, mais n'est pas plus grand élément de A .

Exercice 1.7. Soit $\emptyset \neq A \subset \mathbb{R}$, où $\mathbb{R}$ est muni de la relation d'ordre $\leq$. Montrer que s est la borne supérieure de A si s vérifie :

$$\begin{cases} \forall a \in A, & a \leq s \\ \forall \epsilon > 0, \exists a \in A, & s - \epsilon < a \leq s \end{cases}$$

Exercice 1.8. Montrer que tout ordre dans E est « isomorphe » (en un sens à préciser) à l'inclusion entre des sous-ensembles de E .

1.4 RELATION D'ÉQUIVALENCE, ENSEMBLES QUOTIENTS

Une relation d'ordre introduit un ordonnancement, une hiérarchie, entre les éléments de E . Au contraire, une relation d'équivalence est associée à un regroupement en « classes ».

Définition 1.10. Une relation binaire dans E est une relation d'équivalence si elle a les trois propriétés :

- | | | |
|--|---|----------------|
| (1) $\forall x \in E,$ | $x \mathcal{R} x$ | (réflexivité) |
| (2) $\forall (x, y) \in E \times E,$ | $(x \mathcal{R} y) \Rightarrow (y \mathcal{R} x)$ | (symétrie) |
| (3) $\forall (x, y, z) \in E \times E \times E,$ | $(x \mathcal{R} y \text{ et } y \mathcal{R} z) \Rightarrow (x \mathcal{R} z)$ | (transitivité) |

La notion de relation d'équivalence est étroitement associée à la notion de partition : une famille $(E_i)_{i \in I}$ de sous-ensembles de E est une **partition** de E si :

- $\bigcup_{i \in I} E_i = E$.
- $\forall i, j \in I, \quad (i \neq j) \Rightarrow (E_i \cap E_j = \emptyset)$

On obtient alors la partition associée à une relation d'équivalence par :

Définition 1.11. À tout $a \in E$, on associe sa **classe d'équivalence** :

$$\bar{a} = \{b \in E \mid b \mathcal{R} a\}$$

et les classes d'équivalence forment une partition de E .

La démonstration est immédiate, et le lecteur vérifiera sans peine qu'à toute partition de E , on peut associer une relation d'équivalence. L'ensemble formé par les classes d'équivalence s'appelle le **quotient de E par $\mathcal{R}$** , il est noté $E/\mathcal{R}$. Ce vocabulaire s'explique aisément : on a divisé l'ensemble E en faisant des regroupements, dans le quotient, des éléments équivalents sont considérés comme identiques. Ce passage au

quotient est extrêmement fréquent en algèbre, nous en verrons des exemples dans les chapitres suivants. Commençons par un exemple général, avec E un ensemble muni d'une relation d'équivalence $\mathcal{R}$.

Définition 1.12. Soit $f : E \rightarrow F$ une application. On dit qu'elle est **compatible** avec $\mathcal{R}$ si

$$\forall (a, b) \in E^2, \quad (a \mathcal{R} b) \Rightarrow (f(a) = f(b))$$

Si une application est compatible avec $\mathcal{R}$, elle « passe au quotient » :

Proposition 1.13. Si $f : E \rightarrow F$ est une application compatible avec une relation d'équivalence $\mathcal{R}$, il existe une unique application $\bar{f} : E/\mathcal{R} \rightarrow F$ telle que $f(a) = \bar{f}(\bar{a})$.

La vérification est directe : l'important est que $\bar{f}$ est bien définie par la relation ci-dessus, par définition d'une application compatible. Si on appelle p l'application qui à a associe $\bar{a}$ (p est la **projection**), on a $f = \bar{f} \circ p$.

Enfin, supposons que E soit muni d'une **loi de composition**, c'est-à-dire d'une application :

$$\begin{array}{ccc} E \times E & \longrightarrow & E \\ (x, y) & \longmapsto & x * y \end{array}$$

alors on dit que la loi est **compatible** avec la relation dès que :

$$\forall (a, a', b, b') \in E^4, \quad \begin{cases} a \mathcal{R} a' \\ b \mathcal{R} b' \end{cases} \Rightarrow (a * b) \mathcal{R} (a' * b')$$

On en déduit alors, comme ci-dessus, une loi de composition dans l'ensemble quotient $E/\mathcal{R}$, définie par :

$$\bar{a} * \bar{b} = \overline{a * b}$$

On vérifie que cette définition a un sens.

Exercice 1.9. Que dire d'une relation qui est à la fois d'ordre et d'équivalence ?

Exercice 1.10. Soit T strictement positif. On suppose que $\mathcal{R}$ est la relation définie sur $\mathbb{R}$ par

$$a \mathcal{R} b \iff a - b \in \mathbb{Z}T$$

Que signifie qu'une application de $\mathbb{R}$ dans $\mathbb{R}$ est compatible avec $\mathcal{R}$? Caractériser de la même façon les fonctions paires.

Exercice 1.11. Soit $\mathcal{E}$ l'ensemble des relations définies sur E . On définit dans $\mathcal{E}$ une relation par :

$$\mathcal{R} \prec \mathcal{S} \iff (\forall (a, b) \in E^2, \quad (a \mathcal{R} b) \Rightarrow (a \mathcal{S} b))$$

Montrer que c'est une relation d'ordre, et que $\mathcal{E}$ possède un plus petit élément et un plus grand élément que l'on précisera.

1.5 RAPPELS D'ARITHMÉTIQUE ÉLÉMENTAIRE

Donnons sans démonstration quelques notations et résultats d'arithmétique dans l'ensemble $\mathbb{Z}$ des entiers relatifs.

1.5.1. Divisibilité, congruence

Si a et b sont des éléments de $\mathbb{Z}$, on dit que a divise b , ou que b est un multiple de a s'il existe k dans $\mathbb{Z}$ tel que $b = ak$. Cela s'écrit $a \mid b$. Restreinte à $\mathbb{N}$, c'est une relation d'ordre et alors :

$$a \mid b \iff b\mathbb{Z} \subset a\mathbb{Z}$$

On dit que a et b sont **congrus** modulo n lorsque $a - b$ est divisible par n . Cela s'écrit $a \equiv b \pmod{n}$. La congruence modulo n est une relation d'équivalence et l'ensemble des classes d'équivalence est noté $\mathbb{Z}/n\mathbb{Z}$. Cette relation de congruence est compatible avec les opérations $+$ et $\times$ de $\mathbb{Z}$:

$$\begin{cases} a \equiv a' \pmod{n} \\ b \equiv b' \pmod{n} \end{cases} \Rightarrow a + b \equiv a' + b' \pmod{n} \text{ et } ab \equiv a'b' \pmod{n}$$

1.5.2. Bézout et Gauss

Deux entiers de $\mathbb{Z}$ sont premiers entre eux s'ils n'ont d'autres diviseurs communs que 1 et -1 . On écrit alors $a \wedge b = 1$. Il est équivalent de dire (c'est le théorème de Bézout), que l'équation $ax + by = 1$ admet au moins une solution (x, y) dans $\mathbb{Z}^2$. Le théorème de Gauss relie divisibilité et nombres premiers entre eux :

$$\begin{cases} a \mid bc \\ a \wedge b = 1 \end{cases} \Rightarrow a \mid c$$

1.5.3. p.g.c.d. et p.p.c.m.

Si a et b sont dans $\mathbb{Z}$, leur p.g.c.d. est l'entier naturel d défini par $a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$, et leur p.p.c.m. est l'entier naturel m défini par $a\mathbb{Z} \cap b\mathbb{Z} = m\mathbb{Z}$. Le p.g.c.d. est le plus grand diviseur commun en ce sens que si d' divise à la fois a et b alors d' divise d . De la même façon, m est le plus petit commun multiple. On écrit $a \wedge b = d$ pour désigner le p.g.c.d. et deux nombres sont premiers entre eux si leur p.g.c.d. est égal à 1. On écrit $a \vee b = m$ pour désigner le p.p.c.m. Résultats utiles :

$$a \wedge b = d \iff \begin{cases} \exists a' \in \mathbb{Z}, a = a'd \\ \exists b' \in \mathbb{Z}, b = b'd \\ a' \wedge b' = 1 \end{cases}$$

$$(a \wedge b)(a \vee b) = |ab|$$

1.5.4. Nombres premiers

Un entier naturel p est premier s'il n'a pas d'autres diviseurs dans $\mathbb{N}$ que 1 et lui-même. L'ensemble $\mathcal{P}$ des nombres premiers est assez mystérieux, sujet de nombreuses conjectures ; il y a une infinité de nombres premiers (résultat du à Euclide), ils se raréfient mais on ignore par exemple s'il existe une infinité de nombres premiers jumeaux, c'est-à-dire de la forme $n, n + 2$ comme 11 et 13. Une autre propriété.

Proposition 1.14. Lemme d'Euclide.

Si p est premier, alors $p \mid ab \Rightarrow p \mid a$ ou $p \mid b$.

Démonstration. Si p est premier et si a est dans $\mathbb{Z}$, alors $p \wedge a = 1$ ou $p \wedge a = p$, car p n'a que 1 et p comme diviseur dans $\mathbb{N}$. Si donc p divise ab et ne divise pas a , il est premier à a et, par le théorème de Gauss, il divise b . $\square$

EXERCICES

Exercice 1.12. Dans l'ensemble $\mathcal{E}$ des relations binaires sur E , on définit la composition par :

$$a \mathcal{R} \bullet \mathcal{S} b \iff \exists c \in E, a \mathcal{S} c \text{ et } c \mathcal{R} b$$

C'est alors une opération toujours définie. Examiner ses propriétés : est elle associative, commutative ? Le composé de deux relations réflexives l'est-il ? Et de deux relations symétriques ? Transitives ?

Exercice 1.13. Si on note $A \Delta B$ l'ensemble $(A \setminus B) \cup (B \setminus A)$, vérifier que $(\mathcal{P}(E), \Delta, \cap)$ est un anneau commutatif, (avec Δ dans le rôle de l'addition et $\cap$ dans le rôle de la multiplication).

Exercice 1.14. La relation dite **équivalent** entre deux suites est-elle une relation d'équivalence ? Que dire des suites qui sont équivalentes à une suite stationnaire ?

Exercice 1.15. Soit $\mathcal{R}$ une relation réflexive sur E . On note $\mathcal{R}^{-1}$ la relation définie par $x\mathcal{R}^{-1}y \iff y\mathcal{R}x$. On note $\mathcal{S}$ la relation définie par :

$$x\mathcal{S}y \iff \exists n \in \mathbb{N}, x_0, x_1, \dots, x_n \mid x = x_0\mathcal{R}_1x_1\mathcal{R}_2x_2 \dots \mathcal{R}_nx_n = y$$

avec $\mathcal{R}_i = \mathcal{R}$ ou $\mathcal{R}^{-1}$. Montrer que $\mathcal{S}$ est la plus petite relation d'équivalence qui contienne $\mathcal{R}$. Comment interpréter $\mathcal{S}$ lorsque $\mathcal{R}$ représente « est l'enfant de » ?

Exercice 1.16. f et g sont deux applications de A dans B (resp. de B dans A). On suppose que $f \circ g$ est injective : peut-on affirmer que f ou que g est injective ? Même question avec surjective. On suppose maintenant que $f \circ g = \text{id}_B$ et $g \circ f = \text{id}_A$. Montrer que f et g sont bijectives et réciproques l'une de l'autre.

Exercice 1.17. Examiner les relations suivantes : sont-elles des relations d'ordre total, partiel, y-a-t-il des plus petits ou plus grands éléments, majorants, bornes supérieures pour E , pour une partie de E , des éléments maximaux ou minimaux ?

- Les ensembles de nombres $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{C}$ pour l'ordre naturel (c'est-à-dire habituel...).
- $\mathbb{N}$ pour la relation « divise ».
- L'ensemble $\mathcal{P}(E)$ des parties d'un ensemble pour la relation d'inclusion.
- $\mathbb{N} \times \mathbb{N}$ pour la relation d'ordre **produit**

$$(a, b) \leq (a', b') \iff a \leq a' \text{ et } b \leq b'$$

puis pour l'ordre **lexicographique** :

$$(a, b) \preceq (a', b') \iff a < a' \text{ ou } a = a' \text{ et } b \leq b'$$

PROBLÈME

Le corrigé de ce problème est disponible sur le site de Dunod : www.dunod.com.

1.1. MATRICE D'INCIDENCE, MATRICE DE MÖBIUS

On considère un ensemble ordonné fini P , la relation d'ordre sera notée $\leq$. Soit n le cardinal de P , on notera $a_1, \dots, a_n$ ses éléments. On appelle **matrice d'incidence** de P toute matrice M de $\mathcal{M}_n(\mathbb{C})$ ayant la propriété :

$$\forall i, j, a_i \not\leq a_j \Rightarrow m_{i,j} = 0$$

- (1) Dans cette question, on suppose que $n = 3$. Décrire l'ensemble des matrices d'incidence lorsque l'ordre est défini par

$$a_1 \leq a_2 \leq a_3, \quad \text{puis par} \quad a_2 \leq a_1 \text{ et } a_2 \leq a_3$$

On suppose implicitement qu'il n'y a pas d'autre relations (à part bien sûr $a_i \leq a_i$ pour tout i).

- (2) Montrer que l'ensemble E des matrices d'incidence est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{C})$ et qu'il est stable pour le produit. La matrice de l'identité est-elle une matrice d'incidence ?
- (3) On veut montrer que si une matrice d'incidence est inversible, alors son inverse est aussi une matrice d'incidence. Cette question est plus difficile :
- (a) On commence par supposer que l'on a :

$$\forall (i, j) \in \{1, \dots, n\}^2, \quad a_i \leq a_j \Rightarrow i \leq j$$

Démontrer qu'alors les matrices d'incidence sont triangulaires supérieures.

- (b) Montrer que si une matrice d'incidence est inversible, alors son inverse est aussi une matrice d'incidence. On pourra raisonner par l'absurde.
- (c) Montrer que si $P = \{a_1, a_2, \dots, a_n\}$ est muni d'une relation d'ordre, il existe une permutation σ de $\mathcal{S}_n$ telle que :

$$\forall (i, j) \in \{1, \dots, n\}^2, \quad a_{\sigma(i)} \leq a_{\sigma(j)} \Rightarrow i \leq j$$

En déduire le résultat dans le cas général.

- (4) Soit $Z = (z_{i,j})_{i,j}$ la matrice définie par :

$$\forall (i, j) \in \{1, \dots, n\}^2, \quad a_i \not\leq a_j \Rightarrow z_{i,j} = 0, \quad a_i \leq a_j \Rightarrow z_{i,j} = 1$$

C'est donc une matrice d'incidence. Montrer qu'elle est inversible. On appelle **matrice de Möbius** la matrice inverse Z^{-1} . On notera $m_{i,j}$ l'élément d'indice (i, j) de cette matrice.

- (5) On suppose que P est l'ensemble $\{1, 2, \dots, n\}$ muni de l'ordre naturel, et on note $a_i = i$. Décrire la matrice Z associée ainsi que la matrice de Möbius Z^{-1} .
- (6) On suppose que P est l'ensemble $\{1, 2, \dots, n\}$ muni de l'ordre « divise », et on prend encore $a_i = i$. Décrire la matrice Z associée. Montrer que les coefficients de la forme $m(k, \ell)$ vérifient :

- $m(k, k) = 1$
- $\forall (k, \ell) \in \{1, 2, \dots, n\}, \quad k \nmid \ell \Rightarrow m(k, \ell) = 0$
- $\forall (k, \ell) \in \{1, 2, \dots, n\}, \quad (k \mid \ell, k \neq \ell) \Rightarrow \sum_{k \mid d \mid \ell} m(k, d) = 0$, la somme portant donc sur les tous les d qui sont multiples de k et qui divisent ℓ .

Vérifiez que ces relations suffisent à déterminer les coefficients $\mu(k, \ell)$.

- (7) On définit la fonction de Möbius sur $\mathbb{N}$ par :

$$\mu(1) = 1, \quad \mu(p_1 p_2 \dots p_k) = (-1)^k, \text{ si les } p_i \text{ sont des premiers distincts}$$

et μ est nul dans tous les autres cas. Montrer que

$$\sum_{\delta|\ell} \mu(\delta) = 0$$

lorsque $\ell > 1$. On utilisera la décomposition en facteurs premiers de ℓ et on cherchera ses diviseurs. En déduire que, lorsque $k \mid \ell$ on a $m(k, \ell) = \mu\left(\frac{\ell}{k}\right)$

- (8) Démontrer la « formule d'inversion de Möbius », lorsque (u_n) et (v_n) sont deux suites numériques :

$$u_n = \sum_{d|n} v_d \quad \Longleftrightarrow \quad v_n = \sum_{d|n} \mu\left(\frac{n}{d}\right) u_d$$

- (9) Si on note $\phi(n)$ le cardinal des entiers k de $\{1, 2, \dots, n\}$ qui sont premiers avec n (ϕ est l'indicateur d'Euler), montrer que :

$$\phi(n) = \sum_{d|n} d \mu\left(\frac{n}{d}\right)$$

puis

$$\phi(n) = n \prod_{p \in \mathcal{P}, p|n} \left(1 - \frac{1}{p}\right)$$

où $\mathcal{P}$ désigne l'ensemble des nombres premiers.

- (10) Quel est l'analogue de la formule d'inversion de Möbius lorsqu'on prend l'ordre naturel au lieu de l'ordre divisé ?

SOLUTIONS DES EXERCICES

Solution 1.1. On a facilement $A \cup \emptyset = A$, $A \cap \emptyset = \emptyset$. De même, $\emptyset \times A = \emptyset$. $A^\emptyset = \{\emptyset\}$. En effet, $A^\emptyset$ est un sous-ensemble de

$$\mathcal{P}(A \times \emptyset) = \mathcal{P}(\emptyset) = \{\emptyset\}$$

la seule relation définie dans $A \times \emptyset$ est $\emptyset$. Le résultat est le même que A soit vide ou non. Reste à savoir si cette relation est fonctionnelle : la phrase

$$\forall x \in A, \exists y \in B, x \mathcal{R} y$$

est vraie si A est vide, si A et B sont vides, mais pas si A est non vide et B vide. Ainsi, si A est non vide,

$$A^\emptyset = \{\emptyset\}, \quad \emptyset^\emptyset = \{\emptyset\}, \quad \emptyset^A = \emptyset$$

Solution 1.2.

$$\mathcal{P}(\emptyset) = \{\emptyset\}, \quad \mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\} \quad \mathcal{P}(\{\emptyset, \{\emptyset\}\}) = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$$

et on pourra continuer si on supporte ses innombrables accolades...

Solution 1.3. À une partie A de E on associe l'application de E dans $\{0, 1\}$ appelée **fonction caractéristique**, notée 1_A et définie par :

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

L'application $A \mapsto 1_A$ de $\mathcal{P}(E)$ dans $\{0, 1\}^E$ est bijective car à toute application f de E dans $\{0, 1\}$ correspond une seule partie de E , celle définie par $A = f^{-1}(1)$. On en déduit donc que, si E est fini, $\mathcal{P}(E)$ est de cardinal $2^{\text{Card}(E)}$. On prolonge ce résultat au cas infini : si $\aleph_0$ (lire aleph zéro) est le cardinal de $\mathbb{N}$, alors le cardinal de $\mathcal{P}(\mathbb{N})$ est $2^{\aleph_0}$, que l'on note $\aleph_1$.

Solution 1.4. Notons $A = \mathbb{N}$, $B = \mathbb{N}$. Pour qu'un entier ait un antécédent par f il doit être pair, pour qu'un entier ait un antécédent par g il doit être multiple de trois. Ainsi, A_0 est formé des entiers de la forme $2^k 3^{k'} m$ où $k \geq k'$, (avec m premier à 2 et à 3) tandis que A_1 est formé des entiers de la forme $2^k 3^{k'} m$ où $k < k'$. L'ensemble A_∞ est réduit à 0. Ainsi ϕ est définie par :

$$\begin{cases} 0 \mapsto 0 \\ 2^k 3^{k'} m \mapsto 2^{k+1} 3^{k'} m & \text{si } k \geq k' \\ 2^k 3^{k'} m \mapsto 2^k 3^{k'-1} m & \text{si } k < k' \end{cases}$$

avec toujours m premier à 2 et 3. On peut vérifier que ϕ est bijective, en se contentant par exemple de visualiser le passage de (k, k') à $(k+1, k')$ ou $(k, k'-1)$ sur un quadrillage. L'application réciproque est obtenue en échangeant les rôles de A et B , donc de 2 et 3.

Solution 1.5. Il suffit de trouver une injection f de $\mathbb{N}$ dans $\mathbb{N} \times \mathbb{N}$ et une injection g de $\mathbb{N} \times \mathbb{N}$ dans $\mathbb{N}$. On peut prendre $f(n) = (n, 0)$ et, par exemple, $g(n, n') = 2^n 3^{n'}$. Il est moins immédiat (et c'est inutile par le théorème de Cantor-Bernstein) de décrire une bijection : en utilisant que tout entier s'écrit de façon unique comme une puissance de 2 fois un entier impair : $(n, m) \mapsto 2^n(2m+1) - 1$ (le -1 pour ne pas oublier 0). Il existe de même une injection de $\mathbb{N}$ dans $\mathbb{Q}$, donnée par $n \mapsto \frac{n}{1}$ et une injection de $\mathbb{Q}$ dans $\mathbb{N} \times \mathbb{N}$, donnée par $\frac{p}{q} \mapsto (2p, q)$ si p est positif, $\frac{p}{q} \mapsto (-2p+1, q)$ si p est strictement négatif (en prenant l'écriture irréductible d'un rationnel).

Solution 1.6. Si α est le plus grand élément de $A \subset E$ et β aussi. Alors α et β sont dans A et $\alpha \prec \beta$, puisque β est plus grand élément, de même $\beta \prec \alpha$ et par antisymétrie, $\alpha = \beta$. Si A admet une borne supérieure, elle est donc unique puisque c'est le plus petit élément de l'ensemble des majorants.

Soit maintenant A un sous-ensemble de E qui admet un plus grand élément α . Par définition du plus grand élément, α est un majorant de A . Si m est un majorant de A , il est plus grand que tous les éléments de A , donc de α : α est le plus petit élément de l'ensemble des majorants de A , c'est la borne supérieure de A . L'exemple dans l'ensemble $\mathbb{R}$ des intervalles de la forme $[a, b[$ montre que la borne supérieure peut exister sans qu'il y ait de plus grand élément.

Solution 1.7.

$$\begin{cases} \forall a \in A, & a \leq s \\ \forall \epsilon > 0, \exists a \in A, & s - \epsilon < a \leq s \end{cases}$$

Supposons que s vérifie ces propriétés : la première indique que s est un majorant de A . Soit M un majorant quelconque de A , alors $s \leq M$, sinon, en notant $\epsilon = s - M$, la seconde propriété impliquerait qu'il existe $a \in A$ vérifiant $s - \epsilon = M < a \leq s$, contradiction. On a montré que s est le plus petit des majorants. La réciproque se traite de la même façon. Cette « propriété de la borne supérieure » s'utilise constamment en analyse.

Solution 1.8. Si E est un ensemble ordonné, on peut commencer par définir les **segments initiaux** par :

$$\forall x \in E, \phi(x) = \{y \in E \mid y \prec x\}$$

On a ainsi défini une application ϕ de E dans $\mathcal{P}(E)$, et cette application est injective : si I est un segment initial, il est égal à $\phi(x)$ où x est l'unique plus grand élément de I . Enfin,

$$x \prec y \Rightarrow (\forall z \in E, z \prec x \Rightarrow z \prec y)$$

donc $x \prec y \Rightarrow \phi(x) \subset \phi(y)$. On a ainsi un isomorphisme d'ensemble ordonnés, c'est-à-dire une bijection qui conserve l'ordre.

Solution 1.9. Une relation qui est à la fois d'ordre et d'équivalence est à la fois symétrique et antisymétrique : dès que $x\mathcal{R}y$, on a $y\mathcal{R}x$ donc $x = y$. Comme la relation est non vide (réflexivité), c'est l'égalité.

Solution 1.10. Une fonction f est compatible avec la relation $\mathcal{R}$ lorsque $f(a) = f(b)$ dès que $\exists k \in \mathbb{Z}, a = b + kT$. Cela signifie que f est périodique, T est une période. En particulier, f « passe au quotient », c'est-à-dire que l'on peut définir $\bar{f}$ de $\mathbb{R}/T\mathbb{Z}$ dans $\mathbb{R}$ par $\bar{f}(\bar{a}) = f(a)$. Une fonction paire est une application compatible avec la relation $x\mathcal{R}y \iff x = \pm y$.

Solution 1.11. Cette relation est en fait l'inclusion dans l'ensemble $\mathcal{P}(E \times E)$. Elle a donc un plus petit élément qui est la relation vide (jamais vraie) et un plus grand élément qui est $E \times E$, la relation toujours vraie.

Solution 1.12. Montrons que la composition est associative :

$$\begin{aligned} a((S \bullet \mathcal{R}) \bullet T)b &\iff \exists c \in E, a(S \bullet \mathcal{R})c \text{ et } cTb \\ &\iff \exists c \in E, \exists d \in E, aSd \text{ et } d\mathcal{R}c \text{ et } cTb \end{aligned}$$

et la traduction de $a(S \bullet (\mathcal{R} \bullet T))b$ est identique. Cette opération n'est pas commutative dès que E est assez grand : prendre $E = \{a, b\}$, $\mathcal{R}$ la relation réduite à (a, b) et S la relation réduite à (b, a) . Alors $\mathcal{R} \bullet S$ et $S \bullet \mathcal{R}$ diffèrent.

Le composé de deux relations réflexives l'est car

$$\forall a, \exists c, a\mathcal{R}c \text{ et } cSa$$

il suffit de prendre $c = a$. Si deux relations sont symétriques, leur composé ne l'est pas forcément mais :

$$\begin{aligned} a\mathcal{R} \bullet Sb &\Rightarrow \exists c \in E, a\mathcal{R}c \text{ et } cSb \\ &\Rightarrow \exists c \in E, c\mathcal{R}a \text{ et } bSc \\ &\Rightarrow bS \bullet \mathcal{R}a \end{aligned}$$

Enfin, le composé de deux relations transitives ne l'est pas forcément, il suffit de bâtir un contre-exemple. On pourra prolonger cet exercice en cherchant un élément neutre, en définissant la relation réciproque d'une relation...

Solution 1.13. Pour vérifier les propriétés de ces opérations, il est pratique de se servir des fonctions caractéristiques vues dans l'exercice 1.3 page 18. On a en effet :

$$1_{A \cap B} = 1_A 1_B \quad \text{et} \quad 1_{A \Delta B} = 1_A + 1_B - 2 \times 1_{A \cap B}$$

formules que l'on vérifie en examinant tous les cas possibles. Il est alors plus agréable de décider que les fonctions caractéristiques sont à valeur dans $\mathbb{Z}/2\mathbb{Z}$, d'où $1_{A\Delta B} = 1_A + 1_B$. On montre alors par exemple la distributivité par

$$\begin{aligned} 1_{A\cap(B\Delta C)} &= 1_A 1_{B\Delta C} = 1_A(1_B + 1_C) \\ 1_{(A\cap B)\Delta(A\cap C)} &= 1_A 1_B + 1_A 1_C \end{aligned}$$

L'associativité de la différence symétrique résulte alors du calcul :

$$\begin{aligned} 1_{(A\Delta B)\Delta C} &= 1_{A\Delta B} + 1_C \\ &= 1_A + 1_B + 1_C \end{aligned}$$

et le calcul de $1_{A\Delta(B\Delta C)}$ donne le même résultat. L'élément neutre pour l'addition Δ est $\emptyset$ (dont la fonction caractéristique est constante nulle), l'élément neutre pour le produit $\cap$ est l'ensemble E tout entier. L'opposé de A est lui-même.

Solution 1.14. Soit E l'ensemble des suites à valeurs réelles. Deux suites (u_n) et (v_n) sont équivalentes lorsqu'il existe une suite (ϵ_n) tendant vers 0 telle que :

$$\forall n \in \mathbb{N} \quad v_n = (1 + \epsilon_n)u_n$$

Avec cette définition, l'équivalence n'est pas une relation d'équivalence... Il n'y a pas symétrie dans le cas où v_n est nulle pour certaines valeurs de n sans que u_n le soit. Mais il suffit de remplacer la définition par :

$$\exists n_0 \in \mathbb{N}, \forall n \geq n_0 \in \mathbb{N} \quad v_n = (1 + \epsilon_n)u_n$$

pour avoir la symétrie, n_0 dépend des deux suites. Alors

$$v_n = (1 + \epsilon_n)u_n \Rightarrow u_n = \frac{1}{1 + \epsilon_n} v_n = \left(1 - \frac{\epsilon_n}{1 + \epsilon_n}\right) v_n$$

pour n suffisamment grand (car ϵ_n tendant vers 0 sera différent de -1 pour tous les n suffisamment grands). La réflexivité est immédiate, on prend pour ϵ_n la suite constante nulle. Pour la transitivité :

$$\begin{cases} v_n = (1 + \epsilon_n)u_n \\ w_n = (1 + \epsilon'_n)v_n \end{cases} \Rightarrow w_n = (1 + \epsilon_n + \epsilon'_n + \epsilon_n \epsilon'_n)u_n$$

Si une suite est équivalente à une suite stationnaire de limite non nulle ℓ , alors elle converge vers ℓ . Réciproquement, si une suite converge vers ℓ non nul, alors $u_n = \ell + \epsilon_n = (1 + \frac{\epsilon_n}{\ell})\ell$, et (u_n) est donc équivalente à la suite constante ℓ . Par contre, une suite est équivalente à la suite constante nulle si et seulement si elle est stationnaire nulle.

Solution 1.15. La relation $\mathcal{S}$ contient la relation $\mathcal{R}$, avec ce qu'il faut pour être une relation d'équivalence. Il faut comprendre que $x_1 \mathcal{R} x_2 \mathcal{R} x_3$ abrège $x_1 \mathcal{R} x_2$ et $x_2 \mathcal{R} x_3$. La relation $\mathcal{S}$ est bien d'équivalence : elle est réflexive : on prend $n = 1$, $\mathcal{R}_1 = \mathcal{R}$. Elle est symétrique : si $x \mathcal{S} y$, on a également $y \mathcal{S} x$, on reprend la même succession dans l'ordre inverse, en changeant $\mathcal{R}$ en $\mathcal{R}^{-1}$. Enfin, $\mathcal{S}$ est transitive : si $x \mathcal{S} y$ et $y \mathcal{S} z$,

il suffit de faire se succéder les deux suites pour relier x à z . Une relation d'équivalence contenant $\mathcal{R}$ doit contenir tous les éléments de $\mathcal{S}$: on peut formaliser davantage, en faisant une récurrence sur la longueur des suites. Pour l'exemple proposé (en ajoutant la réflexivité), on pourra interpréter $\mathcal{S}$ par « est de la même famille que ».

Solution 1.16. Supposons $f \circ g$ injective, et prenons a et a' dans A . Si $g(a) = g(a')$ alors $f \circ g(a) = f \circ g(a')$ d'où $a = a'$. On a montré que g est injective. Supposons maintenant $f \circ g$ surjective. Alors un élément quelconque b de B a un antécédent b' par $f \circ g$, $b = f \circ g(b')$, et b admet $g(b')$ comme antécédent par f , f est surjective. Prenons $A = B = \mathbb{N}$, $g(n) = 2n$ et $f(n) = \lfloor \frac{n}{2} \rfloor$ (où $\lfloor x \rfloor$ désigne la partie entière de x). Alors, $f \circ g = \text{id}_{\mathbb{N}}$, g est injective, f est surjective, aucune des deux n'est bijective.

Par contre, si $f \circ g = \text{id}_B$ et $g \circ f = \text{id}_A$, f et g sont bijectives. Si $b \in B$, il a un antécédent unique par f , et comme $b = f \circ g(b)$, cet antécédent est $g(b)$ donc $g = f^{-1}$.

Solution 1.17.

- Pour l'ordre naturel, qui est total, une partie de $\mathbb{N}$ a toujours un plus petit élément, elle admet un plus grand élément si elle est majorée. Toute partie de $\mathbb{Z}$ admet un plus petit élément si elle est minorée, un plus grand si elle est majorée. Par contre, il existe des parties de $\mathbb{Q}$ qui n'ont pas de plus grand élément ni même de borne supérieure bien qu'elles soient majorées ; l'exemple le plus simple est celui de

$$A = \{x \in \mathbb{Q} \mid x^2 < 2\}$$

C'est ce « défaut » qui conduit à la construction des nombres réels. Dans $\mathbb{R}$, toute partie non vide et majorée admet une borne supérieure.

- Pour l'ordre divise, une partie comme $\{2, 3\}$ n'a ni plus petit, ni plus grand élément. Par contre, toute partie finie admet une borne supérieure, son p.p.c.m., et une borne inférieure, son p.g.c.d. Attention, avec notre définition de la relation d'ordre divise, $\mathbb{N}$ admet un plus petit élément 1 et un plus grand élément...0. On n'a donc pas toujours $a \mid b \Rightarrow a \leq b$.
- $\mathcal{P}(E)$ muni de l'inclusion a un plus petit élément $\emptyset$ et un plus grand élément E . L'ordre n'est pas total (sauf si E est de cardinal 0 ou 1), et toute partie non vide de $\mathcal{P}(E)$ a une borne supérieure qui est la réunion de ses éléments, une borne inférieure qui est leur intersection.
- L'ordre produit est de façon immédiate une relation d'ordre, il n'est pas total, car $(1, 2)$ et $(2, 1)$ par exemple ne sont pas comparables. Il est instructif de visualiser cet ordre : on représente les points à coordonnées entières dans le plan, et les couples supérieurs au couple (a, b) sont dans un quart de plan limité par les droites $x = a$ et $y = b$, (en haut à droite) les éléments inférieurs sont dans un rectangle limité par les mêmes droites et les axes. On pourra facilement utiliser cette représentation pour examiner les parties qui ont un plus grand élément, et pour justifier que toute partie majorée admet une borne supérieure, que toute partie admet une borne inférieure.

Parties non vides bien sûr.

En ce qui concerne l'ordre lexicographique, commençons par dire qu'il s'inspire de l'ordre alphabétique : un mot commençant par A se place avant un mot commençant par $B, C, \dots$, et si deux mots commencent par A , on compare les secondes lettres. Le fait que l'ordre lexicographique est un ordre demande des vérifications. Montrons par exemple la transitivité. On suppose que

$$(a, b) \preceq (c, d) \iff (a < c) \quad (1) \quad \text{ou} \quad (a = c \text{ et } b \leq d) \quad (2)$$

$$(c, d) \preceq (e, f) \iff (c < e) \quad (3) \quad \text{ou} \quad (c = e \text{ et } d \leq f) \quad (4)$$

Comme, dans chacune des hypothèses, les possibilités 1 et 2 (resp. 3 et 4) s'excluent, il suffit d'observer que ((1) et (3) ou (1)) et (4) ou ((2) et (3)) impliquent $a < e$, puis que ((2) et (4)) impliquent $a = e$ et $b \leq f$ pour établir la transitivité. L'ordre lexicographique est total, et une preuve « géométrique » apparaît si on utilise une représentation graphique comme pour l'ordre produit. Cette fois, l'ensemble des couples plus grands que (a, b) et l'ensemble des couples plus petits constituent une partition de $\mathbb{N} \times \mathbb{N}$. Pour conclure, c'est aussi un bon ordre, mais différent de celui de $\mathbb{N}$: il existe des sous-ensembles infinis qui ont un plus grand élément, comme $\{0\} \times \mathbb{N} \cup (1, 0)$.

Chapitre 2

Anneaux et corps

L'ensemble $\mathbb{Z}$ des entiers relatifs, muni de l'addition et de la multiplication est le prototype des **anneaux**. Le but de ce chapitre est de chercher comment on peut généraliser à un anneau quelconque les propriétés de divisibilité de $\mathbb{Z}$.

2.1 ANNEAUX, SOUS-ANNEAUX, IDÉAUX

2.1.1. Définitions et exemples

Définition 2.1. *Un anneau $(A, +, \times)$ est un ensemble non vide dans lequel on a défini une addition $+$ et une multiplication $\times$ telles que :*

- *$(A, +)$ est un groupe commutatif d'élément neutre 0_A .*
- *la multiplication est associative, distributive à gauche et à droite sur l'addition, et possède un élément neutre 1_A .*

Si la multiplication est commutative, on dit que A est un anneau commutatif.

Les ensembles de nombres

$$\mathbb{Z} \subset \mathbb{D} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C} \subset \mathbb{H}$$

(entiers relatifs, décimaux, rationnels, réels, complexes, quaternions) sont des anneaux pour les opérations habituelles. Il existe beaucoup d'anneaux entre $\mathbb{Z}$ et $\mathbb{Q}$, construits

comme l'anneau des décimaux ou autrement, et beaucoup entre $\mathbb{Z}$ et $\mathbb{C}$, comme l'anneau des entiers de Gauss $\mathbb{Z}[i]$ que nous étudions dans un problème en fin de chapitre. Comme d'habitude, le produit $a \times b$ sera souvent noté par juxtaposition ab , aa s'écrira a^2 , etc.

Un autre exemple d'anneau est l'anneau nul, qui ne contient que 0, les opérations étant $0 + 0 = 0$ et $0 \times 0 = 0$. Montrons le petit théorème suivant :

Proposition 2.2. *Dans un anneau, $0_A \times a = a \times 0_A = 0_A$ et si A n'est pas l'anneau nul, $0_A \neq 1_A$,*

Démonstration.

$$0_A a = (0_A + 0_A) a = 0_A a + 0_A a \quad \text{donc} \quad 0_A a = 0_A$$

en ajoutant l'opposé de $0_A a$. Par ailleurs, si $1_A = 0_A$, alors pour tout a de l'anneau, $a = 1_A a = 0_A a$ donc $a = 0_A$, et l'anneau A ne contient que 0_A . $\square$

Beaucoup d'énoncés ne seront vrais que pour les anneaux non nuls. Autres exemples d'anneaux :

- les anneaux résiduels de la forme $\mathbb{Z}/n\mathbb{Z}$, anneaux finis de cardinal n .
- les anneaux de matrices $\mathcal{M}_n(K)$ ou $\mathcal{M}_n(A)$, à coefficients dans un corps ou dans un anneau commutatif, les anneaux d'endomorphismes d'un espace vectoriel.
- les anneaux d'applications d'un ensemble E dans un corps (ou un anneau), pour les opérations définies par :

$$f + g : x \mapsto f(x) + g(x) \quad \text{et} \quad f \times g : x \mapsto f(x)g(x)$$

- les anneaux de polynômes, à coefficients dans un anneau ou dans un corps commutatifs.

Ces anneaux partagent beaucoup de propriétés avec l'anneau des entiers, mais... pas toutes. Un des objectifs de ce chapitre est de « comprendre » les propriétés arithmétiques de $\mathbb{Z}$, et de chercher comment et à quel point elles peuvent se généraliser.

2.1.2. Inversibles et diviseurs de zéro

Un élément a d'un anneau est **inversible** s'il existe a^{-1} dans l'anneau tel que $aa^{-1} = a^{-1}a = 1$. On dit également que c'est une **unité**. Dans un anneau non commutatif, un élément peut n'être inversible que à gauche ou que à droite : voir des exemples en exercice. Rappelons également que le produit de deux éléments inversibles a et b est inversible, d'inverse $b^{-1}a^{-1}$. Il est donc immédiat que :

Proposition 2.3. *L'ensemble $A^\times$ des éléments d'un anneau qui sont inversibles est un groupe pour la multiplication.*

Une propriété utile pour les calculs :

Définition 2.4. $a \in A$ est **régulier à gauche** si

$$\forall (x, y) \in A^2, \quad (ax = ay) \Rightarrow (x = y)$$

Cette propriété est moins forte que l'inversibilité car

Proposition 2.5. Si $a \in A$ est inversible à gauche, alors a est régulier à gauche.

Démonstration.

$$(ax = ax') \Rightarrow (a^{-1}(ax) = a^{-1}(ax')) \Rightarrow x = x'$$

□

On peut également définir les éléments réguliers à droite, réguliers des deux côtés (on dira simplement réguliers).

La notion « contraire » est également utile.

Définition 2.6. Un élément a d'un anneau A est un **diviseur de zéro à gauche**, s'il est non nul, et s'il existe b non nul tel que $ab = 0$.

Proposition 2.7. $a \in A$ est non nul et non diviseur de zéro à gauche si et seulement si il est régulier à gauche.

Démonstration. Supposons a non nul et non diviseur de zéro à gauche

$$\forall (b, b') \in A^2, \quad ab = ab' \Rightarrow a(b - b') = 0 \Rightarrow b = b'$$

et dans l'autre sens, si $ab = 0$, on a $ab = a0$, donc si a est régulier à gauche, $b = 0$ et a n'est pas diviseur de zéro à gauche. □

Les mêmes énoncés s'étendent au cas à droite.

L'ensemble des réguliers peut contenir strictement l'ensemble des inversibles, c'est par exemple le cas dans $\mathbb{Z}$. Parmi les anneaux connus :

- (1) Les inversibles de $\mathbb{Z}$ sont 1 et -1 . Les éléments réguliers sont tous les éléments non nuls.
- (2) Les inversibles et les réguliers de $\mathbb{Z}/n\mathbb{Z}$ sont les $\bar{k}$ où k est premier à n .
- (3) Les inversibles et les réguliers de $\mathcal{M}_n(K)$ sont les matrices de déterminant non nul.

La recherche des réguliers de $\mathcal{M}_n(K)$ est proposée en exercice. Pour les inversibles de $\mathbb{Z}/n\mathbb{Z}$, rappelons l'équivalence qui fait tout fonctionner :

$$\exists \bar{k}', \bar{k} \bar{k}' = \bar{1} \pmod{n} \iff \exists k' \in \mathbb{Z}, \exists \ell \in \mathbb{Z}, kk' - \ell n = 1 \iff k \wedge n = 1$$

d'après le théorème de Bezout. L'algorithme d'Euclide permet de trouver une solution explicite rapidement. Par ailleurs, si $k \wedge n = d$ avec d différent de 1 et de n , alors $\overline{k}$ est diviseur de zéro. En effet,

$$k = k'd, \quad n = n'd \quad \text{donc} \quad kn' = k'n \quad \text{soit} \quad \overline{k} \overline{n'} = \overline{0}$$

avec $\overline{n'} \neq \overline{0}$. On peut aussi utiliser l'exercice 2.1.

Terminons ces généralités par deux définitions :

Définition 2.8.

- Un anneau non nul est **intègre** s'il n'a pas de diviseur de zéro.
- Un anneau non nul est un **corps** si tous ses éléments non nuls sont inversibles.

Comme un inversible n'est jamais diviseur de zéro, un corps est intègre. Rappelons le cas bien connu :

Proposition 2.9. *L'anneau $\mathbb{Z}/n\mathbb{Z}$ est un corps si et seulement si n est un nombre premier p , on le note alors $\mathbb{F}_p$*

qui résulte de l'examen des inversibles de $\mathbb{Z}/n\mathbb{Z}$ évoqués ci-dessus.

2.1.3. Sous-anneaux, idéaux

Si A est un anneau, un sous-ensemble non vide B est un **sous-anneau** de A s'il est un sous-anneau pour les mêmes opérations, et avec le même élément neutre $1_A = 1_B$ ¹. On vérifie facilement que :

Proposition 2.10. *B est un sous-anneau de A si et seulement si*

$$\forall (x, y) \in B^2, \quad x + y \in B, \quad xy \in B.$$

$$\forall x \in B, \quad -x \in B.$$

$$1_A \in B.$$

L'intersection d'une famille non vide $(B_i)_{i \in I}$ de sous-anneaux de A est un sous-anneau de A , ce qui permet de donner la définition :

Définition 2.11. *Si A est un anneau et X un sous-ensemble non vide de A , le sous-anneau de A engendré par X est l'intersection des sous-anneaux de A qui le contiennent. C'est aussi le plus petit anneau de A qui contient X .*

Par exemple, si A est l'anneau des fonctions définies sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$, le sous-anneau des fonctions polynômes est le sous-anneau engendré par la fonction $x \mapsto x$,

1. Si A n'est pas l'anneau nul, $\{0_A\}$ n'est donc pas un sous-anneau de A .

le sous-anneau des polynômes trigonométriques est le sous-anneau engendré par les fonctions $x \mapsto \sin x$ et $x \mapsto \cos x$.

2.1.4. Morphismes

Comme on définit des applications linéaires, on peut définir des applications entre anneaux qui préservent les opérations.

Définition 2.12. Soit A et B deux anneaux. Un **morphisme d'anneaux** est une application $f : A \rightarrow B$ telle que, pour tout (a, b) de A^2 :

$$f(a + b) = f(a) + f(b), \quad f(ab) = f(a)f(b), \quad f(1_A) = 1_B$$

On vérifie facilement qu'on a également : $f(0_A) = 0_B$ en calculant $f(0_A + 0_A)$. Un exemple de morphisme de $\mathbb{C}[X]$ dans lui-même : $P(X) \mapsto P(X + a)$. Par contre $P(X) \mapsto P(X)^2$ ne définit pas un morphisme d'anneaux. Remarquons qu'on ne peut déduire $f(1_A) = 1_B$ de $f(ab) = f(a)f(b)$.

Proposition 2.13. Soit f un morphisme d'anneaux de A dans B . Alors :

- Si C est un sous-anneau de A , $f(C)$ est un sous-anneau de B .
- Si D est un sous-anneau de B , $f^{-1}(D)$ est un sous-anneau de A .

Deux cas particuliers, avec les mêmes notations :

Définition 2.14. Le noyau d'un morphisme est $\text{Ker}(f) = f^{-1}(0)$; son image est $\text{Im}(f) = f(A)$.

L'image est un sous-anneau mais pas le noyau car il ne contient pas 1_A (sauf le cas exceptionnel du morphisme nul dans l'anneau nul). De plus :

Proposition 2.15. Un morphisme d'anneaux f de A vers B est injectif si et seulement si $\text{Ker } f$ est réduit à $\{0\}$. Il est surjectif si et seulement si $\text{Im } f = B$.

2.1.5. Idéaux

Il se trouve que la notion de sous-anneau n'est pas la plus riche. La relation d'équivalence $a \equiv b \iff a - b \in B$ n'est en général pas compatible avec le produit lorsque B est un sous-anneau de A . Pour définir des quotients, il faut utiliser des **idéaux** et même des **idéaux bilatères**.

Définition 2.16. Si A est un anneau, un sous-ensemble non vide $\mathcal{I}$ s'appelle un idéal à gauche (resp. à droite) si :

- $\forall x \in \mathcal{I}, \forall y \in \mathcal{I}, \quad x + y \in \mathcal{I} \text{ et } -x \in \mathcal{I}$
- $\forall a \in A, \forall x \in \mathcal{I}, \quad ax \in \mathcal{I}$
- (resp. $\forall a \in A, \forall x \in \mathcal{I}, \quad xa \in \mathcal{I}$)

Définition 2.17. $\mathcal{I} \subset A$ est un idéal **bilatère** si c'est à la fois un idéal à gauche et un idéal à droite de A .

Bien sûr, dans le cas où la multiplication est commutative, ces trois notions coïncident. Notons que $\{0_A\}$ et A lui-même sont des idéaux bilatères. On les appelle parfois **idéaux triviaux**.

Remarque : Un idéal est forcément un sous-groupe pour l'addition. En ce qui concerne la multiplication, la contrainte est différente que pour un sous-anneau, mais on ne demande pas que 1_A soit dans l'idéal. D'ailleurs, si un idéal à gauche $\mathcal{I}$ contient un inversible a , la seconde propriété impose qu'il contienne $a^{-1}a = 1_A$, et en réappliquant cette même propriété, il doit contenir tout $b1_A$ où $b \in A$, $\mathcal{I}$ coïncide avec A .

Pour prolonger la fin de la remarque :

Proposition 2.18. Un anneau $A \neq \{0\}$ est un corps si et seulement si il ne contient aucun idéal à gauche autre que 0 et lui-même.

Démonstration. Si A est un corps, tout idéal non nul contient un inversible, donc coïncide avec A , par la remarque. Réciproquement, soit a un élément non nul de A . Alors Aa est un idéal à gauche, non nul donc coïncidant avec A . Comme A contient 1_A , il existe a' tel que $a'a = 1_A$, a admet un inverse à gauche. Comme a' est non nul, il existe a'' tel que $a''a' = 1_A$. D'où

$$a'' = a''(a'a) = (a''a')a = a$$

a' est aussi l'inverse à droite de a . On montre de même qu'un anneau qui n'a aucun idéal à droite non trivial est un corps. Par contre, il existe des anneaux qui n'ont aucun idéal bilatère, mais qui ne sont pas des corps ; c'est par exemple le cas de l'anneau des matrices carrées ($n \geq 2$) à coefficients dans un corps : voir l'exercice 2.4. $\square$

Les sous-ensembles de la forme $n\mathbb{Z}$ sont des idéaux de $\mathbb{Z}$, la vérification est immédiate. Mais il y a plus :

Proposition 2.19. Les seuls idéaux de $\mathbb{Z}$ sont de la forme $n\mathbb{Z}$.

Démonstration. Nous avons déjà noté que ce sont bien des idéaux (regarder en passant les cas particuliers $n = 0$ et $n = 1$). La réciproque utilise la division euclidienne : soit $\mathcal{I}$ un idéal de $\mathbb{Z}$, qu'on suppose non réduit à $\{0\}$. Alors $\mathcal{I}$ contient des éléments positifs (car s'il contient x , il contient aussi $(-1)x$), et comme $\mathbb{N}$ est bien ordonné, il contient un plus petit élément strictement positif, noté n . Alors $\mathcal{I}$ contient alors $n\mathbb{Z}$, par définition d'un idéal. Et si $m \in \mathcal{I}$, on peut le diviser par n d'où $m = nq + r$ avec $0 \leq r < n$, et alors $r = m - nq \in \mathcal{I}$, donc $r = 0$ par le caractère minimal de n . $\square$

2.1.6. Anneaux quotients

Dans ce paragraphe, nous allons généraliser dans un anneau quelconque la construction faite pour l'anneau $\mathbb{Z}/n\mathbb{Z}$. À tout idéal $\mathcal{I}$ **bilatère** on peut associer une relation $\equiv$ définie par :

$$x \equiv y \iff x - y \in \mathcal{I}$$

Proposition 2.20. $\equiv$ est une relation d'équivalence, compatible avec les opérations.

Démonstration.

- (1) $0 \in \mathcal{I}$, d'où la réflexivité. Si on suppose $x \equiv y$, alors $x - y \in \mathcal{I}$ donc $-(x - y) \in \mathcal{I}$ et $y \equiv x$, la relation est symétrique. Enfin,

$$\left. \begin{array}{l} x \equiv y \\ y \equiv z \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} x - y \in \mathcal{I} \\ y - z \in \mathcal{I} \end{array} \right\} \Rightarrow (x - y) + (y - z) \in \mathcal{I}$$

et $x \equiv z$, la relation est transitive.

- (2) Si $x \equiv x'$ alors $xy \equiv x'y$ et si $y \equiv y'$ alors $x'y \equiv x'y'$ puisque l'idéal est bilatère. Et donc

$$x \equiv x' \text{ et } y \equiv y' \Rightarrow xy \equiv x'y \equiv x'y'$$

On vérifie de même que :

$$x \equiv x' \text{ et } y \equiv y' \Rightarrow x + y \equiv x' + y'$$

$\square$

Il est alors possible de considérer l'ensemble des classes d'équivalence, c'est ce qu'on appelle l'**anneau quotient**, noté $A/\mathcal{I}$, et de le munir d'une structure d'anneau :

Théorème 2.21.

- Les classes d'équivalence sont de la forme $a + \mathcal{I}$, abrégé en $\bar{a}$ s'il n'y a pas d'ambiguïté.
- Si on pose $\bar{a} + \bar{b} = \overline{a + b}$ et $\bar{a}\bar{b} = \overline{ab}$, ces opérations sont bien définies.
- $A/\mathcal{I}$ est alors un anneau d'éléments neutres $\bar{0}$ et $\bar{1}$.
- L'application $\pi : A \rightarrow A/\mathcal{I}$ définie par $a \mapsto \bar{a}$ est appelée **projection**. C'est un morphisme surjectif dont le noyau est $\mathcal{I}$.

Démonstration.

- Si $b \equiv a$ alors il existe $i \in \mathcal{I}$ tel que $b - a = i$ d'où $b \in a + \mathcal{I}$, et réciproquement.
- La proposition précédente montre que :

$$\overline{a} = \overline{a'} \text{ et } \overline{b} = \overline{b'} \Rightarrow \overline{a + b} = \overline{a' + b'}$$

ce qui montre bien la cohérence de la définition de la somme de deux classes. Il en va de même pour la définition du produit.

- C'est une simple écriture : les propriétés d'associativité, de distributivité, etc. de l'anneau A sont transmises au quotient.
- La surjectivité découle de la définition du quotient. De plus, l'image réciproque de $\overline{0}$ est l'ensemble des a tels que $a \equiv 0$, c'est $\mathcal{I}$.

□

On vient de voir que le noyau de cette projection est l'idéal bilatère $\mathcal{I}$; c'est un fait beaucoup plus général, le théorème suivant fait le lien entre morphismes et idéaux.

Théorème 2.22. *Soit $f : A \rightarrow B$ un morphisme d'anneaux. Alors, et les deux premiers énoncés concernent soit les idéaux à droite, soit les idéaux à gauche, soit les idéaux bilatères :*

- Si $\mathcal{I} \subset A$ est un idéal de A , alors $f(\mathcal{I})$ est un idéal de l'anneau $f(A)$.
- Si $\mathcal{J} \subset B$ est un idéal de B , alors $f^{-1}(\mathcal{J})$ est un idéal de A .
- En particulier, le noyau d'un morphisme d'anneaux est toujours un idéal bilatère.

Démonstration. Attention à la différence entre les deux cas, image directe et image réciproque. Regardons le premier cas : Si j et j' sont dans $f(\mathcal{I})$, b dans $f(A)$, alors il existe i et i' dans $\mathcal{I}$ et a dans A tels que :

$$j = f(i), j' = f(i'), b = f(a)$$

Comme f est un morphisme et $\mathcal{I}$ un idéal, on a :

$$j - j' = f(i - i') \in f(\mathcal{I}), bj = f(ai) \in f(\mathcal{I})$$

pour le cas idéal à gauche. Remarquer pourquoi l'image de $\mathcal{I}$ n'est pas forcément un idéal de B . La suite de la démonstration est immédiate. Enfin, le dernier point découle de ce que $\{0\}$ est manifestement un idéal bilatère. □

Nous allons terminer cette étude des anneaux quotients par deux théorèmes importants. Le premier appelé théorème de correspondance permet de décrire les idéaux d'un quotient.

Théorème 2.23. Théorème de correspondance. *Soit $A/\mathcal{I}$ un anneau quotient, où $\mathcal{I}$ est bilatère. Alors, si on note π la projection, les idéaux (à gauche, à droite, bilatères) de $A/\mathcal{I}$ sont les $\pi(\mathcal{J})$, $\mathcal{J}$ décrivant l'ensemble des idéaux (à gauche, à droite, bilatères) de A contenant $\mathcal{I}$.*

Démonstration. Ce sont des idéaux, car π est un morphisme surjectif. Si maintenant $\mathcal{K}$ est un idéal de l'anneau quotient, alors $\pi^{-1}(\mathcal{K})$ est un idéal de A ; il contient $\mathcal{I}$ car $\mathcal{K}$ contient $\bar{0}$. De plus, comme π est surjective, $\mathcal{K} = \pi(\pi^{-1}(\mathcal{K}))$. Enfin, si $\mathcal{J}$ est un idéal de A contenant $\mathcal{I}$, on a $\pi^{-1}(\pi(\mathcal{J})) = \mathcal{J}$: si $a \in \pi^{-1}(\pi(\mathcal{J}))$, alors $\pi(a) \in \pi(\mathcal{J})$, donc il existe $j \in \mathcal{J}$ tel que $a \equiv j$, soit $a - j \in \mathcal{I}$, mais on en déduit que a est somme de deux éléments de $\mathcal{J}$ puisque $\mathcal{J}$ contient $\mathcal{I}$. On a donc $\pi^{-1}(\pi(\mathcal{J})) \subset \mathcal{J}$, l'autre inclusion est immédiate. En définitive, la correspondance décrite est bijective. $\square$

Et le second théorème :

Théorème 2.24. Théorème d'isomorphisme. Si $f : A \rightarrow B$ est un morphisme d'anneau, alors :

$$A/\text{Ker}(f) \simeq \text{Im}(f)$$

($\simeq$ désigne un isomorphisme d'anneaux).

Démonstration. Décrivons cet isomorphisme, noté $\bar{f}$: on pose $\bar{f}(\bar{a}) = f(a)$. Il faut vérifier que cela ne dépend pas du représentant choisi pour la classe de a :

$$\bar{a} = \bar{b} \Rightarrow a - b \in \text{Ker}(f) \Rightarrow f(a - b) = 0 \Rightarrow f(a) = f(b)$$

$$\bar{f}(\overline{a+b}) = f(a+b) = f(a) + f(b) = \bar{f}(\bar{a}) + \bar{f}(\bar{b})$$

idem pour le produit et pour l'image de $1_{A/\text{Ker } f}$: c'est bien un morphisme d'anneaux. Il est surjectif par choix de l'ensemble d'arrivée, et injectif car son noyau est l'ensemble des $\bar{a}$ pour tous les $a \in \text{Ker } f$, c'est donc $\{\bar{0}\}$. $\square$

Donnons un exemple. L'application $P \mapsto P(i)$ de $\mathbb{R}[X]$ dans $\mathbb{C}$ est un morphisme d'anneaux ; il est surjectif (prendre un polynôme P du premier degré). Un polynôme P est dans le noyau si $P(i) = 0$. Mais comme P est à coefficients réels, s'il admet i comme racine, il admet aussi $-i$, conjugué de i comme racine. Il est donc divisible par $(X - i)(X + i) = X^2 + 1$, et réciproquement tout polynôme multiple de $X^2 + 1$ est dans le noyau. L'ensemble des multiples de $X^2 + 1$ est tout simplement noté $(X^2 + 1)$. On en déduit que :

$$\mathbb{R}[X]/(X^2 + 1) \simeq \mathbb{C}$$

Cet exemple est très important, il est à la base de l'étude des corps, et nous aurons l'occasion d'y revenir dans le chapitre suivant.

2.1.7. Caractéristique d'un anneau

Puisque 1_A est un élément de l'anneau A , celui-ci contient aussi $1_A + 1_A$, $1_A + 1_A + 1_A$, ..., et leurs opposés. Ces éléments sont notés 2.1_A , 3.1_A ou par abus 2 , 3 , ..., mais ils ne sont pas forcément distincts. Plus précisément :

Définition 2.25. Si A est un anneau, l'application $k \mapsto k \cdot 1_A$ de $\mathbb{Z}$ dans A est un morphisme d'anneaux. Son noyau est de la forme $n\mathbb{Z}$ et n s'appelle la **caractéristique** de A .

L'application est un morphisme,

$$k \cdot 1_A + k' \cdot 1_A = (k + k') \cdot 1_A \quad \text{et} \quad (k \cdot 1_A)(k' \cdot 1_A) = (kk') \cdot 1_A$$

et ce morphisme est rarement surjectif (quand l'est-il ?). On peut démontrer les propriétés suivantes :

Proposition 2.26. Soit A un anneau de caractéristique n . Alors,

- (i) si $n = 0$, alors A est infini,
- (ii) si A est intègre alors $n = 0$ ou $n = p$ nombre premier.
- (iii) si A est commutatif et $n = p$ premier, alors :

$$\begin{aligned} f : A &\longrightarrow A \\ a &\longmapsto a^p \end{aligned}$$

est un morphisme d'anneaux (appelé « morphisme de Frobenius »).

Démonstration.

- (i) Le théorème d'isomorphisme dit que notre morphisme a une image isomorphe à $\mathbb{Z}$, qui est infini ; on dit que $\mathbb{Z}$ s'injecte dans A .
- (ii) Cette fois, en raisonnant par l'absurde, on voit que si A est de caractéristique non nulle et non première, il contient un sous-anneau isomorphe à $\mathbb{Z}/n\mathbb{Z}$, qui n'est intègre que si n est premier.
- (iii) On commence par le lemme :

$$0 < k < p \Rightarrow p \text{ divise } \binom{p}{k} = \frac{p(p-1) \dots (p-k+1)}{k!}$$

En effet, dans l'égalité $k! \binom{p}{k} = p(p-1) \dots (p-k+1)$, le nombre p divise le second membre, et il est premier à $k!$ lorsque $k < p$. On applique alors le lemme de Gauss.

Si maintenant on considère l'application $a \mapsto a^p$, on a

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

tandis que $(ab)^p = a^p b^p$. Remarquons que pour ces deux propriétés on utilise la commutativité de l'anneau.

□

2.1.8. Anneaux produits

Définition 2.27. Si A et B sont deux anneaux, on appelle **anneau produit** le produit cartésien $A \times B$ muni des opérations :

$$(a, b) + (a', b') = (a + a', b + b') \quad \text{et} \quad (a, b)(a', b') = (aa', bb')$$

Il faut bien sûr vérifier que c'est un anneau. Cela n'offre pas de difficulté, et on peut définir de même le produit d'une famille d'anneaux. Attention, le produit de deux corps n'est pas un corps : chercher les inversibles du produit $\mathbb{R} \times \mathbb{R}$. Dans le même ordre d'idées, $A \times \{0_B\}$ est un idéal bilatère de $A \times B$, a une structure d'anneau (isomorphe à A), mais n'est pas un sous-anneau de $A \times B$, car il n'a pas le même élément neutre pour le produit.

Exercice 2.1. Soit A un anneau et $a \in A$. Montrer que a est régulier à gauche si et seulement si l'application $x \mapsto ax$ est injective. En déduire que, dans les anneaux finis, la notion d'élément régulier coïncide avec la notion d'élément inversible.

Exercice 2.2. Déterminer dans l'anneau $\mathcal{M}_n(K)$ les diviseurs de zéro, à gauche ou à droite. On pourra utiliser des endomorphismes associés. Comparer l'ensemble des inversibles et l'ensemble des réguliers.

Exercice 2.3. Soit A un anneau et a un de ses éléments. On appelle annulateur à droite de a l'ensemble des éléments x de A tels que $ax = 0$. Montrer que c'est un idéal à droite de A . Décrire l'annulateur d'un élément $\bar{k}$ de $\mathbb{Z}/n\mathbb{Z}$, puis l'annulateur à droite d'une matrice $M \in \mathcal{M}_n(K)$, décrire également l'annulateur à gauche.

Exercice 2.4. Dans l'anneau $\mathcal{M}_n(K)$, montrer que les seuls idéaux bilatères sont l'idéal nul et l'anneau tout entier. On pourra utiliser les matrices de base E_{ij} .

Exercice 2.5. Décrire les idéaux de $\mathbb{Z}/n\mathbb{Z}$.

Exercice 2.6. Dans l'anneau $\mathbb{Z}/n\mathbb{Z}$, rechercher s'il existe des éléments **nilpotents**, c'est-à-dire dont une puissance est nulle.

2.2 DIVISIBILITÉ

2.2.1. Vocabulaire

Examinons maintenant la relation de divisibilité dans un anneau commutatif ; à partir de maintenant, on préférera se limiter à un **anneau commutatif intègre** A : certaines notions pourront être définies dans des anneaux plus généraux, mais elles ont moins d'intérêt et des propriétés différentes.

Définition 2.28.

- On dit que a et b sont **associés** s'il existe une unité u telle que $a = bu$. Cette relation est une relation d'équivalence.
- On dit que $a \mid b$ (a divise b) s'il existe c tel que $b = ac$. Ce n'est pas une relation d'ordre : si $a \mid b$ et $b \mid a$, alors a et b sont associés.

- On dit que $a \in A \setminus (A^\times \cup \{0\})$ est **irréductible** si :

$$\forall (b, c) \in A^2, \quad (a = bc) \Rightarrow (b \in A^\times \text{ ou } c \in A^\times)$$

- On dit que $a \in A \setminus (A^\times \cup \{0\})$ est **premier** si :

$$\forall (b, c) \in A^2, \quad (a \mid bc) \Rightarrow (a \mid b \text{ ou } a \mid c)$$

Une affirmation à vérifier dans cette définition : si $a \mid b$ et $b \mid a$, alors a et b sont associés. En effet, il existe alors c et c' tels que $b = ca$ et $a = bc'$. On en déduit $a = acc'$. Si a est nul, b aussi, et a et b sont associés, sinon, par intégrité, $cc' = 1$ et a et b sont associés. Dans $\mathbb{Z}$ un nombre premier est irréductible (c'est la définition) mais aussi premier dans le sens que nous venons de définir : c'est le lemme d'Euclide, voir 1.14 p. 14. C'est heureux.

Remarque : Dans le cas général, les deux notions sont différentes (voir par exemple l'exercice 2.9. Cependant, **un élément premier est toujours irréductible** : si a est premier et $a = bc$, alors $a \mid bc$. Donc, $a \mid b$ (par exemple) et $b = ad$. On en déduit : $a = adc$ et l'intégrité prouve que $1 = dc$, donc c est une unité, et a est irréductible.

La relation de divisibilité peut aussi s'exprimer à l'aide des idéaux.

2.2.2. Idéaux principaux

Dans un anneau commutatif A tous les idéaux sont bilatères : on pourra se contenter de parler d'idéaux, sans plus de précision. Comme l'intersection d'une famille d'idéaux est un idéal, on peut définir l'idéal engendré par une partie S non vide de A : c'est le plus petit idéal contenant les éléments de S , intersection de tous les idéaux qui contiennent S . On peut le décrire explicitement :

Proposition 2.29. *L'idéal engendré par S est l'ensemble des $b = \sum_{i=1}^k a_i s_i$ où (s_i) et (a_i) sont des suites finies quelconques d'éléments de S et de A . On le note souvent (S) .*

Démonstration. Indiquons seulement les étapes : on vérifie que (S) ainsi défini est bien un idéal de A , qu'il contient S , et que tout idéal contenant S contient (S) . Remarquer l'analogie avec les sous-espaces vectoriels et les combinaisons linéaires. $\square$

En particulier, si S est réduit à un élément s , l'idéal qu'il engendre sera appelé **idéal principal**. On le notera (s) ou plus explicitement sA . On retrouve ainsi la notation déjà rencontrée dans $\mathbb{Z}$, par exemple $2\mathbb{Z}$ est l'idéal principal engendré par 2.

Proposition 2.30. *Si A est un anneau commutatif intègre,*

- $a \mid b \iff (b) \subset (a)$.
- a et b sont associés si et seulement si $(a) = (b)$.
- $a \in A \setminus (A^\times \cup \{0\})$ est irréductible si et seulement si (a) est **maximal** dans l'ensemble des idéaux principaux différents de A .
- $a \in A \setminus (A^\times \cup \{0\})$ est premier si et seulement si :

$$\forall (b, c) \in A^2, \quad (bc \in (a)) \Rightarrow (b \in (a) \text{ ou } c \in (a))$$

Démonstration.

- $b = ac \Rightarrow bA = acA \subset aA$ et réciproquement, si $(b) \subset (a)$, $b \in (a)$ donc il existe c dans A tel que $b = ac$.
- On a déjà vu que a et b associés équivaut à $a \mid b$ et $b \mid a$ simultanément, donc équivaut à $(a) = (b)$ par ce qui précède.
- Soit (b) tel que $(a) \subset (b)$. Alors $b \mid a$. Comme a est irréductible, soit b est une unité, soit b est associé à a . Dans le premier cas $(b) = A$, dans le second $(b) = (a)$. La réciproque se traite de la même manière.
- Supposons donc a premier. Alors, si $bc \in (a)$, il existe m dans A tel que $bc = am$ et $a \mid bc$. Comme a est premier, $a \mid b$, par exemple, et donc $b \in (a)$. La démonstration de la réciproque est analogue, c'est un simple jeu de traduction. $\square$

La proposition précédente amène naturellement à deux définitions :

Définition 2.31. *Un idéal $\mathcal{I}$ de A est dit **premier** s'il est différent de A et vérifie la propriété :*

$$\forall (b, c) \in A^2, \quad bc \in \mathcal{I} \Rightarrow (b \in \mathcal{I} \text{ ou } c \in \mathcal{I})$$

Définition 2.32. *Un idéal $\mathcal{I}$ de A est dit **maximal** s'il est maximal dans l'ensemble des idéaux stricts de A .*

Ainsi, si a est premier, l'idéal (a) est un idéal premier. Un idéal (a) maximal est forcément celui d'un irréductible. On peut alors énoncer le théorème :

Théorème 2.33.

- Un idéal $\mathcal{I}$ est maximal si et seulement si $A/\mathcal{I}$ est un corps.
- Un idéal $\mathcal{I}$ est premier si et seulement si $A/\mathcal{I}$ est intègre.
- Tout idéal maximal est aussi un idéal premier.

Démonstration. Dans le premier cas, on se sert du théorème de correspondance (théorème 2.23, page 32) : comme $\mathcal{I}$ est maximal, il n'y a pas d'idéal strictement entre $\mathcal{I}$ et A , donc $A/\mathcal{I}$ ne contient pas d'idéal non trivial, c'est un corps.

Dans le second cas, on fait un petit travail de traduction. L'intégrité dans le quotient, qui s'écrit

$$\forall (\bar{a}, \bar{b}) \in (A/\mathcal{I})^2, \quad (\bar{a}\bar{b} = \bar{0}) \quad \Rightarrow \quad (\bar{a} = \bar{0} \text{ ou } \bar{b} = \bar{0})$$

est équivalente à

$$\forall (a, b) \in A^2, \quad (ab \in \mathcal{I}) \quad \Rightarrow \quad (a \in \mathcal{I} \text{ ou } b \in \mathcal{I})$$

Il faut remarquer qu'un corps et un anneau intègre sont par définition non réduits à $\{0\}$, dans les deux cas $\mathcal{I}$ est différent de A .

La troisième assertion du théorème découle de ce qu'un corps est nécessairement intègre. $\square$

2.2.3. Le théorème chinois

Voici un exemple important d'utilisation des quotients ; on commence par définir une opération dans l'ensemble des idéaux.

Définition 2.34. Si A est un anneau commutatif, $\mathcal{I}$ et $\mathcal{J}$ deux idéaux, leur somme $\mathcal{I} + \mathcal{J}$ est l'ensemble des sommes d'un élément quelconque de $\mathcal{I}$ et d'un élément quelconque de $\mathcal{J}$.

Cette somme, analogue à la somme de sous-espaces vectoriels, est de façon immédiate un idéal.

Théorème 2.35. Théorème chinois. Soit A un anneau commutatif, $\mathcal{I}$ et $\mathcal{J}$ deux idéaux tels que $\mathcal{I} + \mathcal{J} = A$. On a alors un isomorphisme d'anneaux

$$A/(\mathcal{I} \cap \mathcal{J}) \simeq A/\mathcal{I} \times A/\mathcal{J}$$

Démonstration. On part de l'application $a \mapsto (a + \mathcal{I}, a + \mathcal{J})$ de A dans $A/\mathcal{I} \times A/\mathcal{J}$ qui est un morphisme d'anneaux car obtenue à l'aide des projections. Son noyau est formé des éléments de A qui se projettent en $(\bar{0}, \bar{0})$, c'est-à-dire qui sont dans $\mathcal{I}$ et

dans $\mathcal{J}$. C'est donc l'idéal $\mathcal{I} \cap \mathcal{J}$. Reste à montrer la surjectivité. Soit $(\alpha + \mathcal{I}, \beta + \mathcal{J})$, on cherche a de A tel que a vérifie

$$\begin{cases} a \equiv \alpha \pmod{\mathcal{I}} \\ a \equiv \beta \pmod{\mathcal{J}} \end{cases}$$

On utilise $\mathcal{I} + \mathcal{J} = A$ pour écrire $1 = x + y$ où $x \in \mathcal{I}$ et $y \in \mathcal{J}$; on a donc $y \equiv 1 \pmod{\mathcal{I}}$ et $x \equiv 1 \pmod{\mathcal{J}}$ il suffit alors de prendre $a = x\beta + y\alpha$. $\square$

Remarque : La condition $\mathcal{I} + \mathcal{J} = A$ s'exprime en disant que les idéaux sont **étrangers**. Par analogie avec le cas de $\mathbb{Z}$, les relations d'équivalence modulo un idéal s'appellent des congruences. Le théorème chinois revient donc à résoudre un système de congruences simultanées, c'est effectivement ce qu'on trouve dans les mathématiques chinoises. Nous verrons une version particulière de ce théorème dans le cas où A est un anneau principal.

Exercice 2.7. Opérations dans l'ensemble des idéaux. On suppose que A est un anneau commutatif. Si $\mathcal{I}$ et $\mathcal{J}$ sont deux idéaux de A , montrer que :

$$\sqrt{\mathcal{I}} = \{x \in A \mid \exists n \in \mathbb{N}, x^n \in \mathcal{I}\}$$

$$\mathcal{I} : \mathcal{J} = \{x \in A \mid x\mathcal{J} \subset \mathcal{I}\}$$

définissent des idéaux. Comment définir le produit $\mathcal{I}\mathcal{J}$ de deux idéaux ? Examiner le cas particulier des idéaux principaux. Montrer que si $\mathcal{I} + \mathcal{J} = A$, on a

$$\mathcal{I}\mathcal{J} = \mathcal{I} \cap \mathcal{J}$$

Donner des exemples, par exemple dans $\mathbb{Z}$.

Exercice 2.8. Montrer, à l'aide du théorème de Zorn, que dans un anneau non nul, tout idéal est inclus dans un idéal maximal. Ce résultat est parfois appelé théorème de Krull.

2.3 ANNEAUX PRINCIPAUX, EUCLIDIENS ET FACTORIELS

Continuons à étudier des cas particuliers d'anneaux, toujours en rapport avec la notion de divisibilité.

2.3.1. Anneaux euclidiens

Les anneaux euclidiens sont ceux pour lesquels il existe une division semblable à la division euclidienne des entiers.

Définition 2.36. Un anneau A est euclidien s'il est commutatif, intègre et s'il existe $\phi : A \setminus \{0\} \rightarrow \mathbb{N}$ telle que :

- (1) $b \mid a \Rightarrow \phi(b) \leq \phi(a)$
- (2) $\forall a \in A, \forall b \in A \setminus \{0\}, \exists q, r \in A$

$$\begin{cases} a = bq + r \\ r = 0 \text{ ou } \phi(r) < \phi(b) \end{cases}$$

Remarque : L'application ϕ porte le nom de **stathme** (du mot grec signifiant mesure) : on prend la valeur absolue pour $\mathbb{Z}$, le degré pour $K[X]$. Notons que la définition de la division euclidienne ne demande pas forcément l'unicité du couple (q, r) (que l'on appelle **quotient** et **reste**). Avec notre définition, il n'y a pas unicité dans le cas de $\mathbb{Z}$. $17 = 3 \times 5 + 2 = 3 \times 6 - 1$ sont deux divisions euclidiennes de 17 par 3. Il y aura unicité si on impose au reste d'être positif.

Outre $\mathbb{Z}$, un exemple d'anneau euclidien est fourni par $K[X]$ l'anneau des polynômes sur un corps commutatif, nous aurons l'occasion d'y revenir dans le chapitre suivant.

Proposition 2.37. Dans un anneau euclidien, tous les idéaux sont principaux.

Démonstration. On mime la démonstration que l'on fait pour $\mathbb{Z}$ (cf. 2.19). Soit $\mathcal{I}$ un idéal, différent de l'idéal nul, et x_0 un élément non nul de $\mathcal{I}$ de stathme minimal (ne pas oublier que les stathmes sont à valeurs entières). Si maintenant x est un élément quelconque de $\mathcal{I}$, on peut écrire $x = x_0q + r$ avec $r = 0$ ou $\phi(r) < \phi(x_0)$; comme $r = x - x_0q \in \mathcal{I}$, le choix de x_0 impose que $r = 0$. En définitive $x = x_0q$ et $\mathcal{I} = (x_0)$. $\square$

Cette propriété des anneaux euclidiens est essentielle, et conduit à la définition qui suit.

2.3.2. Anneaux principaux

Définition 2.38. Un anneau (qui n'est pas un corps) est **principal** s'il est intègre, commutatif, et si tout idéal est principal, c'est-à-dire engendré par un élément.

Les idéaux triviaux $\{0\} = (0)$ et $A = (1)$ sont en particuliers principaux, et nous venons de voir que tout anneau euclidien est nécessairement principal. Il n'est pas facile de trouver des exemples d'anneaux principaux qui ne sont pas euclidiens : voir néanmoins l'exercice 2.17.

Un premier résultat qui concerne les anneaux principaux.

Théorème 2.39. *Dans un anneau principal, un élément est irréductible si et seulement si il est premier.*

Démonstration. On sait déjà qu'un élément premier est toujours irréductible. Soit A un anneau principal, et p irréductible. Supposons que $p \mid ab$ et soit $\mathcal{I}$ l'idéal engendré par p et a . On peut écrire :

$$\mathcal{I} = \{z \in A \mid \exists x \in A, \exists y \in A, z = xp + ya\}$$

Comme l'anneau est principal, $\mathcal{I}$ est principal, et il existe d tel que $\mathcal{I} = dA$. Donc p et a qui sont dans $\mathcal{I}$ sont des multiples de d . Comme p est irréductible, d est soit une unité, soit un associé de p . Si c'est un associé de p , p divise a et c'est terminé ; si c'est une unité, alors $\mathcal{I} = A$ et $1 = xp + ya$ pour un x et un y de A . En multipliant par b on voit que p divise b . Remarquer la ressemblance de la démonstration avec celle du lemme de Gauss dans $\mathbb{Z}$. $\square$

Dans les anneaux principaux, les notions de premier et d'irréductible coïncident. Nous allons voir que c'est encore vrai dans la catégorie suivante d'anneaux, encore plus générale.

2.3.3. Anneaux factoriels

Rappelons que dans l'ensemble des entiers, tout nombre se décompose de façon unique en produit de nombres premiers. C'est ce qu'on appelle parfois le « théorème fondamental de l'arithmétique ». Les anneaux qui ont cette propriété s'appellent **anneaux factoriels**. Plus précisément :

Définition 2.40. *Un anneau commutatif intègre est factoriel si tout élément non nul et non inversible se décompose de façon unique en produit d'irréductibles.*

L'unicité s'entend à l'ordre près et aux unités près : par exemple dans $\mathbb{Z}$, on ne distingue pas $2 \times 3 \times 3$ et $-3 \times 2 \times (-3)$. Commençons par remarquer :

Proposition 2.41. *Dans un anneau factoriel, un élément est irréductible si et seulement si il est premier.*

Démonstration. On sait déjà que tout premier est un irréductible ; soit A un anneau factoriel et $a \in A$ un élément irréductible. Supposons que $a \mid bc$; cela signifie que dans la décomposition en irréductibles de bc , il y a l'irréductible a (à une unité près) et donc que la décomposition en irréductibles de b ou de c contient a , et donc a divise b ou c . $\square$

Nous terminons par le théorème principal de cette section : on sait déjà que tout anneau euclidien est principal, nous allons démontrer que tout anneau principal est factoriel, ce qui montre la hiérarchie entre les trois notions.

Théorème 2.42. *Tout anneau principal A est factoriel.*

Démonstration. La démonstration est un peu délicate. Faisons un raisonnement par l'absurde : soit $\mathcal{E}$ l'ensemble des éléments x de A , non nuls ni inversibles, qui ne se décomposent pas en irréductibles, et supposons que $\mathcal{E}$ soit non vide. Alors, $\mathfrak{F} = \{(x) \mid x \in \mathcal{E}\}$ est un ensemble non vide. Montrons que cet ensemble, muni de la relation d'inclusion, admet (au moins un) élément maximal. En effet, (nouveau raisonnement par l'absurde), si ce n'était pas vrai, on pourrait construire une chaîne infinie :

$$(x_1) \subsetneq (x_2) \subsetneq \dots \subsetneq (x_n) \subsetneq \dots$$

où les x_i sont dans $\mathcal{E}$. Alors l'ensemble $U = \bigcup_{n \geq 1} (x_n)$ est un idéal : si a et b sont dans U , il existe un certain n tel que a et b soient tous les deux dans (x_n) , donc $a + b \in (x_n) \subset U$, même raisonnement pour l'autre propriété. U est donc principal, c'est-à-dire qu'il existe $x \in A$ tel que $U = \bigcup_{n \geq 1} (x_n) = (x)$.

Mais alors, x est dans $\bigcup_{n \geq 1} (x_n)$, donc $x \in (x_{n_0})$ pour un n_0 , et donc alors $(x) = (x_{n_0})$ et $\bigcup_{n \geq 1} (x_n) = (x_{n_0})$, ce qui est absurde au vu des hypothèses.

Soit donc (x_0) maximal dans $\mathfrak{F}$. Alors x_0 n'est pas inversible, ni irréductible (sinon il se décomposerait en produit d'irréductibles) dont il peut s'écrire $x_0 = ab$ et donc $(x_0) \subset (a)$, $(x_0) \subset (b)$. Les inclusions sont strictes puisque ni a ni b ne sont associés à x_0 . Par maximalité, a et b ne sont pas dans $\mathcal{E}$, donc se décomposent en irréductibles, mais alors x_0 aussi se décompose en irréductibles, il y a contradiction.

Il faut maintenant montrer l'unicité. Commençons par remarquer que si p est irréductible et si u est une unité, alors pu est irréductible. Supposons maintenant que :

$$p_1 p_2 \dots p_s = q_1 q_2 \dots q_r$$

Notre objectif est de montrer que $r = s$ et que les p_i sont égaux, à l'ordre et à des unités près aux q_j . Pour cela nous allons utiliser le théorème qui dit que dans un anneau principal, tout irréductible est premier. On peut alors en effet dire que q_1 divisant le produit $p_1 \dots p_s$, divise un des p_i , et, par irréductibilité des p_i , on peut écrire que $q_1 = p_i u$ où u est une unité. On peut alors simplifier et reprendre le même raisonnement. À la fin, on aura $r = s$ car un irréductible n'est pas un produit d'unités. $\square$

Nous sommes maintenant au sommet de la hiérarchie : les anneaux pour lesquels le théorème fondamental de l'arithmétique est vrai sont les anneaux factoriels. Dans le chapitre suivant, on rencontrera des anneaux qui sont factoriels sans être principaux ; ainsi les trois notions anneaux euclidiens, anneaux principaux, anneaux factoriels sont bien distinctes.

Venons-en maintenant aux propriétés arithmétiques des anneaux principaux et factoriels.

2.3.4. P.g.c.d, p.p.c.m., relation de Bezout

Définition 2.43. Si A est un anneau commutatif intègre, on dit que a et b admettent un p.g.c.d. d si d est un diviseur commun de a et de b et si

$$\forall \delta \in A, \quad (\delta \mid a \text{ et } \delta \mid b) \quad \Rightarrow \quad (\delta \mid d)$$

On dit que a et b admettent un p.p.c.m. m si m est un multiple commun de a et de b et si

$$\forall \mu \in A, \quad (b \mid \mu \text{ et } a \mid \mu) \quad \Rightarrow \quad (m \mid \mu)$$

Remarque :

- Il n'y a pas en général unicité du p.g.c.d. et du p.p.c.m. : si d est un p.g.c.d. de a et b , les autres sont les associés de d , idem pour le p.p.c.m. On note néanmoins $d = a \wedge b$, et $m = a \vee b$, en ayant choisi un représentant ; dans le cas de $\mathbb{Z}$ par exemple on choisit le représentant positif, dans le cas de $K[X]$ on peut choisir le représentant unitaire.
- Le p.g.c.d. et le p.p.c.m. n'existent pas toujours, mais si le p.p.c.m. existe, alors le p.g.c.d. aussi : voir exercice 2.10.
- On définit de la même façon le p.g.c.d. et le p.p.c.m. de plusieurs éléments.

L'existence des p.g.c.d. et des p.p.c.m. est assurée dans les anneaux principaux et factoriels. On pourrait se contenter d'étudier le cas factoriel, mais l'étude du cas principal est intéressante pour elle-même.

Proposition 2.44.

- a et b admettent m comme p.p.c.m. si et seulement si $(a) \cap (b) = (m)$.
- S'il existe d tel que $(a) + (b) = (d)$, alors a et b admettent d comme p.g.c.d.
- En particulier, si A est un anneau principal, il existe toujours p.p.c.m. et p.g.c.d.

Démonstration. Remarquer que les énoncés concernant p.p.c.m. et p.g.c.d. diffèrent. Pour le p.p.c.m, l'ensemble $(a) \cap (b)$ est exactement l'ensemble des multiples communs de a et de b . S'il est de la forme (m) , m est un multiple commun et tout multiple commun est multiple de m , qui est donc bien le p.p.c.m. Réciproquement, l'existence du p.p.c.m. dit exactement que l'ensemble des multiples communs de a et de b est de la forme (m) .

Supposons que $(a) + (b) = (d)$. Alors $a \in (a) + (b) = (d)$, donc a est un multiple de d , de même b est un multiple de d . De plus, $d \in (a) + (b)$ donc il existe deux éléments u et v de A tels que $d = au + bv$ et si δ divise a et divise b , alors δ divise d . Il n'y a pas équivalence : dans l'anneau $\mathbb{Z}[X]$, 2 et X ont pour p.g.c.d. 1 mais $(2) + (X)$ n'est pas égal à $(1) = \mathbb{Z}[X]$, puisqu'il est formé des polynômes dont le terme constant est pair. Cela montre en particulier que l'anneau $\mathbb{Z}[X]$ n'est pas principal. Nous verrons dans le chapitre suivant qu'il est néanmoins factoriel.

Dans un anneau principal, $(a) \cap (b)$ et $(a) + (b)$ sont des idéaux, donc sont principaux et l'existence des p.p.c.m. et p.g.c.d. est assurée. $\square$

Plaçons nous dans un anneau principal. L'égalité $au + bv = d$ de la démonstration précédente s'appelle **relation de Bezout**.

Proposition 2.45. Propriété de Bezout. *Si A est principal, alors*

$$(a \wedge b = d) \quad \Rightarrow \quad (\exists(u, v) \in A^2, \quad au + bv = d)$$

De plus,

$$(a \wedge b = 1) \quad \Longleftrightarrow \quad (\exists(u, v) \in A^2, \quad au + bv = 1)$$

Démonstration. Compte-tenu de la démonstration précédente, il reste à démontrer que, s'il existe u et v tels que $au + bv = 1$, alors $a \wedge b = 1$. Si en effet d est un diviseur commun de a et de b , on a $a = a'd$, $b = b'd$ et $1 = d(a'u + b'v)$, d est donc inversible. Les seuls diviseurs communs sont les inversibles, dont 1 est un représentant. $\square$

Comme dans le cas des entiers, si $a \wedge b = 1$ on dit que a et b sont **premiers entre eux**. On observera que cela équivaut à dire que les idéaux (a) et (b) sont étrangers, voir 2.2.3..

Corollaire 2.46. Lemme de Gauss. *Soit A un anneau principal. Alors*

$$\begin{cases} a \mid bc \\ a \wedge b = 1 \end{cases} \quad \Rightarrow \quad a \mid c$$

Démonstration. On utilise une relation de Bezout $au + bv = 1$ et en multipliant par c , $acu + bcv = c$, a divise le premier membre donc a divise c . $\square$

Regardons maintenant ce qu'il se passe si l'anneau A est factoriel. On a encore :

Proposition 2.47.

- Si A est factoriel, le p.g.c.d. et le p.p.c.m. existent toujours.
- Si A est factoriel, le lemme de Gauss est satisfait.

Démonstration. Commençons par utiliser un vocabulaire très pratique. Dans un anneau factoriel A , si p est un irréductible, on appelle **valuation** associée à p l'application v_p de $A \setminus \{0\}$ dans $\mathbb{N}$ qui à tout $x \in A$ associe l'exposant de p dans la décomposition de x en irréductibles. Notons $\mathcal{P}$ l'ensemble des irréductibles, la factorialité de A permet d'affirmer :

$$\forall(x, y) \in A^2, \quad (x \mid y) \quad \Longleftrightarrow \quad (\forall p \in \mathcal{P}, \quad v_p(x) \leq v_p(y))$$

On peut alors démontrer le théorème.

- On utilise la décomposition en irréductibles de a , b et

$$a = u \prod_{p \in \mathcal{P}} p^{v_p(a)}, \quad b = u' \prod_{p \in \mathcal{P}} p^{v_p(b)}$$

où u et u' sont des unités. P est la réunion des irréductibles intervenant dans la décomposition de a et b . Alors, par unicité de la décomposition en irréductibles, si δ est un diviseur de a , il s'écrit $\delta = u'' \prod_{p \in P} p^{v_p(\delta)}$ où $v_p(\delta) \leq v_p(a)$. Si δ est un diviseur commun, on aura donc $v_p(\delta) \leq \inf(v_p(a), v_p(b))$, et le cas où $v_p(\delta) = \inf(v_p(a), v_p(b))$ pour tout $p \in P$, donne un diviseur commun qui a les propriétés du p.g.c.d. De même, le choix $\sup(v_p(a), v_p(b))$ donne le p.p.c.m.

- Prenons les mêmes conventions de notation. On suppose que $a \mid bc$ et $a \wedge b = 1$. Il faut montrer que $a \mid c$. Comme $a \mid bc$, on a $v_p(a) \leq v_p(b) + v_p(c)$ pour tout p . Si p divise a , $v_p(a) > 0$, mais comme $a \wedge b = 1$, on a alors $v_p(b) = 0$ et l'inégalité précédente implique $v_p(a) \leq v_p(c)$. On a bien $v_p(a) \leq v_p(c)$ pour tous les diviseurs irréductibles de a et a divise c .

□

Dans le cas d'un anneau principal, on a donc deux démonstrations différentes, l'une utilisant les propriétés des idéaux principaux, l'autre utilisant la décomposition en irréductibles.

Il existe une circonstance où l'on peut comparer ces deux approches, c'est le cas où l'anneau est non seulement principal mais euclidien : on sait qu'on peut alors déterminer le p.g.c.d. (et les coefficients d'une relation de Bezout) grâce à l'algorithme d'Euclide, extrêmement efficace (le nombre d'étapes dans l'algorithme d'Euclide pour le couple (a, b) , $a > b$, est de l'ordre de $\ln(b)$). Alors que la décomposition en irréductibles peut être un problème très compliqué, et souvent d'un coût en calcul élevé, même dans le cas de $\mathbb{Z}$.

Exercice 2.9. Soit $\mathbb{Z}[i\sqrt{3}]$ le sous-anneau de $\mathbb{C}$ engendré par $i\sqrt{3}$.

(1) Montrer que

$$\mathbb{Z}[i\sqrt{3}] = \{z \in \mathbb{C} \mid \exists a \in \mathbb{Z}, \exists b \in \mathbb{Z}, z = a + ib\sqrt{3}\}$$

et généraliser aux sous-anneaux engendrés par $\sqrt{d}$ où d est un entier (positif ou négatif) qui n'est pas un carré.

- (2) Montrer que l'application N définie par $N : z = a + ib\sqrt{3} \mapsto |z|^2$ est un morphisme pour le produit, à valeurs dans $\mathbb{N}$. En déduire les unités de $\mathbb{Z}[i\sqrt{3}]$.
- (3) Montrer que 4 admet dans cet anneau deux décompositions distinctes en irréductibles. Vérifier que $1 + i\sqrt{3}$ est irréductible mais pas premier.

Exercice 2.10. Soit A un anneau commutatif intègre. On suppose que a et b non nuls ont un p.p.c.m. m . Montrer qu'ils ont un p.g.c.d. d et que $md = ab$ (à une unité près).

Exercice 2.11. Soit A un anneau principal et $a_1, a_2, \dots, a_n$ des éléments premiers entre eux **deux à deux**. Montrer le théorème chinois (version anneau principal) :

$$A/(a_1 a_2 \dots a_n) \equiv A/(a_1) \times A/(a_2) \times \dots \times A/(a_n)$$

On sera conduit à démontrer que l'hypothèse implique

$$\bigcap_{i=1}^k (a_i) = \left(\prod_{i=1}^k a_i \right)$$

EXERCICES

Exercice 2.12. Soit A l'anneau des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$, muni de l'addition et du produit des fonctions. Déterminer quel est l'ensemble des éléments inversibles, donner un exemple de diviseur de 0 et de fonction qui n'est ni inversible, ni diviseur de zéro.

Exercice 2.13. Soit A un anneau non nul tel que $x^2 = x$ pour tout x de A . On dit que A est un anneau de Boole.

- (1) Montrer que pour tout x de A , on a : $x + x = 0$ (A est de caractéristique 2), puis que A est commutatif.
 - (2) Montrer que A ne peut avoir exactement 3 éléments. Quel peut-être le nombre d'éléments d'un anneau de Boole fini ?
 - (3) Trouver des exemples d'anneaux de Boole.
-

Exercice 2.14. A est un anneau commutatif intègre, S est un sous-ensemble stable pour le produit et ne contenant pas 0 mais contenant 1. On dit alors que S est **multipliatif**, et on définit une relation dans $A \times S$ par

$$(a, s)\mathcal{R}(a', s') \iff as' - a's = 0$$

- (1) Montrer que c'est une relation d'équivalence.
- (2) On note $S^{-1}A$ l'ensemble quotient et $\frac{a}{s}$ la classe de (a, s) . Montrer que $S^{-1}A$ est un anneau pour les lois $+$ et $\times$ définies par :

$$\frac{a}{s} + \frac{b}{t} = \frac{at + bs}{st} \quad \text{et} \quad \frac{a}{s} \times \frac{b}{t} = \frac{ab}{st}$$

(ne pas oublier de vérifier que ces opérations sont bien définies).

- (3) Vérifier que $S = A \setminus \{0\}$ est une partie multiplicative, et qu'alors $S^{-1}A$ est un corps contenant un sous-anneau isomorphe canoniquement à A . Reconnaitre le cas $A = \mathbb{Z}$ et $A = K[X]$. On dit alors que $S^{-1}A$ est le **corps des fractions** de A .
- (4) Montrer que si $\mathcal{I}$ est un idéal premier de A , $S = A \setminus \mathcal{I}$ est une partie multiplicative. Décrire le cas où $A = \mathbb{Z}$ et où $\mathcal{I} = (2)$.

Exercice 2.15. On rappelle que les inversibles de $\mathbb{Z}/n\mathbb{Z}$ sont les classes $\bar{k}$ où k est premier à n . Le nombre des entiers k tels que $1 \leq k \leq n-1$ et $k \wedge n = 1$ est noté $\phi(n)$, c'est l'indicateur d'Euler. Rechercher les inversibles de $\mathbb{Z}/p^k\mathbb{Z}$. En utilisant le théorème chinois (cf. l'exercice 2.11), dénombrer les inversibles de $\mathbb{Z}/n\mathbb{Z}$, et donner une formule explicite pour $\phi(n)$.

Exercice 2.16. Soit A l'anneau $\mathbb{Z}[i\sqrt{5}]$. (cf. l'exercice 2.9). Si $z \in A$, on rappelle que l'application norme $N : z \mapsto |z|^2$ est un morphisme multiplicatif de A dans $\mathbb{N}$.

- (1) Chercher les éléments inversibles (unités).
- (2) Déterminer les diviseurs de 9. Préciser lesquels sont irréductibles.
- (3) Trouver des éléments irréductibles qui ne sont pas premiers. En déduire que A n'est ni principal ni factoriel.
- (4) Trouver deux éléments qui admettent un p.g.c.d. mais qui n'ont pas de p.p.c.m.
- (5) Trouver deux éléments qui n'ont ni p.g.c.d., ni p.p.c.m.

Exercice 2.17. Soit $\beta = \frac{1+i\sqrt{19}}{2}$ et $A = \mathbb{Z}[\beta]$, le sous-ensemble de $\mathbb{C}$ formé des combinaisons linéaires à coefficients entiers de 1 et de β .

- (1) Montrer que

$$\beta^2 - \beta + 5 = 0$$

et que A est un sous-anneau de $\mathbb{C}$ stable par conjugaison.

- (2) Chercher les unités (éléments inversibles) de A .
- (3) On veut montrer que A n'est pas euclidien. Pour cela, on suppose qu'il existe une division euclidienne associée à une application (stathme) ϕ . Montrer qu'il existe z_0 différent de 0 et de ± 1 tel que $\phi(z_0)$ soit minimum.
- (4) Démontrer que $A/(z_0)$ est de cardinal 2 ou 3.
- (5) Montrer que l'équation

$$x^2 - x + 5 = 0$$

n'a pas de racine dans $\mathbb{F}_2$ ou $\mathbb{F}_3$.

En déduire que A n'est pas euclidien.

Remarque : on peut démontrer que A est cependant principal. Voir par exemple [21].

PROBLÈMES

Les corrigés de ces problèmes sont disponibles sur le site de Dunod : www.dunod.com.

2.1. LES ENTIERS DE GAUSS

On note $\mathbf{G}$ l'ensemble $\mathbb{Z}[i] = \{z \in \mathbb{C} \mid \exists a \in \mathbb{Z}, \exists b \in \mathbb{Z}, z = a + bi\}$.

- (1) Vérifier que $\mathbf{G}$ est un anneau commutatif isomorphe au quotient $\mathbb{Z}[X]/(X^2 + 1)$.
- (2) Montrer que $z = a + bi \mapsto N(z) = |z|^2 = a^2 + b^2$ est un morphisme pour le produit de $\mathbf{G}$ dans $\mathbb{N}$. En déduire les unités (éléments inversibles) de $\mathbf{G}$. On pourra se référer à l'exercice 2.9.
- (3) Montrer que $\mathbf{G}$ est un anneau euclidien pour le stathme N . On pourra montrer que si a et $b \neq 0$ sont dans $\mathbf{G}$, il existe q dans $\mathbf{G}$ tel que :

$$\left| \frac{a}{b} - q \right| < 1$$

Puis on posera $r = a - bq$. On pourra s'aider d'une représentation graphique. Y a-t-il unicité du quotient et du reste ?

- (4) Soit $a \in \mathbf{G}^*$. Montrer que si $N(a)$ est un nombre premier, alors a est un irréductible de $\mathbf{G}$.
- (5) Soit $z = \alpha + i\beta$ un irréductible de $\mathbf{G}$, où ni α ni β ne sont nuls. Montrer que $p = \alpha^2 + \beta^2$ est premier dans $\mathbb{Z}$. On pourra utiliser le fait que $p = z\bar{z}$ et vérifier que $\bar{z}$ est aussi irréductible dans $\mathbf{G}$.
- (6) Montrer que les irréductibles de $\mathbf{G}$ sont :
 - Les $\alpha + i\beta$ tels que $\alpha^2 + \beta^2 = p$ soit un nombre premier de $\mathbb{Z}$.
 - Les premiers de $\mathbb{Z}$ qui ne peuvent s'écrire comme somme de deux carrés (il en existe : 3, 7, ...), et leurs associés.
- (7) Montrer l'équivalence des trois énoncés, pour p premier différent de 2 :
 - $p \equiv 1 \pmod{4}$
 - p non premier dans $\mathbf{G}$.
 - -1 carré dans $\mathbb{Z}/p\mathbb{Z}$.
 - p est somme de deux carrés dans $\mathbb{Z}$.
- (8) Déterminer parmi les éléments de $\mathbf{G}$ de partie réelle et de partie imaginaire dans $[0, 3]$, lesquels sont irréductibles. Donner une décomposition des réductibles.

2.2. LES SOUS-ANNEAUX DE $\mathbb{Q}$

Ce problème a pour objectif de décrire tous les sous-anneaux de $\mathbb{Q}$. On note A un tel sous-anneau (non nul), et D l'ensemble des dénominateurs des éléments de A dans leur écriture irréductible. C'est un sous-ensemble de $\mathbb{N}^*$.

- (1) Quel est l'ensemble D lorsque $A = \mathbb{D}$, anneau des nombres décimaux ?
- (2) Montrer que $d \in D$ si et seulement si $\frac{1}{d} \in A$.
- (3) Montrer que D est stable pour le produit (propriété 1), puis montrer que si d est un élément de D , alors tous les diviseurs (dans $\mathbb{N}$) de d sont dans D (propriété 2).
- (4) On suppose que D est une partie de $\mathbb{N}^*$ qui a les propriétés 1 et 2. Vérifier que

$$A = \left\{ x \in \mathbb{Q} \mid \exists p \in \mathbb{Z}, q \in D, \quad p \wedge q = 1 \text{ et } x = \frac{p}{q} \right\}$$

est bien un sous-anneau de $\mathbb{Q}$.

- (5) Soit $\mathcal{P}$ l'ensemble des nombres premiers. Si D est associé à un anneau A , on note $\mathcal{Q}$ l'ensemble des diviseurs premiers des éléments de D . Montrer que $d \in D$ si, et seulement si ses facteurs premiers sont dans $\mathcal{Q}$.
- (6) On note $\mathcal{Q}$ un sous-ensemble non vide de $\mathcal{P}$. Montrer que l'ensemble des entiers dont les facteurs premiers sont dans $\mathcal{Q}$ vérifie les propriétés 1 et 2. Donner la description des sous-anneaux de $\mathbb{Q}$.
- (7) Reconnaître les anneaux associés à $\mathcal{Q} = \{2, 5\}$ puis $\mathcal{Q} = \mathcal{P} \setminus \{2\}$.

SOLUTIONS DES EXERCICES

Solution 2.1. $x \mapsto ax$ est injective équivaut à

$$\forall (x, y) \in A^2, \quad (ax = ay) \quad \Rightarrow \quad (x = y)$$

c'est donc exactement la régularité à gauche de a . Si de plus A est fini, on dispose d'une application injective d'un ensemble fini dans lui-même : elle est forcément bijective. Donc 1_A admet un antécédent unique et a admet un inverse à droite. Un élément régulier est donc inversible, et comme on sait que tout inversible est régulier les deux notions coïncident dans le cas des anneaux finis. Application : un anneau intègre fini est un corps.

Solution 2.2. Si M est un diviseur de zéro, il existe N non nul tel que $MN = 0$. Supposons que M représente un endomorphisme f de K^n , alors N représente g telle que $f \circ g = 0$. Cette égalité équivaut à $\text{Im } g \subset \text{Ker } f$. Comme $g \neq 0$ équivaut à $\text{Im } g \neq 0$, on voit qu'il est nécessaire que $\text{Ker } f \neq 0$. Réciproquement, si cette condition est réalisée, il est facile de construire un endomorphisme non nul dont l'image est incluse dans $\text{Ker } f$, par exemple la projection sur $\text{Ker } f$ suivant un supplémentaire de $\text{Ker } f$. Les diviseurs de zéro à gauche sont donc les matrices de déterminant nul. Cherchons les diviseurs de zéro à droite : le même raisonnement montre qu'il existe f non nul tel que $f \circ g = 0$ seulement si $\text{Im } g$ est un sous-espace vectoriel strict de E , donc si g n'est pas bijective, les diviseurs de zéro à droite sont encore les matrices non inversibles. Dans l'anneau des matrices sur un corps, les notions d'éléments réguliers et inversibles coïncident.

Solution 2.3. Soit $\mathcal{I} = \{x \in A \mid ax = 0\}$. Alors, si x et y sont dans $\mathcal{I}$, $a(x+y) = ax + ay = 0$ et si $z \in A$, $a(xz) = (ax)z = 0$ donc $\mathcal{I}$ est un idéal à droite de A . Bien sûr, il est non trivial (ni nul, ni A) lorsque a est diviseur de zéro à gauche. Si $\bar{k} \in \mathbb{Z}/n\mathbb{Z}$, alors $\bar{k}\bar{x} = \bar{0}$ équivaut à $kx \in n\mathbb{Z}$. Si on introduit le p.g.c.d. de k et de n , cette condition équivaut à x multiple de $\frac{n}{n \wedge k}$. L'annulateur de $\bar{k}$ est l'idéal engendré par la classe de $\frac{n}{n \wedge k}$.

Dans le cas de l'annulateur à droite d'une matrice, il suffit d'utiliser l'exercice précédent : l'annulateur à droite d'une matrice de noyau F et d'image G est formé des matrices dont l'image est incluse dans F , l'annulateur à gauche est formé des matrices dont le noyau contient G . On peut montrer que ces annulateurs sont les seuls types d'idéaux de $\mathcal{M}_n(K)$.

Solution 2.4. Soit $\mathcal{I}$ un idéal bilatère de $\mathcal{M}_n(K)$ et M une matrice de $\mathcal{I}$. Supposons que M ne soit pas nulle, alors un de ses coefficients, par exemple m_{11} n'est pas nul. Comme $\mathcal{I}$ est un idéal bilatère, la matrice $E_{i1}ME_{1j} = m_{11}E_{ij}$ est dans $\mathcal{I}$. Si on multiplie à gauche par $\frac{1}{m_{11}}I_n$, on voit que E_{ij} est dans $\mathcal{I}$ pour tout couple (i, j) , $\mathcal{I}$ est donc l'anneau tout entier.

Solution 2.5. Il suffit d'utiliser le théorème de correspondance. Les idéaux de $\mathbb{Z}/n\mathbb{Z}$ sont en bijection avec les idéaux $d\mathbb{Z}$ de $\mathbb{Z}$ tels que $n\mathbb{Z} \subset d\mathbb{Z} \subset \mathbb{Z}$ et d doit donc être un diviseur de n , l'idéal est engendré par la classe de d modulo n . L'ensemble des idéaux de $\mathbb{Z}/n\mathbb{Z}$ est donc en bijection avec l'ensemble des diviseurs positifs de n , l'idéal nul étant engendré par $\bar{n} = \bar{0}$, l'anneau tout entier est engendré par $\bar{1}$.

Solution 2.6. Soit $\bar{k} \in \mathbb{Z}/n\mathbb{Z}$ un élément nilpotent. Si on suppose $\bar{k}^r = \bar{0}$, cela équivaut à

$$\exists \ell \in \mathbb{N}, \quad k^r = \ell n$$

Tout diviseur premier de n est un diviseur de k^r donc de k , et donc k est multiple de $p_1 p_2 \dots p_s$ où $p_1, p_2, \dots, p_s$ sont les diviseurs premiers de n . Réciproquement, si k est multiple de ce produit, il existe une puissance de k qui sera multiple de n , il suffit de choisir le plus grand des exposants des p_i dans la décomposition de n . En revenant dans $\mathbb{Z}/n\mathbb{Z}$, on constate qu'il n'y aura pas de nilpotent non nul lorsque n est égal à $p_1 p_2 \dots p_s$. On dit alors que n est sans facteur carré.

Solution 2.7. Si x et y sont dans $\sqrt{\mathcal{I}}$, on aura $x^k \in \mathcal{I}$ et $y^\ell \in \mathcal{I}$ pour un certain k et un certain ℓ . Donc

$$(x+y)^n = \sum_{i=0}^n \binom{n}{i} x^{n-i} y^i$$

sera dans $\mathcal{I}$ dès que $n \geq k + \ell$ car alors $n - i \geq k$ ou $i \geq \ell$ pour tout i entre 0 et n . Par ailleurs, si a est dans A , $(ax)^k = a^k x^k$ est dans l'idéal $\mathcal{I}$. On a bien démontré que $\sqrt{\mathcal{I}}$ est un idéal. Si $\mathcal{I}$ est l'idéal nul, cet idéal est formé des éléments **nilpotents** de A . Soit maintenant x et y dans $\mathcal{I} : \mathcal{J}$, alors

$$(x + y)\mathcal{J} \subset x\mathcal{J} + y\mathcal{J} \subset \mathcal{I} + \mathcal{I} = \mathcal{I}$$

et donc $\mathcal{I} : \mathcal{J}$ est un idéal. Si $\mathcal{I}$ est l'idéal nul, on retrouve la notion d'annulateur. Donnons un exemple dans l'anneau $\mathbb{Z}$: si $n = p_1^{a_1} \dots p_k^{a_k}$ où les p_i sont des nombres premiers et où les exposants sont non nuls, on a :

$$\sqrt{n\mathbb{Z}} = p_1 p_2 \dots p_k \mathbb{Z}$$

En effet, si x^k est un multiple de n , tout premier qui divise n doit diviser x . Et si un nombre est multiple de $p_1 p_2 \dots p_k$, sa puissance k , pour k assez grand, sera multiple de n . On montrera de même que $a\mathbb{Z} : b\mathbb{Z} = a'\mathbb{Z}$ où $a' = \frac{a}{a \wedge b}$ (examiner à part les cas où a ou b est nul).

Comment définir le produit de deux idéaux ? En règle générale, si $\mathcal{I}$ et $\mathcal{J}$ sont deux idéaux d'un anneau commutatif A , l'ensemble des produits d'un élément de $\mathcal{I}$ par un élément de $\mathcal{J}$ n'est pas un idéal. On définira $\mathcal{I}\mathcal{J}$ comme l'idéal engendré par les produits, c'est donc

$$\mathcal{I}\mathcal{J} = \{x \in A \mid \exists k \in \mathbb{N}, \exists (x_i) \in \mathcal{I}^k, \exists (y_i) \in \mathcal{J}^k, x = \sum_{i=1}^k x_i y_i\}$$

Cas particulier : si $\mathcal{I} = (a)$ et $\mathcal{J} = (b)$ alors il est immédiat que $\mathcal{I}\mathcal{J} = (ab)$.

Supposons maintenant que les idéaux $\mathcal{I}$ et $\mathcal{J}$ sont étrangers, c'est-à-dire que leur somme est A .

- $\mathcal{I}\mathcal{J} \subset \mathcal{I} \cap \mathcal{J}$ puisque si $z_i = a_i b_i$ où $a_i \in \mathcal{I}$ et $b_i \in \mathcal{J}$, alors $z_i \in \mathcal{I} \cap \mathcal{J}$ d'après la seconde propriété des idéaux. Les sommes d'éléments de la forme z_i sont donc dans $\mathcal{I} \cap \mathcal{J}$. On n'a pas utilisé l'hypothèse.
- $\mathcal{I} \cap \mathcal{J} \subset \mathcal{I}\mathcal{J}$. L'hypothèse $A = \mathcal{I} + \mathcal{J}$ implique qu'il existe $i \in \mathcal{I}$ et $j \in \mathcal{J}$ tels que $1 = i + j$. Si $x \in \mathcal{I} \cap \mathcal{J}$, alors $x = xi + xj$ est somme de deux produits formés d'un élément de $\mathcal{I}$ avec un élément de $\mathcal{J}$, on a montré que x est dans $\mathcal{I}\mathcal{J}$.

Dans $\mathbb{Z}$, les idéaux (a) et (b) sont étrangers si et seulement si a et b sont premiers entre eux. De plus, $(a)(b) = (ab)$ et l'égalité $(a)(b) = (ab) = (a) \cap (b)$ traduit alors que le p.p.c.m. de a et de b est leur produit ab .

Solution 2.8. Soit $\mathcal{J}$ un idéal de A , différent de A . On se place dans l'ensemble E des idéaux différents de A qui contiennent $\mathcal{J}$. E est non vide car il contient au moins $\mathcal{J}$. Soit $\mathcal{C}$ un ensemble totalement ordonné (pour l'inclusion) d'éléments de E . Alors montrons que $U = \bigcup_{\mathcal{I} \in \mathcal{C}} \mathcal{I}$ est un majorant de $\mathcal{C}$; c'est un idéal, car si x et y sont dans U , on a $x \in \mathcal{I}_1$ et $y \in \mathcal{I}_2$, mais comme l'ensemble $\mathcal{C}$ est totalement ordonné, on a par exemple, $\mathcal{I}_1 \subset \mathcal{I}_2$ et $x \in \mathcal{I}_2$, donc $x + y \in \mathcal{I}_2 \subset U$. La seconde propriété se vérifie encore plus directement. Il faut montrer que U est bien dans E , donc que c'est un idéal

différent de A . Sinon, en effet, $1_A \in U$, donc il existe $\mathcal{I} \in \mathcal{C}$ tel que $1_A \in \mathcal{I}$. Mais alors $\mathcal{I}$ serait A , ce qui est contraire à l'hypothèse. Enfin, U contient $\mathcal{I}$ qui contient $\mathcal{J}$. En appliquant le théorème de Zorn, on a montré que E avait au moins un élément maximal $\mathcal{K}$. Une subtilité : il faut montrer que $\mathcal{K}$ est bien un idéal maximal de A . Si en effet l'idéal $\mathcal{L}$ vérifie $\mathcal{K} \subset \mathcal{L} \subsetneq A$, il contient $\mathcal{J}$ comme $\mathcal{K}$ et c'est donc un élément de E . Comme $\mathcal{K}$ est maximal dans E , on a $\mathcal{L} = \mathcal{K}$.

En appliquant ce résultat à $\mathcal{J} = \{0\}$, on obtient que tout anneau non nul admet un idéal maximal. Bien entendu, ces démonstrations s'appliquent aux idéaux à gauche, à droite ou bilatères.

Solution 2.9.

- (1) $\mathbb{Z}[i\sqrt{3}]$ doit contenir 1 et $i\sqrt{3}$, les sommes et les opposés de ces éléments donc tous les $a + bi\sqrt{3}$ où a et b sont dans $\mathbb{Z}$. Mais l'ensemble ainsi construit est un anneau. Il est stable pour la somme, et pour le produit car :

$$(a + bi\sqrt{3})(a' + b'i\sqrt{3}) = aa' - 3bb' + (ab' + a'b)i\sqrt{3}$$

c'est donc un sous-anneau de $\mathbb{C}$, le plus petit qui contienne $i\sqrt{3}$. Le même calcul fonctionne dans le cas de $\sqrt{d}$, où d est un entier positif qui n'est pas un carré, il y a unicité de l'écriture sous la forme $a + b\sqrt{d}$ avec a et b entiers, car $\sqrt{d}$ est irrationnel.

- (2) $|a + ib\sqrt{3}|^2 = a^2 + 3b^2 \in \mathbb{N}$. C'est un morphisme pour le produit, car c'est déjà vrai dans $\mathbb{C}$. Si z est inversible d'inverse z' dans notre anneau, on aura $zz' = 1$, donc $N(z)N(z') = N(1) = 1$. Comme $N(z)$ et $N(z')$ sont dans $\mathbb{N}$, ils sont égaux à 1. Et si $N(z) = 1$, avec $z = a + bi\sqrt{3}$, on a $N(z) = (a + bi\sqrt{3})(a - bi\sqrt{3}) = 1$, l'inverse de z est son conjugué, il est dans $\mathbb{Z}[i\sqrt{3}]$. Le problème se réduit à trouver tous les entiers relatifs tels que $a^2 + 3b^2 = 1$, b est nul et $a = \pm 1$. Seuls 1 et -1 sont inversibles dans notre anneau.

- (3) On a $N(1 + i\sqrt{3}) = 4$. Supposons $z = 1 + i\sqrt{3}$ réductible sous la forme $z = xy$; ni x ni y n'étant inversible, on doit avoir $N(x) = N(y) = 2$. Mais $a^2 + 3b^2$ ne peut prendre la valeur 2 lorsque a et b sont entiers. De même, $\bar{z} = 1 - i\sqrt{3}$ et 2 sont irréductibles, et ces irréductibles sont non associés. Donc

$$4 = 2 \times 2 = (1 + i\sqrt{3})(1 - i\sqrt{3})$$

donne deux décompositions distinctes en irréductibles. $N(a)N(b) = 4$. On voit également que z divise 2×2 mais il ne divise pas 2, car $z = 2x$ impliquerait que $N(x) = 1$ soit $x = \pm 1$.

Solution 2.10. Soit a et b dans A . On suppose qu'il existe un p.p.c.m., noté m . Comme ab est évidemment un multiple commun de a et de b , il existe d tel que $ab = md$. Mais, comme a divise m , da divise $dm = ab$. Donc d est un diviseur de b . En échangeant les rôles, d est aussi un diviseur de a .

Soit maintenant δ un diviseur commun de a et de b . Alors $a = a'\delta$ et $b = b'\delta$ et $a'b'\delta$ est un multiple commun de a et de b . C'est donc un multiple de m , et il existe k tel que $a'b'\delta = km$. Mais, en multipliant par d :

$$a'b'\delta d = kmd = kab = ka'b'\delta^2 \quad \Rightarrow \quad d = k\delta$$

tout diviseur commun est un diviseur de d , on a prouvé l'existence du p.g.c.d. de a et b .

Solution 2.11. On va utiliser la même démonstration que le théorème chinois vu dans le cours. Soit ϕ le morphisme de A dans $A/(a_1) \times A/(a_2) \times \dots \times A/(a_k)$ qui à $a \in A$ associe le k -uple de ses classes. C'est un morphisme d'anneau. De plus,

$$a \in \text{Ker } \phi \iff \forall i, a \in (a_i) \iff a \in \bigcap_{i=1}^k (a_i)$$

Montrons maintenant $\bigcap_{i=1}^k (a_i) = \left(\prod_{i=1}^k a_i \right)$ sous l'hypothèse que les (a_i) sont premiers entre eux deux à deux. C'est vrai pour $k = 2$, en utilisant une relation de Bezout :

$$a_1 u_1 + a_2 u_2 = 1 \Rightarrow (a_1 a_2) u_1 + (a_2 a_2) u_2 = a_2$$

et si a est multiple de a_1 et de a_2 , cette égalité montre qu'il est multiple de leur produit. Et bien sûr, tout multiple de $a_1 a_2$ est multiple commun de a_1 et de a_2 . On a donc :

$$a_1 \wedge a_2 = 1 \Rightarrow (a_1) \cap (a_2) = (a_1 a_2)$$

On pourra bien sûr faire le lien avec l'exercice 2.7.

Utilisons pour continuer le résultat :

$$\begin{cases} a_1 \wedge a_3 = 1 \\ a_2 \wedge a_3 = 1 \end{cases} \Rightarrow a_1 a_2 \wedge a_3 = 1$$

qui s'obtient par le produit de deux relations de Bezout $a_1 u + a_3 v = 1$ et $a_2 u' + a_3 v' = 1$ qui s'écrit $a_1 a_2 u u' + a_3 (a_3 v v' + a_2 u' v + a_1 u v') = 1$. On a donc

$$((a_1) \cap (a_2)) \cap (a_3) = (a_1 a_2) \cap (a_3) = (a_1 a_2 a_3)$$

et ainsi de suite.

Pour conclure en appliquant le théorème d'isomorphisme, il faut montrer la surjectivité de ϕ ; cela se fait également par récurrence. Si $(x, y) \in A \times A$, alors, en utilisant une relation de Bezout $a_1 u + a_2 v = 1$, on voit que $a = a_1 u y + a_2 v x$ a pour image $(x + (a_1), y + (a_2))$, et on traite le cas général de proche en proche.

Solution 2.12. Une fonction f est inversible s'il existe une fonction g telle que $\forall x \in \mathbb{R}, f(x)g(x) = 1$. Il est donc nécessaire que f ne s'annule jamais, et on sait alors que la fonction $g = \frac{1}{f}$ est continue sur $\mathbb{R}$ comme f . Une fonction f sera un diviseur de zéro s'il existe une fonction g continue telle que $\forall x \in \mathbb{R}, f(x)g(x) = 0$. Si, par exemple, f est nulle sur l'intervalle $[0, 1]$, on peut prendre g nulle sur $] -\infty, 0]$

et $[1, +\infty[$ et égale à $x(1 - x)$ sur $[0, 1]$. Si par contre f est nulle seulement en un point isolé x_0 , ce n'est ni un élément inversible, ni un diviseur de zéro. Soit en effet g continue telle que $fg = 0$. Alors, on doit avoir $g(x) = 0$ pour tout $x \neq x_0$, et, par continuité,

$$g(x_0) = \lim_{x \rightarrow x_0} g(x) = 0$$

et g est la fonction constante nulle.

Solution 2.13.

(1) Pour tout x de A ,

$$(x + 1)^2 = x + 1 \iff x^2 + x + x + 1 = x + 1 \quad \text{d'où} \quad x + x = 0,$$

et A est de caractéristique 2. Si x et y sont quelconques dans A ,

$$(x + y)^2 = x + y \iff x + xy + yx + y = x + y \quad \text{et donc} \quad xy = -yx$$

comme A est de caractéristique 2, $-yx = yx$, l'anneau est commutatif.

(2) Supposons que A ait trois éléments distincts, 0, 1 et x . Alors $1 + x$ est distinct de ces trois éléments car $1 + x \neq 0$ puisque $x \neq 1$, $1 + x \neq 1$ puisque $x \neq 0$ et $1 + x \neq x$ puisque $1 \neq 0$. Remarquons au passage que $x(1 + x) = x + x = 0$, un anneau de Boole qui a plus de deux éléments ne peut être intègre.

Prolongeons encore : tout anneau de Boole A peut être muni d'une structure d'espace vectoriel sur $\mathbb{F}_2$, en posant $\bar{0}.x = 0$ et $\bar{1}.x = x$, les vérifications sont faciles, et utilisent que A est de caractéristique 2. Si donc A est fini, il est isomorphe (en tant qu'espace vectoriel) à $\mathbb{F}_2^n$ pour un certain n , et le cardinal de A est de la forme 2^n .

(3) $\mathbb{F}_2^n$ donne des exemples d'anneaux de Boole à 2^n éléments, de même que $\mathcal{P}(E)$, muni de Δ et $\cap$: voir l'exercice 1.13, p. 14, qui montre que ces deux exemples ne sont pas très différents...

Solution 2.14.

(1) $(a, s)\mathcal{R}(a, s) \iff as - as = 0$, la relation est réflexive. Si $(a, s)\mathcal{R}(a', s')$, alors $a's - as' = -(as' - as) = 0$ donc $(a', s')\mathcal{R}(a, s)$, la relation est symétrique, enfin

$$\left. \begin{array}{l} (a, s)\mathcal{R}(a', s') \iff as' - a's = 0 \\ (a', s')\mathcal{R}(a'', s'') \iff a's'' - a''s' = 0 \end{array} \right\} \Rightarrow as's'' = a'ss'' = a''ss'$$

d'où $as'' = a''s$ et $(a, s)\mathcal{R}(a'', s'')$, la relation est transitive. On a utilisé l'intégrité de l'anneau.

(2) Il faut vérifier que les opérations sont compatibles avec les lois. Faisons-le pour la somme : si $(a, s)\mathcal{R}(a', s')$ et $(b, t)\mathcal{R}(b', t')$ alors $(at + bs, st)\mathcal{R}(a't' + b's', s't')$ puisque

$$(at + bs)s't' - (a't' + b's')st = (as' - a's)tt' + (bt' - b't)ss' = 0$$

Le reste des vérifications est immédiat, avec le modèle des fractions en tête. Par exemple, l'élément neutre pour le produit est la classe $\frac{s}{s} = 1$ où s est un élément quelconque de S , l'élément neutre pour l'addition est la classe $\frac{0}{s}$.

- (3) Comme A est intègre, si s et s' sont non nuls, alors ss' aussi, $S' = A \setminus \{0\}$ est multiplicatif. Si a n'est pas nul, $\frac{a}{s} \times \frac{s}{a} = 1$, donc $S^{-1}A$ est un corps. Enfin, $a \mapsto \frac{a}{1}$ est un isomorphisme de A sur un sous-anneau de $S^{-1}A$. Le corps des fractions de $\mathbb{Z}$ est bien sûr $\mathbb{Q}$, le corps des fractions de $K[X]$ est le corps des fractions rationnelles à coefficients dans K , noté $K(X)$.
- (4) Soit $S = A \setminus \mathcal{I}$ où $\mathcal{I}$ est un idéal premier. S ne contient pas 0 puisque $0 \in \mathcal{I}$ et par définition,

$$\forall (a, b) \in A^2, \quad (ab \in \mathcal{I}) \Rightarrow (a \in \mathcal{I} \text{ ou } b \in \mathcal{I})$$

et par contraposition,

$$\forall (a, b) \in A^2, \quad (a \in S \text{ et } b \in S) \Rightarrow (ab \in S)$$

S est donc multiplicatif. On a donc un anneau $S^{-1}A$. Dans le cas où $A = \mathbb{Z}$ et $\mathcal{I} = 2\mathbb{Z}$, cet anneau s'identifie à l'ensemble des rationnels qui s'écrivent $\frac{a}{b}$ où b est impair. C'est donc un sous-anneau de $\mathbb{Q}$; on vérifie en particulier que l'ensemble des éléments non inversibles est l'idéal engendré par 2.

Solution 2.15. Dans l'anneau $\mathbb{Z}/p^k\mathbb{Z}$, $\bar{x}$ est inversible si et seulement si x est premier à p^k , donc x non multiple de p . Pour compter les inversibles, comptons les ℓp tels que $0 \leq \ell p < p^k$. Il vient $0 \leq \ell < p^{k-1}$. On a donc

$$\phi(p^k) = p^k - p^{k-1}$$

Dans le cas général, commençons par observer qu'on a

$$(A \times B)^\times = A^\times \times B^\times$$

car, pour la loi produit, (a, b) est inversible si et seulement si a et b sont inversibles. Ce résultat s'étend à un produit quelconque. Si maintenant $n = p_1^{\alpha_1} p_2^{\alpha_2} \dots p_k^{\alpha_k}$, où les p_i sont des nombres premiers. Le théorème chinois permet d'obtenir l'isomorphisme :

$$\mathbb{Z}/n\mathbb{Z} \simeq \mathbb{Z}/p_1^{\alpha_1}\mathbb{Z} \times \dots \times \mathbb{Z}/p_k^{\alpha_k}\mathbb{Z}$$

et donc

$$\phi(n) = \prod_{i=1}^k (p_i^{\alpha_i} - p_i^{\alpha_i-1}) = \prod_{i=1}^k p_i^{\alpha_i} \left(1 - \frac{1}{p_i}\right) = n \prod_{i=1}^k \left(1 - \frac{1}{p_i}\right)$$

On retrouve un résultat obtenu à la fin du problème du premier chapitre.

Solution 2.16.

- (1) Soit $(a, b) \in \mathbb{Z}^2$. Si $z = a + ib\sqrt{5}$ est inversible dans A , alors

$$N(z)N(z^{-1}) = N(1) = 1$$

Comme les normes des éléments de A sont des entiers positifs, il est nécessaire que $N(z) = a^2 + 5b^2 = 1$. Les seules solutions en entiers de cette équation sont $a = 1, b = 0$ et $a = -1, b = 0$. On a donc $A^\times = \{-1, 1\}$.

- (2) Si $z = a + bi\sqrt{5}$ est un diviseur de 9, $N(z)$ est un diviseur de $N(9) = 81$ dans $\mathbb{N}$. Si $N(z) = 1$, on a les solutions 1 et -1 . $N(z) = 3$ est impossible, et $N(z) = 9$ donne $a = \pm 2$ et $b = \pm 1$, ou bien $a = \pm 3$ et $b = 0$. $N(z) = 27$ ne peut se produire, car le second facteur z' vérifierait $N(z') = 3$. Enfin les diviseurs tels que $N(z) = 81$ sont 9 et -9 puisque le second facteur est de norme 1. En définitive, on a

$$9 = 3 \times 3 = (2 + i\sqrt{5})(2 - i\sqrt{5})$$

à une unité près. Les diviseurs ± 3 et $\pm 2 \pm i\sqrt{5}$ sont irréductibles : leur norme est 9 et il n'y a pas d'élément de norme 3.

- (3) C'est la question précédente qui nous guide : 3 est irréductible, mais 3 divise $(2 + i\sqrt{5})(2 - i\sqrt{5}) = 9$ sans diviser un des facteurs. En effet, si 3 divisait $2 + i\sqrt{5}$ par exemple, le quotient serait de norme 1, donc serait ± 1 , qui ne conviennent pas. Le nombre 3 est irréductible sans être premier, et ce seul fait montre que A n'est ni principal, ni factoriel.

Remarquons également que dans la question précédente on a obtenu deux factorisations différentes de 9 en produit d'irréductibles, ce qui confirme que A n'est pas factoriel.

- (4) 3 et $2 + i\sqrt{5}$ sont premiers entre eux, leurs seuls diviseurs communs sont les unités. Ils ont comme multiple commun 9, mais aussi $3(2 + i\sqrt{5}) = 6 + 3i\sqrt{5}$. Or 9 et $6 + 3i\sqrt{5}$ ont même norme, 81. Ils ne sont pas multiples d'un même élément de A , comme le montre la question précédente.
- (5) Même situation : 9 et $6 + 3i\sqrt{5}$ ont des diviseurs communs, 1, 3, $6 + 3i\sqrt{5}$ et leurs associés. Mais, parmi ces diviseurs, il n'y en a pas de « plus grand » (au sens de la relation divise). Ces deux éléments n'ont pas de p.g.c.d. Ils ne peuvent avoir de p.p.c.m., car un exercice précédent montre que des éléments qui ont un p.p.c.m. ont nécessairement un p.g.c.d.

Solution 2.17.

- (1) On peut résoudre l'équation $x^2 - x + 5 = 0$ dans $\mathbb{C}$. Les solutions sont

$$x_1 = \frac{1 - i\sqrt{19}}{2}, \quad x_2 = \frac{1 + i\sqrt{19}}{2} = \beta$$

Montrons que $A = \mathbb{Z}[\beta]$ est stable pour le produit. Soit $z = a + b\beta$ et $z' = a' + b'\beta$, alors

$$\begin{aligned} zz' &= (a + b\beta)(a' + b'\beta) = aa' + (a'b + ab')\beta + bb'\beta^2 \\ &= aa' - 5bb' + (a'b + ab' + bb')\beta \end{aligned}$$

qui est bien dans A . A est aussi stable pour l'addition et contient 1, c'est un sous-anneau de $\mathbb{C}$. Enfin, β et $\bar{\beta}$ sont les racines de l'équation $x^2 - x + 5$, leur somme est donc 1, ce qui prouve que $\bar{\beta}$ est dans A et donc que A est stable pour la conjugaison.

- (2) La méthode est toujours la même : on considère $N(z) = z\bar{z}$. Si $z = a + b\beta$, on a

$$N(z) = (a + b\beta)(a + b\bar{\beta}) = (a + b\beta)(a + b - b\beta) = a^2 + ab + 5b^2$$

z inversible demande que $N(z) = 1$, puisque $N(z)N(z^{-1}) = 1$ et que ces deux normes sont des entiers positifs. Mais, en multipliant par 4,

$$4N(z) = 4a^2 + 4ab + 20b^2 = (2a + b)^2 + 19b^2 = 4$$

impose $b = 0$ donc $a = \pm 1$. Seuls $z = 1$ et $z = -1$ conviennent.

- (3) Supposons qu'il existe une division euclidienne, associée au stathme ϕ . L'ensemble des valeurs prises par l'application stathme restreinte à $A \setminus \{0, 1, -1\}$ est un sous-ensemble de $\mathbb{N}$ qui a un plus petit élément. Ce plus petit élément est atteint pour un certain z_0 de A .
- (4) Soit alors (z_0) l'idéal principal engendré par z_0 . Comme la division par z_0 ne peut avoir comme reste que 0, 1 ou -1 , ce quotient a deux ou trois éléments, selon que 1 et -1 sont ou non dans la même classe. C'est un anneau : ce ne peut donc être que $\mathbb{F}_2$ ou $\mathbb{F}_3$: il est en effet facile de vérifier qu'il n'existe qu'une structure de groupe à deux (resp. trois) éléments, et qu'on ne peut alors construire qu'une multiplication qui en fasse un anneau.
- (5) Si p est la projection de A sur $A/(z_0)$, l'image $\alpha = p(\beta)$ doit vérifier, comme β , l'équation $\alpha^2 - \alpha + \bar{5} = \bar{0}$. Or cette égalité, qui s'écrit $\alpha^2 + \alpha + \bar{1} = \bar{0}$ dans $\mathbb{F}_2$, et $\alpha^2 - \alpha + \bar{2} = \bar{0}$ dans $\mathbb{F}_3$ n'a pas de solution, dans l'un et l'autre cas. L'anneau A ne peut être euclidien.

On montre néanmoins qu'il est principal, en utilisant une « division » qui ressemble à la division euclidienne : voir par exemple le livre [21].

Chapitre 3

Polynômes

3.1 L'ANNEAU DES POLYNÔMES $A[X]$

3.1.1. Généralités

Si A est un anneau commutatif, on définit l'anneau des polynômes à coefficients dans A comme dans le cas d'un corps : ce sont des suites d'éléments de A « presque tous nuls », avec une opération de somme et de produit.

Définition 3.1. *L'anneau des polynômes $A[X]$ est l'ensemble des suites $P = (a_n)_{n \in \mathbb{N}}$ où les a_n sont dans A , et sont nuls sauf un nombre fini. Il est muni des opérations $+$ et $\times$ définies par :*

$$\begin{aligned}(a_n)_{n \in \mathbb{N}} + (b_n)_{n \in \mathbb{N}} &= (a_n + b_n)_{n \in \mathbb{N}} \\ (a_n)_{n \in \mathbb{N}} \times (b_n)_{n \in \mathbb{N}} &= (c_n)_{n \in \mathbb{N}}\end{aligned}$$

avec

$$c_n = \sum_{i=0}^n a_i b_{n-i}$$

On vérifie que c'est un anneau commutatif. Le sous-ensemble des suites de la forme $(a, 0, 0, \dots)$ est un sous-anneau isomorphe à A , et on l'identifie à A . Il contient les deux éléments neutres 0 et 1. Si on note X la suite $(0, 1, 0, 0, \dots)$, on constate que

$X \times X = (0, 0, 1, 0, 0, \dots)$, et que la suite des polynômes $(X^k)_{k \in \mathbb{N}}$ est génératrice en ce sens que

$$P = \sum_{i=0}^{\infty} a_i X^i$$

Cette somme est en réalité finie puisque la suite des coefficients est presque nulle.

On définit comme dans le cas usuel la notion de degré, et on appelle **coefficient dominant** le coefficient du terme de plus haut degré, et un polynôme est dit **unitaire** si son coefficient dominant est 1. Le degré du polynôme nul est, par convention, $-\infty$. Attention... en toute généralité, le degré du produit de deux polynômes est inférieur ou égal à la somme des degrés. Dans un anneau non intègre comme $\mathbb{Z}/6\mathbb{Z}$, on a par exemple

$$(2X^2 + 1)(3X^3 + X + 2) = 5X^3 + 4X^2 + X + 2$$

Proposition 3.2. *Si A est intègre, $\deg(PQ) = \deg(P) + \deg(Q)$.
 $A[X]$ est intègre si et seulement si A est intègre.*

Démonstration. La première propriété est immédiate, avec des conventions naturelles dans le cas du polynôme nul. Pour la seconde propriété, la condition est nécessaire, à cause des polynômes constants ; réciproquement, si A est intègre, la première propriété du théorème assure que $A[X]$ est intègre. $\square$

3.1.2. Division euclidienne, racines

Définition 3.3. *Soit D un polynôme de $A[X]$ dont le coefficient dominant est inversible. Alors, pour tout $P \in A[X]$, il existe Q et R de $A[X]$ tels que :*

$$P(X) = D(X)Q(X) + R(X), \text{ avec } R = 0 \text{ ou } \deg R < \deg D$$

On dit que c'est la division euclidienne de P par D , la démonstration se fait facilement par récurrence sur le degré de D . La division euclidienne est toujours possible si A est un corps (et si D n'est pas le polynôme nul), et donc

Proposition 3.4. *Si K est un corps commutatif, $K[X]$ est euclidien.*

Si A n'est pas un corps, la division euclidienne n'est pas toujours possible. On ne peut diviser un polynôme unitaire par un polynôme de degré inférieur et dont le coefficient dominant est non inversible : considérer les termes de plus haut degré. On verra même en exercice (cf. exercice 3.4) que lorsque A n'est pas un corps, $A[X]$ n'est pas principal. Néanmoins, la division euclidienne reste un outil fondamental pour l'étude des anneaux $A[X]$.

Si α est un élément de A , on définit $P(\alpha)$ comme étant $P(\alpha) = \sum a_i \alpha^i$, et si $P(\alpha) = 0$, on dit que α est une **racine** du polynôme P . Si B est un anneau contenant A , on peut encore définir $P(\alpha)$ avec $\alpha \in B$.

Proposition 3.5. Dans $A[X]$, $P(a) = 0 \iff X - a \mid P$. Si A est intègre, un polynôme P non nul a au maximum $\deg P$ racines.

Démonstration. Il suffit d'utiliser la division euclidienne par $X - a$ qui est toujours définie et s'écrit

$$P(X) = (X - a)Q(X) + P(a)$$

On démontre la seconde propriété par récurrence sur le degré. Si P est de degré $n + 1$ et s'il a une racine a , alors, pour tout b différent de a , $P(b) = (b - a)Q(b)$, avec $\deg Q = n$, et b est une racine de P si et seulement si c'est une racine de Q puisque A est intègre. $\square$

Attention au cas où A n'est pas intègre. Par exemple, le polynôme X^2 a 4 racines dans $\mathbb{Z}/16\mathbb{Z}$.

On définit les racines multiples, en disant que a est racine multiple d'ordre (exactement) k si

$$P(X) = (X - a)^k Q(X) \text{ et } Q(a) \neq 0$$

Exercice 3.1. Soit $P \in A[X]$, $P(X) = \sum_{i=0}^n a_i X^i$. On définit P' par

$$P'(X) = \sum_{i=1}^n i a_i X^{i-1}$$

et on a les propriétés habituelles de la dérivation. Montrer que a est racine simple de P si et seulement si $P(a) = 0$ et $P'(a) \neq 0$. Généraliser à des racines multiples.

Exercice 3.2. Soit $P(X)$ un polynôme de $\mathbb{R}[X]$. On suppose que la décomposition en irréductibles de P est de la forme $P(X) = \prod_{i=1}^k P_i(X)^{\alpha_i}$, où les P_i sont irréductibles distincts. Comment calculer le polynôme $R(X) = \prod_{i=1}^k P_i(X)$? Bien sûr, cela sans connaître les P_i . On utilisera le p.g.c.d. $P \wedge P'$ et on fera l'application au polynôme P défini par

$$X^{10} - 7X^9 + 2X^8 + 15X^7 + 87X^6 + 140X^5 + 207X^4 + 183X^3 + 146X^2 + 65X + 25$$

Il est conseillé d'utiliser un logiciel de calcul formel.

Exercice 3.3. Soit P un polynôme de $\mathbb{Z}[X]$. On suppose que P a une racine rationnelle. Montrer qu'on peut la trouver avec un nombre fini d'essais. Application : trouver les racines rationnelles du polynôme P défini par

$$P(X) = 8X^5 + 78X^3 + 52X^4 + 3X^2 + 70X - 49$$

et le factoriser en irréductibles. On pourra également chercher à factoriser le polynôme obtenu dans l'exercice précédent.

Exercice 3.4. Soit A un anneau commutatif. Démontrer que $A[X]$ est principal si et seulement si A est un corps. On pourra considérer l'idéal (a, X) engendré par a élément non inversible de A et par X .

3.2 POLYNÔMES IRRÉDUCTIBLES

3.2.1. Corps algébriquement clos

Pour étudier les questions de divisibilité, il est important de connaître les polynômes irréductibles de $A[X]$. La question n'est pas facile à résoudre en général, même lorsque A est un corps, et nous nous limiterons au cas d'un **anneau commutatif intègre**. Commençons par un cas élémentaire :

Proposition 3.6. *Pour tout $a \in A$, $X - a$ est irréductible.*

L'outil de base est le degré : on commence par observer qu'un polynôme est une unité si et seulement si c'est un polynôme constant égal à une unité de A . Ensuite, puisque $\deg QR = \deg Q + \deg R$, les diviseurs de $X - a$ sont de degré 0 ou 1, ce sont les unités ou les associés de $X - a$.

Il est un cas où on obtient ainsi tous les irréductibles.

Définition 3.7. *Le corps K est **algébriquement clos** si tout polynôme non constant admet (au moins) une racine.*

Théorème 3.8. *Si K est algébriquement clos, tous les polynômes de $K[X]$ se factorisent en polynômes du premier degré. On dit qu'ils sont scindés. Les polynômes irréductibles de $K[X]$ sont les polynômes du premier degré.*

Démonstration. La démonstration se fait par récurrence sur le degré. Comme $K[X]$ est euclidien donc factoriel, la factorisation obtenue est la décomposition en irréductibles. $\square$

Voici maintenant un théorème qui prouve qu'il existe des corps algébriquement clos. Nous n'en faisons pas la démonstration (on en trouve dans les ouvrages d'analyse, comme par exemple celui de cette collection).

Théorème 3.9. Théorème de d'Alembert-Gauss. *Le corps $\mathbb{C}$ est algébriquement clos.*

On en déduit :

Théorème 3.10. *Les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré 1 et les polynômes du second degré sans racines réelles.*

Démonstration. Tout polynôme P de $\mathbb{R}[X]$ peut être considéré comme un polynôme de $\mathbb{C}[X]$. De plus, si α est une racine non réelle de P , on a $\overline{P(\alpha)} = P(\overline{\alpha}) = 0$, donc $\overline{\alpha}$ est racine de P avec la même multiplicité. On termine en observant que $(X - \alpha)(X - \overline{\alpha})$ est alors un polynôme du second degré à coefficient réels, sans racine réelle. $\square$

On peut démontrer (c'est le théorème de Steinitz) que tout corps est inclus dans un corps algébriquement clos, le plus petit s'appelant sa clôture algébrique (il est unique à isomorphisme près).

3.2.2. Factorialité des anneaux de polynômes

Si A est un anneau, sans être un corps, $A[X]$ n'est pas euclidien, et $A[X]$ n'est pas principal même si A l'est (cf. l'exercice 3.4). Par contre, la factorialité s'hérite :

Théorème 3.11. *Si A est un anneau factoriel, alors l'anneau des polynômes $A[X]$ est factoriel.*

Démonstration. Soit P un polynôme (non nul, comme dans tout le théorème) de $A[X]$ et $c = \text{p.g.c.d.}(a_0, a_1, \dots, a_n)$ où les a_i sont ses coefficients. On dit parfois que c est le **contenu** de P , il est défini à une unité près. Si $c = 1$ (en fait donc une unité), on dit que le polynôme P est **primitif**; tout polynôme peut s'écrire $P = c(P)P'$ où P' est primitif. Montrons d'abord le lemme de Gauss.

Lemme 3.12. de Gauss

- Si P_1 et P_2 sont primitifs, alors $P_1 P_2$ est primitif.
- Le contenu du produit de deux polynômes est le produit des contenus.

Démonstration. [du lemme] Notons (a_i) , (b_j) et (c_k) les coefficients respectifs de P_1 , P_2 et $P_1 P_2$. Soit p un irréductible. Il peut diviser $a_1, a_2, \dots$ mais pas tous les a_i . Soit h le premier indice tel que p ne divise pas a_h . De même, soit k le premier indice tel que p ne divise pas b_k . Alors, le coefficient $c_{h+k} = \sum_{\ell} a_{\ell} b_{h+k-\ell}$ n'est pas divisible par p puisqu'il s'écrit comme un multiple de p plus $a_h b_k$. En conclusion, aucun irréductible ne divise tous les coefficients de $P_1 P_2$ qui est donc primitif. Pour la deuxième partie du lemme, écrivons $P = c(P)P'$, $Q = c(Q)Q'$, alors $PQ = c(P)c(Q)P'Q'$ et $P'Q'$ est primitif, donc $c(PQ) = c(P)c(Q)$. $\square$

Appelons K le corps des fractions l'anneau intègre A (voir l'exercice 2.14), et démontrons un autre lemme :

Lemme 3.13. *Si $P \in A[X]$ se factorise dans $K[X]$, sous la forme $P = QR$, alors il existe deux polynômes Q' et R' de $A[X]$, de mêmes degré respectif que Q et R , tels que $P = Q'R'$.*

Démonstration. [du nouveau lemme] Commençons par le cas où P est primitif. Dans le corps des fractions de l'anneau A , on peut trouver des «dénominateurs communs» pour les coefficients et donc écrire Q et R sous la forme :

$$Q = \frac{n}{d}Q^* \text{ et } R = \frac{n'}{d'}R^*$$

où Q^* et R^* sont dans $A[X]$ et sont primitifs. Mais on a alors : $dd'P = nn'Q^*R^*$ et donc $dd' = nn'$ en regardant les contenus de chaque membre, et donc $P = Q^*R^*$. Passons maintenant au cas où P est quelconque. On écrit $P = c(P)P^*$, mais alors $P^* = \frac{QR}{c(P)}$ et on applique ce qui précède, P^* se factorise dans A , c'est donc aussi le cas de P . $\square$

On peut remarquer qu'une conséquence de ce lemme est que, pour un polynôme primitif, être irréductible sur A ou être irréductible sur K sont équivalents. Plus précisément, les irréductibles de $A[X]$ sont les constantes irréductibles de A et les polynômes primitifs qui sont irréductibles dans $K[X]$.

Montrons maintenant que $A[X]$ est factoriel. Un polynôme P s'écrit $P = c(P)P^*$ où P^* est primitif. Le contenu est dans A , donc il se factorise de façon unique en irréductibles de A . Le polynôme P^* se factorise dans l'anneau principal $K[X]$: $P^* = P_1P_2 \dots P_k$. Mais les P_i se peuvent écrire : $P_i = \frac{n_i}{d_i}P_i^*$ et le même raisonnement que ci-dessus montre que P^* est produit des P_i^* qui sont irréductibles dans $A[X]$.

Reste à montrer l'unicité de la décomposition. Plutôt que de faire la démonstration, on va montrer que *tout irréductible est premier*, puisque c'est cela qui fait tout marcher. Supposons donc que P soit irréductible et divise QR . Alors il existe un polynôme S tel que $PS = QR$. Si P est constant, il est irréductible donc premier (dans A factoriel) et, en prenant les contenus, il divise soit le contenu de Q , soit le contenu de R . S'il n'est pas constant, on se place dans $K[X]$, et P y divise, par exemple R . Mais alors on fait comme dans le lemme, on écrit la factorisation de R :

$$R = \frac{n}{d}LP$$

où L et P sont primitifs, n et d dans A . En multipliant par d et en prenant les contenus, on voit que n est un multiple de d , d'où le résultat. $\square$

Insistons sur les résultats obtenus au cours de la démonstration, avec A factoriel et K son corps des fractions (par exemple $A = \mathbb{Z}$ et $K = \mathbb{Q}$) :

- Un polynôme de $A[X]$ réductible dans $K[X]$ est aussi réductible dans $A[X]$.
- Les irréductibles de $A[X]$ sont :
 - les polynômes constants $P(X) = a$ où a est irréductible dans A .
 - les polynômes primitifs qui sont irréductibles dans $K[X]$.

3.2.3. Critères d'irréductibilité

Nous allons donner quelques exemples de critères d'irréductibilité pour les polynômes de $\mathbb{Z}[X]$. Le plus simple consiste à «réduire» le polynôme modulo un entier n : si $P \in \mathbb{Z}[X]$, avec $P(X) = \sum_{i=0}^n a_i X^i$, on note $\overline{P}$ le polynôme de $(\mathbb{Z}/n\mathbb{Z})[X]$ dont les coefficients sont les classes modulo n des coefficients de P . Si P est unitaire et réductible dans $\mathbb{Z}[X]$, alors $\overline{P}$ est réductible dans $(\mathbb{Z}/n\mathbb{Z})[X]$, pour tout n : en effet, la factorisation d'un polynôme unitaire de $\mathbb{Z}[X]$ donne forcément deux polynômes unitaires, dont la réduction n'est jamais nulle ou constante. Par contraposition :

Proposition 3.14. *Soit P unitaire dans $\mathbb{Z}[X]$. S'il existe n tel que sa réduction modulo n est irréductible, alors P est irréductible dans $\mathbb{Z}[X]$.*

Ainsi, $X^5 - 2X^4 - 4X^3 + 3X^2 + 6X + 5$ se réduit modulo 2 à $X^5 + X^2 + 1$, qui est irréductible : il y a un nombre fini (et petit) d'essais de factorisation à faire, qui tous échouent.

Notre critère de réduction modulo n est loin de tout régler. Il existe en effet des cas de polynômes irréductibles, mais dont la réduction modulo p est réductible pour tout p premier. Un autre critère :

Proposition 3.15. Critère d'Eisenstein. *Si $P \in \mathbb{Z}[X]$ s'écrit $P(X) = \sum_{i=0}^n a_i X^i$ et s'il existe un nombre premier p tel que :*

$$p \mid a_0, \quad p \mid a_1, \dots, p \mid a_{n-1}, \quad p \nmid a_n, \quad p^2 \nmid a_0$$

alors P est irréductible dans $\mathbb{Q}[X]$.

Démonstration. D'après les hypothèses, la réduction modulo p de $P(X)$ est $\overline{P}(X) = \overline{a_n} X^n$. Si P était réductible sur $\mathbb{Q}$, il serait réductible sur $\mathbb{Z}$ et on aurait $P = QR$, avec Q et R non constants. Donc $\overline{Q}(X) = \overline{b} X^r$ et $\overline{R}(X) = \overline{c} X^s$, puisque l'anneau $\mathbb{F}_p[X]$ est factoriel. Mais alors les termes constants de R et de Q seraient tous les deux divisibles par p , et le terme constant de P , qui est leur produit, serait divisible par p^2 : contraire à l'hypothèse. $\square$

Exemple : $P(X) = X^n - 12$ et, plus généralement $P(X) = X^n - pk$ où p est premier ne divisant pas k , sont des polynômes irréductibles sur $\mathbb{Q}$. C'est un cas où la méthode de réduction ne donne rien. Un exemple moins direct :

Proposition 3.16. *Si p est premier, le polynôme $\Phi_p(X) = 1 + X + X^2 + \dots + X^{p-1}$, p -ième polynôme cyclotomique est irréductible sur $\mathbb{Q}$.*

Démonstration. On considère le polynôme composé $Q(Y) = P(Y + 1)$. Alors si P est réductible sur $\mathbb{Z}$, Q l'est également, et réciproquement puisque $P(X) = Q(X - 1)$.

Or

$$Q(Y) = 1 + (Y+1) + (Y+1)^2 + \dots + (Y+1)^{p-1} = \frac{1}{Y} ((Y+1)^p - 1) = \sum_{i=0}^{p-1} \binom{p}{i} Y^i$$

Or (voir dans la démonstration de la proposition 2.26) tous les coefficients $\binom{p}{i}$, pour $i = 1$ à $p-1$, sont divisibles par p , le dernier étant p n'est pas divisible par p^2 . On peut appliquer le critère d'Eisenstein, $\Phi_p(X)$ est irréductible sur $\mathbb{Q}[X]$. $\square$

Les polynômes cyclotomiques ont beaucoup d'intérêt en arithmétique. On regardera par exemple le premier problème, en fin de chapitre.

3.2.4. Corps finis

Commençons par un théorème difficile, dont nous admettrons la démonstration (voir par exemple [21]).

Théorème 3.17. de Wedderburn. *Tout corps fini est commutatif.*

On sait par contre qu'il existe des corps non commutatifs infinis, comme le célèbre corps des quaternions $\mathbb{H}$.

Un autre résultat, élémentaire cette fois :

Proposition 3.18. *Le cardinal d'un corps fini K est de la forme $q = p^n$ où p est un nombre premier, $n \in \mathbb{N}$.*

Démonstration. Rappelons en effet que K , en tant qu'anneau intègre fini, a pour caractéristique un nombre premier p ; il contient donc un sous-corps identifié à $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$, image du morphisme canonique de $\mathbb{Z}$ dans K . Par ailleurs, l'addition des éléments de K et le produit d'un élément de $\mathbb{F}_p$ par un élément de K font de K un espace vectoriel sur $\mathbb{F}_p$. Comme K est fini, cet espace vectoriel est de dimension finie n , et K est isomorphe (en tant qu'espace vectoriel) à $\mathbb{F}_p^n$. Il a donc pour cardinal celui de $\mathbb{F}_p^n$, soit $q = p^n$. $\square$

Et comment obtient-on des corps de cardinal $q = p^n$? On sait déjà qu'il existe un corps de cardinal p , c'est $\mathbb{F}_p$, mais on peut affirmer beaucoup plus :

Théorème 3.19. *Pour tout p premier et pour tout $n \in \mathbb{N}$, il existe un corps de cardinal $q = p^n$. De plus, ce corps est unique, à isomorphisme près, on le note $\mathbb{F}_q$.*

Nous ne démontrons pas ce théorème, qui demande notamment pour l'unicité de connaître les débuts de la théorie de Galois; cependant, montrons comment on peut, dans des cas concrets, obtenir une description de $\mathbb{F}_q$.

L'idée est la même que celle qui consiste à construire le corps des complexes comme quotient $\mathbb{R}[X]/(X^2 + 1)$.

Proposition 3.20. Soit k un corps, P un polynôme irréductible de $k[X]$ de degré n , alors $k[X]/(P(X))$ est un corps, contenant un sous-corps identifié à k , et c'est un k -espace vectoriel de dimension n .

Démonstration. Le début est bien connu : dans l'anneau principal $k[X]$, l'idéal engendré par un irréductible est maximal, donc le quotient $K = k[X]/(P(X))$ est un corps. Il contient les classes des polynômes constants que l'on identifie à k . Si on note α la classe de X , les éléments de K sont de la forme

$$x_0 + x_1\alpha + \dots + x_{n-1}\alpha^{n-1}$$

puisque tout polynôme est congru à son reste modulo $P(X)$. K est alors un k -espace vectoriel de base $1, \alpha, \dots, \alpha^{n-1}$. Si le polynôme P s'écrit

$$P(X) = X^n + a_{n-1}X^{n-1} + \dots + a_0$$

les calculs de produits dans K se font en utilisant

$$\alpha^n = -a_{n-1}\alpha^{n-1} - a_{n-2}\alpha^{n-2} - \dots - a_0$$

En définitive, tout se passe comme si on ajoutait au corps k un élément α «imaginaire» qui est une racine de P . $\square$

Outre $\mathbb{C}$ déjà évoqué, les exemples sont innombrables, comme

$$\mathbb{Q}[X]/(X^2 - 2) \equiv \{x \in \mathbb{R} \mid \exists a \in \mathbb{Q}, \exists b \in \mathbb{Q}, x = a + b\sqrt{2}\}$$

en «notant» $\sqrt{2}$ la classe de X . Avec un corps fini $\mathbb{F}_p$ et un polynôme de degré n , on obtient donc un corps de cardinal $q = p^n$. Par exemple,

$$\mathbb{F}_4 = \mathbb{F}_2[X]/(1 + X + X^2)$$

car le polynôme $1 + X + X^2$ est irréductible sur $\mathbb{F}_2$ puisqu'il n'a pas de racine. $\mathbb{F}_4$ a quatre éléments qui sont $0, 1, \alpha$ et $1 + \alpha$. Le seul produit à calculer est $\alpha(1 + \alpha) = \alpha + \alpha^2 = 1$ puisque $1 + \alpha + \alpha^2 = 0$ et que l'on est en caractéristique 2. Les éléments non nuls sont bien inversibles, c'est un corps.

Exercice 3.5. Soit K un corps commutatif. Montrer qu'un polynôme irréductible de degré $d > 1$ dans $K[X]$ n'a pas de racine dans K . Montrer qu'un polynôme de degré 2 ou 3 qui n'a pas de racine dans K est irréductible dans $K[X]$. Application : donner la liste des irréductibles de degré 1, 2, 3 et 4 dans $\mathbb{F}_2[X]$.

Exercice 3.6. Décrire un corps à neuf éléments.

3.3 POLYNÔMES À PLUSIEURS INDÉTERMINÉES

3.3.1. Définitions

Soit K un corps commutatif. On construit par récurrence l'anneau des polynômes à n variables sur K par :

$$A = K[X_1, X_2, \dots, X_n] = K[X_1, \dots, X_{n-1}][X_n]$$

On vérifie que l'ordre des variables n'a guère d'importance (un changement dans l'ordre conduit à un anneau canoniquement isomorphe). Attention, dès que $n \geq 2$, c'est un anneau qui n'est ni euclidien ni principal, mais c'est toujours un anneau factoriel, en utilisant le théorème 3.11 et une récurrence immédiate.

Proposition 3.21. *L'anneau $A = K[X_1, X_2, \dots, X_n]$ est un K -espace vectoriel de dimension infinie. Une base est constituée des **monômes** de la forme $X_1^{\alpha_1} X_2^{\alpha_2} \dots X_n^{\alpha_n}$.*

On pourra parfois utiliser la notion de multi-indice ou multi-exposant : si on pose $\alpha = (\alpha_1, \dots, \alpha_n)$, le monôme ci-dessus sera noté X^α . Son **degré total**, noté $|\alpha|$ est par définition la somme des α_i qui sont les degrés partiels. Ainsi, un polynôme de degré total inférieur à n sera de la forme

$$P = \sum_{|\alpha| \leq n} a_\alpha X^\alpha$$

Un petit résultat bien utile.

Proposition 3.22. *$P \in K[X_1, X_2, \dots, X_n]$ est divisible par $X_1 - X_2$ si et seulement si $P(X_1, X_1, X_3, \dots, X_n)$ est le polynôme nul.*

Démonstration. Clairement, si $P = (X_1 - X_2)Q$, alors $P(X_1, X_1, X_3, \dots, X_n)$ est le polynôme nul. Pour la réciproque, proposons deux approches.

Première démonstration : on pose $A = K[X_3, \dots, X_n]$ et on considère P comme élément de $A[X_1, X_2]$. Alors :

$$P(X_1, X_2) = \sum_{(i,j) \in I} a_{i,j} X_1^i X_2^j, \quad P(X_1, X_1) = \sum_{(i,j) \in I} a_{i,j} X_1^i X_1^j = 0$$

où I est un ensemble fini de couples d'indices. Par soustraction :

$$P(X_1, X_2) = \sum_{(i,j) \in I} a_{i,j} X_1^i (X_2^j - X_1^j)$$

et il ne reste plus qu'à observer que

$$X_2^j - X_1^j = (X_2 - X_1) \sum_{k=0}^{j-1} X_2^{j-1-k} X_1^k$$

pour j non nul. On en déduit que P est divisible par $X_1 - X_2$.

Seconde démonstration : on pose $A = K[X_2, X_3, \dots, X_n]$ et on considère P comme élément de $A[X_1]$. On peut alors écrire la division euclidienne de P par $X_1 - X_2$, qui est un polynôme unitaire du premier degré en l'indéterminée X_1 :

$$P = (X_1 - X_2)Q(X_1, \dots, X_n) + R$$

R est nul ou de degré 0 en X_1 , c'est donc un élément de A , c'est la valeur de P lorsque X_1 prend la valeur X_2 . Donc

$$P(X_1, X_2, \dots, X_n) = (X_1 - X_2)Q(X_1, \dots, X_n) + P(X_2, X_2, X_3, \dots, X_n)$$

d'où le résultat. $\square$

3.3.2. Polynômes symétriques

Commençons par définir une «action» de groupe, celle du groupe symétrique $\mathcal{S}_n$ sur l'anneau des polynômes. Soit σ une permutation des n entiers $1, 2, \dots, n$. Si P est un polynôme à n variables, on peut lui associer le polynôme :

$$(\sigma.P)(X_1, \dots, X_n) = P(X_{\sigma(1)}, X_{\sigma(2)}, \dots, X_{\sigma(n)})$$

Par exemple, si σ est la transposition qui échange 1 et 2 (notée $(1, 2)$), on aura :

$$(1, 2).(X_1^2 + X_2 + 3X_3X_1) = X_2^2 + X_1 + 3X_3X_2$$

On dit alors qu'un polynôme est symétrique si :

$$\forall \sigma \in \mathcal{S}_n, \sigma.P = P$$

Par exemple :

$$P(X, Y, Z) = X^2YZ + XY^2Z + XYZ^2$$

On notera S l'ensemble des polynômes symétriques.

Proposition 3.23. S est un sous-anneau de $K[X_1, X_2, \dots, X_n]$.

Il suffit de l'écrire. C'est un sous-anneau strict dès qu'il y a plus d'une variable.

Il existe des polynômes symétriques plus simples que d'autres, ce sont les n polynômes symétriques élémentaires :

Définition 3.24. On appelle k -ième polynôme symétrique élémentaire de n variables le polynôme :

$$\sigma_k = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} X_{i_1} X_{i_2} \dots X_{i_k}$$

En toute rigueur, on devrait écrire $\sigma_{k,n}$ pour faire figurer le nombre n des indéterminées. La somme porte sur **tous** les k -uplets tels que la condition $1 \leq i_1 < i_2 < \dots < i_k \leq n$ soit satisfaite. Ainsi :

$$\begin{aligned}\sigma_1 &= X_1 + X_2 + \dots + X_n, \\ \sigma_2 &= X_1X_2 + X_1X_3 + \dots + X_{n-1}X_n, \\ &\dots \\ \sigma_n &= X_1X_2 \dots X_n\end{aligned}$$

On pose parfois $\sigma_0 = 1$. Une partie de l'intérêt qu'apportent les polynômes symétriques élémentaires est le lien qu'il y a avec la théorie des équations polynômiales. Pour être plus précis, et en introduisant une nouvelle indéterminée que nous appellerons T :

$$\prod_{i=1}^n (T - X_i) = \sum_{i=0}^n (-1)^i \sigma_i T^{n-i}$$

la démonstration se fait par récurrence sur le nombre des indéterminées, en utilisant :

$$\sigma_{i,n+1}(X_1, X_2, \dots, X_{n+1}) = \sigma_{i,n}(X_1, \dots, X_n) + \sigma_{i-1,n}(X_1, \dots, X_n)X_{n+1}$$

pour i vérifiant $0 < i < n$.

En appliquant cette identité à des polynômes en T scindés, on obtient le lien entre les racines de ce polynôme et les coefficients. Notons également que cette relation permet de démontrer que les polynômes symétriques élémentaires sont bien symétriques, car le premier membre est invariant par toute permutation des indéterminées X_i .

Le théorème principal est :

Théorème 3.25. *L'anneau des polynômes symétriques est engendré algébriquement par les polynômes symétriques élémentaires. De plus, ces polynômes symétriques élémentaires sont algébriquement indépendants.*

Ce théorème signifie que tout polynôme symétrique est un polynôme en les polynômes symétriques élémentaires, avec la précision que cette écriture est unique. Précisons également que des polynômes $P_1, P_2, \dots, P_m$ sont algébriquement indépendants si :

$$\forall Q \in K[T_1, T_2, \dots, T_m], \quad Q(P_1, P_2, \dots, P_m) = 0 \Rightarrow Q = 0$$

Et quelle est la signification de $Q(P_1, P_2, \dots, P_n)$? On peut utiliser la notion de composition de polynômes, ou dire que c'est la valeur de la fonction polynôme associée à Q en un élément de l'anneau $K[X_1, X_2, \dots, X_n]^m$.

Avant de faire la démonstration, un ou deux exemples faciles :

$$\begin{aligned}X^2 + Y^2 + Z^2 &= (X + Y + Z)^2 - 2(XY + YZ + ZX) = \sigma_1^2 - 2\sigma_2 \\ XY^2 + X^2Y + YZ^2 + Y^2Z + XZ^2 + X^2Z &= (X + Y + Z)(XY + YZ + ZX) \\ &\quad - 3XYZ \\ &= \sigma_1\sigma_2 - 3\sigma_3\end{aligned}$$

Passons maintenant à la démonstration du théorème.

Démonstration. Commençons par ordonner les monômes d'un polynôme symétrique P en suivant l'ordre dit **lexicographique** ou alphabétique. On dit que

$$aX_1^{\alpha_1} X_2^{\alpha_2} \dots X_n^{\alpha_n} > bX_1^{\beta_1} X_2^{\beta_2} \dots X_n^{\beta_n}$$

si et seulement si $\alpha_m > \beta_m$ où $m = \min\{k \mid \alpha_k \neq \beta_k\}$. Cet ordre est total, voir l'exercice 1.17. Notons que les coefficients a et b n'interviennent pas (et ne sont pas nuls). Soit alors $aX_1^{\alpha_1} X_2^{\alpha_2} \dots X_n^{\alpha_n}$ le plus grand des monômes de P . Alors nécessairement $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n$, car le polynôme est symétrique et P contient tous les monômes de la forme $aX_1^{\alpha_{\sigma(1)}} X_2^{\alpha_{\sigma(2)}} \dots X_n^{\alpha_{\sigma(n)}}$. Et si par exemple on avait $\alpha_2 > \alpha_1$, on aurait une contradiction en prenant la permutation qui échange 1 et 2. Soit alors Q le polynôme :

$$Q = \sigma_1^{\alpha_1 - \alpha_2} \sigma_2^{\alpha_2 - \alpha_3} \dots \sigma_{n-1}^{\alpha_{n-1} - \alpha_n} \sigma_n^{\alpha_n}$$

Il est somme de monômes, et le plus grand de ces monômes est :

$$X_1^{\alpha_1 - \alpha_2} (X_1 X_2)^{\alpha_2 - \alpha_3} (X_1 X_2 X_3)^{\alpha_3 - \alpha_4} \dots (X_1 \dots X_n)^{\alpha_n} = X_1^{\alpha_1} X_2^{\alpha_2} \dots X_n^{\alpha_n}$$

Donc $P - aQ$ contient uniquement des monômes plus petits que le plus grand monôme de P . Il est encore symétrique, on a enlevé un polynôme symétrique à un polynôme symétrique. On réitère le processus avec le polynôme $P - aQ$. Le processus s'arrête en un nombre fini d'étapes : pour le dire de façon abstraite, c'est parce que l'ordre lexicographique est un bon ordre. Le processus décrit est algorithmique en ce sens qu'il décrit un procédé effectif.

Montrons maintenant que les σ_i sont algébriquement indépendants. Soit Ψ l'application linéaire de $K[X_1, X_2, \dots, X_n]$ dans lui-même qui à $X_1^{i_1} X_2^{i_2} \dots X_n^{i_n}$ associe

$$X_1^{i_1 + i_2 + \dots + i_n} X_2^{i_2 + \dots + i_n} \dots X_n^{i_n}$$

Raisonnons par l'absurde, et supposons que Q soit un polynôme reliant les σ_i . Pour fixer les idées, ce serait par exemple : $Q(X_1, X_2, X_3) = \alpha X_1^2 X_2 + \beta X_1 X_2 X_3$, ce qui supposerait $\alpha \sigma_1^2 \sigma_2 + \beta \sigma_1 \sigma_2 \sigma_3 = 0$. On suppose bien sûr que les coefficients ne sont pas nuls. Alors soit alors $aX_1^{i_1} X_2^{i_2} \dots X_n^{i_n}$ le monôme de Q dont l'image par Ψ est la plus grande pour l'ordre lexicographique (dans notre exemple, c'est $\beta X_1 X_2 X_3$ dont l'image est $X_1^3 X_2^2 X_3$). Dans $Q(\sigma_1, \dots, \sigma_n)$, il y le terme $a \sigma_1^{i_1} \sigma_2^{i_2} \dots \sigma_n^{i_n}$. Mais en développant ce terme, le monôme le plus grand est

$$aX_1^{i_1 + i_2 + \dots + i_n} X_2^{i_2 + \dots + i_n} \dots X_n^{i_n} = \Psi(aX_1^{i_1} \dots X_n^{i_n})$$

Mais ce monôme est alors le plus grand des monômes de $Q(\sigma_1, \dots, \sigma_n)$. Comme ce polynôme est nul, le coefficient a est nul, ce qui est absurde. $\square$

Maintenant un exemple. En utilisant la démonstration du théorème, étudions $P(X_1, X_2, X_3) = X_1^3 + X_2^3 + X_3^3$. Le plus grand monôme est X_1^3 , on soustrait donc $\sigma_1^3 = (X_1 + X_2 + X_3)^3$ ce qui donne :

$$P - \sigma_1^3 = -3(X_1^2 X_2 + X_1 X_2^2 + \dots) - 6X_1 X_2 X_3$$

Le plus grand monôme est alors $X_1^2 X_2$, on soustrait donc $-3\sigma_1\sigma_2$ et en définitive $P = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3$.

Un corollaire du théorème est le suivant :

Corollaire 3.26. *L'anneau $K[X_1, X_2, \dots, X_n]$ est isomorphe à l'anneau S .*

C'est donc un exemple de sous-anneau strict isomorphe à l'anneau tout entier.

Démonstration. On définit l'isomorphisme par :

$$\Phi(P)(X_1, X_2, \dots, X_n) = P(\sigma_1, \sigma_2, \dots, \sigma_n)$$

Par exemple, pour $n = 3$, $\Phi(X^2 - 2Y) = X^2 + Y^2 + Z^2$. Il est immédiat que Φ est un morphisme d'anneau : $\Phi(P + Q) = \Phi(P) + \Phi(Q)$, $\Phi(PQ) = \Phi(P)\Phi(Q)$ et $\Phi(1) = 1$. De plus, le théorème annonce bien que Φ est surjectif, et que son noyau est réduit à 0. $\square$

Exercice 3.7. Dans $A = K[X, Y]$ décrire l'idéal engendré par X et Y et justifier qu'il n'est pas principal.

Exercice 3.8. Montrer que tout polynôme symétrique est de même degré partiel en chacune des indéterminées. Comment se caractérisent alors les polynômes symétriques élémentaires ?

Exercice 3.9. Soit $X^3 + 2X^2 - X - 3$ un polynôme de $\mathbb{C}[X]$. On note α , β et γ ses trois racines. Déterminer un polynôme du troisième degré dont les trois racines sont α^2 , β^2 et γ^2 .

Exercice 3.10. À l'aide des polynômes symétriques élémentaires, écrire le polynôme

$$P(X_1, X_2, \dots, X_n) = \sum_{i,j,k \in I} X_i^2 X_j X_k$$

la somme portant sur les triplets (i, j, k) distincts.

EXERCICES

Exercice 3.11. Petite révision de l'équation de Bezout dans l'anneau (euclidien) des polynômes. Chercher dans $\mathbb{Q}[X]$ des polynômes A et B tels que :

$$(X^5 + 2)A(X) + (X^2 + X + 1)B(X) = 1$$

puis trouver les polynômes V tels que :

$$V(X) \equiv 3X^3 + 9 \pmod{X^5 + 2}$$

et

$$V(X) \equiv 3X + 7 \pmod{X^2 + X + 1}$$

Chercher un générateur de l'idéal engendré par $X^4 - 3X + 1$ et $X^2 + 2$.

Exercice 3.12. La méthode de Kronecker Soit P un polynôme de $\mathbb{Z}[X]$, de degré n . On a vu dans un exercice (3.3) qu'il était possible de déterminer toutes les racines rationnelles de P . On va montrer maintenant qu'il est possible, au moins en théorie, de décomposer P en irréductibles de $\mathbb{Z}[X]$. On pose $d = \left\lfloor \frac{n}{2} \right\rfloor$ où $[x]$ désigne la partie entière du réel x , et on suppose que P est réduit.

- (1) Soient $a_0, \dots, a_d$ des entiers relatifs distincts et $b_0, \dots, b_d$ des entiers relatifs. Montrer qu'il existe un seul polynôme Q de $\mathbb{Q}[X]$, de degré inférieur ou égal à d tel que, pour tout i , $Q(a_i) = b_i$.
- (2) Montrer que si P est réductible, il admet un diviseur $D \in \mathbb{Z}[X]$ non constant de degré inférieur ou égal à d .
- (3) Montrer que les listes $(D(a_0), \dots, D(a_d))$ forment un sous-ensemble fini de $\mathbb{Z}^{d+1}$. Conclure.

Exercice 3.13. On dit qu'un polynôme de $K[X_1, \dots, X_n]$ est **antisymétrique** s'il vérifie :

$$\forall \sigma \in \mathcal{S}_n, \sigma.P = \epsilon(\sigma)P$$

où $\epsilon(\sigma)$ est la signature de la permutation σ . Vérifier que le polynôme

$$V(X_1, X_2, \dots, X_n) = \prod_{1 \leq i < j \leq n} (X_j - X_i) \quad (\text{polynôme de Vandermonde})$$

est antisymétrique. Montrer que tout polynôme antisymétrique P peut s'écrire :

$$P(X_1, \dots, X_n) = V(X_1, \dots, X_n)Q(X_1, \dots, X_n)$$

où Q est un polynôme symétrique.

Exercice 3.14. Sommes de Newton. Dans l'anneau $A = K[X_1, X_2, \dots, X_n]$, où K est un corps commutatif, on appelle **sommes de Newton** les polynômes S_k définis par :

$$S_k(X_1, X_2, \dots, X_n) = \sum_{i=1}^n X_i^k$$

Ce sont bien sûr des polynômes symétriques.

- (1) Calculer S_0, S_1, S_2 et S_3 en fonction des polynômes symétriques élémentaires σ_i .

(2) On suppose que $k \geq n$. En introduisant le polynôme

$$\prod_{i=1}^n (U - X_i) = U^n - \sigma_1 U^{n-1} + \sigma_2 U^{n-2} + \dots + (-1)^n \sigma_n$$

Montrer que :

$$0 = S_k - \sigma_1 S_{k-1} + \sigma_2 S_{k-2} + \dots + (-1)^n \sigma_n S_{k-n}$$

(3) Dans le cas où $k < n$, montrer que :

$$S_k - \sigma_1 S_{k-1} + \sigma_2 S_{k-2} + \dots + (-1)^{k-1} \sigma_{k-1} S_1 + (-1)^k k \sigma_k = 0$$

On pourra évaluer les produits $\sigma_p S_{k-p}$.

(4) Montrer que réciproquement on peut calculer les polynômes symétriques élémentaires (σ_i) en fonction des sommes de Newton S_i . Faire le calcul pour $i = 0$ à 4.

Exercice 3.15. Soit M la matrice à coefficients dans $K[X_1, X_2, \dots, X_n]$ définie par

$$M = (X_i^{j-1})_{i=1..n, j=1..n}$$

(1) Calculer le déterminant de M (déterminant de Vandermonde).

(2) Exprimer les coefficients de la matrice ${}^t M M$ en fonction des sommes de Newton.

(3) Application : si x_1, x_2 et x_3 sont les trois racines complexes du polynôme $X^3 + pX + q$, calculer

$$\Delta = (x_2 - x_1)^2 (x_3 - x_1)^2 (x_3 - x_2)^2$$

que l'on appelle **discriminant** du polynôme.

PROBLÈMES

Les corrigés de ces problèmes sont disponibles sur le site de Dunod : www.dunod.com.

3.1. POLYNÔMES CYCLOTOMIQUES

Soit $\mathbb{U}_n$ l'ensemble des racines n -ièmes de l'unité dans $\mathbb{C}$, c'est-à-dire l'ensemble des complexes z tels que $z^n = 1$.

(1) Montrer que $\mathbb{U}_n$ est un groupe cyclique. On appelle **racine primitive** n -ième de l'unité un de ses générateurs.

(2) Déterminer les racines primitives de l'unité pour $n = 1$ à 10.

(3) On note $\Phi_n(X)$ le polynôme dont les racines sont exactement les racines primitives n -ièmes de l'unité. Déterminer Φ_n pour $n = 1$ à 10. Ce polynôme Φ_n est appelé n -ième polynôme cyclotomique.

(4) Quel est le degré de Φ_n ? Déterminer Φ_n pour n premier.

(5) Montrer que

$$X^n - 1 = \prod_{d|n} \Phi_d(X)$$

(6) En déduire

$$\Phi_n(X) = \prod_{d|n} (X^d - 1)^{\mu(\frac{n}{d})}$$

où μ est la fonction de Möbius, présentée dans le problème du premier chapitre. On rappelle que $\mu(1) = 1$, $\mu(p_1 p_2 \dots p_k) = (-1)^k$ si les p_i sont des irréductibles distincts, $\mu(n) = 0$ si n n'est pas produit d'irréductibles distincts.

(7) Montrer que les polynômes Φ_n sont unitaires à coefficients dans $\mathbb{Z}$. On peut démontrer qu'ils sont irréductibles sur $\mathbb{Q}$, pour tout n . Voir le cas où n est premier dans l'application du critère d'Eisenstein page 65.

3.2. POLYNÔMES IRREDUCTIBLES SUR LES CORPS FINIS

Le but de ce problème est d'étudier les polynômes irréductibles de $\mathbb{F}_p[X]$ où p est un nombre premier. On prouvera en particulier qu'il existe toujours au moins un polynôme irréductible de degré n , et donc un corps fini de cardinal $q = p^n$.

- (1) Soit $q = p^n$. On suppose qu'il existe un corps K de cardinal q . Montrer que les éléments de K sont les racines du polynôme $X^q - X \in K[X]$.
- (2) On note $N(n, p)$ le nombre des polynômes unitaires irréductibles de degré n dans $\mathbb{F}_p[X]$. Calculer $N(1, p)$ et $N(2, p)$.
- (3) Soit P un polynôme irréductible de $\mathbb{F}_p[X]$, de degré d . Si $K = \mathbb{F}_p[X]/(P(X))$, montrer que P est un diviseur du polynôme $X^{p^n} - X$, pour tout n multiple de d .
- (4) Réciproquement, montrer que les facteurs irréductibles unitaires de $X^{p^n} - X$ sont exactement les polynômes unitaires irréductibles de $\mathbb{F}_p[X]$ dont le degré est un diviseur de n .
- (5) Montrer que $X^{p^n} - X$ n'a pas de facteurs irréductibles multiples.
- (6) En déduire

$$p^n = \sum_{d|n} dN(d, p)$$

puis

$$N(n, p) = \frac{1}{n} \sum_{d|n} p^d \mu\left(\frac{n}{d}\right)$$

(voir la formule d'inversion de Möbius, vue à la fin du problème du chapitre 1).

- (7) Vérifier cette formule pour $n = 1$ ou 2 , ainsi que pour $p = 2$, $n = 3$ et $n = 4$.
- (8) Montrer que $N(n, p) > 0$.

SOLUTIONS DES EXERCICES

Solution 3.1. Par définition, a est racine simple de P si $P(X) = (X - a)Q(X)$ avec $Q(a) \neq 0$. On a donc $P'(X) = Q(X) + (X - a)Q'(X)$, donc $P'(a) = Q(a)$ d'où la propriété annoncée. Dans le cas des anneaux de polynômes sur un corps, cela se généralise en :

$$a \text{ racine multiple d'ordre } k \iff P(a) = P'(a) = \dots = P^{(k-1)}(a) = 0, \\ P^{(k)}(a) \neq 0$$

et cela se démontre aisément, avec par exemple l'identité de Taylor pour les polynômes. Mais ce résultat ne subsiste pas en caractéristique non nulle. Par exemple, $P(X) = X^2$ admet 0 comme racine double, mais en caractéristique deux, toutes ses dérivées successives sont identiquement nulles. En caractéristique non nulle, on peut néanmoins trouver une relation entre racines multiples et dérivées successives ; il suffit, tout simplement, de changer la définition d'un polynôme dérivé. On pourra par exemple consulter [10].

Solution 3.2. Soit $D = P \wedge P'$. Un diviseur irréductible de D est nécessairement un diviseur irréductible de P . Choisissons un indice i et écrivons $P = P_i^{\alpha_i} Q$ où P_i est un irréductible qui ne divise pas Q . Alors $P' = P_i^{\alpha_i-1}(\alpha_i P_i' Q + P_i Q')$, et comme P_i , irréductible, ne divise ni P_i' ni Q , P_i ne divise pas $\alpha_i P_i' Q + P_i Q'$ donc $P_i^{\alpha_i-1}$ est facteur de D avec l'exposant maximum. Ainsi,

$$D(X) = P(X) \wedge P'(X) = \prod_{i=1}^k P_i(X)^{\alpha_i-1}$$

Le polynôme cherché est donc $R(X) = \frac{P(X)}{P(X) \wedge P'(X)}$. Dans le cas de l'application proposée, on trouve avec l'algorithme d'Euclide :

$$P(X) \wedge P'(X) = X^5 - 3X^4 - 7X^3 - 13X^2 - 9X - 5$$

$$R(X) = X^5 - 4X^4 - 3X^3 - 9X^2 - 4X - 5$$

Le polynôme P a donc cinq racines distinctes.

Solution 3.3. Supposons que le rationnel $\frac{p}{q}$ soit racine de $P(X) = \sum_{i=0}^n a_i X^i$, avec a_n non nul. On peut supposer que c'est aussi le cas de a_0 , si on a au préalable éliminé la racine nulle. En multipliant par q^n ,

$$a_0 q^n + a_1 p q^{n-1} + \dots + a_{n-1} p^{n-1} q + a_n p^n = 0$$

donc

$$q(a_0q^{n-1} + a_1pq^{n-2} + \dots + a_{n-1}p^{n-1}) = -a_np^n$$

et, comme p et q sont premiers entre eux, le théorème de Gauss permet d'affirmer que q est un diviseur de a_n . De même, on voit que p est un diviseur de a_0 . Il y a donc un nombre fini d'essais à faire pour rechercher d'éventuelles racines rationnelles. Ainsi, pour le polynôme

$$P(X) = 8X^5 + 78X^3 + 52X^4 + 3X^2 + 70X - 49$$

on doit tester $p = \pm 1, \pm 7, \pm 49$ et $q = \pm 1, \pm 2, \pm 4, \pm 8$. Avec un peu de patience, on trouve les solutions $\frac{1}{2}$ et $-\frac{7}{2}$. Donc

$$P(X) = (2X - 1)(2X + 7)(2X^3 + 7X^2 + 2X + 7)$$

À nouveau $-\frac{7}{2}$ est racine et $P(X) = (2X - 1)(2X + 7)^2(X^2 + 1)$, on a obtenu une factorisation en irréductibles sur $\mathbb{Z}$. De même, pour le polynôme R obtenu dans l'exercice précédent, $a_0 = -5$ et $a_5 = 1$, on trouve une racine rationnelle qui est 5, d'où

$$R(X) = X^5 - 4X^4 - 3X^3 - 9X^2 - 4X - 5 = (X - 5)(X^4 + X^3 + 2X^2 + X + 1)$$

et, mais là il faut de l'imagination...

$$X^4 + X^3 + X^2 + X^2 + X + 1 = (X^2 + 1)(X^2 + X + 1)$$

on a factorisé en irréductibles.

Solution 3.4. Si A est un corps, $A[X]$ est un anneau euclidien donc un anneau principal. Supposons $A[X]$ principal (donc intègre), alors A est nécessairement intègre. Si A n'est pas un corps, soit $a \neq 0$ un non inversible de A . On va montrer que l'idéal $(a, X) = aA[X] + XA[X]$ n'est pas principal. En effet, si $(a, X) = (P)$, P serait un diviseur de X et de a . Pour des raisons de degré, ce serait un élément b de A , et on aurait $X = bQ(X)$. Mais alors Q serait du premier degré de la forme $cX + d$ et $bc = 1$, b serait inversible. L'idéal (a, X) serait l'anneau tout entier, ce qui est absurde car les polynômes constants de (a, X) sont seulement les éléments de aA , qui ne contiennent aucun inversible.

La démonstration faite prouve en particulier que $(2, X)$ n'est pas principal dans $\mathbb{Z}[X]$; concrètement, c'est l'ensemble des polynômes dont le terme constant est pair.

Solution 3.5. Tout polynôme ayant une racine a et un degré strictement supérieur à 1 est factorisable par $X - a$, avec un facteur non constant, il est donc réductible. Par contraposition, un polynôme irréductible de degré supérieur ou égal à deux n'a pas de racine.

Si P est réductible dans $K[X]$, on a $P = QR$, et $\deg P = \deg Q + \deg R$. De plus, Q et R ne sont pas des unités, donc sont de degré non nul. Si donc $\deg P = 2$ ou $\deg P = 3$, les seules possibilités sont $2 = 1 + 1$ et $3 = 2 + 1$, un des facteurs est donc de degré 1 et P a une racine dans K . Par contre, un polynôme de degré 4 peut

ne pas avoir de racine et être néanmoins réductible. C'est le cas par exemple sur $\mathbb{R}$ du polynôme $(X^2 + 1)^2$ mais aussi, de façon plus cachée :

$$P(X) = X^4 + 1 = (X^2 + X\sqrt{2} + 1)(X^2 - X\sqrt{2} + 1)$$

Sur le corps $\mathbb{F}_2$, les polynômes du premier degré, X et $X + 1$ ¹ sont irréductibles. Les polynômes du second degré sont X^2 , $X^2 + X$, $X^2 + 1$, $X^2 + X + 1$. Les trois premiers sont réductibles car ils ont une ou deux racine(s), par exemple $X^2 + 1 = (X + 1)^2$. Par contre $X^2 + X + 1$ n'a pas de racine, il est irréductible.

Pour le degré 3, les polynômes sans racine sont $X^3 + X^2 + 1$, $X^3 + X + 1$, ils sont irréductibles. Enfin, pour le degré 4, les polynômes sans racines sont

$$X^4 + X^3 + X^2 + X + 1, X^4 + X^3 + 1, X^4 + X + 1, X^4 + X^2 + 1$$

Un polynôme du quatrième degré sans racine ne peut être réductible que s'il est produit de deux polynômes irréductibles du second degré. Or,

$$(X^2 + X + 1)^2 = X^4 + X^2 + 1$$

Il y a donc trois polynômes du quatrième degré irréductibles.

Solution 3.6. Un corps à 9 éléments est nécessairement de caractéristique 3. Cherchons un polynôme du second degré irréductible sur $\mathbb{F}_3[X]$: $X^2 + 1$ convient car il n'a pas de racine. $K = \mathbb{F}_3[X]/(X^2 + 1)$ est donc un corps à 9 éléments. Si on note 0, 1 et -1 les éléments de $\mathbb{F}_3$ et α la classe de X , les neuf éléments seront

$$0, 1, -1, \alpha, -\alpha, \alpha + 1, -\alpha + 1, \alpha - 1, -\alpha - 1$$

les calculs de sommes se font sans difficultés, car $K = \text{vect}_{\mathbb{F}_3}(1, \alpha)$; pour les produits, montrons un exemple

$$\alpha(\alpha + 1) = \alpha^2 + \alpha = \alpha - 1$$

On pourra remarquer que K^* , qui a huit éléments, est un groupe cyclique pour le produit. Il est engendré par exemple par $1 + \alpha$. Ce n'est pas un hasard : voir le théorème 5.16, p. 127.

Solution 3.7. Tout polynôme de $(K[X])[Y] = K[X, Y]$ s'écrit sous la forme $P_0(X) + P_1(X)Y + P_2(X)Y^2 + \dots + P_n(X)Y^n$ avec $P_0(X) = a_0 + a_1X + \dots + a_kX^k$. Il est dans l'idéal (X, Y) si et seulement si a_0 est nul. C'est idéal n'est pas principal... On l'a déjà démontré dans l'exercice 3.4, tout repose sur le fait que X n'est pas inversible dans $K[X]$.

Solution 3.8. Soit $P(X_1, X_2, \dots, X_n)$. Son degré partiel en X_1 est le maximum des degrés partiels en X_1 de chacun de ses monômes. Mais si $aX_1^{\alpha_1}X_2^{\alpha_2}\dots X_n^{\alpha_n}$ est un monôme de P , il y aura aussi $aX_1^{\alpha_2}X_2^{\alpha_1}\dots X_n^{\alpha_n}$, donc le degré partiel en X_2 est égal au degré partiel en X_1 . Les polynômes symétriques élémentaires sont ainsi les polynômes symétriques dont le degré partiel en chacune des variables est 1.

1. On note 1 la classe de 1.

Solution 3.9. Si $P(X) = X^3 + 2X^2 - X - 3$ a pour racines α, β et γ , on a

$$\begin{cases} \alpha + \beta + \gamma = -2 \\ \alpha\beta + \beta\gamma + \gamma\alpha = -1 \\ \alpha\beta\gamma = 3 \end{cases}$$

d'où l'on tire

$$\begin{aligned} \alpha^2 + \beta^2 + \gamma^2 &= (\alpha + \beta + \gamma)^2 - 2(\alpha\beta + \beta\gamma + \gamma\alpha) \\ &= 4 + 2 = 6 \\ \alpha^2\beta^2 + \beta^2\gamma^2 + \gamma^2\alpha^2 &= (\alpha\beta + \beta\gamma + \gamma\alpha)^2 - 2\alpha^2\beta\gamma - 2\alpha\beta^2\gamma - 2\alpha\beta\gamma^2 \\ &= (\alpha\beta + \beta\gamma + \gamma\alpha)^2 - 2\alpha\beta\gamma(\alpha + \beta + \gamma) \\ &= 1 - 2 \times 3 \times (-2) = 13 \\ \alpha^2\beta^2\gamma^2 &= (\alpha\beta\gamma)^2 = 9 \end{aligned}$$

et les trois nombres α^2, β^2 et γ^2 sont les racines de $Q(X) = X^3 - 6X^2 + 13X - 9$

Solution 3.10. L'algorithme décrit dans la démonstration conduit à ôter du polynôme P le produit $\sigma_1\sigma_3$. Or ce produit se développe en :

$$\sigma_1\sigma_3 = \left(\sum_i X_i \right) \left(\sum_{i,j,k} X_i X_j X_k \right) = \sum_{i,j,k} X_i^2 X_j X_k + 4 \sum_{i,j,k,\ell} X_i X_j X_k X_\ell$$

Les sommes portent sur des listes d'indices distincts, et il est important de comprendre d'où provient le second terme du résultat ainsi que le coefficient 4 (choisir un parmi quatre...). Et donc

$$\sum_{i,j,k \in I} X_i^2 X_j X_k = \sigma_1\sigma_3 - 4\sigma_4$$

Bien sûr, cette formule se simplifie s'il n'y a que trois indéterminées, en ce cas σ_4 disparaît. Signalons néanmoins que, même avec notre algorithme, les calculs deviennent rapidement fastidieux : il suffit de prendre X_i^3 au lieu de X_i^2 dans notre exemple pour s'en rendre compte.

Solution 3.11. On écrit l'algorithme d'Euclide :

$$\begin{aligned} X^5 + 2 &= (X^2 + X + 1)(X^3 - X^2 + 1) - X + 1 \\ X^2 + X + 1 &= (-X + 1)(-X - 2) + 3 \end{aligned}$$

Il est alors possible de «remonter» cet algorithme, en écrivant

$$\begin{aligned} 3 &= (-X + 1)(X + 2) + (X^2 + X + 1) \\ &= (X + 2)(X^5 + 2 - (X^2 + X + 1)) - (X^3 - X^2 + 1) + (X^2 + X + 1) \\ &= (X + 2)(X^5 + 2) - (X^4 + X^3 - 2X^2 + X + 1)(X^2 + X + 1) \end{aligned}$$

d'où une solution

$$A(X) = \frac{X+2}{3}, \quad B(X) = -\frac{X^4 + X^3 - 2X^2 + X + 1}{3}$$

Pour la seconde question, cherchons une solution. D'après la question précédente, le polynôme suivant sera solution :

$$V_0(X) = (3X^3 + 9)(X^2 + X + 1)B(X) + (3X + 7)(X^5 + 2)B(X)$$

On trouve

$$V_0(X) = \frac{1}{3}(19 + 8X + 6X^2 - 4X^5 - 3X^3 - 6X^4 + 4X^6 + 3X^7 - 6X^8 - 3X^9)$$

On peut réduire modulo $(X^5 + 2)(X^2 + X + 1)$ et on trouve toutes les solutions sous la forme :

$$V(X) = \frac{1}{3}(1 - 4X + 9X^3 - 13X^5 - 2X^6) + Q(X)(X^5 + 2)(X^2 + X + 1)$$

où Q est un polynôme quelconque.

Quant à la dernière question, ça va plus vite. Les deux polynômes sont premiers entre eux, l'idéal qu'ils engendrent est donc l'anneau tout entier.

Solution 3.12.

- (1) La classique théorie des polynômes de Lagrange permet de résoudre cette question. Si on définit les L_i par

$$L_i(X) = \frac{\prod_{j \neq i} (X - a_j)}{\prod_{j \neq i} (a_i - a_j)}$$

on obtient un polynôme à coefficients rationnels tel que

$$L_i(a_j) = \delta_{i,j}$$

où $\delta_{i,j}$ est le symbole de Kronecker qui vaut 0 en général et 1 lorsque $i = j$. Notre polynôme solution sera ainsi $Q = \sum_{i=0}^d b_i L_i(X)$. Il est de degré inférieur ou égal à d .

- (2) Si P est non constant et réductible sur $\mathbb{Z}[X]$, il admet une factorisation $P = P_1 P_2$ avec $\deg P_1 + \deg P_2 = \deg P$. Il est impossible que l'on ait à la fois $\deg P_1 < \frac{\deg P}{2}$ et $\deg P_2 < \frac{\deg P}{2}$, d'où l'affirmation de l'énoncé.
- (3) Lorsque les a_i sont fixés, les nombres $P(a_0), \dots, P(a_d)$ sont en nombre fini. L'ensemble de leurs diviseurs est également fini. Si donc D est un diviseur de P , les valeurs possibles des listes $(D(a_0), \dots, D(a_d))$ sont en nombre fini. Mais, si on se donne une telle liste de valeurs, on peut trouver par la méthode de Lagrange un seul polynôme de degré inférieur ou égal à d qui prenne exactement ces valeurs en les a_i . Il y a donc un nombre fini de diviseurs D à essayer : le problème de la factorisation d'un polynôme dans $\mathbb{Z}[X]$ est donc algorithmiquement résoluble. Au moins en théorie. En pratique, on utilise des algorithmes beaucoup plus performants, basés sur une réduction modulo p . Voir par exemple l'ouvrage [8].

Solution 3.13. Si V est le polynôme de Vandermonde, et $\sigma = (i, j)$ une transposition, alors

$$\sigma.V(X_1, \dots, X_n) = \prod_{k>\ell} (X_{\sigma(k)} - X_{\sigma(\ell)})$$

change de signe un nombre impair de fois, et est donc multiplié par la signature de cette transposition. Pour le cas général, il suffit de dire qu'une permutation est composée de transpositions, et que la signature est un morphisme.

Supposons maintenant P antisymétrique. Alors

$$P(X_1, \dots, X_i, \dots, X_j, \dots, X_n) = -P(X_1, \dots, X_j, \dots, X_i, \dots, X_n)$$

et donc

$$P(X_1, \dots, X_i, \dots, X_i, \dots, X_n) = 0$$

ce qui prouve que P est divisible par $X_j - X_i$. On écrit alors $P = (X_j - X_i)Q$, et on répète le même argument pour toutes les autres transpositions, d'où le résultat (on peut être plus formel, et faire une démonstration par récurrence sur le nombre des indéterminées.)

Solution 3.14.

(1) On a immédiatement :

$$\begin{aligned} S_0 &= X_1^0 + X_2^0 + \dots + X_n^0 = n\sigma_0 \\ S_1 &= X_1 + X_2 + \dots + X_n = \sigma_1 \\ S_2 &= X_1^2 + X_2^2 + \dots + X_n^2 = \sigma_1^2 - 2\sigma_2 \\ S_3 &= X_1^3 + X_2^3 + \dots + X_n^3 = \sigma_1^3 - 3\sigma_1\sigma_2 + 3\sigma_3 \end{aligned}$$

(2) On commence par substituer X_i à U :

$$0 = X_i^n - \sigma_1 X_i^{n-1} + \sigma_2 X_i^{n-2} + \dots + (-1)^n \sigma_n$$

puis on multiplie par X_i^{k-n} :

$$0 = X_i^k - \sigma_1 X_i^{k-1} + \sigma_2 X_i^{k-2} + \dots + (-1)^n \sigma_n X_i^{k-n}$$

enfin on somme pour tous les i de 1 à n :

$$0 = S_k - \sigma_1 S_{k-1} + \sigma_2 S_{k-2} + \dots + (-1)^n \sigma_n S_{k-n}$$

(3) Soit k fixé inférieur à n . On a

$$\begin{aligned} \sigma_1 S_{k-1} &= \left(\sum X_i \right) \left(\sum X_i^{k-1} \right) \\ &= \sum X_i^k + \sum X_i^{k-1} X_j \end{aligned}$$

les indices étant distincts. Plus généralement, Si p est un entier inférieur à $k - 2$,

$$\begin{aligned} \sigma_p S_{k-p} &= \left(\sum X_{i_1} \dots X_{i_k} \right) \left(\sum X_i^{k-p} \right) \\ &= \sum X_{i_1}^{k-p+1} X_{i_2} \dots X_{i_k} + \sum X_{i_1}^{k-p} X_{i_2} \dots X_{i_{k+1}} \end{aligned}$$

mais il y a un cas particulier si la somme de Newton est S_1 :

$$\begin{aligned}\sigma_{k-1}S_1 &= \left(\sum X_{i_1} \dots X_{i_{k-1}}\right) \left(\sum X_i\right) \\ &= \sum X_{i_1}^2 X_{i_2} \dots X_{i_{k-1}} + k \sum X_{i_1} X_{i_2} \dots X_{i_k} \\ &= \sum X_{i_1}^2 X_{i_2} \dots X_{i_{k-1}} + k\sigma_k\end{aligned}$$

La formule de Newton demandée s'obtient alors en sommant, on trouve

$$S_k - \sigma_1 S_{k-1} + \sigma_2 S_{k-2} + \dots (-1)^k \sigma_k = 0$$

il y a «télescopage».

- (4) La forme même des formules de Newton montre que l'on peut calculer les polynômes symétriques élémentaires en fonction des polynômes de Newton. Ainsi

$$\begin{aligned}\sigma_1 &= S_1 \\ \sigma_2 &= \frac{1}{2} (S_1^2 - S_2) \\ \sigma_3 &= \frac{1}{6} (2S_3 + S_1^3 - 3S_1 S_2)\end{aligned}$$

Solution 3.15. On peut considérer qu'on se place sur le corps des fractions rationnelles en les n indéterminées.

- (1) En développant par rapport à la dernière ligne, on voit que le déterminant de M est un polynôme $V(X_1, \dots, X_n)$ de degré $n-1$ en X_n , les coefficients étant dans $\mathbb{Z}[X_1, X_2, \dots, X_n]$. De plus, le coefficient de X_n^{n-1} est $V(X_1, X_2, \dots, X_{n-1})$. Enfin, on a $V(X_1, \dots, X_i, \dots, X_i) = 0$ pour tout $i < n$ par les propriétés des déterminants (deux lignes identiques). On a donc :

$$V(X_1, X_2, \dots, X_n) = V(X_1, X_2, \dots, X_{n-1}) \prod_{i=1}^{n-1} (X_n - X_i)$$

et une récurrence donne alors :

$$V(X_1, X_2, \dots, X_n) = \prod_{1 \leq i < j \leq n} (X_j - X_i)$$

- (2) Soit P la matrice ${}^t M M$. SI on appelle $p_{i,j}$ son coefficient général :

$$p_{i,j} = \sum_{k=1}^n X_k^{i-1} X_k^{j-1} = S_{i+j-2}$$

les coefficients de la matrice sont donc des sommes de Newton.

- (3) Dans le cas où $n = 3$, on a donc

$$\det {}^t M M = V(X_1, X_2, X_3)^2 = (X_2 - X_1)^2 (X_3 - X_1)^2 (X_3 - X_2)^2 = \begin{vmatrix} 3 & S_1 & S_2 \\ S_1 & S_2 & S_3 \\ S_2 & S_3 & S_4 \end{vmatrix}$$

Si on applique cette relations aux racines du polynôme $X^3 + pX + q$, on commence par observer que

$$\begin{aligned}S_1 &= \sigma_1 = 0 \\S_2 &= (\sigma_1)^2 - 2\sigma_2 = -2p \\S_3 &= -pS_1 - 3q = -3q \\S_4 &= -pS_2 - 3qS_1 = 2p^2\end{aligned}$$

et donc

$$\Delta = (x_2 - x_1)^2(x_3 - x_1)^2(x_3 - x_1)^2 = \begin{vmatrix} 3 & 0 & -2p \\ 0 & -2p & -3q \\ -2p & -3q & 2p^2 \end{vmatrix} = -4p^3 - 27q^2$$

c'est ce qu'on appelle le **discriminant** de l'équation. Il est nul si et seulement si il y a une racine multiple, mais il a d'autres propriétés intéressantes...

Chapitre 4

Algèbre linéaire

Dans ce chapitre, on va donner une réponse « définitive » à la question : quand deux matrices sont-elles semblables ? Dans le cas du corps des nombres complexes, la réponse sera : elles ont même réduite de Jordan. Nous donnerons également quelques indications sur une autre méthode, la réduction de Frobenius. Signalons que notre approche sera souvent « géométrique », en ce sens que nous utiliserons plutôt des propriétés des endomorphismes que des calculs matriciels.

4.1 DIAGONALISATION ET TRIGONALISATION

4.1.1. Polynôme caractéristique et polynôme minimal

Commençons par des rappels rapides. Dans tout ce qui suit K est un corps commutatif et E est un K -espace vectoriel de dimension finie, l'ensemble des endomorphismes de E est noté $\mathcal{L}(E)$. Il est muni d'une structure de K -algèbre. Si on choisit une base $\mathcal{B}$ de E , il y a un isomorphisme

$$\begin{aligned}\mathcal{L}(E) &\longrightarrow \mathcal{M}_n(K) \\ u &\longmapsto M_{\mathcal{B}}(u)\end{aligned}$$

de l'algèbre $\mathcal{L}(E)$ dans l'algèbre des matrices carrées $\mathcal{M}_n(K)$. Rappelons également qu'un scalaire λ de K est **valeur propre** de l'endomorphisme u s'il existe un vecteur x non nul tel que

$$u(x) = \lambda x$$

On dit alors que x est un **vecteur propre** associé à la valeur propre λ .

On notera $\text{Spec}(u)$ (lire spectre de u) l'ensemble des valeurs propres de u et E_λ désignera

$$E_\lambda = \text{Ker}(u - \lambda \text{id}_E) = \{x \in E \mid u(x) = \lambda x\}$$

il contient les vecteurs propres associés à la valeur propre λ ainsi que le vecteur nul.

Définition 4.1. Soit $u \in \mathcal{L}(E)$.

- $\chi_u(X) = \det(u - X \text{id}_E)$ est un polynôme de degré n en X , appelé **polynôme caractéristique** de u .
- L'application définie par

$$\begin{aligned} \phi_u : K[X] &\longrightarrow \mathcal{L}(E) \\ P(X) = \sum_i a_i X^i &\longmapsto P(u) = \sum_i a_i u^i \end{aligned}$$

est un morphisme de K -algèbres. Le générateur unitaire de son noyau s'appelle **polynôme minimal** de u , on le note $\mu_u(X)$.

Remarques

- Commençons par observer que pour définir le polynôme caractéristique, on utilise une matrice contenant l'indéterminée X . On peut convenir qu'on travaille dans $\mathcal{M}_n(K(X))$ où $K(X)$ est le corps des fractions rationnelles à coefficients dans K .
- Un point important également : en dimension finie le polynôme minimal n'est pas nul. Il suffit d'appliquer le théorème d'isomorphisme : l'image de ϕ_u , que l'on note $K[u]$ est incluse dans $\mathcal{L}(E)$ donc est de dimension finie. Elle est isomorphe au quotient $K[X]/(\mu_u(X))$, donc l'idéal engendré par μ_u n'est pas nul. Cet idéal est formé des polynômes P pour lesquels $P(u) = 0$, on les appelle **polynômes annulateurs**. En dimension infinie, un endomorphisme peut avoir un polynôme minimal nul.
- $K[u]$ est donc une sous-algèbre commutative de $\mathcal{L}(E)$. Comme elle est isomorphe à $K[X]/(\mu_u(X))$, c'est un espace vectoriel de dimension $\deg \mu_u$.

Les premières propriétés de ces polynômes sont réunies dans ce théorème :

Théorème 4.2.

- Les racines du polynôme caractéristique sont les valeurs propres de u , elles forment le spectre de u .
- Toute valeur propre est racine d'un polynôme annulateur.
- χ_u est un multiple de μ_u (c'est le théorème de Cayley-Hamilton).
- χ_u et μ_u ont les mêmes facteurs irréductibles.

Démonstration.

- On utilise les équivalences :

$$(\exists x \neq 0, u(x) = \lambda x) \iff \text{Ker}(u - \lambda \text{id}) \neq 0 \iff \det(u - \lambda \text{id}) = 0$$

qui prouvent que les valeurs propres sont les racines du polynôme caractéristique.

- Par ailleurs, si $u(x) = \lambda x$, alors pour tout k de $\mathbb{N}$, $u^k(x) = \lambda^k x$ par une récurrence immédiate. On en déduit par linéarité que $P(u)(x) = P(\lambda)x$, et toute valeur propre est forcément racine d'un polynôme annulateur.
- Il faut démontrer que χ_u est dans l'idéal engendré par le polynôme minimal, autrement dit que c'est un polynôme annulateur. Fixons un vecteur non nul x , on va montrer que $\chi_u(u)(x) = 0$. Ce sera vrai pour tout x , et donc le résultat sera démontré. On est en dimension n , la famille $(x, u(x), \dots, u^n(x))$ est liée. Soit k le plus grand des indices i tels que $(x, u(x), \dots, u^i(x))$ soit libre, on a $0 \leq k < n$. Soit alors F le sous-espace engendré par $(x, u(x), \dots, u^k(x))$, (il est de dimension $k+1$) et soit G un supplémentaire de F . On complète $(x, u(x), \dots, u^k(x))$ par une base de G , et dans cette base, la matrice de u est du type :

$$\begin{pmatrix} C & B \\ 0 & A \end{pmatrix}$$

où C est une matrice carrée de taille $(k+1) \times (k+1)$, B une matrice rectangle, A une matrice carrée de taille $(n-k-1) \times (n-k-1)$, 0 représente une matrice bloc nulle. De plus, la matrice C a la forme

$$\begin{pmatrix} 0 & \dots & & 0 & \alpha_0 \\ 1 & 0 & \dots & \vdots & \vdots \\ 0 & 1 & 0 & & \\ \vdots & & \ddots & \ddots & \\ \vdots & & & 1 & 0 & \alpha_{k-1} \\ 0 & 0 & \dots & 0 & 1 & \alpha_k \end{pmatrix}$$

puisque $u(u^i(x)) = u^{i+1}(x)$ pour $i < k$ et que $u^{k+1}(x)$ est de la forme $\sum_{i=0}^k \alpha_i u^i(x)$. On a donc :

$$\chi_u(X) = \det(C - XI) \det(A - XI)$$

Le premier déterminant se calcule par récurrence ou, plus directement, en développant par rapport à la dernière colonne et :

$$\chi_u(X) = (-1)^{k+1} \left(X^{k+1} - \sum \alpha_i X^i \right) \det(A - XI) = G(X)F(X)$$

où $G(X) = (-1)^{k+1} (X^{k+1} - \sum \alpha_i X^i)$ et $F(X) = \det(A - XI)$ mais alors

$$\chi_u(u) = G(u) \circ F(u) = F(u) \circ G(u) \quad \text{et} \quad \chi_u(u)(x) = F(u)(G(u)(x))$$

et

$$G(u)(x) = (-1)^{k+1} \left(u^{k+1}(x) - \sum \alpha_i u^i(x) \right) = 0$$

On a bien $\chi_u(u)(x) = 0$. Comme le même calcul peut être fait pour tout vecteur x , le théorème de Cayley-Hamilton est démontré.

- Commençons par l'observation suivante : si u est un endomorphisme de E , on peut le représenter par sa matrice M dans une base de E . Le polynôme caractéristique de u est $\det(M - \lambda I)$. Si on considère un corps L contenant K , ce polynôme ne change pas lorsqu'on considère la matrice M comme élément de $\mathcal{M}_n(L)$.

Par ailleurs, le polynôme minimal de u peut être vu comme le polynôme unitaire de plus bas degré tel que $\mu_u(M) = 0$. Il n'est pas alors évident que ce polynôme ne change pas si on considère la matrice M comme élément de $\mathcal{M}_n(L)$. Appelons Q le polynôme unitaire de plus bas degré de $L[X]$ tel que $Q(M) = 0$. Alors Q divise μ_u , car Q est le générateur de l'idéal des polynômes annulateurs de $L[X]$. Mais le degré de Q est aussi le rang du système $I, M, \dots, M^{n-1}$, dans l'espace vectoriel $\mathcal{M}_n(L)$, le degré de μ_u est le rang de ce même système dans $\mathcal{M}_n(K)$. Si on utilise la base canonique, ce rang se calcule comme taille maximal d'un déterminant extrait non nul, et ce calcul est le même que les coefficients soient vus comme éléments de K ou comme éléments de L . On en déduit que le polynôme minimal ne change pas si on « agrandit » le corps de base.

La démonstration du théorème suit : on se place dans le cas d'un corps algébriquement clos ; μ_u et χ_u sont scindés, et μ_u divise χ_u . L'ensemble des racines de μ_u est inclus dans l'ensemble des valeurs propres, mais comme toute valeur propre doit être racine du polynôme annulateur μ_u , ces deux ensembles coïncident, on peut écrire :

$$\mu_u(X) = \prod_{i=1}^k (X - \lambda_i)^{\beta_i} \quad \text{et} \quad \chi_u(X) = (-1)^n \prod_{i=1}^k (X - \lambda_i)^{\alpha_i}$$

où $1 \leq \beta_i \leq \alpha_i$ pour tout i . Les deux polynômes ont mêmes facteurs irréductibles. Dans le cas général, on peut montrer que de même les deux polynômes ont mêmes facteurs irréductibles.

□

Remarque : Une matrice de la forme de C s'appelle **matrice compagnon**, elle « accompagne » le polynôme $P(X) = X^k - \sum_{i=0}^{k-1} \alpha_i X^i$. On la notera $\mathcal{C}(P)$.

4.1.2. Diagonalisation

Si λ_i est une valeur propre de u , on a noté $E_{\lambda_i} = \text{Ker}(u - \lambda_i \text{id})$ le sous-espace propre correspondant. On va montrer que les sous-espaces propres sont en somme directe, et on rappelle que u est dit **diagonalisable** lorsque cette somme directe est égale à E .

Donnons une généralisation de cette définition.

Définition 4.3. Soit u un endomorphisme de E et $k \geq 2$. S'il existe k sous-espaces E_i non nuls stables par u et tels que :

$$E = E_1 \oplus E_2 \oplus \dots \oplus E_k$$

on dit que u est diagonalisable par blocs. Si on note u_i la restriction de u à E_i , on écrira :

$$u = u_1 \oplus u_2 \oplus \dots \oplus u_k$$

Ce vocabulaire se justifie facilement ; si $\mathcal{B}_i$ est une base de E_i , alors la succession des $\mathcal{B}_i$ est une base $\mathcal{B}$ de E , et la matrice de u dans cette base est diagonale par blocs, de la forme

$$M_{\mathcal{B}}(u) = \begin{pmatrix} M_1 & & & \\ & M_2 & & \\ & & \ddots & \\ & & & M_k \end{pmatrix} = \text{diag}(M_1, M_2, \dots, M_k)$$

Le reste de la matrice est constitué de blocs nuls. Revenons au cas ordinaire : un endomorphisme u est donc diagonalisable s'il est somme directe d'homothéties.

Proposition 4.4. Soient $\lambda_1, \lambda_2, \dots, \lambda_k$ des valeurs propres distinctes d'un endomorphisme u , alors les sous-espaces propres sont en somme directe :

$$\sum_{i=1}^k E_{\lambda_i} = \bigoplus_{i=1}^k E_{\lambda_i}$$

On dit parfois que ces sous-espaces propres sont **indépendants**. Rappelons que, par définition cela signifie que dans la somme, la décomposition d'un vecteur en somme de vecteurs propres est unique. Un critère pour cela est que, pour tout i :

$$E_{\lambda_i} \cap \sum_{j \neq i} E_{\lambda_j} = \{0\}$$

Démonstration. Procédons par récurrence sur k . Si $k = 2$, soit $x \in E_{\lambda_1} \cap E_{\lambda_2}$. Alors $u(x) = \lambda_1 x = \lambda_2 x$ et donc $x = 0$.

Supposons la propriété vraie pour $k - 1$ valeurs propres distinctes et prenons $x \in E_{\lambda_i} \cap \sum_{j \neq i} E_{\lambda_j}$. Alors si $x = \sum_{j \neq i} x_j$, en appliquant u on trouve :

$$u(x) = \lambda_i x = \lambda_i \sum_{j \neq i} x_j \quad \text{mais aussi} \quad u(x) = \sum_{j \neq i} u(x_j) = \sum_{j \neq i} \lambda_j x_j$$

Par soustraction,

$$\sum_{j \neq i} (\lambda_i - \lambda_j) x_j = 0$$

Par hypothèse de récurrence et définition de la somme directe, tous les x_j sont nuls, et x aussi. $\square$

Nous savons que les valeurs propres sont les racines du polynôme caractéristique ; on va préciser ce qu'on peut dire a priori de la dimension des sous-espaces propres et énoncer un premier critère de diagonalisabilité.

Théorème 4.5. Soit $u \in \mathcal{L}(E)$ et χ_u son polynôme caractéristique. Alors :

(i)

$$\lambda \text{ est racine de } \chi_u \text{ de multiplicité } m_\lambda \Rightarrow \dim(E_\lambda) \leq m_\lambda$$

(ii)

$$u \text{ diagonalisable} \iff (\chi_u \text{ scindé sur } K \text{ et } m_\lambda = \dim(E_\lambda) \text{ pour tout } \lambda \in \text{Spec}(u))$$

On rappelle qu'un polynôme est scindé s'il est factorisable en polynômes du premier degré (distincts ou non), et chaque racine a alors une multiplicité.

Démonstration.

(i) Supposons que la dimension de E_λ est k ; on peut alors considérer une de ses bases $e_1, e_2, \dots, e_k$ et la compléter en une base $\mathcal{B}$ de E . La matrice de u dans cette base s'écrit alors :

$$\begin{pmatrix} \lambda I_k & A_1 \\ 0 & A_2 \end{pmatrix}$$

où I_k est la matrice de l'identité de $\text{vect}(e_1, \dots, e_k)$ et A_1, A_2 des matrices. Le polynôme caractéristique de u s'écrit : $\chi_u(X) = (\lambda - X)^k \chi_{A_2}(X)$, ce qui prouve que $k \leq m_\lambda$.

(ii) Pour l'implication $\Rightarrow$, il suffit d'écrire la matrice de u dans une base de diagonalisation : le nombre d'occurrences de chaque valeur propre est à la fois la dimension du sous-espace propre associé et la multiplicité de la valeur propre dans le polynôme caractéristique.

Pour l'autre implication, on sait que

$$\sum_{\lambda \in \text{Spec}(u)} E_\lambda = \bigoplus_{\lambda \in \text{Spec}(u)} E_\lambda$$

donc la dimension de $\bigoplus_{\lambda \in \text{Spec}(u)} E_\lambda$ est la somme des m_λ . À cause de l'hypothèse sur le polynôme caractéristique cette somme est égale à n , dimension de E . On a montré que E est somme directe de sous-espaces propres.

□

Corollaire 4.6. Si χ_u est un polynôme ayant n racines distinctes, où n est la dimension de E , alors u est diagonalisable.

Démonstration. La dimension d'un sous-espace propre est inférieure ou égale à 1, donc est égale à 1 car elle est strictement positive. Il y a n sous-espaces propres indépendants et de dimension 1, leur somme directe est donc E tout entier. □

Ce cas est « très » fréquent, dans le cas d'un corps algébriquement clos comme $\mathbb{C}$. On peut en effet vérifier que l'ensemble des matrices ayant n valeurs propres distinctes est dense dans l'ensemble des matrices.

Remarque : Sur $\mathbb{R}$, le polynôme caractéristique peut ne pas être scindé (cf. $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$) et l'endomorphisme ne sera pas diagonalisable, mais il existe des endomorphismes dont le polynôme caractéristique est scindé et qui ne sont pas diagonalisables ; l'exemple le plus simple est $\begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$: il y a une valeur propre a de multiplicité 2 mais le sous-espace propre associé est de dimension 1.

Un premier résultat non immédiat est que la propriété de diagonalisabilité passe aux restrictions.

Théorème 4.7. *Soit $u \in \mathcal{L}(E)$ diagonalisable et F un sous-espace vectoriel stable par u . Alors la restriction de u à F est diagonalisable.*

Démonstration. On va montrer que $F = \bigoplus_{i=1}^k (E_{\lambda_i} \cap F)$ où les λ_i sont les valeurs propres. On peut d'abord remarquer que la somme est bien directe, avec le critère rappelé ci-dessus, et cette somme est incluse dans F . Prenons $x \in F$, décomposons-le dans la somme directe $E = \bigoplus_{i=1}^k E_{\lambda_i}$ et appliquons $k-1$ fois l'endomorphisme u . On obtient

$$\begin{aligned} x &= x_1 + x_2 + \cdots + x_k \\ u(x) &= \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_k x_k \\ u^2(x) &= \lambda_1^2 x_1 + \lambda_2^2 x_2 + \cdots + \lambda_k^2 x_k \\ &\vdots \\ u^{k-1}(x) &= \lambda_1^{k-1} x_1 + \lambda_2^{k-1} x_2 + \cdots + \lambda_k^{k-1} x_k \end{aligned}$$

Ce système a un déterminant non nul (déterminant de Vandermonde), on peut donc le résoudre et exprimer chacun des x_i comme combinaison linéaire de $u^i(x)$. Comme x est dans F et que F est u -stable, les x_i sont dans F : on a démontré l'autre inclusion, u restreint à F est diagonalisable. Attention, certains des ces sous-espaces peuvent être réduits au nul, cela n'a pas d'importance dans la démonstration. $\square$

Encore un théorème sur les sous-espaces propres.

Théorème 4.8. *Soient deux éléments u et v de $\mathcal{L}(E)$ qui commutent. Alors tout sous-espace propre de u est v -stable.*

Démonstration. Soit E_λ un sous-espace propre de u , et x un vecteur de E_λ :

$$u(v(x)) = v(u(x)) = v(\lambda x) = \lambda v(x)$$

et cette égalité prouve que $v(x)$ est soit nul, soit vecteur propre pour u (avec la même valeur propre), il est donc dans le sous-espace E_λ . $\square$

On peut se demander si un sous-espace u -stable quelconque est aussi toujours v -stable (quand u et v commutent) : prendre $u = \text{id}_E$ pour un contre-exemple.

4.1.3. Trigonalisation

Définition 4.9. Soit u un endomorphisme de E , espace vectoriel de dimension n et k un entier tel que $2 \leq k \leq n$. On dit que u est **trigonalisable par blocs** s'il existe une suite

$$\{0\} \subsetneq F_1 \subsetneq F_2 \subsetneq \dots \subsetneq F_k = E$$

formée de sous-espaces vectoriels stables par u .

Si $k = n$, on dit que cette suite est un drapeau de E et que u est **trigonalisable**.

Expliquons d'abord le dernier cas : si la suite contient n espaces vectoriels, c'est que chaque F_i est de dimension i , puisque la suite des dimensions est une suite strictement croissante d'entiers. On peut construire une base $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ telle que, pour tout i , $F_i = \text{vect}(e_1, \dots, e_i)$; si M est la matrice de u dans la base $\mathcal{B}$, elle est triangulaire supérieure, puisque $u(e_i) \in \text{vect}(e_1, \dots, e_i)$. Dans le cas général, si on construit une base $\mathcal{B}$ telle que

$$\text{vect}(e_1, \dots, e_{k_i}) = F_i$$

alors la matrice de u dans cette base sera de la forme

$$\begin{pmatrix} M_1 & * & * & * \\ 0 & M_2 & * & \\ \vdots & \ddots & \ddots & \\ 0 & \dots & 0 & M_k \end{pmatrix}$$

où les M_i sont des matrices carrées les $*$ représentent des matrices quelconques.

Le résultat important est le suivant.

Théorème 4.10. Soit E un K -espace vectoriel de dimension finie. Un endomorphisme u de E est trigonalisable si et seulement si le polynôme caractéristique est scindé.

Démonstration.

- Supposons u trigonalisable. Alors sa matrice dans une base $\mathcal{B}$ de trigonalisation est triangulaire supérieure et le polynôme caractéristique qui s'écrit :

$$\chi_u(X) = \det \begin{pmatrix} \lambda_1 - X & * & * & * \\ 0 & \lambda_2 - X & * & \\ \vdots & \ddots & \ddots & \\ 0 & \dots & 0 & \lambda_n - X \end{pmatrix} = \prod_{i=1}^n (\lambda_i - X)$$

est scindé.

- L'hypothèse est que le polynôme caractéristique est scindé. On procède par récurrence sur n . Si la dimension est 1, la matrice est triangulaire supérieure (!). Supposons que la propriété du théorème soit vraie en dimension $n - 1$ et soit λ une racine du polynôme caractéristique ; il lui correspond un vecteur propre e_1 . Si on complète e_1 en une base $\mathcal{B} = (e_1, e_2, \dots, e_n)$ de E , la matrice de u est de la forme :

$$\begin{pmatrix} \lambda & U \\ 0 & M \end{pmatrix}$$

et $\chi_u(X) = (\lambda - X)\chi_M(t)$ prouve que M a un polynôme caractéristique scindé. En utilisant (hypothèse de récurrence) une base de trigonalisation pour M , dans $\text{vect}(e_2, \dots, e_n)$ on obtient une trigonalisation de la matrice de f . On peut être plus précis, en écrivant que :

$$\begin{pmatrix} 1 & 0 \\ 0 & P^{-1} \end{pmatrix} \begin{pmatrix} \lambda & L \\ 0 & M \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & P \end{pmatrix} = \begin{pmatrix} \lambda & LP \\ 0 & P^{-1}MP \end{pmatrix}$$

□

Exercice 4.1. Déterminer les endomorphismes u qui commutent avec tous les autres, c'est ce qu'on appelle le **centre** de l'algèbre $\mathcal{L}(E)$.

Exercice 4.2. Soit $A = (a_{ij})$ une matrice de $\mathcal{M}_n(\mathbb{C})$. Montrer que l'on a

$$\text{Spec}(A) \subset \bigcup_{i=1}^n \overline{B}(a_{ii}, \sum_{j \neq i} |a_{ij}|)$$

En déduire que si, pour tout i de 1 à n on a :

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|$$

alors la matrice A est inversible. C'est le théorème de Hadamard.

Exercice 4.3. Soient u et v deux endomorphismes de E , espace vectoriel de dimension finie. Démontrer que $\chi_{u \circ v} = \chi_{v \circ u}$. On pourra commencer par le cas où l'un des deux est inversible, et utiliser les matrices A et B de u et v dans une base. Montrer par un contre-exemple qu'on n'a pas, en général, $\mu_{u \circ v} = \mu_{v \circ u}$.

Exercice 4.4. Soit u un endomorphisme et $E = E_1 \oplus E_2$ une décomposition de E en sous-espaces supplémentaires qui sont u -stables. On note u_1 et u_2 les restrictions de u à E_1 et E_2 . Montrer que :

$$\chi_u(X) = \chi_{u_1}(X)\chi_{u_2}(X) \quad \text{et} \quad \mu_u(X) = \text{p.p.c.m.}(\mu_{u_1}(X), \mu_{u_2}(X))$$

Exercice 4.5. Montrer que si u et v sont deux endomorphismes de E qui commutent et sont diagonalisables, alors ils sont simultanément diagonalisables : c'est-à-dire qu'il existe une base commune de diagonalisation. Même question avec trigonalisable à la place de diagonalisable.

4.2 DÉCOMPOSITION DE DUNFORD ET RÉDUCTION DE JORDAN

4.2.1. Le lemme des noyaux

Dans ce paragraphe, nous allons proposer des «réductions» d'endomorphismes, et donc écrire des endomorphismes sous forme de sommes directes de restrictions à des sous-espaces stables. Commençons par observer le petit résultat suivant :

Proposition 4.11. Soit $u \in \mathcal{L}(E)$ et $P \in K[X]$. Alors $\text{Ker } P(u)$ et $\text{Im } P(u)$ sont des sous-espaces u -stables.

Démonstration. Si $x \in \text{Ker } P(u)$, alors

$$P(u) \circ u(x) = u \circ P(u)(x) = u(0) = 0$$

donc $u(x) \in \text{Ker } P(u)$. De même, si $x \in \text{Im } P(u)$, il existe y tel que $x = P(u)(y)$ et

$$u(x) = u \circ P(u)(y) = P(u) \circ u(y) \in \text{Im } P(u)$$

□

Remarquons qu'il peut exister des sous-espaces u -stables qui ne sont pas de cette forme, prendre par exemple $u = \text{id}_E$.

Passons maintenant au lemme des noyaux ; c'est une tradition de parler de «lemme», alors que le résultat que nous allons énoncer mériterait certainement d'être promu au rang de théorème...

Lemme 4.12. des noyaux. Soit $u \in \mathcal{L}(E)$ et $(P_i)_{i=1}^k$ une suite de polynômes de $K[X]$. On suppose que les P_i sont premiers entre eux **deux à deux**. Alors :

$$\text{Ker} \left(\prod_{i=1}^k P_i \right) (u) = \bigoplus_{i=1}^k \text{Ker } P_i(u)$$

En particulier, si $P = \prod_{i=1}^k P_i$ est un polynôme annulateur

$$E = \bigoplus_{i=1}^k \text{Ker } P_i(u)$$

Démonstration. On définit des polynômes Q_i par

$$Q_i(X) = \prod_{j \neq i} P_j(X)$$

et par l'hypothèse, les polynômes Q_i sont premiers entre eux dans leur ensemble. Il existe donc des polynômes A_i tels que $\sum_{i=1}^k A_i Q_i = 1$, et donc

$$\sum_{i=1}^k A_i(u) \circ Q_i(u) = \text{id}$$

Si un vecteur est dans l'ensemble $\sum_{i=1}^k \text{Ker } P_i(u)$, on a directement, en calculant son image par $\prod_{i=1}^k P_i(u)$ qu'il est dans $\text{Ker} \left(\prod_{i=1}^k P_i \right) (u)$. Réciproquement, soit x un vecteur de $\text{Ker} \left(\prod_{i=1}^k P_i \right) (u)$. D'après la relation précédente, on peut écrire

$$x = \sum_{i=1}^k A_i(u) \circ Q_i(u)(x) = \sum_{i=1}^k x_i \quad \text{où} \quad x_i = A_i(u) \circ Q_i(u)(x)$$

Mais alors, pour un entier j fixé quelconque,

$$P_j(u)(x_j) = P_j(u) \circ A_j(u) \circ Q_j(u)(x) = A_j(u) \circ \prod_{i=1}^k P_i(u)(x) = 0$$

et donc $x_j \in \text{Ker } P_j(u)$. On a donc montré que $\text{Ker} \left(\prod_{i=1}^k P_i \right) (u)$ est égal à $\sum_{i=1}^k \text{Ker } P_i(u)$, il reste à montrer que cette somme est directe.

Soit en effet une égalité $x_1 + \dots + x_k = 0$ où les x_i sont respectivement dans $\text{Ker } P_i(u)$. D'après la relation de Bézout,

$$\begin{aligned} x_1 &= \sum_{i=1}^k A_i(u) \circ Q_i(u)(x_1) \\ &= A_1(u) \circ Q_1(u)(x_1) \\ &= A_1(u) \circ Q_1(u) \left(- \sum_{i=2}^k x_i \right) \\ &= - \sum_{i \geq 2}^k A_1(u) \circ Q_1(u)(x_i) \\ &= 0 \end{aligned}$$

car $Q_1(u)(x_i) = 0$ pour tous les i supérieurs ou égal à 2. Ce calcul peut se faire pour tous les autres indices, on a démontré que la somme est directe. $\square$

Donnons maintenant un nouveau critère de diagonalisabilité, qui utilise les polynômes annulateurs. Ce critère se révèle être redoutablement efficace.

Théorème 4.13.

- u est diagonalisable si et seulement si il existe $P \in K[X]$ annulateur pour u et scindé à racines simples.
- u est diagonalisable si et seulement si μ_u est scindé à racines simples.

Dans cet énoncé, **toutes** les racines doivent être simples, on dira, pour aller plus vite, que le polynôme est scindé simple. C'est le cas par exemple des deux polynômes $X^2 - 1$, $X^2 - X$, qui annulent respectivement les symétries et les projections. Toute symétrie, toute projection est donc diagonalisable.

Démonstration. [du théorème] Commençons par la première équivalence :

$\Rightarrow$ Si $\lambda_1, \dots, \lambda_k$ sont les valeurs propres distinctes l'hypothèse u diagonalisable donne $E = \bigoplus_{i=1}^k E_{\lambda_i}$ et si on pose $P(X) = \prod_{i=1}^k (X - \lambda_i)$, alors $P(u)(x) = 0$ car en décomposant x dans la somme directe, toutes ses composantes sont annulées par $P(u)$. On a donc un polynôme annulateur, qui par construction est scindé simple.

$\Leftarrow$ Si $P(X) = \prod_{i=1}^k (X - \lambda_i)$ où les λ_i sont deux à deux distincts, est un polynôme annulateur pour u , alors $\text{Ker } P(u) = E = \bigoplus_{i=1}^k \text{Ker}(u - \lambda_i \text{id})$ par application du lemme des noyaux.

Passons à la seconde équivalence :

$\Rightarrow$ Si u est diagonalisable, il existe un polynôme scindé simple annulant u , ce polynôme est dans l'idéal engendré par le polynôme minimal, qui est donc scindé simple.

$\Leftarrow$ Si le polynôme minimal est scindé simple, c'est un polynôme annulateur de u , qui est donc diagonalisable.

□

Remarque : On peut maintenant faire une nouvelle démonstration du théorème 4.7 p. 91 : si u est diagonalisable et si μ_u est son polynôme minimal, alors toute restriction de u a pour polynôme minimal un diviseur de μ_u , qui est donc également scindé simple, toute restriction de u est diagonalisable.

4.2.2. Sous-espaces caractéristiques

Revenons au polynôme caractéristique, notamment dans le cas où il est scindé mais non forcément scindé simple.

Théorème 4.14. Soit $u \in \mathcal{L}(E)$, E étant de dimension finie. Supposons que $\chi_u(X)$ soit scindé, de la forme :

$$\chi_u(X) = \prod_{i=1}^k (X - \lambda_i)^{\alpha_i}$$

où les λ_i sont deux à deux distincts ;

alors :

- $E = \bigoplus_{i=1}^k \text{Ker}(u - \lambda_i \text{id})^{\alpha_i}$
- Les $\text{Ker}(u - \lambda_i \text{id})^{\alpha_i}$ sont u -stables, de dimension α_i , et si on note u_i la restriction de u à ces sous-espaces, on a :

$$\chi_{u_i}(X) = (-1)^{\alpha_i} (X - \lambda_i)^{\alpha_i}$$

Les sous-espaces $\text{Ker}(u - \lambda_i \text{id})^{\alpha_i}$ s'appellent les **sous-espaces caractéristiques** de u . À titre d'abréviation, on n'hésitera pas à parler de s.e.c.

Démonstration.

- Par le théorème de Cayley-Hamilton, le polynôme caractéristique de u est annulateur. En appliquant le lemme des noyaux, et compte-tenu de ce que les $(X - \lambda_i)^{\alpha_i}$ sont premiers entre eux deux à deux, on obtient le résultat souhaité.
- La stabilité est immédiate, car un s.e.c. est de la forme $\text{Ker } P(u)$. On peut ensuite écrire la matrice de u sous forme diagonale par blocs, et on a alors $\chi_u = \prod_{i=1}^k \chi_{u_i}$. Pour montrer la fin du théorème, on peut se placer dans le cadre suivant : si $\chi_u = P_1 P_2$ où $P_1 \wedge P_2 = 1$, en notant u_i la restriction de u à $\text{Ker } P_i(u)$, on va montrer que $\chi_{u_i} = P_i$ (à un scalaire multiplicatif près), et une récurrence immédiate donnera alors le résultat.

On a donc $E = \text{Ker } P_1(u) \oplus \text{Ker } P_2(u)$, décomposition en sous-espaces u -stables, donc $\chi_u(X) = P_1(X)P_2(X) = \chi_{u_1}(X)\chi_{u_2}(X)$ en utilisant une matrice diagonale par blocs. Soit maintenant μ_{u_1} le polynôme minimal de u_1 . Alors $P_1(u_1) = 0$ implique $\mu_{u_1} \mid P_1$, et donc $\mu_{u_1} \wedge P_2 = 1$. Mais alors $\chi_{u_1} \wedge P_2 = 1$ puisque le polynôme minimal et le polynôme caractéristique ont mêmes facteurs irréductibles. Le lemme de Gauss permet alors d'affirmer que $\chi_{u_1} \mid P_1$. De même $\chi_{u_2} \mid P_2$. Mais comme $P_1 P_2 = \chi_{u_1} \chi_{u_2}$, on conclut

$$\chi_{u_1}(X) = P_1(X) \quad \text{et} \quad \chi_{u_2}(X) = P_2(X)$$

(à un scalaire près).

□

4.2.3. Décomposition de Dunford

Il est maintenant possible de donner un théorème important de décomposition, qui offre de nombreuses applications théoriques et pratiques, c'est la décomposition de Dunford.

Théorème 4.15. *Pour tout $u \in \mathcal{L}(E)$ dont le polynôme caractéristique est scindé, il existe un **unique** couple (d, n) d'endomorphismes de E tel que :*

$$u = d + n, \quad n \circ d = d \circ n, \quad n \text{ est nilpotent, } d \text{ est diagonalisable.}$$

Bien sûr, l'unicité affirmée dans le théorème laisse entendre que si u est diagonalisable, on aura $n = 0$ et si u est nilpotent on aura $d = 0$.

Démonstration.**(1) Existence**

On écrit le polynôme caractéristique $\chi_u(X) = \prod_{i=1}^k (X - \lambda_i)^{\alpha_i}$ où les λ_i sont distincts, et E est donc somme des sous-espaces caractéristiques $C_i = \text{Ker}(u - \lambda_i \text{id})^{\alpha_i}$. Pour tout i , soit u_i la restriction de u à C_i et $n_i = u_i - \lambda_i \text{id}_{C_i}$. Soit alors

$$d = \bigoplus \lambda_i \text{id}_{C_i}$$

et soit de même n la « somme directe » des n_i . Rappelons, pour être plus précis, que d et n sont définis par :

$$n(x) = \sum_{i=1}^k n_i(x_i), \quad d(x) = \sum_{i=1}^k \lambda_i x_i$$

où $x = \sum_{i=1}^k x_i$ est la décomposition de x dans la somme directe des sous-espaces caractéristiques. Mais alors d est diagonalisable comme somme directe d'homothéties sur les sous-espaces caractéristiques qui sont supplémentaires. De plus n est nilpotent. En effet, on a

$$n^2(x) = \sum_{i=1}^k n_i^2(x_i)$$

puisque $n(x) = \sum_i n_i(x_i)$ est l'écriture de $n(x)$ dans la somme directe. On peut calculer de même n^ℓ et on en déduit que n est nilpotent.

Enfin

$$n \circ d(x) = \sum_{i=1}^k n_i(\lambda_i x_i) = \sum_{i=1}^k \lambda_i n_i(x_i) = d \circ n(x)$$

(2) Unicité

Soit $u = d' + n'$ une autre décomposition, où d' commute avec n' . Alors d' commute avec $u = d' + n'$, donc avec $(u - \lambda_i \text{id})^{\alpha_i}$ et

$$d'(C_i) = d'(\text{Ker}(u - \lambda_i \text{id})^{\alpha_i}) \subset C_i$$

il suffit de l'écrire. Mais comme d' est diagonalisable, d' est diagonalisable sur le sous-espace C_i qu'il stabilise. Soit alors x un vecteur propre pour d' pris dans C_i , et λ la valeur propre associée :

$$d'(x) = \lambda x \Rightarrow u(x) = (d' + n')(x) = \lambda x + n'(x)$$

Si on prend l'entier ℓ le plus grand tel que $n'^\ell(x) \neq 0$, alors :

$$u \circ n'^\ell(x) = n'^\ell \circ u(x) = n'^\ell \circ (d' + n')(x) = n'^\ell \circ d'(x) = \lambda n'^\ell(x)$$

et donc λ est valeur propre de u . Mais comme C_i est stable par u et d' , il est stable

par n' , donc le vecteur $n'^{\ell}(x)$ est dans $C_i = \text{Ker}(u - \lambda_i)^{\alpha_i}$ et la valeur propre λ est λ_i (voir le polynôme caractéristique de u restreint à C_i). On en conclut que $d' = d$ et donc $n' = n$.

□

La décomposition de Dunford a de nombreuses applications pratiques : par exemple le calcul de $u^k = (d + n)^k$ est aisé, puisque

- On peut appliquer la formule du binôme de Newton car d et n commutent.
- Les puissances de d se calculent facilement dans une base de diagonalisation.
- Les puissances de l'endomorphisme nilpotent n sont rapidement nulles...

On peut également calculer l'exponentielle de u , et donc, par exemple, résoudre des systèmes différentiels à coefficients constants.

4.2.4. Réduite de Jordan

Nous allons maintenant utiliser la décomposition de Dunford pour obtenir une nouvelle réduction des endomorphismes dont le polynôme caractéristique est scindé.

Théorème 4.16. Soit $u \in \mathcal{L}(E)$. On suppose que χ_u est scindé. Il existe des bases de E dans laquelle la matrice de u est diagonale par blocs, les blocs diagonaux étant du type :

$$J = (J_{i,j})_{1 \leq i,j \leq p} = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & & \vdots \\ \vdots & & \ddots & \ddots & \\ 0 & \dots & & & \lambda \end{pmatrix}$$

avec donc $J_{i,i} = \lambda$ pour tout i et $J_{i,i+1} = 1$ pour tout $i < p$, les autres coefficients sont nuls. De plus, il y a unicité d'une telle décomposition, à l'ordre près des blocs.

Définition 4.17. Une matrice de la forme donnée dans le théorème est une réduite élémentaire de Jordan. On la note $J(\lambda, p)$ où p est le nombre de lignes et de colonnes. Le nombre $p = 1$ est autorisé. Une matrice diagonale par blocs, dont les blocs sont des réduites élémentaires est appelée réduite de Jordan.

Tout repose sur un lemme, qui envisage le cas d'un endomorphisme nilpotent.

Lemme 4.18. Soit u un endomorphisme nilpotent d'un espace vectoriel de dimension n . Si $u^k = 0$ où $k \in \mathbb{N}^*$ (u est donc non nul), alors il existe une base de E dans laquelle la matrice de u est diagonale par blocs, les blocs étant de la forme $J(0, p)$, avec $p \leq k$.

Démonstration. [du lemme] Supposons que k soit l'ordre de nilpotence (c'est-à-dire que $u^k = 0$, tandis que $u^{k-1} \neq 0$). Pour commencer, remarquons que les « noyaux itérés » $\text{Ker } u^i$ sont emboîtés, et forment une suite croissante. Soit alors S_k un supplémentaire de $\text{Ker } u^{k-1}$. Alors :

$$\begin{cases} u(S_k) \subset \text{Ker } u^{k-1} \\ u(S_k) \cap \text{Ker } u^{k-2} = \{0\} \\ \dim S_k = \dim u(S_k) = \dots = \dim u^{k-1}(S_k) \end{cases}$$

La première inclusion est due à la k -nilpotence de u . Montrons la seconde propriété : Si $x \in u(S_k) \cap \text{Ker } u^{k-2}$, alors il existe $y \in S_k$ tel que $x = u(y)$. Mais alors

$$0 = u^{k-2}(x) = u^{k-1}(y) \quad \text{donc} \quad y \in \text{Ker } u^{k-1} \cap S_k = \{0\}$$

y donc x est nul. Enfin, u restreinte à S_k est bien sûr bijective de S_k sur $u(S_k)$ car S_k ne contient aucun vecteur de $\text{Ker}(u)$, et de même pour les autres sous-espaces.

On a donc dans $\text{Ker } u^{k-1}$, deux sous-espaces supplémentaires, et l'on peut compléter par un sous-espace S_{k-1} (éventuellement nul cette fois) de sorte que :

$$\text{Ker } u^{k-1} = S_{k-1} \oplus u(S_k) \oplus \text{Ker } u^{k-2}$$

À l'étape suivante, on vérifie immédiatement que $u(S_{k-1})$, $u^2(S_k)$ et $\text{Ker } u^{k-3}$ sont dans $\text{Ker } u^{k-2}$. Il y a donc un sous-espace S_{k-2} pour compléter. Par récurrence, on construit donc une suite de sous-espaces vectoriels $S_{k-1}, S_{k-2}, \dots, S_0$ tels que :

$$\text{Ker } u^{i+1} = \left(\bigoplus_{j=0}^{k-i-1} u^j(S_{i+j+1}) \right) \oplus \text{Ker } u^i$$

En définitive, on a :

$$\begin{aligned} E &= S_k \oplus \text{Ker } u^{k-1} \\ \text{Ker } u^{k-1} &= u(S_k) \oplus S_{k-1} \oplus \text{Ker } u^{k-2} \\ \text{Ker } u^{k-2} &= u^2(S_k) \oplus u(S_{k-1}) \oplus S_{k-2} \oplus \text{Ker } u^{k-3} \\ &\dots \\ \text{Ker } u &= u^{k-1}(S_k) \oplus u^{k-2}(S_{k-1}) \oplus \dots \oplus u(S_2) \oplus S_1 \end{aligned}$$

D'où la décomposition de E lui même :

$$E = \bigoplus_{j=1}^k \bigoplus_{i=0}^j u^i(S_j)$$

que l'on peut écrire :

$$\begin{aligned} E &= S_k && \oplus && S_{k-1} && \oplus && \dots \oplus S_2 \oplus S_1 \\ &\oplus u(S_k) && \oplus && u(S_{k-1}) && \oplus && \dots \oplus u(S_2) \\ &\dots && && && && \\ &\oplus u^{k-2}(S_k) && \oplus && u^{k-2}(S_{k-1}) && && \\ &\oplus u^{k-1}(S_k). && && && && \end{aligned}$$

Si maintenant on prend un vecteur e_n non nul de S_k et ses images successives, $e_{n-1} = u(e_n)$, $e_{n-2} = u(e_{n-1})$, $\dots$, $e_{n-k+1} = u(e_{n-k+2})$, alors $(e_{n-k+1}, \dots, e_n)$ est libre et la restriction de u à ce sous-espace a pour matrice $J(0, k)$. Si on continue avec (s'il en existe) un vecteur de S_k indépendant de e_n , on obtiendra un autre sous-espace u -stable, avec une matrice $J(0, k)$. De même, des vecteurs indépendants de S_{k-1} donneront des blocs $J(0, k-1)$.

Remarquons que : la taille maximum d'un bloc de Jordan est k . Dans chaque base correspondant à un bloc, il y a un seul vecteur du noyau de u , et ces bases sont « indépendantes », donc la dimension du noyau est égale au nombre des blocs.

En utilisant des bases des S_i et de leurs images, on obtient une réduction de Jordan. $\square$

Et maintenant, démontrons le théorème.

Démonstration. Si on suppose χ_u scindé, alors d'après le lemme des noyaux, E peut s'écrire :

$$E = \sum_{i=1}^k \text{Ker}(u - \lambda_i \text{id})^{\alpha_i}$$

et dans chacun de ces sous-espaces, la restriction de $u - \lambda_i \text{id}$ est nilpotente. On peut donc « Jordaniser » la matrice de u . On obtiendra des blocs de Jordan de la forme $J(\lambda_i, p)$. $\square$

Voici maintenant un théorème de bilan et d'application de cette réduction de Jordan, que nous énoncerons en terme de matrices.

Théorème 4.19.

- Toute matrice de $\mathcal{M}_n(\mathbb{C})$ est semblable à une réduite de Jordan.
- Si A et B deux matrices de $\mathcal{M}_n(\mathbb{C})$ sont semblables, alors elles ont même polynôme caractéristique et même polynôme minimal.
- Deux matrices A et B de $\mathcal{M}_n(\mathbb{C})$ sont semblables si et seulement si elles ont même réduite de Jordan (à permutation des blocs près).
- Pour toute matrice A de $\mathcal{M}_n(\mathbb{C})$, A et ${}^t A$ sont semblables.

Pour commencer, montrons que la réciproque de la deuxième affirmation est fausse. Pour cela, il faut un contre-exemple de dimension suffisante. Prenons U matrice nilpotente d'ordre 3 et construisons deux matrices blocs A et B de la façon suivante :

$$U = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad A = \begin{pmatrix} U & 0 \\ 0 & U \end{pmatrix} \quad B = \begin{pmatrix} U & 0 \\ 0 & 0 \end{pmatrix}$$

Le polynôme caractéristique de U est X^3 , et c'est aussi le polynôme minimal de U , puisque $U^2 \neq 0$. Quant à A et B , ces matrices ont même polynôme caractéristique (X^6), et même polynôme minimal (X^3), et pourtant elles ne sont pas semblables (car, par exemple, elles n'ont pas le même rang.)

Donnons maintenant la démonstration du théorème :

Démonstration.

- (1) Tout polynôme à coefficients complexes est scindé.
- (2) C'est déjà fait pour le polynôme caractéristique. Et on voit facilement que si P annule une matrice A , alors P annule toute matrice semblable à A .
- (3) En définitive, il s'agit de vérifier une unicité, à permutation des blocs près. Tout repose sur le fait que les blocs sont déterminés par les dimensions des noyaux itérés : leur taille et leur nombre dépendent seulement des propriétés géométriques de l'endomorphisme, non de la matrice qui les représente.
- (4) Il suffit de montrer que c'est vrai pour un bloc ; facile, il suffit d'inverser les vecteurs de base, avec par exemple la matrice de passage

$$\begin{pmatrix} 0 & \dots & \dots & 0 & 1 \\ & & & 1 & \\ & & & & \\ 1 & 0 & \dots & \dots & 0 \end{pmatrix}$$

□

Exercice 4.6. Soit A une matrice de $\mathcal{M}_n(K)$, triangulaire supérieure. On l'écrit sous la forme $A = D + T$, où D est la matrice diagonale formée des éléments diagonaux de A . Montrer que T est nilpotente, mais que cette écriture n'est pas forcément la décomposition de Dunford de A .

Exercice 4.7. Soit E de dimension 2, 3, 4, 5 ou 6 : combien y a-t-il de classes de similitude d'endomorphismes nilpotents ?

4.3 RÉDUCTION DE FROBENIUS

Cette nouvelle réduction s'applique dans tous les cas, même si le polynôme caractéristique n'est pas scindé. Notre théorie pourra donc s'adapter à n'importe quel corps, pas forcément algébriquement clos.

4.3.1. Sous-espace stable et quotient

On est souvent placé devant la situation suivante : F est un sous-espace u -stable de E qui n'admet pas de supplémentaire u -stable. On ne peut donc écrire une diagonalisation par blocs de la matrice de u , mais seulement une trigonalisation par blocs. L'utilisation de sous-espaces vectoriels quotients est adaptée à ce cas.

Si E est un K -espace vectoriel et F un sous-espace de E , on peut définir le quotient E/F issu de la relation d'équivalence $x \equiv y \pmod{F} \iff x - y \in F$. Ce quotient est naturellement muni d'une structure de K -espace vectoriel, si on pose

$$\bar{x} + \bar{y} = \overline{x + y} \quad \text{et} \quad \lambda \bar{x} = \overline{\lambda x}$$

où $\bar{x}$ désigne la classe d'équivalence de x .

Si u est un endomorphisme et si F est u -stable, on peut dire davantage.

Proposition 4.20. *Soit $u \in \mathcal{L}(E)$ et F un sous-espace vectoriel u -stable. Alors il existe un seul endomorphisme $\bar{u}$ de E/F tel que :*

$$\bar{u}(\bar{x}) = \overline{u(x)}$$

De plus, si $\mathcal{B}_1 = (e_1, \dots, e_k)$ est une base de F , que l'on complète en une base $\mathcal{B} = (e_1, \dots, e_n)$ de E , alors $\bar{\mathcal{B}} = (\bar{e}_{k+1}, \dots, \bar{e}_n)$ est une base de E/F . Enfin, la matrice de u s'écrit dans la base $\mathcal{B}$ sous la forme :

$$M_{\mathcal{B}}(u) = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

où

$$A = M_{\mathcal{B}_1}(u) = A \quad \text{et} \quad C = M_{\bar{\mathcal{B}}}(\bar{u})$$

Démonstration. Il faut vérifier que si $\bar{x} = \bar{y}$ alors $\overline{u(x)} = \overline{u(y)}$. C'est le cas puisque F est u -stable et :

$$x - y \in F \Rightarrow u(x) - u(y) = u(x - y) \in F$$

La linéarité se montre sans difficulté. Si x est un vecteur quelconque,

$$x = \sum_{i=1}^n x_i e_i \quad \text{donc} \quad \bar{x} = \sum_{i=1}^n x_i \bar{e}_i = \sum_{i=k+1}^n x_i \bar{e}_i$$

De plus, si cette somme est nulle, le vecteur $\sum_{i>k} x_i e_i$ est dans F et dans $\text{vect}(e_{k+1}, \dots, e_n)$, il est donc nul. Enfin, l'affirmation concernant la matrice A est due à la stabilité de F , l'affirmation concernant C résulte de

$$u(e_j) = \sum_{i=1}^n m_{i,j} e_i \quad \text{donc} \quad \bar{u}(\bar{e}_j) = \sum_{i=1}^n m_{i,j} \bar{e}_i = \sum_{i=k+1}^n m_{i,j} \bar{e}_i$$

□

4.3.2. Sous-espaces cycliques

Soit u un endomorphisme de E , K -espace vectoriel de dimension finie. On va étudier une décomposition de E en sous-espaces stables par u , différente de celle vue dans le paragraphe précédent, et qui permet cependant de « classer » les endomorphismes.

Si $a \in E$ est un vecteur quelconque, l'application $\phi_{u,a}$

$$\begin{aligned} K[X] &\longrightarrow E \\ P(X) = \sum_i a_i X^i &\longmapsto P(u)(a) = \sum_i a_i u^i(a) \end{aligned}$$

est linéaire. Le générateur unitaire de son noyau s'appelle **polynôme minimal de u en a** , on le note $\mu_{u,a}(X)$. Pour donner un exemple élémentaire, si a est un vecteur propre de u associé à la valeur propre λ , son polynôme minimal est $X - \lambda$. Une première propriété s'obtient facilement.

Proposition 4.21. *Pour tout $a \in E$, $\mu_{u,a} \mid \mu_u$.*

Démonstration. En effet, puisque $\mu_u(u) = 0$ (endomorphisme nul), on a $\mu_u(u)(a) = 0$ et le polynôme $\mu_u(X)$ appartient à $\text{Ker } \phi_{u,a}$. $\square$

Le résultat suivant est moins banal, et moins facile à obtenir.

Théorème 4.22. *Pour tout endomorphisme de E , il existe un vecteur $a \in E$ tel que $\mu_u = \mu_{u,a}$.*

Démonstration. Commençons par supposer que μ_u est primaire, c'est-à-dire puissance d'un irréductible P_i : $\mu_u(X) = P_i^k(X)$. Alors, si on prend a dans $E \setminus \text{Ker } P_i^{k-1}(u)$ (qui est non vide, par définition du polynôme minimal), le théorème est satisfait. En effet, $\mu_{u,a}(X)$ est un diviseur du polynôme minimal $P_i^k(X)$ et c'est donc une puissance de P_i puisque P_i est irréductible. Par ailleurs, par définition de a , $P_i^{k-1}(u)(a) \neq 0$, et donc $\mu_{u,a} = P_i^k$.
Si maintenant le polynôme minimal s'écrit

$$\mu_u(X) = \prod_{i=1}^k P_i^{k_i}(X)$$

où les P_i sont irréductibles, on peut choisir a_i dans $\text{Ker } P_i^{k_i}(u) \setminus \text{Ker } P_i^{k_i-1}(u)$, et $a = a_1 + \dots + a_k$ vérifiera $\mu_{u,a} = \mu_u$. $\square$

Soit toujours u un endomorphisme. Si a est un vecteur de E , on appellera **sous-espace cyclique** engendré par a , le sous-espace vectoriel $\text{vect}(a, u(a), \dots, u^k(a), \dots)$. Ce sous-espace sera noté $\langle a \rangle$ s'il n'y a pas risque de confusion. Si $\mu_{u,a}$ est le polynôme minimal de u en a , il permet d'obtenir la plus courte relation de liaison entre les $u^k(a)$. Le degré de ce polynôme minimal est donc la dimension de $\langle a \rangle$, dont une base est $(a, u(a), \dots, u^{k-1}(a))$ où k est le degré de $\mu_{u,a}$. De plus

Proposition 4.23.

- Tout sous-espace cyclique $\langle a \rangle$ est u -stable. Si u_a est la restriction de u à $\langle a \rangle$, alors $\mu_{u_a} = \mu_{u,a}$ et $\chi_{u_a} = \pm \mu_{u,a}$.
- Si F est un sous-espace u -stable tel que $\mu_{u|_F} = \pm \chi_{u|_F}$, alors F est cyclique.
- F est cyclique si et seulement si il existe une base $\mathcal{C}$ de F telle que la matrice de $u|_F$ dans cette base soit une matrice compagnon.

Démonstration.

- Si x est un vecteur de $\langle a \rangle$, il est de la forme $P(u)(a)$ où P est un polynôme. Alors, en posant $R(X) = XP(X)$, on a $u(x) = u \circ P(u)(a) = R(u)(a)$ et $u(x)$ est aussi dans $\langle a \rangle$ qui est ainsi stable par u . Si Q est un polynôme tel que $Q(u_a) = 0$, il vérifie $Q(u) \circ P(u)(a) = 0$ pour tout polynôme P , donc Q est un multiple de $\mu_{u,a}$, et réciproquement. De plus, la dimension de $\langle a \rangle$ est le degré de $\mu_{u,a}$, donc le polynôme caractéristique χ_{u_a} , qui est de même degré et qui est dans l'idéal (μ_{u_a}) , est proportionnel à μ_{u_a} .
- Si F est un sous-espace u -stable et si $\mu_{u|_F}$ est le polynôme minimal de $u|_F$, il existe a de F tel que son polynôme minimal (dans F) soit $\mu_{u|_F}$. Mais alors le sous-espace engendré par les $P(u)(a)$ a même dimension que le degré de $\mu_{u|_F} = \chi_{u|_F}$, donc même dimension que F , c'est F .
- Si $F = \langle a \rangle$, F de dimension k , alors une base de F est $(a, u(a), \dots, u^{k-1}(a))$ et si $\mu_{u,a}(X) = X^k - \sum_{i=1}^{k-1} \alpha_i X^i$ alors la matrice de $u|_F$ est de la forme :

$$C(P) = \begin{pmatrix} 0 & \dots & 0 & \alpha_0 \\ 1 & 0 & \vdots & \vdots \\ 0 & 1 & 0 & \\ \vdots & & & \\ 0 & 0 & \dots & 1 & \alpha_k \end{pmatrix}$$

qui est bien la matrice compagnon (voir la remarque page 88) de $\mu_{u,a}(X)$.

Enfin, la forme même d'une matrice compagnon montre que l'espace engendré par $\mathcal{C} = (e_1, e_2, \dots, e_k)$ est de la forme $\text{vect}(e_1, u(e_1), \dots, u(e_{k-1}))$, donc est cyclique. $\square$

On dit alors que l'endomorphisme u est **cyclique** s'il existe un vecteur a tel que $\langle a \rangle = E$. D'après la proposition précédente, E est cyclique pour u si et seulement si le polynôme minimal de u est égal (au signe près) au polynôme caractéristique de u . On sait que u est cyclique si et seulement si sa matrice est de la forme compagnon dans une certaine base. Donnons un autre exemple.

Proposition 4.24. *Tout endomorphisme dont la matrice dans une base $\mathcal{B}$ est un bloc de Jordan, est cyclique.*

Démonstration. Si $M \in \mathcal{M}_k(K)$ est de la forme

$$M = \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & & \vdots \\ \vdots & & \ddots & & \\ & & & \ddots & 1 \\ 0 & \dots & & & \lambda \end{pmatrix}$$

alors un calcul facile montre que $M - \lambda I$ est nilpotente d'indice k . Le polynôme minimal de M est de la forme $(X - \lambda)^k$, de même degré que la dimension, il est égal au polynôme caractéristique. $\square$

4.3.3. Décomposition en sous-espaces cycliques

Le théorème de décomposition de Jordan permet d'affirmer que tout endomorphisme est somme directe d'endomorphismes cycliques, mais dans le cas d'un polynôme caractéristique scindé. Nous allons donner une autre décomposition en sous-espaces cycliques et préciser des conditions qui permettront d'assurer l'unicité.

Théorème 4.25. *Soit $u \in \mathcal{L}(E)$, où E est un K -espace vectoriel de dimension finie. Il existe des sous-espaces cycliques (non nuls) $E_1 = \langle a_1 \rangle$, $E_2 = \langle a_2 \rangle, \dots, E_k = \langle a_k \rangle$ tels que, si on note u_i la restriction de u à E_i on ait :*

$$E = E_1 \oplus E_2 \oplus \dots \oplus E_k \quad \text{et} \quad u = u_1 \oplus \dots \oplus u_k$$

avec

$$\mu_{u_k} \mid \mu_{u_{k-1}} \mid \dots \mid \mu_{u_1} = \mu_u$$

et

$$\chi_u = \prod_{i=1}^k \mu_{u_i}$$

Unicité : les polynômes μ_{u_i} sont entièrement déterminés par u , on les appelle diviseurs élémentaires de u .

Démonstration. Certaines parties de la démonstration sont délicates. Distinguons les étapes.

- On procède par récurrence sur la dimension de E , le résultat étant banal en dimension 1.
- Il existe un sous-espace cyclique de dimension maximale. En effet, si μ_u est le polynôme minimal, il existe $a_1 \in E$ tel que $\mu_{u, a_1} = \mu_u$, et le sous-espace $E_1 = \langle a_1 \rangle$ est cyclique de dimension maximale, égale au degré de μ_u .
- Comme E_1 est u -stable, il existe un morphisme quotient $\bar{u}$ de E/E_1 dans lui-même : voir la proposition 4.20. Par hypothèse de récurrence,

$$E/E_1 = \bar{E}_2 \oplus \bar{E}_3 \oplus \dots \oplus \bar{E}_k$$

où les $\bar{E}_i$ sont $\bar{u}$ -cycliques, de la forme $\langle \bar{y}_i \rangle$, et où les polynômes minimaux des $\bar{u}$ restreints aux $\bar{E}_i$ vérifient les propriétés du théorème. On va chercher des « bons » représentants des $\bar{y}_i$. Choisissons un indice $i \geq 2$, et notons, pour simplifier P le polynôme minimal de $\bar{y}_i$, qui est aussi le polynôme minimal de $\bar{u}$ restreinte à $\bar{E}_i$. Alors : P divise μ_u . En effet, $\mu_u(u) = 0$ implique, par définition de $\bar{u}$, que $\mu_u(\bar{u}) = 0$. Il existe donc un polynôme Q tel que $\mu_u(X) = Q(X)P(X)$. Soit maintenant x_i un représentant quelconque de $\bar{y}_i$. Alors $P(\bar{u})(\bar{y}_i) = 0$ implique $P(u)(x_i) \in E_1$, et, par définition de E_1 , il existe un polynôme R tel que

$$P(u)(x_i) = R(u)(a_1)$$

Mais on a alors

$$0 = \mu_u(u)(x_i) = Q(u) \circ P(u)(x_i) = Q(u) \circ R(u)(a_1)$$

et le polynôme $Q(X)R(X)$, annulateur par u de a_1 est multiple de

$$\mu_u(X) = Q(X)P(X).$$

On en déduit que R est multiple de P . Si on pose $R(X) = P(X)S(X)$, on aura alors :

$$P(u)(x_i) = R(u)(a_1) = P(u) \circ S(u)(a_1) \quad \text{d'où} \quad P(u)(x_i - S(u)(a_1)) = 0$$

Posons $a_i = x_i - S(u)(a_1)$. Ce vecteur est bien sûr dans la même classe que x_i modulo E_1 .

On a alors $\mu_{u,a_i} = P$. En effet, nous venons d'écrire que $P(u)$ annule a_i , donc $\mu_{u,a_i} \mid P$. Ensuite,

$$\mu_{u,a_i}(u)(a_i) = 0 \quad \text{implique} \quad \mu_{u,a_i}(\bar{u})(\bar{a}_i) = 0$$

donc μ_{u,a_i} divise P .

Si on pose $E_i = \langle a_i \rangle$, pour $i \geq 2$, on vient de voir que les polynômes minimaux de u restreints à ces sous-espaces ont la propriété du théorème. Reste à montrer qu'ils sont en somme directe, et que cette somme est un supplémentaire de E_1 .

Soit donc F la somme $\sum_{i=2}^k E_i$. Alors, il existe un morphisme de $F = \sum_{i=2}^k E_i$ dans $\bigoplus_{i=2}^k \bar{E}_i = E/E_1$ défini par la somme des projections :

$$\sum_{i=2}^k S_i(u)(a_i) \mapsto \sum_{i=2}^k S_i(\bar{u})(\bar{a}_i)$$

où les S_i sont des polynômes quelconques. Il est surjectif par définition des $\bar{a}_i$. C'est un morphisme, et il est injectif car si $\sum_{i=2}^k S_i(\bar{u})(\bar{a}_i) = 0$ alors $S_i(\bar{u})(\bar{a}_i) = 0$, puisque les $\bar{E}_i$ sont en somme directe, donc pour tout $i \geq 2$, le polynôme minimal de $\bar{u}$ restreint à $\bar{E}_i$ divise S_i . Or ce polynôme minimal est aussi μ_{u,a_i} , polynôme minimal de u restreint à E_i . On en déduit que $S_i(u)(a_i) = 0$ pour tout $i \geq 2$, donc que l'application est bijective. En conclusion, les E_i sont en somme directe.

Le lecteur aura bien sûr remarqué que les μ_{u_i} sont des diviseurs de μ_u , puisque ce sont des polynômes minimaux d'éléments de E .

- Nous ne prouverons pas l'unicité des polynômes que l'on appelle facteurs invariants de u . Il est clair que celui de plus haut degré, multiple de tous les autres, est bien déterminé puisque c'est le polynôme minimal de u . De même le produit des facteurs invariants de u est bien déterminé puisque c'est le polynôme caractéristique de u , mais cela n'implique pas, en général, l'unicité de la liste des autres facteurs invariants.

Terminons en remarquant que les (a_i) eux ne sont pas uniquement déterminés par u : si $u = \text{id}$, les polynômes sont

$$\mu_1(X) = \mu_2(X) = \dots = \mu_n(X) = X - 1$$

et on peut prendre pour (a_i) n'importe quelle base de E .

□

Les polynômes ainsi définis s'appellent donc les **diviseurs élémentaires** de u . Si A est une matrice carrée, on peut la considérer comme matrice d'un endomorphisme par exemple d'un endomorphisme de K^n dans la base canonique, et on parlera alors des diviseurs élémentaires de la matrice. Une conséquence de ce théorème, mais qui utilise l'unicité que nous n'avons pas prouvée, est donnée par le corollaire suivant :

Corollaire 4.26. *Deux matrices sont semblables si et seulement si elles ont les mêmes diviseurs élémentaires.*

Démonstration. Il suffit d'observer que notre approche montre que les diviseurs élémentaires d'une matrice sont ceux d'un endomorphisme qu'elle représente. Or deux matrices sont semblables lorsqu'elles représentent le même endomorphisme dans deux bases différentes, si donc elles ont les mêmes diviseurs élémentaires. $\square$

Remarque : On trouvera une démonstration de l'unicité par exemple dans [15]. On trouvera également dans cet ouvrage, et d'autres, des algorithmes permettant de calculer rapidement les diviseurs élémentaires d'une matrice, par de simples calculs de déterminants et de p.g.c.d de polynômes.

Exercice 4.8. Soit u de matrice $\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ et a un vecteur quelconque. Déterminer les sous-espaces $\langle a \rangle$. Donner la décomposition de Frobenius de u .

Exercice 4.9. On suppose que u a pour matrice un bloc de Jordan $J(\lambda, n)$. Trouver, dans une base de Jordanisation, un vecteur a tel que $E = \langle a \rangle$.

Exercice 4.10. On suppose que u est diagonalisable, de valeurs propres $\lambda_1, \lambda_2, \dots, \lambda_k$. Quels sont ses diviseurs élémentaires ? Quand u est-il cyclique ?

Exercice 4.11. On suppose que u est cyclique. Montrer que l'ensemble des endomorphismes v de E qui commutent avec u coïncide avec l'ensemble des polynômes en u . Est-ce une propriété caractéristique des endomorphismes cycliques ?

EXERCICES

Exercice 4.12. Sous-espaces stables, dualité. Soit $u \in \mathcal{L}(E)$, où E est un K -espace vectoriel de dimension finie. On note E^* le dual de E , et on rappelle que ${}^t u$ est l'endomorphisme de E^* défini par : ${}^t u(\phi) = \phi \circ u$.

- (1) Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base de E , on note $\mathcal{B}^* = (e_1^*, \dots, e_n^*)$ la base duale. Si M est la matrice de u dans la base $\mathcal{B}$, déterminer la matrice de ${}^t u$ dans la base $\mathcal{B}^*$.
- (2) Lorsque H est un sous-espace vectoriel de E , on note H° son orthogonal défini par :

$$H^\circ = \{\phi \in E^* \mid \forall x \in H, \phi(x) = 0\}$$

On note également, lorsque G est un sous-espace de E^* :

$$G^\circ = \{x \in E \mid \forall \phi \in G, \phi(x) = 0\}$$

Montrer alors que $H^{\circ\circ} = H$. On pourra commencer par démontrer que $\dim H^\circ = \dim E - \dim H$.

- (3) Montrer que H est stable par u si et seulement si H° est stable par ${}^t u$.
- (4) On suppose que $K = \mathbb{C}$. Montrer qu'un endomorphisme u a des hyperplans stables et donner un moyen pratique de les obtenir. Application : trouver tous les sous-espaces stables de l'endomorphisme u dont la matrice dans une base $\mathcal{B}$ s'écrit :

$$M = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 3 & -1 \\ 0 & 3 & 0 \end{pmatrix}$$

Exercice 4.13. On suppose que le polynôme caractéristique de u est scindé. Montrer que si $u = d + n$ est sa décomposition de Dunford, d et n peuvent s'exprimer sous forme de polynômes en u . On utilisera la démonstration du « lemme des noyaux ».

Exercice 4.14. Soit u un endomorphisme de E , de dimension finie, et soit $(\text{Ker}(u)^k)_{k \in \mathbb{N}}$ la suite des **noyaux itérés**. Montrer qu'elle est d'abord strictement croissante puis stationnaire. Montrer qu'il existe un morphisme injectif naturel de

$$\text{Ker}(u^{k+2}) / \text{Ker}(u^{k+1}) \quad \text{dans} \quad \text{Ker}(u^{k+1}) / \text{Ker}(u^k)$$

et en déduire que la suite (d_i) où $d_i = \dim \text{Ker } u^i$ croît « de moins en moins vite » en ce sens que $d_{k+2} - d_{k+1} \leq d_{k+1} - d_k$.

Soit maintenant u un endomorphisme nilpotent. On note j_k le nombre des blocs de taille k dans sa décomposition de Jordan. Quelle relation y a-t-il entre la suite (d_i) et la suite (j_k) ?

Exercice 4.15. Soit u un endomorphisme dont le polynôme caractéristique est scindé. On considère une valeur propre λ et $(\text{Ker}(u - \lambda \text{id})^k)_{k \in \mathbb{N}}$. On suppose que la multiplicité de λ dans le polynôme caractéristique χ_u est α et que la multiplicité de λ dans le polynôme minimal μ_u est β . Montrer que la suite des noyaux itérés est stationnaire à partir de l'indice β , et que la dimension des derniers noyaux itérés est α .

PROBLÈME

Le corrigé de ce problème est disponible sur le site de Dunod : www.dunod.com.

4.1. QUESTIONS DE TOPOLOGIE

On suppose que le corps de base est $\mathbb{R}$ ou $\mathbb{C}$ et que l'espace vectoriel E est muni d'une norme. On rappelle qu'en dimension finie, toutes les normes sont équivalentes.

- (1) Soit $\| \cdot \|$ une norme de E , on note S la sphère unité, c'est-à-dire l'ensemble des vecteurs de norme 1. Si u est un endomorphisme de E , montrer que

$$\|u\| = \sup_{x \in S} \|u(x)\|$$

définit une norme sur $\mathcal{L}(E)$. On dit que c'est la norme subordonnée à la norme de E dont elle est issue.

- (2) Montrer que toute norme subordonnée vérifie

$$\forall (u, v) \in \mathcal{L}(E), \|u \circ v\| \leq \|u\| \|v\|$$

On dit que cette norme est sous-multiplicative. On se placera toujours dans ce cas, pour la fin du problème.

- (3) On suppose qu'une base de E a été fixée, il y a alors isomorphisme entre E et $\mathcal{M}_{n,1}(K)$, entre $\mathcal{L}(E)$ et $\mathcal{M}_n(K)$. À toute norme dans $\mathcal{M}_{n,1}(K)$ on peut ainsi associer une norme subordonnée dans $\mathcal{M}_n(K)$. Décrire les normes subordonnées associées aux trois normes suivantes :

Si ${}^tX = (x_1, x_2, \dots, x_n)$, alors :

$$\|x\|_1 = \sum_{i=1}^n |x_i|, \quad \|x\|_2 = \left(\sum_{i=1}^n |x_i|^2 \right)^{\frac{1}{2}}, \quad \|x\|_\infty = \sup_{i=1}^n |x_i|$$

- (4) On a choisit une norme sous-multiplicative quelconque. Soit λ une valeur propre réelle ou complexe de u . Montrer que $|\lambda| \leq \|u\|$.
- (5) Montrer que $\mathbf{GL}(E)$ est un ouvert dense de $\mathcal{L}(E)$.
- (6) On suppose $K = \mathbb{C}$. Montrer que l'ensemble des endomorphismes diagonalisables est dense dans $\mathcal{L}(E)$. Que dire dans le cas réel ?
- (7) On suppose toujours $K = \mathbb{C}$. Montrer que l'ensemble des endomorphismes ayant n valeurs propres distinctes est dense dans $\mathcal{L}(E)$.
- (8) En utilisant des arguments de densité, prouver que $u \circ v$ et $v \circ u$ ont même polynôme caractéristique.

SOLUTIONS DES EXERCICES

Solution 4.1. Soit $(e_i)_{i \in I}$ une base de E , (finie ou infinie) et u qui commute avec tous les endomorphismes. Alors, pour tout j , $u(e_j) = \sum_{i \in I} a_{ij} e_i$, avec $\text{Card}\{i \mid a_{ij} \neq 0\} < \infty$. Soit alors g l'endomorphisme défini par :

$$\begin{cases} g_i(e_j) = 0 & \text{si } i \neq j \\ g_i(e_i) = e_i \end{cases}$$

(Si on est en dimension finie, la matrice de g_i est la matrice E_{ii} de la base canonique). Alors, $u \circ g_i = g_i \circ u$ implique $u(e_i) = a_{ii} e_i$ pour tout i , donc tous les a_{ij} pour $i \neq j$ sont nuls. Reste à étudier les coefficients a_{ii} . Si $\text{Card } I > 2$, il existe deux indices distincts i et j , et si l'on définit g_{ij} par :

$$\begin{cases} g_{ij}(e_i) = e_j \\ g_{ij}(e_j) = e_i \\ g_{ij}(e_k) = 0 \end{cases}$$

alors

$$u \circ g_{ij} = g_{ij} \circ u \Rightarrow u \circ g_{ij}(e_i) = g_{ij} \circ u(e_i) \Rightarrow u(e_j) = g_{ij}(a_{ii} e_i)$$

et donc $a_{jj} e_j = a_{ii} e_j$ et, pour tout i, j l'égalité de $a_{i,i}$ et de $a_{j,j}$, et u est de la forme λId . En dimension finie, on en déduit que les matrices qui commutent avec toutes les autres sont les matrices de la forme λI , où I est la matrice de l'identité. Ce sont les matrices dites **scalaires**.

Solution 4.2. Dire que λ est valeur propre, c'est dire qu'il existe x non nul dans $\text{Ker}(A - \lambda \text{Id})$. Si x a pour coordonnées $(z_1, z_2, \dots, z_n)$

$$\forall i, \quad \sum_{j \neq i} a_{ij} z_j = (\lambda - a_{ii}) z_i$$

Soit alors i tel que $|z_i| = \max\{|z_k| \mid k = 1, 2, \dots, n\}$ alors l'égalité précédente, permet d'écrire :

$$|\lambda - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|$$

Les valeurs propres se trouvent donc à l'intérieur de n disques.

Si maintenant $|a_{ii}| > \sum_{j \neq i} |a_{ij}|$, pour tout i , alors la matrice est inversible. Dans ce cas en effet, 0 ne peut être valeur propre. Comme l'hypothèse de ce corollaire le fait comprendre, on parle alors de matrice à diagonale fortement dominante.

Solution 4.3. On suppose que $\dim E = n$; ayant choisi une base, on note A et B les matrices respectives de u et de v . Si par exemple A est inversible, on peut écrire

$$BA = A^{-1}(AB)A$$

et les matrices AB et BA sont semblables, elles ont même polynôme caractéristique (et également même polynôme minimal). Pour les cas où ni A ni B ne sont inversibles, on peut procéder ainsi : dans $\mathcal{M}_{2n}(K)$,

$$\begin{pmatrix} -\lambda I & -A \\ -\lambda B & -\lambda I \end{pmatrix} = \begin{pmatrix} AB - \lambda I & -A \\ 0 & -\lambda I \end{pmatrix} \begin{pmatrix} I & 0 \\ B & I \end{pmatrix} = \begin{pmatrix} I & 0 \\ B & I \end{pmatrix} \begin{pmatrix} -\lambda I & -A \\ 0 & BA - \lambda I \end{pmatrix}$$

Les déterminants sont égaux et valent respectivement $(-\lambda)^n \chi_{AB}$ et $(-\lambda)^n \chi_{BA}$. Par ailleurs, il est possible que AB et BA ne soient pas semblables : il suffit de prendre $AB = 0$ et $BA \neq 0$, par exemple

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

On fabrique cet exemple en prenant l'image de v incluse dans le noyau de u et l'image de u non incluse dans le noyau de v . Dans cette situation, les polynômes minimaux de AB et de BA sont distincts, ce sont respectivement X et X^2 .

Solution 4.4. $\mathcal{B}_1$ et $\mathcal{B}_2$ sont des bases de E_1 et E_2 , les matrices de u_1 et u_2 sont notées M_1 et M_2 . La réunion (ordonnée) de ces deux bases est une base $\mathcal{B}$ de E , puisque E_1 et E_2 sont supplémentaires. On a donc

$$M_{\mathcal{B}}(u) = \begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix}$$

d'où

$$\chi_u(X) = \begin{vmatrix} M_1 - XI & 0 \\ 0 & M_2 - XI \end{vmatrix} = \chi_{u_1}(X) \chi_{u_2}(X)$$

On a utilisé des déterminants de « matrices-blocs » : le lecteur vérifiera par exemple par un raisonnement par récurrence que si

$$M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

où A et C sont des matrices carrées, on a bien $\det M = (\det A)(\det C)$.

Pour ce qui est du polynôme minimal, il faut raisonner plus géométriquement. Si $x = x_1 + x_2$ est la décomposition de x dans la somme directe, et si P est un polynôme,

$$P(u)(x) = P(u)(x_1 + x_2) = P(u)(x_1) + P(u)(x_2) = P(u_1)(x_1) + P(u_2)(x_2)$$

Puisque E_1 et E_2 sont u -stables, cette écriture est la décomposition de $P(u)(x)$ dans la somme directe, et donc $P(u)$ est l'endomorphisme nul si et seulement si $P(u_1) = 0$ et $P(u_2) = 0$. On en déduit bien :

$$\mu_u(X) = \text{p.p.c.m.}(\mu_{u_1}(X), \mu_{u_2}(X))$$

Solution 4.5. On utilise le résultat du théorème 4.8, page 91. Soit E_λ un sous-espace propre de u . Il est stable par v . La restriction de v à E_λ est diagonalisable, puisque v est diagonalisable sur E . Il existe donc une base de E_λ qui est une base de diagonalisation pour v , mais comme u restreinte à E_λ est l'homothétie de rapport λ , cette base est aussi une base de diagonalisation pour u restreint à E_λ . En réunissant de tels bases pour tous les sous-espaces propres de u , on forme une base de diagonalisation simultanée pour u et v .

Le cas trigonalisable est basé sur les mêmes idées. On observe d'abord que u et v ont un vecteur propre commun : on sait en effet que u admet au moins une valeur propre λ puisque son polynôme caractéristique est scindé. Si $\mathcal{B}_1 = (e_1, e_2, \dots, e_k)$ est une base de $E_{u,\lambda}$, on peut la compléter en une base $(e_1, \dots, e_n)$ de l'espace tout entier ; d'autre part, on sait que $E_{u,\lambda}$ est stable par v puisque u et v commutent. Notons u_1 et v_1 les restrictions de u et de v à ce sous-espace. La matrice de v dans la base complétée est de la forme :

$$\begin{pmatrix} M_{\mathcal{B}_1}(v_1) & * \\ 0 & A \end{pmatrix}$$

d'où on déduit que $\chi_v(X) = \chi_{v_1}(X)\chi_A(X)$. Comme v est trigonalisable, son polynôme caractéristique est scindé, et celui de sa restriction aussi : il existe donc un vecteur propre pour v dans $E_{u,\lambda}$, c'est notre vecteur propre commun. Ayant pris ce vecteur propre commun comme premier vecteur de base, les matrices de u et de v ont comme allure :

$$\begin{pmatrix} \lambda & * \\ 0 & M_1 \end{pmatrix} \text{ et } \begin{pmatrix} \mu & * \\ 0 & M_2 \end{pmatrix}$$

et on remarque que si u et v commutent, alors M_1 et M_2 commutent aussi, on peut appliquer une récurrence.

Solution 4.6. T est une matrice triangulaire supérieure de diagonale nulle. Elle est nilpotente ; cela résulte d'un calcul matriciel : on voit que pour T^2 , les coefficients $t_{i,i+1}$ sont nécessairement nuls, puis pour T^3 , les coefficients $t_{i,i+2}$, etc. Mais plus efficacement, le polynôme caractéristique de T est $(-X)^n$ et on applique le théorème de Cayley-Hamilton. On n'a pas forcément $DT = TD$, donc ce n'est pas forcément la décomposition de Dunford de A . Pour un contre-exemple, prendre A triangulaire supérieure à éléments diagonaux tous distincts. Alors elle est diagonalisable, et sa décomposition de Dunford est $A = A + 0$.

Solution 4.7. Les blocs de Jordan sont de la forme $J(0, p)$, mais la somme de leur dimension doit être n . Donc,

- $n = 2$, $2 = 1 + 1$, c'est l'endomorphisme nul, $2 = 2$, c'est la classe de $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

- $n = 3$, $3 = 1 + 1 + 1$, endomorphisme nul, $3 = 1 + 2$, l'ordre de nilpotence est 2, et $3 = 3$, l'ordre de nilpotence est 3.
- $n = 4$, $4 = 1 + 1 + 1 + 1$, $4 = 1 + 1 + 2$, $4 = 1 + 3$, $4 = 2 + 2$, $4 = 4$. Cette fois il y a deux classes différentes pour lesquelles les endomorphismes ont même ordre de nilpotence (donc même polynôme minimal).
- Pour $n = 5$, on trouve 7 classes, pour $n = 6$ il y en a 11. C'est ce qu'on appelle le nombre des partitions de l'entier n , nombre rapidement croissant avec n (par exemple, pour $n = 20$, on en trouve 627).

Solution 4.8. Notons (e_1, e_2) la base. Si $a = 0$, $\langle a \rangle = \{0\}$. Si a est un vecteur propre, donc $a \in \text{vect}(e_1)$ ou $a \in \text{vect}(e_2)$, alors $\langle a \rangle = \text{vect}(a)$. Sinon, $\langle a \rangle$ est donc nécessairement E tout entier. L'endomorphisme est cyclique. Si on prend la base $(a, u(a))$ où a n'est pas propre (par exemple $e_1 + e_2$), la matrice de u est :

$$\begin{pmatrix} 0 & -2 \\ 1 & 3 \end{pmatrix} \quad \text{matrice compagnon du polynôme } X^2 - 3X + 2$$

Solution 4.9. La solution se trouve dans l'examen (attentif) de la démonstration du théorème 4.16. Si $(e_1, \dots, e_n)$ est une base de Jordan de u , on a $\langle e_n \rangle = E$. En effet, $u(e_n) = e_{n-1} + \lambda e_n$ donc $\text{vect}(e_n, u(e_n)) = \text{vect}(e_n, u(e_n))$ et ainsi de suite.

Solution 4.10. Le polynôme minimal est scindé simple, c'est $\prod_{i=1}^k (X - \lambda_i)$. Le polynôme caractéristique est $\prod_{i=1}^k (X - \lambda_i)^{\alpha_i}$. Si $\mu_\ell \mid \dots \mid \mu_2 \mid \mu_1$ est la suite des diviseurs élémentaires, on a $\mu_1(X) = \prod_{i=1}^k (X - \lambda_i)$. Notons $E_1 \oplus E_2 \oplus \dots \oplus E_\ell$ la suite des sous-espaces cycliques, μ_1 est le polynôme minimal de la restriction de u à E_1 . On cherche maintenant μ_2 polynôme minimal de la restriction u_2 de u à $E_2 \oplus \dots \oplus E_\ell$ qui est bien sûr u -stable, donc u_2 est diagonalisable (puisque u est diagonalisable). Ce polynôme minimal est donc scindé simple, et comme

$$\chi_{u_2}(X) = \frac{\chi_u(X)}{\mu_u(X)} = \prod_{i=1}^k (X - \lambda_i)^{\alpha_i - 1}$$

on a $\mu_2(X) = \prod_{i \in I_1}^k (X - \lambda_i)$ où I_1 est l'ensemble des indices des valeurs propres de multiplicité supérieure à 1. On continue de même, et on aura $\mu_2(X) = \prod_{i \in I_2}^k (X - \lambda_i)$, le produit portera sur les valeurs propres de multiplicité supérieure à 2. Bien sûr, u sera cyclique si et seulement si il y a n valeurs propres distinctes.

Solution 4.11. Notons $C(u)$ l'ensemble des endomorphismes qui commutent avec u , c'est une sous-algèbre de $\mathcal{L}(E)$, qui contient $K[u]$, algèbre des polynômes en u . Supposons $E = \langle a \rangle$ et soit $g \in C(u)$. Alors $g(a) \in E$ peut s'écrire sous la forme

$g(a) = P(u)(a)$ où P est un polynôme. Mais alors, si x est quelconque dans E , il peut aussi s'écrire $x = Q(u)(a)$ où Q est un polynôme et

$$g(x) = (g \circ Q(u))(a) = (Q(u) \circ g)(a) = Q(u) \circ P(u)(a) = P(u) \circ Q(u)(a) = P(u)(x)$$

et on a démontré que $g = P(u)$, donc que $C(u) = K[u]$.

La réciproque est vraie, mais la démonstration est moins directe : on suppose que $C(u) = K[u]$ et on considère E_1 et $E_2 \oplus \dots \oplus E_k$ de la décomposition en sous-espaces cycliques. Soit alors p la projection sur le second sous-espace, suivant E_1 . Alors p commute avec u . C'est un polynôme $P(u)$, mais $p = P(u)$ restreint à E_1 est nul donc P doit être multiple du polynôme minimal de u restreint à E_1 . Mais ce polynôme est aussi le polynôme minimal de u pour E tout entier. Donc $P(u) = p = 0$, la seule possibilité est $E_1 = E$, donc u est cyclique.

Solution 4.12. Rappelons comment « fonctionne » la base duale (e_i^*). On a par définition $e_i^*(e_j) = \delta_{i,j}$, symbole de Kronecker qui vaut 1 si $i = j$, 0 sinon.

(1) Notons $M = (m_{i,j})$ et $M' = \mathcal{M}_{B^*}({}^t u) = (m'_{i,j})$. Alors,

$$m'_{i,j} = {}^t u(e_j^*)(e_i) = e_j^*(u(e_i)) = e_j^*\left(\sum_{k=1}^n m_{k,i} e_k\right) = m_{j,i}$$

et donc $M' = {}^t M$... C'était prévisible.

(2) Soit $(e_1, \dots, e_k)$ une base de H , complétée en $(e_1, \dots, e_n)$ base de E . Alors, en posant $\phi = \sum_{i=1}^n \phi_i e_i^*$,

$$\phi \in H^\circ \iff \forall i \leq k, \phi(e_i) = 0 \iff \forall i \leq k, \phi_i = 0 \iff \phi \in \text{vect}(e_{k+1}^*, \dots, e_n^*)$$

donc $\dim H^\circ = \dim E - \dim H$.

Si maintenant $x \in H$ et $\phi \in H^\circ$, alors $\phi(x) = 0$ prouve que $x \in H^{\circ\circ}$. On a donc $H \subset H^{\circ\circ}$. Mais ces espaces vectoriels ont même dimension, il y a donc égalité. Les mêmes arguments montrent que E et E^{**} sont canoniquement isomorphes : à $x \in E$ correspond la forme linéaire de E^* définie par $\phi \mapsto \phi(x)$. Ainsi, ${}^t({}^t u) = u$.

(3) Soit $\phi \in H^\circ$ et $x \in H$, quelconque. Si H est u -stable,

$${}^t u(\phi)(x) = \phi(u(x)) = 0$$

et H° est ${}^t u$ -stable. Si maintenant ${}^t u(H^\circ) \subset H^\circ$ alors ${}^t({}^t u)(H^{\circ\circ}) \subset H^{\circ\circ}$ soit $u(H) \subset H$.

(4) Soit u endomorphisme d'un $\mathbb{C}$ -espace vectoriel. Un hyperplan H est u -stable si et seulement si son orthogonal H° est ${}^t u$ -stable. Mais H° est de dimension 1, il est donc stable si c'est une droite propre de ${}^t u$. Rechercher les hyperplans u -stables revient donc à chercher les valeurs propres de ${}^t u$. Dans l'exemple proposé, M admet trois valeurs propres distinctes, il y a donc trois droites stables. Mais ${}^t M$ admet alors également trois valeurs propres distinctes, il y a donc trois plans stables. On pourra vérifier, bien sûr, que chacun de ces plans contient deux des droites propres.

Solution 4.13. Dans la démonstration du lemme des noyaux, on obtient une décomposition d'un vecteur x dans la somme directe sous la forme :

$$x = \sum_{i=1}^k x_i \quad \text{où} \quad x_i = A_i(u) \circ Q_i(u)(x)$$

et cela s'applique aux cas des sous-espaces caractéristiques. On en déduit que les projections sur un sous-espace caractéristique suivant la somme des autres sont des polynômes en u (ce sont les projecteurs spectraux). Dans la décomposition de Dunford, on a :

$$d = \sum_{i=1}^k \lambda_i p_i \quad \text{et} \quad n = u - d$$

donc d et n sont des polynômes en u . On peut souvent en déduire rapidement la décomposition de Dunford d'un endomorphisme.

Solution 4.14. $\text{Ker } u^k \subset \text{Ker } u^{k+1}$ car $u^k(x) = 0$ implique $u(u^k(x)) = u(0) = 0$. De plus, si on suppose $\text{Ker } u^k = \text{Ker } u^{k+1}$ et que $x \in \text{Ker } u^{k+2}$, alors

$$0 = u^{k+2}(x) = u^{k+1}(u(x)) \quad \text{ce qui prouve que} \quad u(x) \in \text{Ker } u^{k+1} = \text{Ker } u^k.$$

Donc $u^k(u(x)) = 0$ et $x \in \text{Ker } u^{k+1}$. On a montré que $\text{Ker } u^{k+1} = \text{Ker } u^{k+2}$, la suite des noyaux itérés est donc définitivement stationnaire. Remarquons également qu'en dimension finie, la suite des dimensions est majorée par n , elle est donc nécessairement stationnaire à partir d'un certain rang.

Commençons par considérer le morphisme u restreint à $\text{Ker } u^{k+2}$. Si x est dans cet espace, $u(x)$ est dans $\text{Ker } u^{k+1}$, on peut donc restreindre l'ensemble d'arrivée à $\text{Ker } u^{k+1}$. Soit alors $\bar{u} : x \mapsto \bar{x}$, classe modulo le sous-espace vectoriel $\text{Ker } u^k$. Un vecteur x est dans le noyau de $\bar{u}$ lorsque $u(x) \in \text{Ker } u^k$ c'est-à-dire lorsque $x \in \text{Ker } u^{k+1}$. On en déduit, par le théorème d'isomorphisme, qu'il y a un morphisme injectif de

$$\text{Ker}(u^{k+2}) / \text{Ker}(u^{k+1}) \quad \text{dans} \quad \text{Ker}(u^{k+1}) / \text{Ker}(u^k)$$

et, en prenant les dimensions, $d_{k+2} - d_{k+1} \leq d_{k+1} - d_k$.

Supposons maintenant que u soit nilpotent d'indice p , dans un espace de dimension n . Notons $d_0 = 0, d_1, \dots, d_p = n$ la suite des dimensions des noyaux, puis $e_1 = d_1 - d_0, e_2 = d_2 - d_1, \dots, e_p = d_p - d_{p-1}$ la suite (décroissante) des « écarts » entre les dimensions des noyaux. On sait que le nombre de k blocs est la dimension des sous-espaces notés S_k dans la démonstration du théorème. Donc $j_p = e_p, j_{p-1} = e_{p-1} - e_p$, etc... On pourra vérifier par exemple que

$$n = pj_p + (p-1)j_{p-1} + \dots + j_1$$

ce qui est « logique ».

Solution 4.15. On considère la restriction u_i de u au sous-espace caractéristique $C_i = \text{Ker}(u - \lambda_i \text{id})^\alpha$. On sait que $(X - \lambda)^\alpha$ est le polynôme caractéristique de u_i , le degré α de ce polynôme est donc la dimension de C_i . Mais, par ailleurs, la restriction de $u - \lambda_i \text{id}$ à C_i est nilpotente, et l'ordre de nilpotence est le plus petit exposant β tel que $\text{Ker}(u - \lambda_i \text{id})^\beta = C_i$. La suite des noyaux itérés est donc stationnaire à partir de cet entier β , qui est l'ordre de multiplicité de λ_i dans le polynôme minimal.

Chapitre 5

Groupes

La notion de groupe a été rencontrée à de multiples reprises les années précédentes. Ce chapitre reprend rapidement les notions générales, puis propose de s'intéresser au cas particulier des groupes commutatifs finis. Il se termine en exposant les théorèmes de Sylow, première étape vers une classification des groupes finis. Par convention, et sauf cas particuliers, le composé de deux éléments a et b sera noté ab , l'élément neutre sera e , aa sera écrit a^2 , etc. On supposera connus les exemples classiques : groupe additif $\mathbb{Z}/n\mathbb{Z}$ des entiers modulo n , groupe symétrique S_n et groupe alterné A_n .

5.1 SOUS-GROUPES, MORPHISMES

5.1.1. Sous-groupes

Soit G un groupe. Si H est un sous-ensemble non vide de G , on dit que c'est un sous-groupe de G lorsque c'est aussi un groupe pour la même loi : cela se produit si et seulement si H est stable pour le produit et pour la prise d'inverse.

$$\forall (x, y) \in H^2, xy \in H \text{ et } x^{-1} \in H$$

On écrit alors $H \leq G$.

L'intersection de sous-groupes de G est un sous-groupe de G , ce qui permet de définir le sous-groupe engendré par une partie A de G : c'est le plus petit sous-groupe de G qui contient A , et c'est l'intersection de tous les sous-groupes de G qui contiennent A . On le notera $\langle A \rangle$.

Définition 5.1. Si H est un sous-groupe de G , la relation $\mathcal{R}$ définie dans G par :

$$x\mathcal{R}y \iff x^{-1}y \in H$$

est une relation d'équivalence. Les classes d'équivalence sont les ensembles de la forme xH , l'ensemble quotient de G par cette relation est noté G/H .

La vérification est en tout point similaire à ce qui a été fait avec les anneaux et les idéaux et $x\mathcal{R}y \iff y \in xH$, avec bien sûr $xH = \{g \in G \mid \exists h \in H \ g = xh\}$. On peut définir une autre relation d'équivalence, par $x\mathcal{R}'y \iff yx^{-1} \in H$, et les classes sont alors les ensembles de la forme Hx , appelées classes à droite. On peut alors énoncer le premier théorème de la théorie des groupes.

Théorème 5.2. *théorème de Lagrange.* Toute classe à gauche ou à droite est en bijection avec H . Si G est fini, le cardinal d'un sous-groupe H et le cardinal du quotient G/H sont des diviseurs associés du cardinal de G :

$$\text{Card}(G) = \text{Card}(H) \times \text{Card}(G/H)$$

Démonstration. Il suffit de vérifier que l'application $h \mapsto xh$ est une bijection : c'est parce qu'il y a une application réciproque $g \mapsto x^{-1}g$. Et donc toutes les classes (à gauche ou à droite) sont en bijection avec H . G est la réunion disjointe des classes à gauche ; si donc G est fini, son cardinal est la somme des cardinaux des classes, d'où la formule. $\square$

Remarque : Dans le cas où G est infini, le cardinal de H ou le cardinal de G/H peuvent être finis. On note parfois $[G : H]$ (**indice** de H dans G), le cardinal de G/H . Dans le cas fini, le cardinal d'un sous-groupe H est donc un diviseur du cardinal n du groupe G . Réciproquement, si $d \mid n$, il n'existe pas toujours un sous-groupe de G qui a ce cardinal. Ainsi le groupe alterné $\mathcal{A}_4$ a pour cardinal 12, il n'a pas de sous-groupe à 6 éléments.

5.1.2. Morphismes

Définition 5.3. Une application f du groupe G dans le groupe H est un morphisme de groupes si

$$\forall (g, g') \in G^2, \quad \phi(gg') = \phi(g)\phi(g').$$

Il est alors immédiat que l'image d'un sous-groupe de G est un sous-groupe de H , (c'est le cas en particulier de $\text{Im } f = f(G)$) et il en va de même pour l'image réciproque d'un sous-groupe de H . On appelle **noyau** du morphisme l'image réciproque de l'élément neutre :

$$\text{Ker } f = \{x \in G \mid f(x) = e_H\}$$

et on introduit, comme en algèbre linéaire le vocabulaire : endomorphisme, isomorphisme, automorphisme de groupe. Comme pour les anneaux ou les espaces vectoriels :

Proposition 5.4. *Soit f un morphisme du groupe G dans le groupe H . Alors f est surjectif si $f(G) = \text{Im}(f) = H$, f est injectif si $f^{-1}(e_H) = \text{Ker } f = \{e_G\}$. Si f est un morphisme bijectif, son application réciproque est aussi un morphisme de groupes. On parle d'isomorphisme de groupes.*

5.1.3. Ordre, groupes cycliques

L'**ordre** de x élément d'un groupe G est le plus petit entier non nul n tel que $x^n = e$; si un tel entier n'existe pas, on dit que x est d'**ordre infini**.

Proposition 5.5.

(i) Si $x \in G$ est d'ordre n fini alors : $\forall m \in \mathbb{N}^*, \quad (x^m = e \Rightarrow n \mid m)$.

(ii) Si x est d'ordre n fini, le groupe $\langle x \rangle$ engendré par x est

$$\langle x \rangle = \{e, x, x^2, \dots, x^{n-1}\}$$

et $\langle x \rangle$ est isomorphe au groupe $(\mathbb{Z}/n\mathbb{Z}, +)$.

Si x est d'ordre infini, le groupe engendré par x est isomorphe à $(\mathbb{Z}, +)$.

(iii) Si G est fini, tout élément de G est d'ordre fini et cet ordre est un diviseur de $\text{Card } G$.

(iv) Si $x \in G$ est d'ordre n , et si k est un entier, alors x^k est d'ordre $\frac{n}{n \wedge k}$.

Démonstration.

(i) Le morphisme $\phi : k \mapsto x^k$ de $(\mathbb{Z}, +)$ dans G a pour noyau un sous-groupe de $\mathbb{Z}$. D'après la définition de l'ordre de x ce noyau est $n\mathbb{Z}$. Et donc tous les entiers m alors tels que $x^m = e$ sont des multiples de n .

(ii) Si $x^i = x^j$, alors $i - j$ est un multiple de n . Si donc $\text{Ker } \phi = \{0\}$, $n = 0$, l'ensemble des puissances positives et négatives de x est formé d'éléments distincts; ces éléments forment un groupe isomorphe à $(\mathbb{Z}, +)$. Si $\text{Ker } \phi = n\mathbb{Z}$, $e, x, x^2, \dots, x^{n-1}$ sont distincts, et toute puissance de x de la forme x^m peut s'écrire $x^m = x^{nq+r} = x^r$ où r est le reste de la division euclidienne de m par n et vaut $0, 1, \dots$ ou $n - 1$.

L'ensemble $\{e, x, \dots, x^{n-1}\}$ forme un groupe isomorphe à $(\mathbb{Z}/n\mathbb{Z}, +)$. L'isomorphisme est défini par $x^k \mapsto \bar{k}$ où $\bar{k}$ est la classe de k modulo n .

(iii) C'est un cas particulier du théorème de Lagrange, $\langle x \rangle$ est un sous-groupe de G : son ordre, qui est l'ordre de x , divise l'ordre de G .

(iv) Posons $n = n'd$ et $k = k'd$ où $d = n \wedge k$ est le p.g.c.d. de n et k , et où n' et k' sont premiers entre eux. Alors

$$(x^k)^{n'} = x^{k'dn'} = (x^n)^{k'} = e$$

donc l'ordre de x^k divise n' et par ailleurs

$$\left((x^k)^m = e\right) \text{ donc } n = n'd \text{ divise } km = k'dm \text{ et donc } n' \mid k'm$$

et donc m est un multiple de $n' = \frac{n}{n \wedge k}$ par le théorème de Gauss.

□

Remarques De la dernière partie de ce théorème, il ressort que pour tout groupe cyclique C_n , et pour tout diviseur d de n , il existe un sous-groupe d'ordre d (et on montre facilement un seul). À comparer avec les remarques qui ont suivi le théorème de Lagrange.

De la proposition, il ressort également que tous les groupes engendrés par un élément d'ordre n sont isomorphes. On dit qu'il s'agit d'un **groupe cyclique** d'ordre n . On notera C_n un groupe cyclique d'ordre n quelconque. Un modèle de C_n , en notation additive, est le groupe additif $\mathbb{Z}/n\mathbb{Z}$ des entiers modulo n . Un autre modèle est le sous-groupe multiplicatif $\mathbb{U}_n$ des racines n -ièmes de l'unité dans $\mathbb{C}$.

Exercice 5.1. Démontrer qu'un groupe d'ordre p premier est cyclique.

Exercice 5.2. Soit H et K des sous-ensembles d'un groupe G . On note HK et H^{-1} les ensembles définis par :

$$HK = \{g \in G \mid \exists h \in H, \exists k \in K, g = hk\} \quad \text{et} \quad H^{-1} = \{h^{-1} \mid h \in H\}$$

Montrer que H est un sous-groupe de G si et seulement si $HH = H$ et $H^{-1} = H$.

Exercice 5.3. Montrer qu'un groupe G est commutatif si et seulement si $x \mapsto x^{-1}$ est un automorphisme de G .

Exercice 5.4. Soit ϕ un morphisme du groupe G dans le groupe H . Si $x \in G$ est d'ordre n , quel dire de l'ordre de $\phi(x)$? Et si ϕ est injectif?

Exercice 5.5. Soit $\mathbb{D}_8$ le groupe engendré par les matrices

$$a = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{et} \quad b = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Montrer qu'il a huit éléments. Déterminer ses sous-groupes. On l'appelle groupe diédral d'ordre huit.

5.2 GROUPE QUOTIENT ET GROUPE PRODUIT

5.2.1. Sous-groupe distingué

Soit G un groupe et H un sous-groupe. Nous avons déjà défini l'ensemble des classes à gauche noté G/H . Il est possible, sous certaines conditions, que cet ensemble quotient soit muni d'une structure de groupe héritée en quelque sorte de la structure de groupe de G . Il faut et il suffit pour cela que H soit d'un type particulier, qu'il soit **distingué** dans G . On retrouve ici la même type de situation que pour les anneaux : le quotient d'un anneau par un idéal est muni d'une structure d'anneau lorsque l'idéal est bilatère.

Définition 5.6. On dit que H est un sous-groupe distingué de G lorsque

$$\forall x \in G, xH = Hx$$

On écrit alors : $H \triangleleft G$.

Autrement dit, les classes à gauche et à droite de n'importe quel élément de G coïncident. Bien sûr, si G est commutatif, tout sous-groupe est distingué, comme pour les anneaux : dans un anneau commutatif tous les idéaux sont bilatères.

Théorème 5.7. Si $H \triangleleft G$, l'ensemble G/H est un groupe pour la loi définie par

$$xH \cdot yH = xyH$$

L'application p de G sur G/H qui à x associe sa classe xH est alors un morphisme surjectif de groupe, de noyau H . On l'appelle projection de G sur son quotient G/H .

Démonstration. Comme pour le quotient d'un anneau par un idéal bilatère, tout repose sur le fait que la congruence modulo H est compatible avec l'opération de groupe, lorsque H est distingué dans G . En effet, Si $x \equiv x'$ et $y \equiv y'$, alors $xy \equiv x'y'$ puisque $(xy)^{-1}x'y' = y^{-1}x^{-1}x'yy^{-1}y' \in H$ car $x^{-1}x' \in H$ et que H est distingué dans G .

Le fait que la projection soit un morphisme résulte de la définition de l'opération dans le quotient : $\overline{xy} = \overline{x}\overline{y}$ où l'on a noté $\overline{x}$ la classe d'un élément modulo H , au lieu de xH . $\square$

La classe $\overline{e}$ de l'élément neutre de G est $eH = H$, c'est l'élément neutre du quotient G/H et H est le noyau du morphisme p . Plus généralement :

Proposition 5.8. Un sous-groupe H d'un groupe G est distingué dans G si et seulement si c'est le noyau d'un morphisme de source G .

Démonstration. Tout sous-groupe H distingué dans G est le noyau de la projection p . Réciproquement, soit $\phi : G \rightarrow K$ un morphisme de groupes. Alors, si g est quelconque dans G et $h \in \text{Ker } \phi$, on a

$$\phi(ghg^{-1}) = \phi(g)\phi(h)\phi(g^{-1}) = \phi(g)e_K\phi(g^{-1}) = \phi(gg^{-1}) = \phi(e_K) = e_K$$

$\square$

Comme dans le cas des anneaux quotients, on peut énoncer un théorème d'isomorphisme et un théorème de correspondance.

Théorème 5.9. *Théorème d'isomorphisme.* Soit $\phi : G \rightarrow H$ un morphisme de groupes. Il existe un isomorphisme canonique $\bar{\phi}$ de $G/\text{Ker } \phi$ sur $\text{Im } \phi$.

Théorème 5.10. *Théorème de correspondance.* Soit $H \triangleleft G$. La projection p de G sur G/H induit une bijection de l'ensemble des sous-groupes K tels que $H \leq K \leq G$ dans l'ensemble des sous-groupes de G/H .

Les démonstrations sont en tous points semblables à celles faites dans le cas des anneaux, nous ne les reprendrons pas. Signalons de plus que la bijection du second théorème transforme les groupes distingués dans G contenant H en les groupes distingués dans G/H .

5.2.2. Groupe produit

Soient H et K deux groupes. Dans l'ensemble $H \times K$, il est possible de définir une structure de groupe particulièrement simple.

Proposition 5.11. L'ensemble $H \times K$ est un groupe pour la loi définie par :

$$(h, k)(h', k') = (hh', kk')$$

où h et h' sont dans H , k et k' dans K . On dit que c'est le produit direct des groupes H et K .

Démonstration. Les vérifications sont immédiates. On constate en particulier que l'élément neutre est le couple (e_H, e_K) et que $(h, k)^{-1} = (h^{-1}, k^{-1})$. $\square$

À titre d'exemple, on peut bien sûr considérer le produit cartésien $C_n \times C_m$ de deux groupes cycliques. En reprenant l'étude faite dans le chapitre sur les anneaux, on vérifiera que ce produit de groupes cycliques est cyclique si et seulement si $m \wedge n = 1$ et qu'il est alors isomorphe à C_{mn} . On dira que c'est le théorème chinois «version groupe», par opposition avec la version anneau. Le plus petit de ces produits est le groupe $C_2 \times C_2$, qui est commutatif, non cyclique, de cardinal 4. On l'appelle groupe de Klein, ou, pour des raisons géométriques, groupe du rectangle : il est isomorphe au groupe des isométries qui conservent un rectangle non carré (l'identité, les réflexions par rapports aux axes de symétrie, la symétrie centrale de centre le centre du rectangle). Toujours aussi élémentaire l'isomorphisme $\mathbb{C}^* \simeq \mathbb{R}^* \times \mathbb{U}$ (pour la loi produit) est la traduction des propriétés de l'écriture trigonométrique d'un complexe non nul. Il est bien sûr possible de généraliser, on peut définir le produit direct d'un nombre fini de groupes, ou même le produit direct d'une famille de groupes.

On retrouve dans $G = H \times K$ une copie de H et une copie de K . Plus précisément.

Proposition 5.12. Si G est le produit direct de H et K alors

$$\tilde{H} = H \times \{e_K\} \quad \tilde{K} = \{e_H\} \times K$$

sont des sous-groupes de G isomorphes respectivement à H et à K . De plus :

$$\tilde{H} \triangleleft G, \quad \tilde{K} \triangleleft G, \quad \tilde{H} \cap \tilde{K} = e_G, \quad G = \tilde{H}\tilde{K} = \tilde{K}\tilde{H}$$

avec encore une précision

$$\forall g \in G, \exists ! h \in \tilde{H}, \exists ! k \in \tilde{K}, \quad g = hk = kh$$

Démonstration. Les isomorphismes sont définis par $h \mapsto (h, e_K)$ et $k \mapsto (e_H, k)$. De plus, d'après les règles de calcul

$$(h, k)(h', e_K)(h, k)^{-1} = (hh'h^{-1}, e_K) \in \tilde{H} \text{ donc } \tilde{H} \triangleleft G$$

et de même pour $\tilde{K}$. Si (h, k) est dans $\tilde{H}$, alors $k = e_K$, et s'il est aussi dans $\tilde{K}$, c'est $e_G = (e_H, e_K)$. Enfin, $(h, k) = (h, e_K)(e_H, k) = (e_H, k)(h, e_K)$, donc tout élément de G s'écrit comme produit d'un élément de $\tilde{H}$ et d'un élément de $\tilde{K}$. L'unicité de l'écriture résulte du même calcul. $\square$

Soit maintenant G un groupe, H et K deux sous-groupes. Il est intéressant de savoir si on peut « décomposer » G à l'aide de H et de K , de sorte que G soit isomorphe au produit direct $H \times K$. Il est bien sûr nécessaire que H et K vérifient au moins les propriétés de $\tilde{H}$ et $\tilde{K}$ dans la proposition précédente. On peut (et il y a plusieurs possibilités) montrer que certaines de ces conditions sont suffisantes. L'analogie avec des sous-espaces vectoriels supplémentaires est un guide, et voici un exemple de ce que l'on peut énoncer.

Proposition 5.13. Si $HK = G$, si H et K sont distingués dans G et si $H \cap K = \{e\}$, alors G est isomorphe à $H \times K$.

Démonstration. Supposons $H \triangleleft G$, $K \triangleleft G$, et $G = HK$. Commençons par observer que, h et k étant respectivement dans H et dans K ,

$$hkh^{-1}k^{-1} = (hkh^{-1})k^{-1} = h(kh^{-1}k^{-1}) \in H \cap K$$

puisque H et K sont distingués dans G . On a donc, pour tout couple (h, k) de $H \times K$,

$$hkh^{-1}k^{-1} = e_G \quad \Rightarrow \quad hk = kh$$

Considérons maintenant $\phi : (h, k) \mapsto hk$ de $H \times K$ dans G . Par hypothèse, c'est une application surjective entre deux groupes. Montrons que c'est un morphisme :

$$\phi((h, k)(h', k')) = hh'kk' \text{ et } \phi(h, k)\phi(h', k') = hkh'k' = hh'kk'$$

Examinons l'injectivité : $hk = e$ implique $h = k^{-1}$ et cet élément est dans $H \cap K$ donc est le neutre, on a $h = k = e_G$. Le noyau de ϕ est donc réduit au neutre, ϕ est bien un isomorphisme. $\square$

Attention : le fait que $hk = kh$ pour $(h, k) \in H \times K$ n'implique pas que le groupe G soit commutatif... ne serait-ce que parce que ni H ni K n'ont nécessité à l'être.

On peut trouver d'autres critères et il est possible de généraliser à un nombre fini de sous-groupes $(H_i)_{i=1 \text{ à } n}$: comme pour des sous-espaces vectoriels en somme directe, il faut prendre l'hypothèse

$$H_i \cap \prod_{j \neq i} H_j = \{e\}$$

Exercice 5.6. Montrer que le centre $\mathcal{Z}(G)$ d'un groupe (c'est-à-dire l'ensemble des éléments de G qui commutent avec tous les éléments de G), est un sous-groupe distingué dans G .

Exercice 5.7. Montrer qu'on peut avoir $K \triangleleft H \triangleleft G$ mais $K \ntriangleleft G$. On cherchera un contre-exemple dans le groupe $\mathcal{A}_4$.

Exercice 5.8. Soient H et K deux sous-groupes de G . Montrer que HK est un sous-groupe de G si et seulement si $HK = KH$.

On suppose que l'un des deux est distingué dans G . Montrer que cette condition est réalisée.

Exercice 5.9. Soit G le groupe diédral (cf. 5.5), et $H = \langle a \rangle$, $K = \langle b \rangle$. Montrer que $H \cap K = \{e\}$, $H \triangleleft G$, $HK = KH = G$. Pourquoi G n'est-il pas isomorphe au produit direct de H par K ?

5.3 GROUPES COMMUTATIFS FINIS

Nous allons maintenant étudier les groupes commutatifs finis. Il se trouve que cette étude est assez simple, (disons les résultats, plus que les démonstrations...) et le lecteur attentif remarquera des ressemblances certaines avec des théorèmes obtenus dans le cadre de l'algèbre linéaire.

5.3.1. Exposant d'un groupe commutatif fini

Si G est un groupe fini de cardinal n , le théorème de Lagrange permet d'assurer que tout élément a un ordre d qui est un diviseur de n . On appelle **exposant** le plus grand ordre des éléments de G .

Proposition 5.14. Si G est commutatif d'exposant e , l'ordre de tout élément est un diviseur de e .

Démonstration. Commençons par un petit résultat

Lemme 5.15. Si a et b commutent et sont d'ordre n et m premiers entre eux, alors ab est d'ordre nm .

Démonstration. [du lemme] On a tout de suite $(ab)^{nm} = a^{nm}b^{nm} = e$. Par ailleurs,

$$(ab)^k = e \Rightarrow (ab)^{kn} = e \Rightarrow b^{kn} = e \Rightarrow m \mid kn$$

et donc, par le théorème de Gauss, k est un multiple de m . On démontre de même que k est un multiple de n , donc du p.p.c.m. de ces deux nombres, soit mn . L'ordre de ab est donc exactement mn . $\square$

Soit donc maintenant $x \in G$ d'ordre maximum e . Supposons qu'il existe y d'ordre d , inférieur à e mais non diviseur de e . Alors le p.p.c.m. μ de d et de e est un multiple strict de e donc est plus grand que e . De plus, il existe deux entiers d' et e' diviseurs respectifs de d et e tels que $\mu = d'e'$: cela se voit en utilisant la décomposition en facteurs premiers de μ : on sait que si

$$d = \prod_{i \in I} p_i^{d_i} \quad \text{et} \quad e = \prod_{i \in I} p_i^{e_i}$$

alors μ est le produit des mêmes premiers, avec pour exposant le maximum de d_i et e_i . Il suffit d'appeler J l'ensemble des i tels que $d_i \geq e_i$ et on aura :

$$d' = \prod_{i \in J} p_i^{d_i} \quad \text{et} \quad e' = \prod_{i \in I \setminus J} p_i^{e_i}$$

Et si on prend $a = x^{e/e'}$ et $b = y^{d/d'}$, ab est d'ordre μ . Il y a contradiction, et donc l'ordre de tout élément de G est un diviseur de e . $\square$

Un application intéressante, que nous avons déjà signalée (et utilisée) dans le chapitre sur les polynômes :

Théorème 5.16. Si K est un corps fini (donc commutatif) de cardinal q , alors $K^* = K \setminus \{0\}$ est un groupe multiplicatif cyclique.

Démonstration. Soit e l'exposant de K^* . D'après le théorème, c'est un multiple de l'ordre de tous les éléments de K^* , donc $x^e - 1 = 0$ pour tous les éléments de K^* . Mais dans un corps, une équation de degré e a au maximum e racines, C'est donc que $e = q - 1$, et il existe un élément d'ordre $q - 1$, K^* est cyclique. On peut remarquer qu'on n'utilise que l'intégrité. $\square$

5.3.2. Facteurs invariants

Bien sûr, si G est un groupe commutatif fini d'ordre n et d'exposant n , il est cyclique. Le théorème qui vient prouve que tout groupe commutatif fini est isomorphe à un produit de groupes cycliques.

Théorème 5.17. *Si G est un groupe commutatif fini, il existe une suite d'entiers $a_1, a_2, \dots, a_k$ telle que :*

$$a_1 \mid a_2 \mid \dots \mid a_k \text{ et } G \simeq C_{a_1} \times C_{a_2} \times \dots \times C_{a_k}$$

*La suite des (a_i) est unique et les (a_i) sont appelés **facteurs invariants** du groupe G . De plus, a_k est l'exposant de G et le cardinal de G est le produit des a_i .*

Démonstration. Avertissement, cette démonstration est un peu laborieuse. Elle peut être omise en première lecture.

Il est clair que a_k doit être l'exposant m de G car c'est l'ordre maximal d'un élément du groupe produit donné par le théorème. On considère donc $H = \langle x \rangle$ le groupe cyclique engendré par un élément d'ordre $m = a_k$. On va montrer qu'il existe un sous-groupe K de G tel que $G = HK \simeq H \times K$.

Soit donc K un sous-groupe d'ordre maximal parmi ceux pour lesquels $H \cap K = \{e\}$, raisonnons par l'absurde en supposant que HK , qui est isomorphe à $H \times K$ d'après le théorème 5.13, est différent de G ; c'est la partie la plus « technique » de la démonstration.

Commençons par observer qu'il existe un élément g de G dont la classe est d'ordre premier p dans le quotient G/HK : si on part d'un élément d'ordre non premier, une de ses puissances a pour ordre un diviseur premier de cet ordre (cf. théorème 5.5, (iv)). On a donc $g \notin HK$ et $g^p \in HK$ donc :

$$g^p = x^r k$$

avec x^r dans H et k dans K . Le nombre p est un diviseur de l'exposant m de G , sinon, par Bezout on aurait $pu + mv = 1$ et $g = g^{pu+mv} = g^{pu} \in HK$. Donc, il existe p' tel que $pp' = m$ et $e = g^{pp'} = x^{rp'} k^{p'}$. Comme $H \cap K = \{e\}$, $x^{rp'} = e$, et comme x est d'ordre $m = pp'$, $r = \ell p$. On a obtenu :

$$g^p = x^{\ell p} k$$

Posons $y = gx^{-\ell}$, alors $y^p = k$. Si y est dans $H = \langle x \rangle$, alors g aussi, contradiction. Si y n'est pas dans HK , le groupe engendré par y et K rencontre H ailleurs qu'en e (par maximalité de K), et il existe donc un entier q tel que $y^q = k'h'$, avec $h' \neq e$ et q non multiple de p ($k'h' = e$ équivaut à $k' = h' = e$). En définitive, $y^q \in HK$, $y^p \in HK$, par Bezout on en déduit $y \in HK$ donc $g \in HK$ contradiction...

On a donc $G = \langle x \rangle \times G'$, et on peut itérer le processus. L'ordre de tout élément de G' est un diviseur de m , exposant de G' , et comme on part d'un groupe fini, on construit ainsi un isomorphisme de G avec un produit de groupes cycliques. Nous n'allons pas montrer l'unicité de la suite des facteurs invariants ; si

$$b_1 \mid b_2 \mid \dots \mid b_{k'} \text{ et } G \simeq C_{b_1} \times C_{b_2} \times \dots \times C_{b_{k'}}$$

alors l'exposant de ce produit de groupe cyclique est clairement $b_{k'}$ donc également a_k . Il faudrait continuer, en montrant que a_{k-1} est égal à $b_{k'-1}$, mais la démonstration complète est un peu longue. $\square$

Remarquons pour conclure ce paragraphe qu'il en résulte un critère d'isomorphisme pour des groupes commutatifs finis. Il suffit qu'ils aient la même suite de facteurs invariants.

5.3.3. Diviseurs élémentaires

La décomposition obtenue dans le paragraphe précédent est la plus économique, en ce sens qu'il n'est pas possible de trouver un isomorphisme avec un produit de groupes cycliques contenant moins de facteurs (voir les exercices en fin de chapitre). Par contre, il existe une façon intéressante de décomposer un groupe commutatif fini, en produit de groupes cycliques « primaires ».

Théorème 5.18. *Soit G un groupe commutatif fini. Il existe une unique suite de groupes cycliques de la forme $C_{p_i^{a_i}}$, où p_i est un nombre premier, telle que :*

$$G \simeq \prod_{i \in J} C_{p_i^{a_i}}$$

Il y a unicité de cette suite. Un groupe cyclique de la forme $C_{p_i^{a_i}}$ est appelé groupe primaire. Les entiers $p_i^{a_i}$ sont les diviseurs élémentaires du groupe G .

Démonstration. L'existence repose uniquement sur le théorème chinois (version groupe). En effet, si n est un entier dont la décomposition en facteurs premiers est

$$n = \prod_{i=1}^{k_i} p_i^{a_i}$$

alors on a l'isomorphisme

$$C_n = \prod_{i=1}^{k_i} C_{p_i^{a_i}}$$

et, en utilisant la décomposition en groupes cycliques vue auparavant, on obtient une décomposition en groupes cycliques primaires. Et si on part d'une telle décomposition, on peut lui associer une seule suite de facteurs invariants : l'exposant du groupe produit est l'ordre maximum d'un élément, c'est le produit des puissances d'exposant maximum, pour chacun des facteurs premiers. En prenant successivement les exposants immédiatement inférieurs, on obtient la suite des facteurs invariants. Mais il est plus clair de raisonner sur un exemple : supposons que la suite des diviseurs élémentaires est 2, 2, 2, 2³, 2⁴, 3, 3, 3², 5², 5², on construit la table :

2	2	2	2 ³	2 ⁴
1	1	3	3	3 ²
1	1	1	5 ²	5 ²

et les facteurs invariants sont les produits des colonnes successives :

$$2, 2, 6, 600, 3600$$

L'unicité admise de la suite des facteurs invariants permet d'en déduire l'unicité de la suite des diviseurs élémentaires. $\square$

On aura fait le lien avec la décomposition de Frobenius faite dans le chapitre précédent. Ce n'est pas un hasard, les deux développements sont des cas particuliers d'une

théorie qui les « chapeaute », celle des modules de type fini sur un anneau principal. Donnons quelques éléments de traduction :

- Le groupe commutatif fini — l'espace E muni d'un endomorphisme u .
- L'ordre d'un élément de G — le polynôme minimal $\mu_{u,a}$ d'un vecteur.
- L'exposant du groupe — le polynôme minimal de u .
- Le cardinal du groupe — le polynôme caractéristique de u .
- Un groupe cyclique primaire — un sous-espace cyclique.

Exercice 5.10. Dans un groupe G commutatif, x est d'ordre 30 et y d'ordre 12. Donner un élément d'ordre 60.

Exercice 5.11. Déterminer les facteurs invariants et diviseurs élémentaires de $C_{10} \times C_{144} \times C_{18}$.

Exercice 5.12. Soit $\mathcal{V} = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ que l'on appelle groupe du rectangle ou groupe de Klein. Montrer qu'il existe trois sous-groupes A , B et C distincts et cycliques d'ordre 2 tels que $\mathcal{V} = A \times B = B \times C = A \times C$.

Exercice 5.13. Déterminer le nombre des groupes commutatifs ayant 60 éléments, à isomorphisme près.

Exercice 5.14. Soit G un groupe commutatif, n un entier. On pose :

$$G_n = \{g \in G \mid g^n = e\}, \quad G^n = \{g \in G \mid \exists h \in G, g = h^n\}$$

Sont-ce des sous-groupes de G ? Les déterminer dans le cas où on connaît les facteurs invariants de G .

5.4 ACTIONS DE GROUPES

5.4.1. Définitions générales

Il s'agit en quelque sorte d'un retour à la source : les groupes ont été introduits en tant que groupes de permutations des racines d'une équation. Et c'est aussi dans le cadre de la géométrie que s'est développée la théorie des groupes. On va donc préciser ce qu'on appelle l'action d'un groupe sur un ensemble. Soit G un groupe, X un ensemble non vide.

Définition 5.19. On dit que G agit sur X si on a défini une application

$$\begin{aligned} G \times X &\longrightarrow X \\ (g, x) &\longmapsto g.x \end{aligned}$$

et que, pour tout (g, g') de G^2 et x dans X ,

$$g.(g'.x) = (gg').x \quad e_G.x = x$$

D'une façon intuitive, g agit sur un élément x en le transformant en $g.x$. Donnons une définition équivalente : on note S_X le groupe des bijections de X dans lui-même.

Définition 5.20. Une action du groupe G sur l'ensemble X est un morphisme Φ de G dans S_X .

Et donc $\Phi(g)(x) = g.x$. Commençons par donner des exemples.

- G agit sur lui même par translation à gauche : $g.x = gx$, ou par translation à droite : $g.x = xg^{-1}$. Il s'agit bien d'actions. Par exemple, pour la translation à droite, on écrit

$$g.(g'.x) = g.(xg'^{-1}) = xg'^{-1}g^{-1} = x(gg')^{-1} = (gg').x$$

- Un mélange des deux : $g.x = gxg^{-1}$; c'est l'action par conjugaison, que nous rencontrerons souvent.
- Si $G = \mathbf{GL}(E)$ est le groupe linéaire d'un espace vectoriel, alors G agit sur E , mais aussi sur $\mathcal{D}$ l'ensemble des droites vectorielles, ou sur $\mathcal{B}$ l'ensemble des bases. On posera $g.x = g(x)$, $g.D = g(D)$ et, dans le dernier cas :

$$g.B = g.(e_1, e_2, \dots, e_n) = (g(e_1), g(e_2), \dots, g(e_n))$$

Les vérifications nécessaires sont immédiates.

Une action peut avoir différentes propriétés.

Définition 5.21.

(1) Une action est fidèle si :

$$\forall g \in G, \quad (\forall x \in X, (g.x = x)) \Rightarrow (g = e_G)$$

autrement dit, le noyau de Φ est réduit à e_G .

(2) Une action est transitive si :

$$\forall (x, y) \in X^2, \exists g \in G, \quad y = g.x$$

Si une action n'est pas fidèle, on peut faire agir $G/\text{Ker } \Phi$ sur X , et l'action induite sera fidèle. Si par contre une action est fidèle, et ce sera souvent le cas en géométrie, Φ identifie G à un sous-groupe du groupe des bijections de X dans X . On peut dire que G est un groupe de transformations de X .

Par ailleurs, le défaut de transitivité se mesure ainsi.

Définition 5.22. On appelle orbite d'un x dans X l'ensemble des $g.x$ où g parcourt G .

L'orbite de x est notée $\mathcal{O}_x$ ou parfois Gx . Une action est transitive si et seulement si il n'y a qu'une orbite. Être dans la même orbite est une relation d'équivalence, et les orbites partitionnent l'ensemble X .

Enfin, si x est un élément de X , on définit son **stabilisateur** par :

Définition 5.23.

$$G_x = \{g \in G \mid g.x = x\}$$

C'est un sous-groupe de G comme on le vérifie facilement. Et voici le premier, et seul théorème de toute cette introduction.

Théorème 5.24.

$$\mathcal{O}_x \simeq G/G_x$$

où $\simeq$ représente une bijection. Cette bijection n'a, a priori, aucune raison d'être un morphisme, puisque ni $\mathcal{O}_x$ ni G/G_x n'ont de structure de groupe en général.

Démonstration. On pose $g.G_x \mapsto g.x$. C'est bien défini car si $g.G_x = g'.G_x$, on a $g^{-1}g' \in G_x$ et $g.x = g'.x$, c'est surjectif par choix de l'ensemble d'arrivée, et le calcul qu'on vient de faire montre que c'est injectif. $\square$

Dans le cas où X est fini, on en déduit la formule des classes :

Théorème 5.25.

$$\text{Card}(X) = \sum_{i=1}^n \text{Card}(G/G_{x_i})$$

où (x_i) sont des représentants des n orbites.

Démonstration. C'est la conséquence de ce que des classes d'équivalence forment une partition de X . $\square$

5.4.2. Applications aux groupes

Examinons l'action par translation.

Théorème 5.26. Théorème de Cayley. Tout groupe G est isomorphe à un sous-groupe du groupe symétrique S_G .

Démonstration. L'action par translation est fidèle, puisque $gx = x$ pour tout x (et même pour seulement un x) implique que $g = e_G$. Il y a donc un morphisme injectif de G dans S_G . $\square$

Dans le cas fini, un groupe de cardinal n est donc isomorphe à un sous-groupe du groupe symétrique $\mathcal{S}_n$. En théorie, la connaissance des groupes $\mathcal{S}_n$ et de leurs sous-groupes suffit à connaître tous les groupes finis. En pratique... il n'est pas si « pratique » que ça d'étudier le groupe $\mathcal{S}_4$ à $4! = 24$ éléments pour étudier les groupes à quatre éléments.

Continuons avec l'action par conjugaison. Si a est dans G , alors l'application définie par $g \mapsto a.g = aga^{-1}$ est non seulement une bijection de G dans G , mais aussi un morphisme de groupe donc un automorphisme de G . On dit que c'est un **automorphisme intérieur** de G . Si on examine un peu plus précisément comment fonctionnent ces automorphismes intérieurs, on obtient la proposition

Proposition 5.27. *L'ensemble $\text{Int}(G)$ des automorphismes intérieurs est un sous-groupe distingué du groupe $\text{Aut}(G)$ des automorphismes de G . De plus, $\text{Int}(G) \simeq G/\mathcal{Z}(G)$ où $\mathcal{Z}(G)$ est le centre de G .*

Démonstration. Si on note i_a l'application définie par $i_a(x) = axa^{-1}$, on sait déjà que c'est une bijection de G , c'est un morphisme :

$$i_a(gg') = agg'a^{-1} = aga^{-1}ag'a^{-1} = i_a(g)i_a(g')$$

De plus, l'application $a \mapsto i_a$ est un morphisme du G dans $\text{Aut}(G)$:

$$i_a \circ i_b(g) = a(bgb^{-1})a^{-1} = (ab)g(ab)^{-1} = i_{ab}(g)$$

On en déduit que l'image de ce morphisme, ensemble des automorphismes intérieurs, est un sous-groupe de $\text{Aut}(G)$. De plus, $aga^{-1} = g \iff ag = ga$ et le noyau de ce morphisme est donc $\mathcal{Z}(G)$, le théorème d'isomorphisme donne $\text{Int}(G) \simeq G/\mathcal{Z}(G)$. Enfin, si ϕ est un automorphisme de G

$$\phi \circ i_a \circ \phi^{-1} = i_{\phi(a)}$$

et donc $\text{Int}(G) \triangleleft \text{Aut}(G)$. □

Une application importante concerne les groupes finis de cardinal p^n , que l'on appelle p -groupes.

Théorème 5.28. *Le centre d'un p -groupe n'est jamais trivial.*

Démonstration. On applique la formule des classes, en distinguant les orbites réduites à un élément : si on note $\mathcal{Z}(G)$ le centre de G , alors tout x de $\mathcal{Z}(G)$ vérifie $gxg^{-1} = x$ pour tout g ; il est donc seul dans son orbite, et

$$\text{Card}(G) = \text{Card}(\mathcal{Z}(G)) + \sum_i \text{Card}(\mathcal{O}_{x_i})$$

où les orbites $\mathcal{O}_{x_i}$ sont celles qui ont plus d'un élément ; elles ont pour cardinal celui de G/G_{x_i} , qui est une puissance de p puisque G est de cardinal une puissance de p et que tout sous-groupe a pour cardinal une puissance de p . On en déduit que le cardinal du centre est divisible par p , et donc que ce centre contient un autre élément que l'élément neutre (en réalité donc, il contient au moins p éléments). □

En particulier un groupe de cardinal p^n ($n > 1$) n'est jamais simple : il a au moins un sous-groupe distingué, son centre.

Exercice 5.15. Soit dans $\mathcal{S}_n$ un k -cycle σ . Montrer que l'ensemble des conjugués de σ est l'ensemble des k -cycles. Généraliser à l'étude des conjugués d'une permutation quelconque. Combien y a-t-il de classes de conjugaison dans $\mathcal{S}_n$ pour $n = 3, 4, 5, \dots$?

Exercice 5.16. Un groupe G agit sur un ensemble X . Montrer que si x et y sont dans la même orbite, alors leurs stabilisateurs sont conjugués.

Exercice 5.17. On dit qu'une action est **simplement transitive** si

$$\forall x \in X, \forall y \in X, \exists ! g \in G, \quad y = g.x$$

Montrer que si l'action est fidèle, transitive et si G est commutatif, alors l'action est simplement transitive. On rencontrera des actions simplement transitives dans les chapitres de géométrie.

5.5 THÉORÈMES DE SYLOW

Les théorèmes de Sylow font partie des théorèmes les plus célèbres de la théorie des groupes finis ; leur importance est due à ce qu'il s'agit de théorèmes généraux, qui s'appliquent à n'importe quel groupe fini, et affirment l'existence de sous-groupes dont l'ordre est une puissance d'un nombre premier. C'est une réponse partielle au problème soulevé par le théorème de Lagrange : si n est le cardinal d'un groupe et si d est un diviseur de n , existe-t-il un sous-groupe de G de cardinal d ? Mais c'est surtout un premier pas vers la classification des groupes finis. Cette classification est maintenant achevée, plus de cent ans après la démonstration des théorèmes de Sylow.

5.5.1. Conjugués d'un sous-groupe

Si G est un groupe, il agit sur l'ensemble de ses sous-groupes par conjugaison. Les conjugués d'un sous-groupe H sont de la forme gHg^{-1} où g est dans G . Ils sont isomorphes à H puisque tout automorphisme intérieur est.. un automorphisme. Et un sous-groupe est distingué dans G s'il coïncide avec tous ses conjugués, il est seul dans son orbite. On peut compléter ces observations par la proposition :

Proposition 5.29. Si H est un sous-groupe de G , l'ensemble des g tels que $gHg^{-1} = H$ est un sous-groupe de G contenant H , on l'appelle **normalisateur** de H dans G , et on le note $\mathcal{N}_G(H)$. Le nombre des conjugués de H est égal à l'indice de son normalisateur.

Démonstration. On utilise l'action par conjugaison de G sur l'ensemble des sous-groupes de G : le normalisateur de H est le stabilisateur de H pour cette action, c'est donc un sous-groupe de G et il contient bien sûr H . Le cardinal de l'ensemble des conjugués de H , c'est-à-dire de son orbite, est l'indice du normalisateur. On a également $H \triangleleft \mathcal{N}_G(H) \leq G$, et si H est distingué dans G , c'est que $\mathcal{N}_G(H) = G$. $\square$

5.5.2. Les théorèmes de Sylow

Nous allons les rassembler en un seul énoncé :

Théorème 5.30. Soit G un groupe fini de cardinal n . On suppose que p est un nombre premier diviseur de n et on pose $n = p^\alpha m$ où $m \wedge p = 1$.

- G admet des sous-groupes d'ordre p^α , que l'on appelle p -Sylow de G .
- Tout sous-groupe de G d'ordre une puissance de p (plus rapidement tout p -sous-groupe de G) est inclus dans un p -Sylow. Tous les p -Sylow de G sont conjugués.
- Le nombre s des p -Sylow de G vérifie : $s \equiv 1 \pmod{p}$ et $s \mid m$.

Démonstration.

- Soit $\mathcal{X}$ l'ensemble des sous-ensembles de G ayant p^α éléments. Le groupe G agit sur $\mathcal{X}$ par translation à gauche. Le cardinal de $\mathcal{X}$ n'est pas divisible par p . Une façon de le voir est de l'écrire ainsi :

$$\binom{p^\alpha m}{p^\alpha} = \frac{p^\alpha m}{p^\alpha} \frac{p^\alpha m - 1}{p^\alpha - 1} \cdots \frac{p^\alpha m - p^\alpha + 1}{1}$$

et chaque fraction se simplifie : pour la première par p^α , mais $p^\alpha m - i \equiv p^\alpha - i \pmod{p^\alpha}$, donc les autres fractions se simplifient par la même puissance de p , de sorte que le coefficient restant est un entier non divisible par p . Comme les orbites forment une partition de $\mathcal{X}$, il existe au moins une orbite dont le cardinal n'est pas divisible par p . Soit $\mathcal{O}$ une telle orbite. Si X est un élément de $\mathcal{O}$, il est formé de p^α éléments du groupe G . Sous l'action, son stabilisateur S est un sous-groupe de G dont l'indice n'est pas divisible par p (c'est le cardinal de $\mathcal{O}$) ; le cardinal de S est donc de la forme $p^\alpha m'$ où m' est un diviseur de m . On va montrer que S est un p -Sylow de G . Considérons en effet X . C'est un ensemble de p^α éléments de G , $x_1, \dots, x_{p^\alpha}$. Et si $g \in S$, on a $gX = X$ donc $gx_1 = x_i$ soit $g = x_i x_1^{-1}$, où l'indice i est entre 1 et p^α . Il y a donc au maximum p^α possibilités pour g , donc le cardinal de S est inférieur ou égal à p^α . Compte-tenu de ce qui précède, ce cardinal est égal à p^α , on a obtenu un p -Sylow.

- Soit maintenant H un groupe dont le cardinal est une puissance de p . On peut considérer l'action de H sur l'ensemble G/S , définie par $h.gS = (hg)S$. Dans cette action, les orbites ont pour cardinal une puissance de p mais G/S est de cardinal m , non divisible par p . Il y a donc au moins une orbite ayant un seul élément, disons gS . On a donc, pour tout $h \in H$, $hgS = gS$, et donc $H \subset gSg^{-1}$.

Tout p -groupe est inclus dans un conjugué du p -Sylow S . En particulier, tout p -Sylow est un conjugué de S .

- Commençons par la remarque : d'après la question précédente, si un p -Sylow d'un groupe est distingué dans ce groupe, il est le seul p -Sylow. Faisons ensuite agir S sur l'ensemble des p -Sylow de G , par conjugaison : les orbites ont encore comme cardinal une puissance de p , et S est seul dans son orbite. Mais si S' est un autre p -Sylow seul dans son orbite, $XS'X^{-1} = S'$ pour tout x de S , donc S est un p -Sylow du normalisateur de S' , d'après la remarque, $S = S'$. L'ensemble des p -Sylow est donc réunion d'orbites dont une a un seul élément, les autres ont pour cardinal une puissance de p , on a donc $s \equiv 1 \pmod{p}$. Par ailleurs, le nombre des p -Sylow est donc le nombre des conjugués de S . C'est donc l'indice du normalisateur de S , et c'est donc un diviseur du cardinal de G , et donc c'est un diviseur de m .

□

La démonstration est un peu délicate... Mais le résultat est important : au vu du seul cardinal d'un groupe, on peut faire certaines déductions sur ses éventuels sous-groupes. Donnons quelques exemples simples :

- Soit G un groupe ayant 42 éléments. Alors il admet au moins un sous-groupe distingué. En effet, G admet un 7-Sylow. Le nombre de ses conjugués, qui sont les éventuels autres p -Sylow est un diviseur de 6 qui est congru à 1 modulo 7, c'est 1. Ce 7-Sylow est donc distingué dans G .
- Le groupe $\mathbb{Z}/24\mathbb{Z} \times \mathbb{Z}/10\mathbb{Z}$ n'a qu'un seul sous-groupe à 16 éléments. En effet, un 2-Sylow a 16 éléments, et puisque le groupe est commutatif, il coïncide avec ses conjugués, il n'y a pas d'autre 2-Sylow.

Exercice 5.18. Vérifier que le normalisateur de H est le plus grand groupe dans lequel H est distingué. Chercher le normalisateur des sous-groupes à quatre éléments et à deux éléments du groupe diédral $\mathbb{D}_8$ (cf. exercice 5.5). Si deux sous-groupes d'un groupe G sont isomorphes, sont-ils conjugués ?

Exercice 5.19. Chercher les 2-Sylow de S_4 .

Exercice 5.20. Montrer que si p divise le cardinal de G , alors il existe au moins un élément d'ordre p (c'est le théorème de Cauchy).

EXERCICES

Exercice 5.21. Décrire l'ensemble des morphismes de C_n dans lui-même ; lesquels sont des isomorphismes ?

Exercice 5.22. On note $\mathbb{H}_8$ le groupe engendré par les matrices

$$a = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{et} \quad b = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

Montrer qu'il a huit éléments et déterminer ses sous-groupes. Vérifier qu'il n'est pas isomorphe au groupe diédral $\mathbb{D}_8$. On l'appelle groupe quaternionique.

Exercice 5.23. Soit G un groupe fini dont tous les éléments (hormis le neutre) sont d'ordre p premier. Démontrer que G est isomorphe au produit direct de groupes cycliques d'ordre p . On montrera que G est muni d'une structure d'espace vectoriel sur le corps $\mathbb{F}_p$.

Exercice 5.24.

- (1) Soit G un groupe, H un sous-groupe, On considère l'action de H sur G/H définie par

$$h.gH = hgH$$

Montrer que le noyau de cette action est le plus grand sous-groupe distingué de G inclus dans H .

- (2) On suppose que G est fini et que p le plus petit entier divisant le cardinal de G . On suppose qu'il existe H sous-groupe de G d'indice p . Montrer que H est distingué dans G .
- (3) Démontrer directement que si H est d'indice 2 dans G alors H est distingué dans G .
-

Exercice 5.25. Montrer que si G est un groupe fini de cardinal $p^\alpha m$ où p est premier et $m \wedge p = 1$, alors il existe un sous-groupe d'ordre p^β pour tout entier $\beta \leq \alpha$.

PROBLÈMES

Les corrigés de ces problèmes sont disponibles sur le site de Dunod : www.dunod.com.

5.1. PROBLÈME

Soit G un groupe. Si x et y sont des éléments de G , on appelle **commutateur** de x et de y l'élément $xyx^{-1}y^{-1}$ que l'on note $[x, y]$. Le nom de cet élément s'éclaircit si on remarque que

$$[x, y] = e \iff xy = yx$$

Le but de ce problème est d'étudier certaines propriétés des commutateurs. On appelle **groupe dérivé** le sous-groupe de G engendré par les commutateurs. On le note G' .

- (1) Montrer que le conjugué d'un commutateur est un commutateur. En déduire que G' est distingué dans G .
- (2) Soit H distingué dans G . Montrer que G/H est commutatif si et seulement si H contient G' .
- (3) D'après la question précédente, G/G' est le plus « grand » quotient commutatif de G . On l'appelle l'**abélianisé** de G . Déterminer le groupe dérivé de S_n . Déterminer l'abélianisé du groupe diédral $\mathbb{D}_8$, du groupe quaternionique.
- (4) Montrer que tout sous-groupe H vérifiant $G' \leq H \leq G$ est distingué dans G .

5.2. PROBLÈME

Le groupe symétrique S_n est le groupe des permutations de $\mathbb{I}_n = \{1, 2, \dots, n\}$ dans lui-même. Nous allons retrouver des résultats connus sur ce groupe symétrique en utilisant les idées de style « action de groupe ».

- (1) Soit $\sigma \in S_n$. Le groupe $\langle \sigma \rangle$ agit naturellement sur $\mathbb{I}_n$. on dit que σ est un r -cycle si il y a une orbite de cardinal r , (que l'on appelle **support** de σ) et si toutes les autres orbites sont de cardinal 1. Montrer que les éléments du support peuvent être ordonnés en une suite $(a_i)_{i=1 \text{ à } r}$ telle que

$$\sigma(a_i) = a_{i+1} \text{ pour } i < r, \sigma(a_r) = a_1$$

On écrit alors $\sigma = (a_1, a_2, \dots, a_r)$.

- (2) Montrer que deux cycles de supports disjoints commutent, et que toute permutation se décompose de façon unique en produit de cycles de support disjoints.
- (3) On suppose que $\sigma = c_1 c_2 \dots c_k$ où c_i est un r_i -cycle, les supports étant disjoints. Montrer que l'ordre de σ est le p.p.c.m. des r_i .
- (4) Si σ est une permutation quelconque, on note $\epsilon(\sigma) = (-1)^{n-k}$, où k est le nombre total des orbites de l'action de $\langle \sigma \rangle$ sur $\mathbb{I}_n$. Ce nombre est la signature de σ . Calculer la signature d'un r -cycle.

- (5) On rappelle qu'une transposition $\tau = (i, j)$ est un 2-cycle. Montrer qu'un r -cycle est composé de r transpositions, et donc que les transpositions engendrent le groupe $\mathcal{S}_n$.
- (6) En déduire que la signature est un morphisme de $\mathcal{S}_n$ dans le groupe multiplicatif $\{1, -1\}$. Le noyau de ce morphisme est le groupe alterné $\mathcal{A}_n$ des permutations paires.
- (7) Y a-t-il d'autres morphismes de $\mathcal{S}_n$ dans le groupe multiplicatif $\{1, -1\}$?

SOLUTIONS DES EXERCICES

Solution 5.1. Soit G un groupe d'ordre p , premier et $x \in G$, différent du neutre. Alors $\langle x \rangle$ est un sous-groupe de G , non réduit au neutre et d'ordre diviseur de p : cet ordre est donc égal à p et x est un générateur de G , qui est donc cyclique. Il existe à isomorphisme près, un seul groupe d'ordre p , le groupe cyclique C_p . Cette situation est assez exceptionnelle : il existe par exemple une quinzaine de groupes d'ordre 24, un seul d'ordre 15 ou 33, des centaines de milliers d'ordre 512, à isomorphisme près, bien sûr.

Solution 5.2. Commençons par observer que :

$$\forall (h, h') \in H^2, hh' \in H \iff HH \subset H$$

et

$$\forall h \in H, h^{-1} \in H \iff H^{-1} \subset H$$

donc, si $HH = H$ et $H^{-1} = H$, alors H est un sous-groupe de G . Réciproquement, si H est un sous-groupe de G , on a donc $HH \subset H$, mais comme $1_G \in H$, $h = 1_G h$ implique que $H \subset HH$. Par ailleurs, $H^{-1} \subset H$, mais $h = (h^{-1})^{-1}$ prouve que $H \subset H^{-1}$ lorsque H est un sous-groupe de G .

Solution 5.3. Commençons par observer que dans un groupe G , l'application $\phi : x \mapsto x^{-1}$ est bien définie et bijective : son application réciproque est elle-même car $(x^{-1})^{-1} = x$. Supposons G commutatif : $(xy)^{-1} = y^{-1}x^{-1} = x^{-1}y^{-1}$, pour tous x et y . Cela s'écrit $\phi(xy) = \phi(x)\phi(y)$, et ϕ est bien un morphisme. La réciproque résulte du même argument :

$$xy = \phi(x^{-1})\phi(y^{-1}) = \phi(x^{-1}y^{-1}) = \phi((yx)^{-1}) = yx$$

Solution 5.4. Si $x \in G$ est d'ordre p et si ϕ est un morphisme de source G , alors $\phi(x)^p = \phi(x^p) = \phi(e_G) = e_H$. On en déduit que $\phi(x)$ est d'ordre fini, diviseur de p . Mais si ϕ est injectif,

$$\phi(x)^m = e_H \Rightarrow \phi(x^m) = \phi(e_G) \Rightarrow x^m = e_G$$

prouve qu'alors l'ordre de $\phi(x)$ est un multiple de p : un morphisme injectif conserve l'ordre des éléments. Même type de résultat si x est d'ordre infini.

Solution 5.5. Un calcul simple montre que a est d'ordre 4, et que b est d'ordre 2. Par ailleurs, $ab = ba^3$, $a^2b = ba^2$, $a^3b = ba$, ces trois matrices étant distinctes. Ces formules montrent également que tous les produits successifs de puissances de a par des puissances de b se ramènent à huit éléments distincts :

$$\mathbb{D}_8 = \{e, a, a^2, a^3, b, ab, a^2b, a^3b\}$$

Les quatre premiers éléments forment un groupe cyclique, qui contient un sous-groupe d'ordre 2, $\{e, b^2\}$; les quatre autres éléments sont d'ordre 2, ils engendrent chacun un groupe d'ordre 2. On peut se demander s'il existe d'autres sous-groupes : si c'est le cas, ce sont des groupes non cycliques, d'ordre diviseur strict de huit : ce ne peut être que des groupes d'ordre 4. On constate qu'il y en a 2, $\langle a^2, b \rangle$ et $\langle a^2, ab \rangle$. Ces groupes sont isomorphes à $C_2 \times C_2$ (voir le paragraphe suivant). L'étude du groupe diédral est facilitée quand on en fait une interprétation géométrique : il est isomorphe au groupe des isométries qui conservent un carré.

Solution 5.6. Le centre n'est jamais vide car il contient toujours l'élément neutre. Si z et z' sont dans le centre de G , alors, pour tout x de G :

$$(zz')x = z(z'x) = z(xz') = (zx)z' = x(zz')$$

et

$$z^{-1}x = (x^{-1}z)^{-1} = (zx^{-1})^{-1} = xz^{-1}$$

donc $\mathcal{Z}(G)$ est un sous-groupe de G . De plus, $xzx^{-1} = z$ donc $\mathcal{Z}(G)$ est distingué dans G . Le centre d'un groupe est une sorte de mesure de sa commutativité, par exemple, si $\mathcal{Z}(G) = G$, alors G est commutatif. Il existe des groupes pour lesquels le centre se réduit au neutre, c'est le cas de $\mathcal{S}_n$, pour $n \geq 3$.

Solution 5.7. Rappelons que $\mathcal{A}_4$ est le groupe des permutations paires de 4 éléments. Il contient les doubles transpositions $u = (1, 2)(3, 4)$, $v = (1, 3)(2, 4)$ et $w = (1, 4)(2, 3)$. Avec le neutre, ces quatre doubles transpositions forment un groupe, $\mathcal{V}$, qui est distingué dans $\mathcal{A}_4$: on vérifie en effet que, si σ est une permutation,

$$\sigma \circ (1, 2)(3, 4) \circ \sigma^{-1} = (\sigma(1), \sigma(2))(\sigma(3), \sigma(4))$$

De plus, comme $\mathcal{V}$ est commutatif, son sous-groupe $\langle u \rangle$ est distingué dans $\mathcal{V}$. Mais $\langle u \rangle$ n'est pas distingué dans $\mathcal{A}_4$, il suffit de calculer le conjugué de u par $(1, 2, 3)$, par exemple. La relation $\triangleleft$ n'est pas transitive, alors que c'est le cas de la relation $\leq$ (est un sous-groupe).

Solution 5.8. On suppose que H et K sont des sous-groupes et que $HK = KH$. Alors $HKHK = HHKK = HK$ et $(HK)^{-1} = K^{-1}H^{-1} = KH = HK$. Donc HK est un sous-groupe. Supposons maintenant que H , K et HK sont des sous-groupes de G . Alors

$$HK = (HK)^{-1} = K^{-1}H^{-1} = KH$$

On a bien prouvé $HK = KH$.

Si par exemple H est distingué dans G , alors, quelque soient h et k dans H et K respectivement, $khk^{-1} = h' \in H$ et $KH \subset HK$. On a aussi $k^{-1}hk = h'' \in H$ et $HK \subset KH$. On peut aussi utiliser que HK est la réunion des classes à droites modulo H , et KH est la réunion des classes à gauche.

Solution 5.9. On a d'abord $H = \langle a \rangle = \{e, a, a^2, a^3\}$ et $K = \langle b \rangle = \{e, b\}$. L'étude faite de ce groupe dans l'exercice 5.5 permet de vérifier facilement les propriétés indiquées. Pourquoi G n'est-il pas isomorphe au produit direct de H par K ? On peut argumenter en disant que K n'est pas distingué dans G où, plus rapidement, que H et K étant commutatifs, on vérifie immédiatement que leur produit direct est commutatif. On peut trouver dans l'ensemble $H \times K$ une structure de groupe telle que ce groupe soit isomorphe à G : voir la notion de produit semi-direct, par exemple dans [6].

Solution 5.10. C'est une simple application des idées de la démonstration du théorème 5.14. Si x est d'ordre 30 et y d'ordre 12, le p.p.c.m. est 60 ; mais $60 = 15 \times 4$ donc x^2 qui est d'ordre 15 et y^3 qui est d'ordre 4 commutent et ont des ordres premiers entre eux, l'ordre de x^2y^3 est 60.

Solution 5.11. D'après le théorème chinois :

$$C_{10} \times C_{144} \times C_{18} \simeq (C_2 \times C_5) \times (C_{16} \times C_9) \times (C_2 \times C_9)$$

Les facteurs invariants sont $2, 2, 2^4, 3^2, 3^2, 5$ et les diviseurs élémentaires sont $2 \mid 6 \mid 720$. Le groupe de départ est isomorphe à $C_2 \times C_6 \times C_{720}$.

Solution 5.12. Il suffit d'observer qu'il y a trois éléments d'ordre 2, $(\bar{1}, \bar{0})$, $(\bar{0}, \bar{1})$ et $(\bar{1}, \bar{1})$ et de prendre pour A , B et C les groupes engendrés par chacun de ces trois éléments : il n'y a pas unicité de la décomposition en sous-groupes primaires, mais unicité de la liste des facteurs invariants.

Solution 5.13. $60 = 2^2 \times 3 \times 5$, il y a seulement 2 types de groupes commutatifs à 60 éléments, celui de facteurs invariants $2^2, 3, 5$ et celui de facteurs invariants $2, 2, 3, 5$. Le nombre de possibilités sera plus important s'il existe des exposants élevés dans la décomposition en facteurs premiers du cardinal du groupe. On pourra faire le lien avec le nombre de façons d'écrire un entier comme somme d'autres entiers... et se reporter à la situation analogue en algèbre linéaire.

Solution 5.14. Soit x et y dans G_n . Alors $(xy)^n = x^n y^n = e$ et $(x^{-1})^n = (x^n)^{-1} = e$. G_n est un sous-groupe, il est formé des éléments dont l'ordre est un diviseur de n . Si par exemple $G = \mathbb{C}^*$ muni du produit, c'est le groupe cyclique des racines de l'unité. Supposons G commutatif fini. Pour que G_n ne soit pas réduit au neutre, il est nécessaire que n et le cardinal de G ne soient pas premiers entre eux.

De même, G^n est un sous-groupe de G , puisque si x et y en font partie, on a $x = h^n$ et $y = k^n$ donc $xy = (hk)^n$ et $x^{-1} = (h^{-1})^n$.

Si G est un produit de groupes commutatifs, $G = G_1 \times \dots \times G_k$, alors la définition de la loi dans le groupe produit permet de vérifier facilement que $G_n = G_{1,n} \times \dots \times G_{k,n}$ et $G^n = G_1^n \times \dots \times G_k^n$. Il suffit donc de déterminer G_n et G^n lorsque G est un groupe cyclique de la forme C_{p^α} . Mais les éléments de ce groupe sont d'ordre un diviseur de p^α et, en notant x un générateur de C_{p^α} , on aura $G_{p^\beta} = \langle x^{p^{\alpha-\beta}} \rangle$.

Solution 5.15. Si $\sigma = (i_1, i_2, \dots, i_k)$, alors on vérifie directement que

$$s \circ \sigma \circ s^{-1} = (s(i_1), s(i_2), \dots, s(i_k))$$

Comme le choix de σ est libre, l'ensemble des conjugués d'un k -cycle est l'ensemble des k -cycles. On sait que toute permutation se décompose en cycles de supports disjoints : les mêmes calculs prouvent que l'ensemble des conjugués d'une permutation est l'ensemble des permutations qui ont même structure (type de décomposition en cycle). Ainsi, dans $\mathcal{S}_3$ il y a trois classes de conjugaison : celle qui contient l'identité, celle qui contient les transpositions, celle qui contient les 3-cycles, cela correspond aux décompositions en cycles du type :

$$3 = 1 + 1 + 1 = 1 + 2 = 3 \quad \text{et de même}$$

$$4 = 1 + 1 + 1 + 1 = 2 + 1 + 1 + 1 = 2 + 2 = 3 + 1$$

permet de dire qu'il y a quatre classes de conjugaison, et plus généralement le nombre des classes de conjugaison de $\mathcal{S}_n$ est le nombre des partitions de n , déjà rencontré.

Solution 5.16. Supposons que x et y sont dans la même orbite. Il existe g tel que $y = g.x$. En notant G_x et G_y les stabilisateurs de x et de y , on a :

$$g \in G_y \iff g.y = y \iff g.(h.x) = h.x \iff h^{-1}gh \in G_x \iff g \in hG_xh^{-1}$$

et donc G_y est un sous-groupe conjugué de G_x . Tous les stabilisateurs sont conjugués, donc en particulier sont isomorphes.

Solution 5.17. Soit x et y deux éléments quelconques de X . Comme l'action est transitive, il existe $g \in G$ tel que : $y = g.x$. Supposons que $y = h.x$, et soit z quelconque ; il existe k tel que $z = k.x$ et

$$g.z = (gk).x = (kg).x = k.y = (kh).x = (hk).x = h.z$$

(en utilisant la commutativité). On a donc, pour tout z , $(h^{-1}g).z = z$, et, puisque l'action est fidèle, $h^{-1}g = e$ et en définitive $h = g$. L'action est simplement transitive. La situation décrite se rencontre fréquemment en géométrie, par exemple pour l'action par translation des vecteurs sur un espace affine.

Solution 5.18. C'est pratiquement la définition : le normalisateur de H est l'ensemble de tous les g tels que $gHg^{-1} = H$, et donc H est inclus dans son normalisateur $\mathcal{N}_G(H)$ et il y est distingué. Et tout groupe dans lequel H est normal doit être formé d'éléments g tels que $gHg^{-1} = H$, il est inclus dans le normalisateur. Si bien sûr le normalisateur est G , H est distingué dans G ; il arrive que H soit égal à son normalisateur, on parle de sous-groupe « autonormalisant ». Dans le cas du groupe diédral $\mathbb{D}_8$, le groupe $\langle a \rangle$ est distingué dans $\mathbb{D}_8$: son normalisateur est donc lui-même. Les sous-groupes $\langle a^2, b \rangle$ et $\langle a^2, ab \rangle$ qui ont également quatre éléments sont également distingués, ils ont quatre éléments mais ne sont pas cycliques.

Occupons-nous des sous-groupes à deux éléments. Le sous-groupe $\langle a^2 \rangle$ est le centre, il est distingué. Le sous-groupe $\langle b \rangle$ n'est pas distingué, et comme b ne commute qu'avec lui-même et le neutre, il est égal à son normalisateur. Il a donc quatre conjugués, lui-même et les trois derniers sous-groupes à 2 éléments, $\langle ab \rangle$, $\langle a^2b \rangle$, $\langle a^3b \rangle$.

Si deux sous-groupes de G sont conjugués, alors ils sont isomorphes, puisque $x \mapsto gxg^{-1}$ est un isomorphisme. Mais deux sous-groupes de G qui sont isomorphes ne sont pas forcément conjugués : l'exemple ci-dessus des sous-groupes à deux éléments du groupe diédral le prouve.

Solution 5.19. Le groupe $\mathcal{S}_4$ a 24 éléments. Ses 2-Sylow ont donc 8 éléments, et d'après le théorème de Sylow, il y en a 1 ou 3. Les éléments d'un 2-Sylow sont d'ordre 1, 2 ou 4 (il n'y a pas d'élément d'ordre 8). Si par exemple on part d'un élément d'ordre 4 comme $\sigma = (1, 2, 3, 4)$, on obtient un sous-groupe d'ordre 8 en prenant les composés avec un élément d'ordre 2 comme $\tau = (1, 3)$. On a alors $\tau\sigma\tau = \sigma^{-1}$ et on peut vérifier que le groupe $\langle \sigma, \tau \rangle$ est isomorphe au groupe diédral $\mathbb{D}_8$. On vérifiera qu'il y a d'autres Sylow (deux autres donc) construit sur le même modèle. Ces trois 2-Sylow sont autonormalisant (voir l'exercice précédent).

Solution 5.20. Il suffit d'utiliser le théorème de Sylow : si p divise le cardinal de G , il existe un p -Sylow S , qui est d'ordre $p\alpha$. Tout élément x de S est d'ordre une puissance de p . Et dans le groupe (cyclique) engendré par x , il existe donc un élément d'ordre exactement p . (cf. la remarque à la fin du théorème 5.5).

Cauchy n'a pas démontré son théorème à l'aide du théorème de Sylow, ne serait-ce que pour des raisons chronologiques... Par contre, certaines démonstrations du théorème de Sylow utilisent le théorème de Cauchy (qu'il faut donc démontrer autrement).

Solution 5.21. Comme C_n est monogène, un morphisme est entièrement déterminé si on se donne l'image du générateur x . On aura donc n morphismes ϕ_k définis par $x \mapsto x^k$ et donc $x^\ell \mapsto x^{k\ell}$. De plus ϕ_k sera bijectif si x^k engendre C_n , donc si k est premier à n . On vérifiera facilement que $\phi_k \mapsto \bar{k}$ est un morphisme de l'ensemble des ϕ_k (muni de la loi $\circ$, dans le groupe additif $(\mathbb{Z}/n\mathbb{Z}, +)$). Et qu'alors l'ensemble des automorphismes de C_n est isomorphe à l'ensemble des inversibles de l'anneau $\mathbb{Z}/n\mathbb{Z}$.

Solution 5.22. Il faut calculer... a est d'ordre 4, $b^2 = a^2$, $ba = a^3b$ et les éléments du groupe engendré par a et b sont huit :

$$e, a, a^2, a^3, b, ab, a^2b, a^3b$$

Ce groupe n'est pas isomorphe au groupe diédral car il a un seul élément d'ordre 2, $a^2 = b^2$. Il a donc un seul sous-groupe d'ordre 2. On vérifiera que a , b et ab engendrent les trois groupes d'ordre 4 (qui sont donc cycliques). On l'appelle groupe quaternionique car il est relié au corps $\mathbb{H}$ des quaternions.

Solution 5.23. Montrons que G peut être muni d'une structure d'espace vectoriel : c' est un groupe commutatif pour la loi du groupe (que l'on notera additivement). Si $\bar{k} \in \mathbb{F}_p$ (on rappelle que c' est le corps $\mathbb{Z}/p\mathbb{Z}$), on définit $\bar{k}.x = x + x + \cdots + x$ avec k termes dans la somme. Cette définition est cohérente, puisque $\bar{k}.x = \bar{k}'.x$ si et seulement si $k \equiv k' \pmod{p}$, car x est d'ordre p . On vérifie alors facilement que G est un $\mathbb{F}_p$ -espace vectoriel. Comme il est fini, il est isomorphe à une puissance finie de $\mathbb{F}_p$, en tant qu'espace vectoriel mais donc aussi en tant que groupe.

Solution 5.24.

- (1) $h \in H$ est dans le noyau de l'action lorsque, pour tout g de G , $hgH = gH$ ce qui équivaut à $g^{-1}hg \in H$, donc $h \in gHg^{-1}$. L'ensemble K des h est donc l'intersection des conjugués de H , c'est un sous-groupe distingué de H (car noyau d'un morphisme). Si $K' \subset H$ est un sous-groupe distingué de G , pour tout g de G on a $gK'g^{-1} = K' \subset gHg^{-1}$ et donc K' est inclus dans l'ensemble des conjugués de H . K contient donc tous les sous-groupes distingués de G qui sont inclus dans H . Il joue un rôle analogue à celui du normalisateur dans G de H .

- (2) Soit K comme dans la question précédente. Soit ϕ le morphisme de G dans le groupe symétrique de G/H défini par l'action de translation à gauche. K est le noyau de ce morphisme et par le théorème d'isomorphisme G/K est isomorphe à un sous-groupe du groupe symétrique de G/H . Or ce groupe symétrique a pour cardinal $p!$, et G/K a pour ordre un diviseur du cardinal de G , c'est donc que G/K a pour ordre 1 (exclu car K est inclus dans H), la seule solution est que G/K soit de cardinal p , que K soit égal à H et donc que H soit distingué dans G .
- (3) On peut appliquer le théorème précédent, mais il est plus rapide de remarquer que si H est d'indice 2 dans G , il y a seulement deux classes à gauche, H et une classe de la forme xH , où $x \notin H$, et ces classes forment une partition de G . Mais comme il y a seulement deux classes à droite, H et Hx , on a nécessairement $xH = Hx$ ce pour tout x . Le sous-groupe H est distingué dans G .

Solution 5.25. Soit S un p -Sylow de G , de cardinal p^n . Un argument déjà utilisé : il existe dans S un élément d'ordre p , car tout élément est d'ordre une puissance de p . Mais on peut utiliser le théorème qui dit que le centre de S (qui est un p -groupe) est non trivial. Il existe donc dans le centre un élément x d'ordre p . Puisque x est dans le centre, $\langle x \rangle$ est distingué dans S . Alors $S/\langle x \rangle$ est un groupe d'ordre p^{n-1} , l'hypothèse de récurrence montre qu'il a des sous-groupes d'ordre p^β , qui sont les images de sous-groupes d'ordre $p^{\beta+1}$, avec β est inférieur à $\alpha - 1$. Le théorème est démontré.

Chapitre 6

Algèbre bilinéaire

Nous revenons à l'algèbre linéaire, et ce chapitre est consacré aux formes bilinéaires. Il s'agit donc d'associer à un couple de vecteurs un « scalaire » (c'est pour cela qu'on parle de « forme »). Comme cas particulier, nous retrouverons le produit scalaire de la géométrie. Dans tout le début du chapitre, on se place dans le cadre général d'un corps commutatif K , de caractéristique différente de 2.

6.1 FORMES BILINÉAIRES

6.1.1. Définitions

Une forme bilinéaire sur un K -espace vectoriel E est une application ϕ de $E \times E$ dans K qui vérifie les propriétés :

$$\begin{aligned} \forall (x, x', y, y') \in E^4, \forall \lambda \in K, \quad \phi(x + \lambda x', y) &= \phi(x, y) + \lambda \phi(x', y) \\ \phi(x, y + \lambda y') &= \phi(x, y) + \lambda \phi(x, y') \end{aligned}$$

Lorsque E est le corps K , les seules formes bilinéaires possibles sont données par $\phi(x, y) = axy$. Dans l'espace vectoriel des fonctions continues sur $[0, 1]$, l'application $(f, g) \mapsto \int_0^1 f(t)g(t) dt$ est une forme bilinéaire.

Théorème 6.1. *L'ensemble noté $\mathcal{L}_2(E, K)$ des formes bilinéaires sur le K -espace vectoriel E est un K -espace vectoriel. En dimension finie n , il est isomorphe à l'espace vectoriel $\mathcal{M}_n(K)$.*

Démonstration. La structure d'espace vectoriel est celle qu'ont tous les ensembles d'applications à valeurs dans K . Si $(e_i)_{i \in I}$ est une base de E , la bilinéarité implique qu'une forme bilinéaire ϕ est déterminée par les scalaires $\phi(e_i, e_j)$. En dimension n , notons $\mathcal{B} = (e_1, e_2, \dots, e_n)$ une base. La matrice $A = (\phi(e_i, e_j))_{i,j} \in \mathcal{M}_n(K)$ suffit donc à déterminer ϕ . On la notera $M_{\mathcal{B}}(\phi)$, et on a immédiatement :

$$M_{\mathcal{B}}(\phi + \lambda\psi) = M_{\mathcal{B}}(\phi) + \lambda M_{\mathcal{B}}(\psi)$$

□

Si on note alors X (resp. Y) la matrice des coordonnées de x (resp. de y) dans la base $\mathcal{B}$, on aura :

$$\phi(x, y) = \phi\left(\sum_i x_i e_i, \sum_j y_j e_j\right) = \sum_{i,j} x_i \phi(e_i, e_j) y_j = {}^t X A Y$$

On a bien sûr décidé de considérer qu'une matrice 1×1 se confond avec un scalaire de K .

6.1.2. Rang

À partir de ce paragraphe, nous nous limiterons au cas de la dimension finie. C'est dans le cours d'analyse que l'étude de la dimension infinie sera reprise.

La bilinéarité d'une forme ϕ peut se formuler autrement. On définit les deux applications partielles :

$$\begin{array}{ccc} \phi(x, \cdot) : E & \longrightarrow & K \\ y & \longmapsto & \phi(x, y) \end{array} \quad \text{et} \quad \begin{array}{ccc} \phi(\cdot, y) : E & \longrightarrow & K \\ x & \longmapsto & \phi(x, y) \end{array}$$

Alors ϕ est bilinéaire si, et seulement si ces deux applications partielles sont des formes linéaires, donc éléments du dual E^* . Et on peut alors introduire

$$\begin{array}{ccc} L_{\phi} : E & \longrightarrow & E^* \\ x & \longmapsto & \phi(x, \cdot) \end{array} \quad \text{et} \quad \begin{array}{ccc} R_{\phi} : E & \longrightarrow & E^* \\ y & \longmapsto & \phi(\cdot, y) \end{array}$$

deux applications qui sont encore, grâce à la bilinéarité, des morphismes d'espaces vectoriels.

Théorème 6.2. Lorsque E est de dimension finie, L_{ϕ} et R_{ϕ} sont de même rang, que l'on appelle **rang** de la forme bilinéaire ϕ .

On dit que ϕ est **non dégénérée** si l'une des applications L_{ϕ} ou R_{ϕ} est bijective ¹.

Démonstration. Soit A la matrice de ϕ par rapport à une base $\mathcal{B}$. Nous allons commencer par rechercher la matrice de L_{ϕ} et de R_{ϕ} par rapport aux bases $\mathcal{B}$ et $\mathcal{B}^*$ (base duale) : $L_{\phi}(e_j)(e_i) = \phi(e_j, e_i)$ et ce nombre est la i -ème coordonnée de $L_{\phi}(e_j)$ donc c'est le coefficient d'indice (i, j) de la matrice de L_{ϕ} . De même, $R_{\phi}(e_j)(e_i) = \phi(e_i, e_j)$. On en déduit que les matrices de L_{ϕ} et de R_{ϕ} sont respectivement ${}^t A$ et A , d'où le résultat, car une matrice et sa transposée ont même rang. □

1. Plus précisément, si l'une est bijective, l'autre l'est également.

Une forme bilinéaire est donc non dégénérée lorsque sa matrice A est inversible. Les deux applications L_ϕ et R_ϕ ont donc même rang, mais elles sont bien distinctes, en particulier leurs noyaux sont en général distincts.

6.1.3. Changements de base

Si E est rapporté à la base $\mathcal{B}$, et si ϕ est une forme bilinéaire de matrice A par rapport à cette base, quelle est la matrice A' de ϕ lorsque on se place dans une base $\mathcal{C}$?

Proposition 6.3. Si P est la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{C}$, on a

$$A' = {}^tPAP$$

(formule de changement de base pour les formes bilinéaires)

Démonstration. Si $\mathcal{C} = (e'_i)$ alors, par définition de $P = (p_{i,j})$, $e'_j = \sum_i p_{i,j} e_i$ et :

$$\phi(e'_j, e'_k) = \sum_{i,\ell} p_{i,j} \phi(e_i, e_\ell) p_{\ell,k} = \sum_{i,\ell} p'_{j,i} \phi(e_i, e_\ell) p_{\ell,k}$$

On a noté $p'_{i,j}$ les coefficients de la matrice tP . D'où le produit matriciel annoncé. $\square$

6.1.4. Matrice congruentes, formes congruentes

Comme dans le cas des applications linéaires ou des endomorphismes, on peut maintenant se poser un problème de classification des formes bilinéaires.

- Une forme bilinéaire est donnée : peut-on trouver une base dans laquelle l'écriture de la forme bilinéaire est particulièrement simple ?
- Comment reconnaître si deux matrices représentent (par rapport à deux bases) la même forme bilinéaire ?
- Quand deux formes bilinéaires ont-elles la même « signification » géométrique ?

Les trois questions sont liées, donnons des précisions de vocabulaire :

Définition 6.4. Deux matrices symétriques (ou antisymétriques) A et A' sont **congruentes** s'il existe une matrice P de $\mathbf{GL}(n, K)$ telle que :

$$A' = {}^tPAP$$

Autrement dit deux matrices congruentes représentent la même forme bilinéaire dans des bases différentes.

Deux formes bilinéaires ϕ et ϕ' sont **congruentes** s'il existe un automorphisme de E tel que :

$$\forall x, y \in E, \phi(x, y) = \phi'(u(x), u(y))$$

ϕ et ϕ' sont congruentes lorsque leur matrice, par rapport à la même base, sont congruentes, et la congruence est une relation d'équivalence. On voit tout de suite que deux matrices congruentes ont même rang.

Pour continuer l'étude des formes bilinéaires, nous allons nous spécialiser dans deux cas particuliers essentiels, et c'est dans ces cas que l'on obtiendra des résultats de classification.

6.2 FORMES BILÉAIRES SYMÉTRIQUES ET ANTISYMÉTRIQUES

6.2.1. Généralités

Définition 6.5. Soit ϕ une forme bilinéaire. On dit que c'est une **forme bilinéaire symétrique** si

$$\forall x, y \in E, \phi(x, y) = \phi(y, x)$$

On dit que c'est une **forme bilinéaire antisymétrique** lorsque :

$$\forall x, y \in E, \phi(x, y) = -\phi(y, x)$$

Voici quelques exemples :

- En dimension 1, $(x, y) \rightarrow xy$ est symétrique. La seule forme antisymétrique est constante nulle.
- Dans l'espace vectoriel $\mathbb{K}^2$, $((x_1, y_1), (x_2, y_2)) \mapsto a(x_1y_2 - x_2y_1)$ où a est un scalaire quelconque, est une forme antisymétrique. Les formes symétriques sont données par :

$$((x_1, x_2), (y_1, y_2)) \mapsto ax_1y_1 + b(x_1y_2 + x_2y_1) + cx_2y_2$$

où a, b et c sont des scalaires.

- Une forme bilinéaire est symétrique si sa matrice A dans n'importe quelle base est une matrice symétrique : $A = {}^tA$. Elle est antisymétrique si sa matrice est antisymétrique, $A = -{}^tA$.
- Comme nous sommes en caractéristique différente de 2, on a l'équivalence :

$$\forall x, y \in E, \phi(x, y) = -\phi(y, x) \iff \forall x \in E, \phi(x, x) = 0$$

Montrons la première implication :

$$\phi(x, x) = -\phi(x, x) \implies 2\phi(x, x) = 0 \implies \phi(x, x) = 0$$

en tenant compte de la caractéristique. Pour l'autre,

$$\begin{aligned} \phi(x+y, x+y) = 0 &\iff \phi(x, x) + \phi(x, y) + \phi(y, x) + \phi(y, y) = 0 \\ &\iff \phi(y, x) = -\phi(x, y) \end{aligned}$$

6.2.2. Noyau d'une forme bilinéaire symétrique ou antisymétrique

Dans le cadre des formes bilinéaires symétriques ou antisymétriques,

$$\phi(x, y) = 0 \iff \phi(y, x) = 0$$

et les applications L_ϕ et R_ϕ ont même noyau :

Définition 6.6. Soit ϕ une forme bilinéaire symétrique ou antisymétrique. On appelle **noyau** de ϕ le noyau commun de L_ϕ et de R_ϕ :

$$\text{Ker } \phi = \{x \in E \mid \forall y \in E, \phi(x, y) = 0\}$$

et une forme bilinéaire symétrique est non dégénérée si et seulement si son noyau est réduit au vecteur nul.

Remarques.

- Dans le cas où ϕ est dégénérée, on peut introduire dans l'espace vectoriel quotient $E / \text{Ker } \phi$ la forme bilinéaire $\bar{\phi}$ définie par :

$$\bar{\phi}(\bar{x}, \bar{y}) = \phi(x, y)$$

Après avoir vérifié la cohérence de la définition, on montre que $\bar{\phi}$ est non dégénérée.

- Il est remarquable que dans le cas non dégénéré : L_ϕ (ou R_ϕ) permet de réaliser un isomorphisme entre E et son dual : toute forme linéaire f est représentée par un unique vecteur a tel que :

$$\forall x \in E, f(x) = \phi(a, x)$$

Voir par exemple la notion de **gradient**, ou le théorème de représentation de Riesz, en analyse.

6.2.3. Orthogonalité

L'orthogonalité se définit comme dans le cas du produit scalaire de la géométrie élémentaire.

Définition 6.7. Soit ϕ une forme bilinéaire symétrique ou antisymétrique sur E . Alors :

- Si $\phi(x, y) = 0$, on dit que x est orthogonal à y , et on écrit parfois $x \perp y$.
- Si F est inclus dans E ,

$$F^\perp = \{x \in E \mid \forall y \in F, \phi(x, y) = 0\}$$

est appelé l'orthogonal de F .

La bilinéarité permet de prouver facilement que les orthogonaux sont des sous-espaces vectoriels et on peut aussi remarquer que $S^\perp = (\text{vect}(S))^\perp$.

Par ailleurs,

$$\text{Ker } \phi = E^\perp$$

et donc on pourra dire qu'une forme est non dégénérée si et seulement si le vecteur nul est le seul qui est orthogonal à tous les autres.

Donnons quelques propriétés de l'orthogonalité des sous-espaces.

Théorème 6.8. *Si F et G sont des sous-espaces vectoriels de E , alors on a :*

- $F \subset G \Rightarrow G^\perp \subset F^\perp$
- $F \subset (F^\perp)^\perp$
- $(F + G)^\perp = F^\perp \cap G^\perp$
- $F^\perp + G^\perp \subset (F \cap G)^\perp$
- $\dim F + \dim F^\perp = \dim E + \dim(F \cap \text{Ker } \phi)$

Dans le cas où ϕ est non dégénérée, on a :

- $\dim F + \dim F^\perp = \dim E$
- $F = (F^\perp)^\perp$
- $F^\perp + G^\perp = (F \cap G)^\perp$

Démonstration.

- Soit $x \in G^\perp$ quelconque, il est orthogonal à tout élément de G , en particulier aux éléments de F , il est donc dans $F^\perp$.
- Si $x \in F$ et $y \in F^\perp$, alors $\phi(x, y) = 0$ peut s'interpréter comme $x \in (F^\perp)^\perp$.
- Une inclusion résulte de la première propriété démontrée :

$$F \subset F + G, G \subset F + G \Rightarrow (F + G)^\perp \subset F^\perp \cap G^\perp$$

Et si $z \in F^\perp \cap G^\perp$, alors $\langle z, f + g \rangle = \langle z, f \rangle + \langle z, g \rangle = 0$ prouve l'autre inclusion.

- On a encore, en utilisant la première propriété, $F^\perp \subset (F \cap G)^\perp$ et $G^\perp \subset (F \cap G)^\perp$ et, comme $(F \cap G)^\perp$ est un sous-espace vectoriel, le plus petit sous-espace vectoriel qui contient $F^\perp$ et $G^\perp$, c'est-à-dire $F^\perp + G^\perp$, est inclus dans $(F \cap G)^\perp$.
- Soit $(e_1, \dots, e_k)$ une base de $F \cap \text{Ker } \phi$. On la complète en une base $(e_{k+1}, \dots, e_p)$ de F . Alors un vecteur x est dans $F^\perp$ si et seulement si, pour tout i de $k+1$ à p , $\phi(x, e_i) = 0$, puisque $\phi(x, e_i) = 0$ est automatique lorsque $i \leq k$. Or ces $p - k$ équations linéaires sont indépendantes, car :

$$\sum_{i=k+1}^p \lambda_i \phi(\cdot, e_i) = 0 \Rightarrow \sum_{i=k+1}^p \lambda_i e_i \in \text{Ker } \phi$$

Et donc les λ_i sont nuls. Donc x est dans $F^\perp$ si et seulement si x vérifie ces $p - k$ équations indépendantes :

$$\begin{aligned} \dim F^\perp &= \dim E - (p - k) \Rightarrow \dim F + \dim F^\perp = \dim E + k \\ &= \dim E + \dim F \cap \text{Ker } \phi \end{aligned}$$

Passons maintenant aux trois dernières affirmations.

- Si ϕ est non dégénérée, $\text{Ker } \phi$ se réduit à $\{0\}$ donc la formule précédente devient $\dim F + \dim F^\perp = \dim E$.

- On a alors

$$\dim(F^\perp)^\perp = \dim E - \dim F^\perp = \dim E - (\dim E - \dim F) = \dim F$$

compte-tenu de l'inclusion $F \subset (F^\perp)^\perp$, on obtient l'égalité $F = (F^\perp)^\perp$.

- Appliquons la relation $(F + G)^\perp = F^\perp \cap G^\perp$ à $F^\perp$ et à $G^\perp$:

$$(F^\perp + G^\perp)^\perp = F \cap G \Rightarrow F^\perp + G^\perp = (F \cap G)^\perp$$

□

Dans certains cas, les résultats précédents se simplifient un peu.

Définition 6.9. Soit ϕ une forme bilinéaire symétrique ou antisymétrique. Un vecteur x est **isotrope** lorsque $\phi(x, x) = 0$.

Un sous-espace F de E est **non isotrope** si la restriction de ϕ à F est non dégénérée. Il est **isotrope** si cette restriction est dégénérée, et **totalelement isotrope** si elle est nulle.

Contrairement au cas des endomorphismes, on peut toujours définir la restriction d'une forme bilinéaire à un sous-espace. Et il faut bien noter que la restriction d'une forme non dégénérée peut être dégénérée. Exemple : si

$$\phi((x_1, x_2), (y_1, y_2)) = x_1 y_1 - x_2 y_2$$

la restriction aux droites $\text{vect}(1, 1)$ et $\text{vect}(1, -1)$ est nulle. Ces droites sont donc totalement isotropes, elle sont engendrées par des vecteurs isotropes. Remarquons également que, en reprenant la définition :

$$\text{Ker } \phi|_F = F \cap F^\perp$$

et donc qu'on peut réécrire :

$$F \text{ est non isotrope} \iff F \cap F^\perp = \{0\}$$

$$F \text{ est totalelement isotrope} \iff F \subset F^\perp$$

Dans le cas d'un sous-espace non isotrope, on a donc :

Proposition 6.10. F est non isotrope si et seulement si $E = F \oplus F^\perp$.

Démonstration.

Si F est non isotrope, $F \cap \text{Ker } \phi \subset F \cap F^\perp = \{0\}$, et on a donc $\dim F + \dim F^\perp = \dim E$. On en déduit que F et son orthogonal sont supplémentaires. La réciproque est immédiate, puisque le fait qu'il y ait somme directe implique $F \cap F^\perp = \{0\}$. □

Exercice 6.1. Montrer que $\mathcal{LS}_2(E)$, ensemble des formes bilinéaires symétriques et $\mathcal{LA}_2(E)$, ensemble des formes bilinéaires antisymétriques sont deux sous-espaces supplémentaires de $\mathcal{L}_2(E)$. Si la dimension de E est n , donner leur dimension.

Exercice 6.2. Démontrer que si ϕ est une forme bilinéaire non dégénérée, avec E de dimension supérieure ou égale à 2, telle que :

$$\forall x \in E, \forall y \in E, \phi(x, y) = 0 \iff \phi(y, x) = 0$$

alors ϕ est symétrique ou antisymétrique.

Exercice 6.3. On considère la forme ϕ définie sur $\mathbb{R}^3$ par la matrice $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

Déterminer l'orthogonal d'une droite vectorielle quelconque. Préciser quelles droites sont isotropes. Déterminer de même l'orthogonal d'un plan, et décrire les différentes possibilités.

Exercice 6.4. Montrer qu'en dimension impaire, une forme bilinéaire antisymétrique est toujours dégénérée.

Exercice 6.5. Montrer, par un contre-exemple, que l'inclusion $F^\perp + G^\perp \subset (F \cap G)^\perp$ peut être stricte.

Soit F une droite vectorielle. Montrer que si elle est non isotrope, alors elle est totalement isotrope. Donner un exemple de sous-espace qui est non isotrope sans être totalement isotrope. Montrer qu'un sous-espace non isotrope contient un sous-espace totalement isotrope (non réduit à $\{0\}$). Est-ce que la réciproque est vraie ?

6.3 FORMES QUADRATIQUES. CLASSIFICATIONS DES FORMES BILINÉAIRES SYMÉTRIQUES RÉELLES ET COMPLEXES

6.3.1. Forme quadratique, cône isotrope

Nous allons dans ce paragraphe, nous limiter aux formes bilinéaires symétriques sur un corps K , et bien sûr encore à la dimension finie. L'étude de l'application $x \mapsto \phi(x, x)$ est essentielle.

Définition 6.11. On appelle **forme quadratique** associée à la forme bilinéaire symétrique ϕ l'application $q : E \rightarrow K$ définie par $q(x) = \phi(x, x)$. Pour tout vecteur x et tout scalaire λ , $q(\lambda x) = \lambda^2 q(x)$ et $q(0) = 0$.

Il est remarquable que la donnée de la forme quadratique suffise pour déterminer la forme bilinéaire symétrique ϕ , c'est ce qu'on appelle la **polarisation** :

Proposition 6.12.

$$\phi(x, y) = \frac{1}{2}(q(x + y) - q(x) - q(y))$$

Démonstration. Calcul immédiat :

$$q(x + y) = \phi(x + y, x + y) = q(x) + 2\phi(x, y) + q(y)$$

□

Par contre, toute application q de E dans K qui vérifie $q(\lambda x) = \lambda^2 q(x)$ n'est pas une forme quadratique ; voir par exemple q défini dans $\mathbb{R}$ par $q(x) = |x|$.

L'ensemble des vecteurs qui annulent la forme quadratique q n'est pas, en général, un sous-espace vectoriel. Il a cependant la propriété remarquable d'être stable par le produit par un scalaire. On dit que c'est un cône.

Définition 6.13. Un vecteur x est dit **isotrope** si $q(x) = 0$. L'ensemble des vecteurs isotropes s'appelle le **cône isotrope** de q (ou de ϕ). Il contient le noyau de ϕ . On le notera $C(q)$.

L'inclusion du noyau dans le cône peut s'exprimer ainsi : les vecteurs du noyau sont orthogonaux à tous les vecteurs, les vecteurs du cône sont orthogonaux à eux-mêmes.

Il arrive que le cône isotrope ne contienne que le vecteur nul, on le verra dans le cas réel pour la forme quadratique associée à un produit scalaire, mais cela peut se produire dans d'autres cas.

Définition 6.14. Une forme quadratique (et la forme bilinéaire associée) est **définie** si le cône isotrope est réduit au vecteur nul.

Et bien sûr, si q est définie, elle est non dégénérée.

Qu'est-ce que « réduire » une forme bilinéaire symétrique ? L'idée est, comme pour les endomorphismes, de se ramener à une matrice diagonale.

Définition 6.15. On dit qu'une base $\mathcal{B} = (e_i)_{1 \leq i \leq n}$ est **orthogonale** si

$$\forall (i, j), \quad i \neq j \Rightarrow \phi(e_i, e_j) = 0$$

Cela équivaut à dire que la matrice de ϕ par rapport à une base orthogonale est diagonale.

Théorème 6.16. Toute forme bilinéaire symétrique ϕ admet une base orthogonale.

Démonstration. Si ϕ est nulle, toute base est orthogonale. Si ϕ n'est pas nulle, la proposition 6.12 prouve que la forme quadratique q n'est pas identiquement nulle non plus, il existe e_1 tel que $q(e_1) \neq 0$. Alors la droite engendrée par e_1 est non isotrope, et $F = e_1^\perp$ est un supplémentaire de $\text{vect}(e_1)$. On répète le même argument avec la restriction de ϕ à F . $\square$

Remarque : Ce théorème permet de montrer que toute matrice symétrique est congruente à une matrice diagonale. Il est cependant moins profond que tout les théorèmes qui concerne les endomorphismes. En tout cas, il est obtenu très facilement...

6.3.2. Classification, théorème de Sylvester

On a donc obtenu une base telle que la matrice de ϕ soit diagonale. Remarquons qu'alors les termes diagonaux de la matrice de ϕ sont les $q(e_i)$. Comme $q(\lambda e) = \lambda^2 q(e)$, on peut aller plus loin dans la réduction, par exemple dans le cas d'un corps algébriquement clos comme $\mathbb{C}$.

Théorème 6.17. Soit ϕ une forme bilinéaire symétrique d'un $\mathbb{C}$ -espace vectoriel, q la forme quadratique associée. Il existe une base orthogonale (e_i) telle que, si r est le rang de ϕ ,

$$q(e_i) = 1, \text{ pour } 1 \leq i \leq r, \quad q(e_i) = 0, \text{ pour } i > r$$

Les formes bilinéaires symétriques complexes sont classées par leur rang.

Démonstration. On part du théorème précédent pour dire que ϕ admet une base orthogonale (e_i) . Comme dans $\mathbb{C}$, tout complexe est un carré, on peut remplacer chacun des e_i qui n'est pas isotrope par un vecteur colinéaire e'_i tel que $q(e'_i) = 1$. Quitte à renuméroter, on obtient la forme indiquée et le nombre des vecteurs de base pour lesquels $q(e'_i) = 1$ est une constante puisque c'est le rang de ϕ . Résultat supplémentaire, on a prouvé que toutes les formes bilinéaires symétriques de même rang sont équivalentes ; il y a donc $n+1$ classes d'équivalences de formes bilinéaires symétriques, puisque deux formes de rangs différents ne peuvent être congruentes. $\square$

Dans le cas réel, la forme quadratique prend des valeurs positives ou négatives : si elle ne prend que des valeurs positives sur une base **orthogonale**, alors elle ne prend que des valeurs positives (utiliser la matrice). Dans les cas intermédiaires, on peut classer les formes bilinéaires symétriques sur $\mathbb{R}$ par le théorème suivant :

Théorème 6.18. Théorème d'inertie de Sylvester. *Pour toute forme bilinéaire symétrique sur un $\mathbb{R}$ -espace vectoriel, il existe une base orthogonale (e_i) et un entier s inférieur au rang s de ϕ*

$$\begin{cases} q(e_i) = 1 \text{ pour } 1 \leq i \leq s \\ q(e_i) = -1 \text{ pour } s+1 \leq i \leq r \\ q(e_i) = 0 \text{ pour } i > r \end{cases}$$

De plus, les entiers s et $p = r - s$ ne dépendent que de ϕ , et non de la base orthogonale trouvée. Le couple (s, p) s'appelle la **signature** de la forme bilinéaire symétrique réelle.

Les formes bilinéaires symétriques réelles sont classées par leur signature.

Démonstration. La première partie du théorème est immédiate. Partant d'une base orthogonale (e_i) , comme les $q(e_i)$ sont des réels positifs, négatifs ou nuls, on peut les multiplier par des constantes adéquates.

Il reste à démontrer que la signature ne dépend que de ϕ et non de la base. Supposons donc que (e_i) et (e'_i) sont deux bases orthogonales telles que

$$q(e_i) = 1 \text{ pour } 1 \leq i \leq s, \quad q(e_i) = -1 \text{ pour } s < i \leq s + p = r$$

$$q(e'_i) = 1 \text{ pour } 1 \leq i \leq s', \quad q(e'_i) = -1 \text{ pour } s' < i \leq s' + p' = r$$

On a utilisé que le rang est constant. Soit alors $F_+ = \text{vect}(e_1, \dots, e_s)$ et $G_- = \text{vect}(e'_{s'}, \dots, e'_n)$. Alors $F_+ \cap G_- = \{0\}$ car la restriction de q à F_+ ne peut prendre que des valeurs strictement positives pour un vecteur non nul, et pour G_- que des valeurs négatives ou nulles. On en déduit que ces deux sous-espaces sont en somme directe, et donc que $s + (n - s') \leq n$, soit $s \leq s'$. En échangeant les rôles, $s' \leq s$, d'où $s = s'$ et $p = p'$ (puisque $s + p = s' + p' = r$). $\square$

La classification des formes bilinéaires symétriques n'est pas toujours aussi simple : dans le cas de $\mathbb{Q}$ en particulier, elle conduit à des problèmes d'arithmétique assez fins.

6.3.3. Réduction de Gauss d'une forme quadratique

La méthode de Gauss permet, dans le cas d'une forme bilinéaire symétrique, d'obtenir de façon automatique (plus précisément algorithmique) une base orthogonale. Elle n'est que la systématisation de la méthode dite de la « forme canonique » pour les polynômes du second degré.

Théorème 6.19. Réduction de Gauss. Si q est une forme quadratique de rang r sur un corps K , il existe r formes linéaires indépendantes (ℓ_i) telles que :

$$q(x) = \sum_{i=1}^r \lambda_i \ell_i(x)^2$$

où les λ_i sont dans K .

Démonstration. Donnons une description de la méthode. Notons (x_i) les coordonnées d'un vecteur x , on suppose que $q(x)$ est exprimée en fonction des coordonnées, lesquelles seront éventuellement renumérotées.

- Si un terme de la forme x_i^2 apparaît avec un coefficient a non nul, on peut supposer que c'est pour l'indice 1, et on met a en facteur dans tous les termes de $q(x)$ qui contiennent x_1 ; on peut alors faire apparaître le carré d'une première forme linéaire :

$$a(x_1^2 + bx_1x_2 + cx_1x_3 + \dots) = a \left(x_1 + \frac{b}{2}x_2 + \frac{c}{2}x_3 + \dots \right)^2 + \psi(x)$$

où ψ est une forme quadratique qui ne contient pas x_1 : il suffit de corriger en tenant compte des termes en $x_2, x_3, \dots$

- Si $q(x)$ ne contient aucun terme en x_i^2 et si, par exemple, x_1x_2 apparaît avec un coefficient a non nul, on met a en facteur dans tous les termes qui contiennent x_1 et x_2 et on écrit :

$$a(x_1 + bx_3 + \dots)(x_2 + cx_3 + \dots) + \psi(x) = a\ell_1(x)\ell_2(x) + \psi(x)$$

où ψ est une forme quadratique qui ne contient ni x_1 , ni x_2 , et il suffit alors d'écrire :

$$\ell_1(x)\ell_2(x) = \frac{1}{4} [(\ell_1(x) + \ell_2(x))^2 - (\ell_1(x) - \ell_2(x))^2]$$

et d'observer que les deux formes linéaires sont indépendantes.

On répète le processus, et toutes les formes ainsi obtenues sont indépendantes, puisque ψ ne contient plus x_1 . On peut alors se placer dans une base telle que les r premières coordonnées soient $X_1 = \ell_1(x), \dots, X_r = \ell_r(x)$. On posera ensuite $X_{r+1} = x_{r+1}, \dots$ et la matrice de q dans la nouvelle base ainsi définie est diagonale, ses r premiers coefficients étant non nuls. Le nombre r des formes linéaires est bien égal au rang de la forme quadratique. $\square$

Exercice 6.6. Donner la signature des formes quadratiques suivantes :

- (1) $q(x, y, z) = x^2 + y^2 + z^2 - 2(xy + yz + zx)$ dans $\mathbb{R}^3$.
- (2) $q(x) = x_1x_2 + x_2x_3 + \dots + x_nx_1$, dans $\mathbb{R}^n$, avec $x = (x_1, x_2, \dots, x_n)$.

Exercice 6.7. Soit ϕ une forme bilinéaire symétrique dans un $\mathbb{R}$ -espace vectoriel. Montrer que la forme quadratique a un signe constant si, et seulement si, le cône isotrope coïncide avec le noyau.

Exercice 6.8. Soit ϕ une forme bilinéaire symétrique non dégénérée sur un corps K (de caractéristique différente de 2). On suppose qu'il existe un vecteur isotrope $v \neq 0$. Montrer que, pour tout $a \in K$, il existe un vecteur w tel que $q(w) = a$.

Exercice 6.9. Montrer que les deux formes quadratiques définies sur $\mathbb{Q}^2$ par

$$q_1(x_1, x_2) = x_1^2 + x_2^2, \quad q_2(x_1, x_2) = x_1^2 + 2x_2^2$$

ne sont pas congruentes.

6.4 ESPACES VECTORIELS EUCLIDIENS ET HERMITIENS

6.4.1. Produit scalaire

C'est un aboutissement : la situation qui va nous intéresser est la plus proche de la géométrie. On suppose que l'espace vectoriel E est muni d'une forme bilinéaire symétrique.

Définition 6.20. Une forme bilinéaire ϕ symétrique telle que, pour tout x , $\phi(x, x) \geq 0$ est dite positive. Si le cône isotrope est réduit à $\{0\}$, $\phi(x, x) = 0$ si et seulement si x est nul et ϕ est définie. Une forme bilinéaire symétrique définie positive est appelée **produit scalaire**.

E est appelé **espace vectoriel euclidien** (on dit préhilbertien réel en dimension infinie). Notations :

$$\phi(x, y) = \langle x, y \rangle, \quad \|x\| = \sqrt{\langle x, x \rangle} = \sqrt{q(x)}$$

$\|x\|$ est la norme euclidienne du vecteur x .

6.4.2. Inégalité de Cauchy-Schwarz

On peut majorer la valeur du produit scalaire, c'est la célèbre inégalité de Cauchy-Schwarz, qu'on peut d'ailleurs énoncer dans un cadre un peu plus général.

Théorème 6.21. Inégalité de Cauchy-Schwarz. Soit ϕ une forme bilinéaire symétrique positive, q la forme quadratique associée

$$\forall (x, y) \in E^2, \quad (\phi(x, y))^2 \leq q(x)q(y)$$

Si ϕ est un produit scalaire, on écrit :

$$\forall (x, y) \in E^2, \quad \langle x, y \rangle \leq \|x\| \times \|y\|$$

De plus, dans le cas du produit scalaire, il y a égalité si et seulement si les vecteurs sont liés.

Démonstration.

- Soit ϕ une forme bilinéaire symétrique positive. Alors, pour tout couple (x, y) de vecteurs et pour tout réel λ , on a :

$$0 \leq q(x + \lambda y) = q(x) + 2\lambda\phi(x, y) + \lambda^2 q(y)$$

Si $q(y) = 0$, cette inégalité n'est possible pour tout λ que si $\phi(x, y) = 0$. Sinon, ce polynôme en λ ne peut avoir un signe constamment positif que si son discriminant est négatif. Dans les deux cas, on a donc

$$\phi(x, y)^2 \leq q(x)q(y)$$

ce qui donne, dans le cas du produit scalaire, l'inégalité demandé.

- Supposons de plus ϕ définie. Alors en supposant y non nul, $q(y) \neq 0$ et le polynôme en λ a un discriminant nul ; il a une racine double, λ_0 , donc $q(x + \lambda_0 y) = 0$ ce qui implique $x + \lambda_0 y = 0$, x et y sont liés. Si $y = 0$, bien sûr, x et y sont liés, et, par ailleurs, si x et y sont liés, l'inégalité de Cauchy-Schwarz est vérifiée. □

Un produit scalaire est forcément non dégénéré : comme il n'y a pas de vecteur isotrope, le noyau est réduit au vecteur nul. Et grâce à l'inégalité de Cauchy-Schwarz, on va montrer qu'une forme bilinéaire positive non dégénérée est un produit scalaire.

Corollaire 6.22. Soit ϕ une forme bilinéaire symétrique sur le $\mathbb{R}$ -espace vectoriel E . Alors

$$\phi \text{ définie et positive} \iff \phi \text{ non dégénérée positive}$$

Démonstration. Comme dit ci-dessus, si ϕ est définie, elle est forcément non dégénérée. Si par contre x est dans le cône isotrope et y quelconque, l'inégalité de Cauchy-Schwarz assure que $|\phi(x, y)| \leq q(x)q(y) = 0$, donc x est orthogonal à y et est dans le noyau de ϕ . □

Une application de l'inégalité de Cauchy-Schwarz est l'inégalité de Minkowski.

Théorème 6.23. Inégalité de Minkowski. Si E est un $\mathbb{R}$ -espace vectoriel muni d'un produit scalaire, alors

$$\forall (x, y) \in E^2, \quad \|x + y\| \leq \|x\| + \|y\|$$

et il n'y a égalité que si x et y sont positivement liés.

La dernière assertion signifie qu'il existe λ réel positif tel que $y = \lambda x$ ou que x est nul.

Démonstration. L'inégalité à démontrer équivaut à $\langle x + y, x + y \rangle \leq (\|x\| + \|y\|)^2$, en développant cela donne $2\langle x, y \rangle \leq 2\|x\|\|y\|$ qui est impliquée par l'inégalité de Cauchy-Schwarz, puisque pour tout réel a , $a \leq |a|$. Pour qu'il y ait égalité, il est nécessaire que x et y soient liés. En remplaçant, on voit que pour avoir égalité dans l'inégalité de Minkowski, il faut et suffit que x et y soit **positivement** liés. □

Application : ces propriétés font que l'application $x \mapsto \|x\|$ a bien les propriétés d'une norme, c'est la **norme euclidienne** associée au produit scalaire.

6.4.3. Produit scalaire hermitien

a) Sesquilinearité

Dans le cas où le corps de base est le corps des complexes, les formes bilinéaires symétriques et les formes quadratiques associées ne sont pas les bonnes généralisations du produit scalaire : une forme quadratique est à valeurs complexes et ne peut permettre de définir une norme. La bonne généralisation est celle des formes **sesquilinéaires hermitiennes** et des **formes hermitiennes** associées, où le carré du module remplace le carré.

Définition 6.24. Soit E un $\mathbb{C}$ -espace vectoriel. Une forme sesquilinéaire hermitienne sur E est une application ϕ de $E \times E$ dans $\mathbb{C}$ qui vérifie :

- (i) $\forall (x, y, y') \in E^3, \forall \lambda \in \mathbb{C}, \quad \phi(x, y + \lambda y') = \phi(x, y) + \lambda \phi(x, y')$
- (ii) $\forall (x, x', y) \in E^3, \forall \lambda \in \mathbb{C}, \quad \phi(x + \lambda x', y) = \phi(x, y) + \lambda \phi(x', y)$
- (iii) $\forall (x, y) \in E^2, \quad \phi(y, x) = \overline{\phi(x, y)}$

L'application q définie par $q(x) = \phi(x, x)$ est la forme hermitienne associée à ϕ .

La seconde propriété s'appelle **semilinéarité**, d'où le nom « sesquilinéaire », ce qui signifie « une fois et demi » linéaire... La troisième propriété est parfois appelée **symétrie hermitienne**.

Nous n'allons pas étudier longuement les formes sesquilinéaires hermitiennes, leur étude est très semblable à celle des formes bilinéaires symétriques : on peut par exemple définir le rang, étudier la notion d'orthogonalité... Il est possible de récupérer une forme sesquilinéaire hermitienne à partir de la forme hermitienne associée par une formule de polarisation un peu plus compliquée que dans le cas réel :

$$\phi(x, y) = \frac{1}{4} (q(x + y) - q(x - y) + iq(x - iy) - iq(x + iy))$$

Donnons également l'écriture matricielle :

Proposition 6.25. Si ϕ est une forme sesquilinéaire hermitienne sur E rapporté à la base $\mathcal{B} = (e_i)$, la matrice de ϕ est la matrice $M = (\phi(e_i, e_j))$, et on a :

$$\phi(x, y) = {}^t \overline{X} M Y$$

De plus, la matrice M vérifie ${}^t \overline{M} = M$. On dit qu'elle est hermitienne.

La vérification est immédiate. Les matrices vérifiant la propriété du théorème sont dite **matrices hermitiennes**, elles sont faciles à construire et donnent donc des

exemples de produit scalaire hermitien. On introduira désormais la notation :

$${}^t\overline{M} = M^*$$

et on dira que cette matrice est l'**adjointe** de la matrice M . Dans le cas où M est à coefficients réels, cela désignera donc la transposée de M , et bien sûr les matrices symétriques réelles sont des exemples de matrices hermitiennes. Le changement de base, avec P matrice de passage, conduit à la relation :

$$M' = P^* M P$$

lorsque M désigne la matrice dans la première base, M' dans la seconde.

b) Espace vectoriel hermitien

Comme dans le cas réel, on va définir des produits scalaires. Mais on commence par remarquer que

Lemme 6.26. *Si ϕ est une forme sesquilinéaire hermitienne, la forme hermitienne associée prend ses valeurs dans $\mathbb{R}$.*

Démonstration. Par symétrie hermitienne, $\overline{q(x)} = \overline{\phi(x, x)} = \phi(x, x) = q(x)$ donc $q(x)$ est réel. $\square$

On a alors, avec la même démonstration, un pendant du théorème de Sylvester : dans une base orthogonale, la matrice de ϕ est diagonale avec pour coefficients diagonaux $\lambda_1, \lambda_2, \dots, \lambda_n$ où les λ_i sont réels. Si le nombre des λ_i positifs est noté s , le nombre des λ_i négatifs est noté p et le couple (s, p) est la signature de la forme sesquilinéaire. Il est indépendant de la base orthogonale choisie.

On peut ensuite définir :

Définition 6.27. Une forme sesquilinéaire hermitienne est **positive** si la forme hermitienne associée ne prend que des valeurs positives. Une forme sesquilinéaire hermitienne est **définie** si le seul vecteur qui annule $q(x)$ est le vecteur nul. Un **produit scalaire hermitien** est une forme sesquilinéaire hermitienne définie positive. On notera $\phi(x, y) = \langle x, y \rangle$ puis $q(x) = \|x\|^2$, et $\|x\|$ désignera la **norme hermitienne** du vecteur x .

Un produit scalaire hermitien satisfait l'inégalité de Cauchy-Schwarz, et l'inégalité de Minkowski, $x \mapsto \|x\|$ fait bien de E un $\mathbb{C}$ -espace vectoriel normé. Les démonstrations sont très semblables ; par exemple pour l'inégalité de Cauchy-Schwarz, on pose

$$P(t) = q(x + \lambda y) = |\lambda|^2 q(y) + \lambda \phi(x, y) + \overline{\lambda} \phi(y, x) + q(x) \geq 0$$

Si $\phi(x, y) = 0$, l'inégalité est satisfaite. Sinon, on écrit $\phi(x, y) = \rho e^{i\theta}$ (ρ réel) et on prend pour λ la valeur $\lambda = s e^{-i\theta}$ (s réel). Alors :

$$P(s e^{i\theta}) = s^2 q(y) + 2s\rho + q(x) \geq 0$$

pour tout réel s . On regarde alors deux cas. Si $q(y) = 0$, l'expression ne peut être toujours positive que si $\rho = 0$ donc $\phi(x, y) = 0$. Si que $q(y) \neq 0$ le discriminant du polynôme doit être négatif. L'inégalité est démontrée.

Un $\mathbb{C}$ -espace vectoriel de dimension finie muni d'un produit scalaire hermitien s'appelle un **espace hermitien**. Dans le cas de la dimension infinie, on parle d'espace **préhilbertien complexe**, mais nous nous cantonnerons ici à la dimension finie.

6.4.4. Bases orthonormales

Nous traitons maintenant en même temps le cas des produits scalaires euclidiens ou hermitiens.

Définition 6.28. On appelle **base orthonormale** une base orthogonale dont tous les vecteurs sont unitaires, c'est-à-dire de norme 1.

Il existe « beaucoup » de bases orthonormales.

Proposition 6.29. *Tout espace vectoriel euclidien ou hermitien admet des bases orthonormales. Si (e_i) est une telle base et si $X = {}^t(x_1, \dots, x_n)$ et $Y = {}^t(y_1, \dots, y_n)$ désignent les matrices de coordonnées des vecteurs x et y dans ces bases, on a :*

$$\langle x, y \rangle = {}^tX Y = \sum_{i=1}^n x_i y_i \quad (\text{cas euclidien})$$

$$\langle x, y \rangle = X^* Y = \sum_{i=1}^n \bar{x}_i y_i \quad (\text{cas hermitien})$$

De plus, $x_i = \langle e_i, x \rangle$

Démonstration. L'existence d'une base orthogonale est assurée dans le cas général d'une forme bilinéaire symétrique. Dans le cas euclidien ou hermitien, le rang est égal à la dimension, la forme quadratique ne prend que des valeurs strictement positives sur les vecteurs non nuls, ce qui permet d'obtenir une base orthonormale à partir d'une base orthogonale. Le reste suit, en observant néanmoins que dans le cas hermitien, $\langle e_i, x \rangle$ et $\langle x, e_i \rangle$ sont conjugués. $\square$

Si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases orthonormales d'une espace vectoriel hermitien, alors la matrice de passage de $\mathcal{B}$ vers $\mathcal{B}'$ est dite **unitaire**. Dans le cas euclidien, on dit **orthogonale**.

Proposition 6.30.

- (1) Une matrice Ω de $\mathcal{M}_n(\mathbb{R})$ est orthogonale si et seulement si ${}^t\Omega\Omega = I$. L'ensemble des matrices orthogonales est un groupe pour le produit, appelé groupe orthogonal, et noté $O(n, \mathbb{R})$.
- (2) Une matrice U de $\mathcal{M}_n(\mathbb{C})$ est unitaire si et seulement si $U^*U = I$. L'ensemble des matrices unitaires est un groupe pour le produit, appelé groupe unitaire, et noté $U(n, \mathbb{C})$.

Exercice 6.10. Généraliser la méthode de Gauss au cas des formes hermitiennes : il s'agit de prouver que toute forme sesquilinéaire hermitienne peut se décomposer en une combinaison (à coefficients réels) de carrés de module de formes linéaires indépendantes. On observera que

$$\ell_1 \overline{\ell_2} + \overline{\ell_1} \ell_2 = \frac{1}{2} (|\ell_1 + \ell_2|^2 - |\ell_1 - \ell_2|^2)$$

Traiter le cas de :

$$q(z_1, z_2) = |z_1|^2 - iz_1 \overline{z_2} + i \overline{z_1} z_2 + |z_2|^2$$

et de

$$q(z_1, z_2, z_3) = z_1 \overline{z_2} + \overline{z_1} z_2 - z_2 \overline{z_3} - \overline{z_2} z_3$$

Exercice 6.11. Dans un espace euclidien ou hermitien, muni d'une base **orthonormale** $\mathcal{B} = (e_i)$, montrer que

$$x = \sum_{i=1}^n \langle e_i, x \rangle e_i$$

Si u est un endomorphisme de matrice $(u_{i,j})$ dans la base $\mathcal{B}$, alors

$$t_{ij} = \langle e_i, u(e_j) \rangle$$

6.5 ADJOINT, DIAGONALISATION DES ENDOMORPHISMES AUTOADJOINTS

6.5.1. Adjoint d'un endomorphisme

Soit E un espace vectoriel euclidien ou hermitien. On va définir une sorte de symétrie entre les endomorphismes de E , liée au produit scalaire.

Théorème 6.31. Soit E un espace vectoriel euclidien ou hermitien, $u \in \mathcal{L}(E)$.

- Il existe un et un seul endomorphisme u^* vérifiant

$$\forall (x, y) \in E^2, \quad \langle u(x), y \rangle = \langle x, u^*(y) \rangle$$

- Si u et v sont des endomorphismes et λ un scalaire,
 $(u + v)^* = u^* + v^*$, $(\lambda u)^* = \overline{\lambda} u^*$, $u^{**} = u$, $(u \circ v)^* = v^* \circ u^*$.
- Si $\mathcal{B}$ est une base **orthonormale**, alors

$$M_{\mathcal{B}}(u^*) = M_{\mathcal{B}}^*(u)$$

On dit alors que u^* est l'**adjoint** de u .

Démonstration. Pour définir f^* , on peut utiliser une base :

$$\langle e_i, u^*(e_j) \rangle = \langle u(e_i), e_j \rangle = \overline{\langle e_j, u(e_i) \rangle}$$

la première égalité montre que les coordonnées des vecteurs $u^*(e_j)$ sont bien déterminées, la seconde permet de prouver que la matrice de u^* , dans la base (e_i) est l'adjointe de la matrice de u (dans la même base). Ayant ainsi défini u^* à l'aide d'une base, on vérifie qu'il satisfait la définition, grâce à la sesquilinearité (ou la bilinéarité). Les autres propriétés se démontrent à l'aide de la définition ou grâce aux matrices. $\square$

Donnons une propriété importante de l'adjoint d'un endomorphisme f .

Proposition 6.32. *Si F est un sous-espace vectoriel de E , alors*

$$u(F) \subset F \Rightarrow u^*(F^\perp) \subset F^\perp$$

Démonstration. On commence par observer que

$$\forall (x, y) \in E^2, \quad \langle u^*(x), y \rangle = \langle x, u(y) \rangle$$

puisque $u^{**} = u$. Si on prend x dans $F^\perp$ et y dans F :

$$\langle u^*(x), y \rangle = \langle x, u(y) \rangle = 0$$

car $u(y)$ est dans F , donc $u^*(x)$ est dans $F^\perp$. $\square$

Cas particulier : si F est u -stable, son orthogonal $F^\perp$ est u^* -stable.

Parmi les endomorphismes, certains ont la propriété de coïncider avec leur adjoint. On les appelle **endomorphismes autoadjoints**. Dans le cas réel, on parle plutôt d'endomorphismes symétriques, et signalons que l'on emploie souvent le mot **opérateur** à la place d'endomorphisme. Les exemples d'endomorphismes autoadjoints sont faciles à trouver, puisque le théorème que nous venons de démontrer permet d'assurer :

Proposition 6.33. *f est autoadjoint si et seulement si sa matrice, dans une base orthonormale est égale à son adjointe. On dit qu'elle est auto-adjointe (hermitienne dans le cas complexe, symétrique dans le cas réel).*

À ce stade de notre étude, on peut observer la situation suivante. Si E est un espace vectoriel euclidien, une matrice symétrique peut s'interpréter de deux façons :

- Soit comme la matrice d'une forme bilinéaire symétrique, par rapport à une base.
- Soit comme la matrice d'un endomorphisme autoadjoint par rapport à une base orthonormale.

Et dans le cas hermitien, une matrice hermitienne est soit la matrice d'une forme sesquilineaire hermitienne, soit la matrice d'un endomorphisme autoadjoint dans une base orthonormale. Cette double interprétation peut se décrire sans l'intermédiaire des matrices, en résumé :

Proposition 6.34. Soit E un espace vectoriel euclidien (resp. hermitien). À toute forme bilinéaire symétrique (resp. sesquilinéaire hermitienne) ϕ , on peut associer un endomorphisme autoadjoint unique u par :

$$\forall x, y \in E, \quad \phi(x, y) = \langle x, u(y) \rangle$$

Réciproquement, cette même formule permet d'associer une forme bilinéaire (resp. sesquilinéaire) à un endomorphisme autoadjoint.

On peut alors utiliser ce point de vue pour définir les endomorphismes autoadjoints positifs ou définis positifs :

Définition 6.35. Un endomorphisme autoadjoint u d'un espace vectoriel euclidien ou hermitien E est positif (resp. défini positif) si la forme quadratique $x \mapsto \langle x, f(x) \rangle$ est positive (resp. définie positive). L'ensemble des matrices de ces endomorphismes, dans une base orthonormale, est noté $\mathbf{S}_n^+$ (resp. $\mathbf{S}_n^{++}$ dans le cas euclidien). Dans le cas complexe, on notera $\mathbf{H}_n^+$ (resp. $\mathbf{H}_n^{++}$).

On dit parfois strictement positif au lieu de défini positif.

6.5.2. Diagonalisation des endomorphismes autoadjoints

Le résultat qui nous intéresse le plus, pas immédiat à démontrer, est énoncé ainsi :

Théorème 6.36. Théorème spectral. Soit E un endomorphisme autoadjoint dans un espace vectoriel hermitien (resp. euclidien) E . Alors ses valeurs propres sont réelles, il est diagonalisable, et on peut trouver une base orthonormale de vecteurs propres.

Démonstration. Il y a plusieurs étapes, et nous commençons par le cas hermitien.

- Si λ est une valeur propre de u , alors λ est réel. En effet, en notant x un vecteur propre,

$$\langle u(x), x \rangle = \overline{\lambda} \langle x, x \rangle = \langle x, u(x) \rangle = \lambda \langle x, x \rangle$$

La valeur propre λ est égale à son conjugué, donc est réelle.

- D'après le théorème 6.32, si F est u -stable alors $F^\perp$ est u -stable puisque $u = u^*$.
- Raisonnons maintenant par récurrence sur la dimension. En dimension 1, la matrice d'un endomorphisme autoadjoint est un réel, le théorème est vrai. Supposons le théorème vrai pour tout endomorphisme autoadjoint en dimension strictement inférieure à n , et soit u autoadjoint en dimension n . Alors, comme le corps de base est $\mathbb{C}$, u admet une valeur propre, qui est donc réelle. Si F est le sous-espace propre associé, on peut choisir une base orthonormale dans F . Comme F est u -stable, son orthogonal est aussi u -stable et on peut appliquer l'hypothèse de récurrence à la restriction de u à $F^\perp$. (En effet, cette restriction reste autoadjointe). En réunissant ces deux bases, on obtient une base orthonormale de vecteurs propres.

- Pour le cas réel, raisonnons matriciellement : une matrice réelle symétrique A est aussi hermitienne si on la considère comme matrice à coefficients complexes. Il existe donc une matrice P telle que $P^{-1}AP = D$ où D est diagonale réelle. Le problème est qu'à priori, P est une matrice de passage à coefficients **complexes**. Notons $P = S + iT$ où S et T sont réelles. Il est possible qu'aucune de ces deux matrices ne soit inversible, mais il existe $t \in \mathbb{R}$ telle que $\det(S + tT)$ soit non nul, puisque ce déterminant est un polynôme en t^2 et que $\mathbb{R}$ est infini. On a alors :

$$P^{-1}AP = D \Rightarrow AP = DP \Rightarrow AS = SD, AT = TD$$

et donc

$$A(S + tT) = (S + tT)D$$

d'où le résultat. Ce n'est pas terminé : on dispose d'une base de vecteur propres, mais nous savons que les sous-espaces propres sont en somme directe. De plus, deux sous-espaces propres distincts sont orthogonaux ; on peut donc construire une base orthonormale de vecteurs propres. Remarquons que notre méthode permet aussi d'énoncer : deux matrices réelles qui sont semblables sur $\mathbb{C}$ sont aussi semblables sur $\mathbb{R}$. □

Ce théorème a beaucoup d'applications. Donnons-en quelques unes :

- Si u est autoadjoint, sa signature est (s, p) où s est le nombre des valeurs propres strictement positives, p le nombre des valeurs propres strictement négatives.
- En particulier, une matrice symétrique est positive (resp. définie positive) si et seulement si ses valeurs propres sont positives (resp. strictement positives). Ce résultat vaut également pour les matrices hermitiennes.
- Si ϕ et ψ sont deux formes bilinéaires symétriques définies sur E , espace vectoriel réel. On suppose que ϕ est définie positive. Alors il existe une base orthonormale pour ϕ qui est orthogonale pour ψ .

Pour le dernier alinéa, il suffit de munir l'espace vectoriel de la structure euclidienne définie par ϕ .

6.5.3. Diagonalisation des endomorphismes normaux

Plaçons nous maintenant dans le cas hermitien seulement. Nous allons alors généraliser le théorème fondamental du paragraphe précédent. Commençons par définir une nouvelle catégorie d'endomorphismes.

Définition 6.37. *Un endomorphisme est normal lorsqu'il commute avec son adjoint.*

Bien sûr, tout endomorphisme autoadjoint est normal, mais il y a d'autres types d'endomorphismes normaux.

2. Il est non identiquement nul car non nul en i .

Théorème 6.38. Soit E un $\mathbb{C}$ espace vectoriel hermitien, alors un endomorphisme est diagonalisable dans une base orthonormale si et seulement s'il est normal.

Démonstration. Supposons qu'il existe une base orthonormale (e_i) formée de vecteurs propres pour u . La matrice de u dans cette base est diagonale, celle de u^* est aussi diagonale, et ces deux matrices commutent donc.

Supposons maintenant que u soit normal. On raisonne par récurrence sur la dimension, le résultat étant immédiat en dimension 1. Puisque nous sommes sur le corps de base $\mathbb{C}$, il existe une valeur propre λ . Soit E_λ le sous-espace propre associé. Si c'est E tout entier, u est une homothétie et c'est fini. Sinon, son supplémentaire orthogonal $E^\perp$ est stable par u . En effet, il est stable par u^* , par propriété de l'adjoint (cf. la proposition 6.32), et, comme u^* et u commutent, il est stable par u , rappelons l'argument.

$$u \circ u^*(x) = u^* \circ u(x) = \lambda u^*(x)$$

pour tout vecteur propre x , donc $u^*(x) \in E_\lambda$. Le même raisonnement montre que u stabilisant E_λ , son adjoint u^* stabilise $E_\lambda^\perp$, on peut donc définir les restrictions de u et de u^* à $E_\lambda^\perp$. De plus, en considérant les définitions,

$$u^*|_{E_\lambda^\perp} = \left(u|_{E_\lambda^\perp} \right)^*$$

et donc sur $E_\lambda^\perp$ la restriction de u est normale : la récurrence marche. $\square$

Les principaux endomorphismes normaux sont :

- les endomorphismes autoadjoints, $u = u^*$, qui ont des valeurs propres réelles.
- les endomorphismes antihermitiens, $u^* = -u$, qui ont des valeurs propres imaginaires pures.
- les endomorphismes unitaires, $u^{-1} = u^*$, qui ont des valeurs propres de module 1.

Il est facile de vérifier que ces propriétés sur les valeurs propres les caractérisent parmi les endomorphismes normaux.

Mise en garde : le théorème ne subsiste pas dans le cas réel. On le verra dans les chapitres suivant pour les rotations, par exemple.

Exercice 6.12. Soit E un espace hermitien. Montrer que tout $u \in \mathcal{L}(E)$ s'écrit de façon unique $u = u_1 + u_2$ où u_1 est auto-adjoint et u_2 est « anti » auto-adjoint (c'est-à-dire $u_2^* = -u_2$). Montrer qu'alors u est normal si et seulement si u_1 et u_2 commutent. Que devient cet exercice en dimension 1 ?

Exercice 6.13. Soit E un espace hermitien, u un endomorphisme de E . Montrer que

- (1) Si $\forall x \in E, \langle u(x), x \rangle = 0$ alors u est nul. (Indication : déduire $\langle u(x), y \rangle = 0$ en appliquant l'hypothèse à $x + y$ et $x + iy$). Ce résultat subsiste-t-il dans le cas réel ?
- (2) $u = u^* \iff \langle u(x), x \rangle \in \mathbb{R}$ pour tout $x \in E$.

EXERCICES

Exercice 6.14. Soit E un espace vectoriel de base $\mathcal{B} = (e_1, e_2, \dots, e_n)$. On note (e_i^*) la base duale. Si f et g sont deux formes linéaires, on note $f \otimes g$ la forme définie par :

$$(f \otimes g)(x, y) = f(x)g(y)$$

- (1) Montrer que $f \otimes g$ est une forme bilinéaire. Est-elle symétrique ? Quel est son rang ?
- (2) Montrer que les n^2 formes $e_i^* \otimes e_j^*$ forment une base de l'espace vectoriel des formes bilinéaires. Faire le lien avec les matrices et proposer une base pour l'espace vectoriel des formes bilinéaires symétriques.
- (3) Si f est une forme linéaire et v un vecteur, on note $f \otimes v$ l'application $(f \otimes v)(x) = f(x)v$, où v est un vecteur. Montrer que c'est un endomorphisme. Quel est son rang ? En déduire également une base de $\mathcal{L}(E)$ et faire le lien avec les matrices.

Exercice 6.15. Soit $E = \mathbb{C}$ considéré comme $\mathbb{R}$ -espace vectoriel.

- (1) Montrer que

$$\langle z|w \rangle = \Re(\bar{z}w)$$

définit un produit scalaire euclidien sur E ; et donner un isomorphisme d'espace vectoriel euclidien entre E et $\mathbb{R}^2$ (muni du produit scalaire canonique).

- (2) Soit m_γ défini par :

$$m_\gamma(z) = \gamma z$$

Montrer que m_γ est linéaire. Chercher l'adjoint de m_γ . Quand m_γ est-il symétrique, orthogonal ?

Exercice 6.16. Soit A une matrice de $\mathcal{M}_n(\mathbb{C})$.

- (1) On suppose qu'il existe une matrice **unitaire** U telle que :

$$A^* = AU$$

Montrer que $AU = UA$ puis que A est une matrice normale.

- (2) Réciproquement, on suppose que la matrice A est normale. Montrer qu'il existe une matrice unitaire U telle que $A^* = AU$. On pourra commencer par traiter le cas où A est diagonale.

Exercice 6.17.

- (1) Soit A une matrice de $\mathcal{M}_n(\mathbb{R})$. Montrer que tAA est symétrique positive. Quand est-elle définie positive ?
- (2) Soit S une matrice symétrique positive. Montrer qu'il existe une matrice A telle que $S = {}^tAA$.
- (3) Reprendre le même exercice dans le cas hermitien.

PROBLÈMES

Les corrigés de ces problèmes sont disponibles sur le site de Dunod : www.dunod.com.

6.1. PROBLÈME

Dans ce problème, on s'intéresse aux formes bilinéaires antisymétriques. Soit ϕ une forme bilinéaire antisymétrique définie sur un K -espace vectoriel E , de dimension finie. On suppose que K est un corps de caractéristique différente de 2.

- (1) On suppose que E est de dimension 2. Montrer que, si ϕ n'est pas nulle, il existe une base (u, v) telle que la matrice de ϕ dans cette base est :

$$T = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

- (2) On se place maintenant en dimension finie n . Montrer qu'il existe une base $\mathcal{B}$ de la forme

$$\mathcal{B} = (u_1, v_1, u_2, v_2, \dots, u_k, v_k, w_1, \dots, w_{n-2k})$$

telle que la matrice de ϕ dans cette base soit diagonale par blocs, avec k blocs de la forme T , le reste de la matrice étant nul.

- (3) En déduire que le rang d'une forme bilinéaire antisymétrique est pair. Déterminer également les classes de congruence d'une forme bilinéaire antisymétrique.
- (4) Montrer que le déterminant d'une matrice antisymétrique est toujours un carré.
- (5) On suppose que $A = (a_{ij})$ est une matrice antisymétrique quelconque de taille $2n$. Montrer que le déterminant de A est le carré d'un polynôme en les a_{ij} , où on ne prend que les indices tels que $i < j$. Déterminer ce polynôme (que l'on appelle le Pfaffien), lorsque $n = 1$, $n = 2$ et $n = 3$.

6.2. PROBLÈME

Soit E un espace vectoriel de dimension finie sur corps K de caractéristique $\neq 2$. Soit $f : E \times E \longrightarrow K$ une forme bilinéaire symétrique et soit q la forme quadratique associée ($q(x) = f(x, x)$). On note $C(q)$ l'ensemble des vecteurs tels que $q(x) = 0$.

- (1) Un sous-espace vectoriel F de E est dit **totalelement isotrope** si pour tout x de F on a $q(x) = 0$. Montrer que F est totalelement isotrope si, et seulement si, $F \subset F^\perp$.
- (2) Supposons que F soit un sous-espace vectoriel totalelement isotrope. Soit $x \in F^\perp \cap C(q)$. Montrer que $F + Kx$ est un sous-espace vectoriel totalelement isotrope.

- (3) On dit que F est un sous-espace totalement isotrope maximal s'il n'existe aucun sous-espace vectoriel totalement isotrope G tel que $F \subset G$ et $F \neq G$. Montrer que si F est un sous-espace vectoriel totalement isotrope maximal, alors :

$$F = F^\perp \cap C(q)$$

- (4) Soient F_1 et F_2 sont deux sous-espaces vectoriels totalement isotropes maximaux. On pose $F = F_1 \cap F_2$. Soit S_1 un sous-espace vectoriel supplémentaire de F dans F_1 :

$$F_1 = F \oplus S_1$$

Soit de même S_2 un supplémentaire de F dans F_2 :

$$F_2 = F \oplus S_2$$

On se propose de montrer que $S_1 \cap S_2^\perp = \{0\}$. Pour cela, soit :

$$x_0 \in S_1 \cap S_2^\perp$$

- (a) Montrer que $x_0 \in F_2^\perp$.
 (b) Montrer successivement $x_0 \in C(q)$, $x_0 \in F_2$, $x_0 \in F$.
 (c) Conclure en constatant que $x_0 \in F \cap S_1$.
- (5) Pour tout sous-espace vectoriel U de E , montrer que :

$$\dim U^\perp \geq \dim E - \dim U$$

(si la forme f est non dégénérée, on a égalité). À l'aide de la question précédente, en déduire $\dim S_1 = \dim S_2$, puis $\dim F_1 = \dim F_2$.

- (6) D'après la question précédente, tous les sous-espaces totalement isotropes maximaux ont la même dimension. Ce nombre s'appelle l'indice de la forme bilinéaire f (ou de la forme quadratique q).

Soit en particulier $E = \mathbb{R}^3$. Sur E , on définit les six formes quadratiques suivantes :

- $q_1(x) = x_1^2 + x_2^2 + x_3^2$
- $q_2(x) = x_1^2 + x_2^2 - x_3^2$
- $q_3(x) = x_1^2 + x_2^2$
- $q_4(x) = x_1^2 - x_2^2$
- $q_5(x) = x_1^2$
- $q_6(x) = 0$

Calculer l'indice de ces six formes quadratiques.

SOLUTIONS DES EXERCICES

Solution 6.1. Si ϕ et ψ sont des formes bilinéaires symétriques, alors $\phi + \psi$ et $\lambda\phi$ sont encore des formes bilinéaires symétriques. Cela résulte de calculs (ou d'écritures...) évidents, comme

$$\begin{aligned} (\phi + \psi)(x, y + \mu y') &= \phi(x, y + \mu y') + \psi(x, y + \mu y') \\ &= \phi(x, y) + \psi(x, y) + \mu(\phi(x, y') + \psi(x, y')) \\ &= (\phi + \psi)(x, y) + \mu(\phi + \psi)(x, y) \end{aligned}$$

qui prouve que la somme de deux applications linéaires pour leur deuxième argument est elle-même linéaire pour le second argument.

Les ensembles $\mathcal{LS}_2(E)$ et $\mathcal{LA}_2(E)$ sont donc des sous-espaces de l'espace des formes bilinéaires $\mathcal{L}_2(E)$. Une forme qui serait à la fois symétrique et antisymétrique vérifierait

$$\forall (x, y) \in E^2, \quad \phi(x, y) = \phi(y, x) = -\phi(x, y)$$

et serait identiquement nulle. Par ailleurs, si f est une forme bilinéaire, si on désire écrire $f = s + a$ où s est symétrique et a antisymétrique, on trouve que la seule solution est

$$s(x, y) = \frac{1}{2}(f(x, y) + f(y, x)) \quad a(x, y) = \frac{1}{2}(f(x, y) - f(y, x))$$

On peut donc conclure que l'espace vectoriel des formes bilinéaires est somme directe du sous-espace vectoriel des formes bilinéaires symétriques et du sous-espace vectoriel des formes bilinéaires antisymétriques. En dimension finie n , l'examen des matrices montre que les dimensions de ces sous-espaces vectoriels sont respectivement $\frac{n(n+1)}{2}$ et $\frac{n(n-1)}{2}$.

Solution 6.2. Si ϕ est symétrique $\phi(x, y) = \phi(y, x)$ et s'il est antisymétrique $\phi(x, y) = -\phi(y, x)$ donc la nullité de l'un implique celle de l'autre. Supposons donc que ϕ soit une forme bilinéaire telle que, pour tout couple (x, y)

$$\phi(x, y) = 0 \quad \Rightarrow \quad \phi(y, x) = 0$$

Alors $x \mapsto \phi(x, y)$ et $x \mapsto \phi(y, x)$ sont deux formes linéaires qui ont même noyau. Il existe donc une constante, a priori dépendant de y telle que, pour tout x , $\phi(x, y) = \lambda_y \phi(y, x)$.

On va montrer que λ_y est en fait indépendant de y . Soient en effet L et M les applications de E dans E^* définies par $y \mapsto \phi(\cdot, y)$ et $y \mapsto \phi(y, \cdot)$ (voir le paragraphe 6.1.2.). Alors L et R sont bijectives puisque ϕ est non dégénérée, et l'égalité ci-dessus prouve :

$$\forall y \in E, \quad L^{-1} \circ R(y) = \lambda_y y$$

On sait qu'en dimension supérieure ou égale à 2, cela implique que $L^{-1} \circ R$ est une homothétie, (voir l'exercice 4.1.) donc λ_y est une constante, indépendante de y . Mais on peut alors écrire :

$$\phi(x, y) = \lambda \phi(y, x) = \lambda^2 \phi(x, y)$$

d'où $\lambda = \pm 1$ (en choisissant x et y de sorte que $\phi(x, y)$ ne soit pas nul)

Solution 6.3. Soit $D = \text{vect}(ae_1 + be_2 + ce_3)$. Alors un vecteur de coordonnées x_1, x_2, x_3 est orthogonal à D si et seulement si : $ax_1 + bx_2 - cx_3 = 0$. L'orthogonal de D est donc un plan P (car a, b et c ne sont pas tous nuls.) Une droite est isotrope si son intersection avec son orthogonal n'est pas réduite à $\{0\}$, donc si elle est incluse dans ce plan, donc si $a^2 + b^2 - c^2 = 0$. Ce vecteur est alors isotrope, la droite est totalement isotrope. Il y a une infinité de solutions qui, géométriquement, sont les génératrices du cône défini par l'équation précédente. L'orthogonal du plan P d'équation $ux_1 + vy_1 + wz_1 = 0$ est la droite $\text{vect}(ue_1 + ve_2 - we_3)$. Il est isotrope si cette droite est une génératrice du cône. En géométrie, on peut démontrer que le plan est alors tangent au cône.

Solution 6.4. Soit ϕ antisymétrique. Sa matrice A est antisymétrique, ${}^tA = -A$. Donc

$$\det(A) = \det({}^tA) = (-1)^n \det A$$

où n est la dimension de l'espace vectoriel. Si n est impair, on en déduit que $\det A$ est nul, le noyau de A qui est aussi le noyau de la forme bilinéaire n'est pas réduit au vecteur nul.

Solution 6.5. Pour trouver un contre-exemple, il faut une forme bilinéaire symétrique dégénérée. Plaçons-nous sur $\mathbb{R}^2$, et prenons la forme bilinéaire définie par :

$$\phi((x_1, x_2), (y_1, y_2)) = x_1 y_1$$

Si F et G sont deux droites distinctes, différentes du noyau $K = \text{vect}((0, 1))$. Alors $F^\perp = G^\perp = K$ et $(F \cap G)^\perp = \mathbb{R}^2$. L'inclusion est stricte.

D est une droite isotrope si $D \cap D^\perp \neq \{0\}$. Comme D est de dimension un, cette intersection est D elle-même, donc $D \subset D^\perp$ et D est totalement isotrope. Pour dire la même chose de façon différente, en dimension 1, une forme bilinéaire dégénérée est forcément nulle.

Pour trouver un exemple de sous-espace isotrope sans être totalement isotrope, il faut au moins prendre un plan. Dans l'exercice 6.3, on a bien un plan (tangent au cône) qui est isotrope sans être totalement isotrope.

Si un sous-espace F est isotrope, il contient un vecteur non nul qui est à la fois dans F et dans $F^\perp$: ce vecteur est orthogonal à lui-même, il est isotrope. Si un sous-espace contient un vecteur isotrope, il n'est pas forcément isotrope. Prendre l'exemple de E tout entier, lorsque la forme bilinéaire admet des vecteurs isotropes (cône isotrope non nul) mais est non dégénérée (noyau réduit à $\{0\}$).

Solution 6.6. Pour le premier exemple, la méthode de Gauss conduit à :

$$q(x, y, z) = (x - y - z)^2 + (y - 2z)^2 - 3z^2$$

et la signature est $(2, 1)$.

On considère ensuite

$$\begin{aligned} q(x_1, x_2, x_3) &= x_1x_2 + x_2x_3 + x_3x_1 \\ &= (x_1 + x_3)(x_2 + x_3) - x_3^2 \\ &= \frac{1}{4}(x_1 + x_2 + 2x_3)^2 - \frac{1}{4}(x_1 - x_2)^2 - x_3^2 \end{aligned}$$

le signature est $(1, 2)$, puis

$$\begin{aligned} q(x_1, x_2, x_3, x_4) &= x_1x_2 + x_2x_3 + x_3x_4 + x_4x_1 \\ &= (x_1 + x_3)(x_2 + x_4) - x_3x_4 + x_3x_4 \\ &= \frac{1}{4}(x_1 + x_2 + x_3 + x_4)^2 - \frac{1}{4}(x_1 - x_2 + x_3 - x_4)^2 \end{aligned}$$

la signature est $(1, 1)$. Le lecteur poursuivra.

Solution 6.7. Si q est de signe constant, on peut suivre la démonstration de l'inégalité de Cauchy-Schwarz, énoncée dans le cas positif. On a donc $\phi(x, y)^2 \leq q(x)q(y)$ pour tout (x, y) et donc si x est dans le cône isotrope, il est dans le noyau. Comme l'autre inclusion est automatique, le cône isotrope est égal au noyau.

Si maintenant le cône isotrope est égal au noyau, supposons que $q(x) < 0$ et $q(y) > 0$, et x et y sont nécessairement indépendants (puisque $q(ax) = a^2q(x)$). Alors $q(x + \lambda y)$ est un polynôme du second degré en λ qui a deux racines distinctes, il existe dans le plan engendré par x et y deux vecteurs indépendants isotropes. Mais alors ces deux vecteurs sont dans le noyau, la restriction de la forme à ce plan est nulle, ce qui est absurde vu le choix de x et y .

Solution 6.8. Si ϕ est non dégénérée et admet un vecteur v isotrope, la dimension est au moins égale à 2. Soit v' un vecteur indépendant de v et non orthogonal à v (il en existe car ϕ n'est pas dégénérée). Si $\lambda \in K$, on pose $w = v' + \lambda v$. Alors $q(w) = q(v') + 2\lambda\phi(v', v)$ et il existe λ tel que $q(w) = \alpha$. On a montré que q est surjective, même en se restreignant au plan $\text{vect}(v, v')$. En particulier, il existe une seconde droite de vecteurs isotropes dans ce plan.

Solution 6.9. Le corps de base est celui des rationnels, montrons que les deux formes ne sont pas congruentes. Si elles l'étaient, leurs matrices seraient congruentes et il existerait une matrice de passage P à coefficients rationnels tels que ${}^tPP = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$. En particulier, $\det(P)^2 = 2$, ce qui est impossible puisque $\det P$ est un rationnel.

Solution 6.10. On part des termes en z_1 :

$$\begin{aligned} q(z_1, z_2) &= (z_1 - iz_2)(\bar{z}_1 + i\bar{z}_2) \\ &= Z_1 \bar{Z}_1 \end{aligned}$$

la forme est de rang 1 et de signature $(1, 0)$. Traitons le cas suivant :

$$\begin{aligned} q(z_1, z_2, z_3) &= z_2(\bar{z}_1 - \bar{z}_3) + \bar{z}_2(z_1 - z_3) \\ &= \frac{1}{4}(|z_1 + z_2 - z_3|^2 - |-z_1 + z_2 - z_3|^2) \end{aligned}$$

La forme est de signature $(1, 1)$.

Solution 6.11. Si on pose $x = \sum_{i=1}^n x_i e_i$, alors

$$\langle e_j, x \rangle = \sum_{i=1}^n x_i \langle e_j, e_i \rangle = x_j$$

d'où le résultat. On a tenu compte de la linéarité à droite et on a donc $x = \sum_{i=1}^n \langle e_i, x \rangle e_i$. Attention, dans le cas hermitien, il faut tenir compte de l'ordre. Soit maintenant u un endomorphisme. Si $(u_{i,j})$ est sa matrice dans une base orthonormale,

$$u(e_j) = \sum_{i=1}^n u_{i,j} e_i = \sum_{i=1}^n \langle e_i, u(e_j) \rangle e_i$$

et donc la matrice de u est la matrice $(\langle e_i, u(e_j) \rangle)$.

Solution 6.12. On cherche u_1 et u_2 tels que $u = u_1 + u_2$, $u_1^* = u_1$ et $u_2^* = -u_2$. Nécessairement, $u^* = u_1^* + u_2^* = u_1 - u_2$ et donc

$$u_1 = \frac{1}{2}(u + u^*) \quad \text{et} \quad u_2 = \frac{1}{2}(u - u^*)$$

Réciproquement, ces deux endomorphismes vérifient les conditions de départ. On ne peut parler de somme directe, car l'ensemble des endomorphismes auto-adjoints n'est pas un $\mathbb{C}$ -espace vectoriel. Par ailleurs,

$$u \circ u^* - u^* \circ u = (u_1 + u_2) \circ ((u_1 - u_2) - (u_1 - u_2) \circ (u_1 + u_2)) = 2(u_2 \circ u_1 - u_2 \circ u_2)$$

d'où la seconde affirmation.

En dimension 1, bien sûr, on retrouve la décomposition d'un complexe en un réel et en imaginaire pur.

Solution 6.13.

(1) On a $\langle u(x+y), x+y \rangle = 0$ pour tout x, y , donc

$$\langle u(x), x \rangle + \langle u(y), x \rangle + \langle u(x), y \rangle + \langle u(y), y \rangle = 0 \quad \text{d'où} \quad \langle u(y), x \rangle + \langle u(x), y \rangle = 0$$

De même, on a $\langle u(x+iy), x+iy \rangle = 0$

$$\langle u(x), x \rangle - i\langle u(y), x \rangle + i\langle u(x), y \rangle + \langle u(y), y \rangle = 0 \quad \text{d'où} \quad \langle u(y), x \rangle - \langle u(x), y \rangle = 0$$

On en déduit que, pour tout (x, y) , $\langle u(x), y \rangle = 0$, et donc, puisque qu'un produit scalaire est non dégénéré, $\forall x \in E$, $u(x) = 0$, u est l'endomorphisme nul.

Le résultat n'est pas vrai dans le cas euclidien : prendre pour u l'endomorphisme dont la matrice dans une base orthonormale est $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Le lecteur géomètre reconnaît la matrice d'une rotation d'angle droit.

(2) D'après les définition

$$\langle u(x), x \rangle = \langle x, u^*(x) \rangle = \overline{\langle u^*(x), x \rangle}$$

et donc $\langle u(x), x \rangle \in \mathbb{R}$ équivaut à

$$\langle u(x), x \rangle = \langle u^*(x), x \rangle \iff \langle (u - u^*)(x), x \rangle = 0$$

et ce sera vrai pour tout x si et seulement si u est autoadjoint.

Solution 6.14.

(1) On a

$$(f \otimes g)(x + \lambda x', y) = (f(x) + \lambda f(x'))g(y) = f \otimes g(x, y) + \lambda f \otimes g(x', y)$$

les autres vérifications sont immédiates. Ce sera une forme symétrique lorsque, pour tout (x, y) , $f(x)g(y) = f(y)g(x)$. Si l'une ou l'autre des forme est nulle, le composé est nul donc symétrique. Comme f et g ne sont pas nul, il existe un y_0 tel que ni $g(y_0)$, ni $f(y_0)$ ne sont nuls, et donc $f(x) = \frac{f(y_0)}{g(y_0)}g(x)$, les deux formes sont nécessairement proportionnelles. Et si $g = kf$ où k est un scalaire, $f \otimes g(x, y) = kf(x)f(y)$ est symétrique. Pour terminer, le rang est 1, lorsque ni f ni g n'est nul. En effet,

$$\{x \in E \mid \forall y \in E, f(x)g(y) = 0\} = \text{Ker } f$$

est un hyperplan

(2) Si x a pour coordonnées (x_i) , y a pour coordonnées (y_j) , alors

$$(e_i^* \otimes e_j^*)(x, y) = x_i y_j.$$

Et donc pour toute forme bilinéaire ϕ

$$\phi(x, y) = \sum_{i,j} \phi(e_i, e_j)(e_i^* \otimes e_j^*)(x, y) = x_i y_j$$

les coordonnées de ϕ dans cette base sont les $\phi(e_i, e_j)$, donc sont les coefficients de la matrice de ϕ . Si on est en caractéristique différente de 2, une base de l'espace vectoriel des formes bilinéaires symétriques sera formée des $\frac{1}{2}(e_i^* \otimes e_j^* + e_j^* \otimes e_i^*)$.

(3) Les vérifications sont similaires. On obtient alors une base de $\mathcal{L}(E)$ en considérant les endomorphismes de la forme $e_i^* \otimes e_j$ définis par :

$$(e_i^* \otimes e_j)(x) = x_i e_j$$

Les coordonnées d'un endomorphisme dans cette base sont les coefficients de sa matrice.

Cet exercice est le point de départ d'une étude unifiée des applications linéaires, bilinéaires, multilinéaires, au moyen de ce qu'on appelle le produit tensoriel.

Solution 6.15.

- (1) La bilinéarité est facile à prouver par la définition, ou bien en écrivant $\langle z, w \rangle = z_1 x_1 + z_2 w_2$ où $z = z_1 + iz_2$ et $w = w_1 + iw_2$. C'est un produit scalaire pour lequel $(1, i)$ est une base orthonormale. L'isomorphisme est naturellement $z \mapsto (z_1, z_2)$.
- (2) Il s'agit de $\mathbb{R}$ -linéarité : $m_\gamma(z + \lambda z') = \gamma(z + \lambda z') = \gamma z + \lambda \gamma z'$, avec λ réel. Il est noté qu'ici il y a également $\mathbb{C}$ -linéarité. Cherchons l'adjoint de m_γ . Par définition,

$$\langle m_\gamma(z), z' \rangle = \langle z, m_\gamma^*(z') \rangle \iff \Re(\overline{\gamma} z z') = \Re(\overline{z} m_\gamma^*(z'))$$

On en déduit

$$m_\gamma^*(z) = \overline{\gamma} z \quad \text{et} \quad m_\gamma^* = m_{\overline{\gamma}}$$

Et donc m_γ est symétrique si et seulement si $\gamma = \overline{\gamma}$, donc si γ est réel. De même, il est orthogonal lorsque $\overline{\gamma} = \gamma^{-1}$ donc si γ est de module 1.

Solution 6.16.

- (1) Puisque $A^* = AU$, on en déduit $A^{**} = U^* A^*$ soit $A = U^{-1} AU$ et donc $UA = AU$. On a alors $AA^* = AAU = AUA = A^* A$, A est une matrice normale.
- (2) Commençons par observer que si $\lambda = |\lambda|e^{i\theta}$ alors $\overline{\lambda} = \lambda e^{-2i\theta}$. Supposons donc que $D = \text{diag}(\lambda_k)$, alors

$$D^* = \text{diag}(\overline{\lambda}_k) = \text{diag}(\lambda_k) \text{diag}(e^{-2i\theta_k}) = DU_1$$

où U_1 est unitaire (on a évidemment $U_1^* = U_1^{-1}$).

Soit maintenant A une matrice normale. Elle est diagonalisable dans une base orthonormale, et il existe donc V unitaire telle que $D = V^* AV$, avec D diagonale. On en déduit $D^* = V^* AVU^*$ et donc $A^* = A(VUV^*)$, la proposition est démontrée.

Solution 6.17.

- (1) ${}^t({}^t AA) = {}^t A {}^t t A = {}^t AA$, la matrice est donc symétrique. Si X est la matrice colonne des coordonnées d'un vecteur x , la forme quadratique définie par ${}^t AA$ est :

$$q(x) = {}^t X {}^t AA X = {}^t (AX) AX = \sum_i y_i^2$$

si on note (y_i) les coefficients de la matrice colonne AX . La forme quadratique q est donc positive, et si A est inversible elle est définie positive.

- (2) Montrons qu'il y a une solution avec A symétrique. On utilise que S est semblable à une matrice diagonale D , avec une matrice de passage O qui est orthogonale. On a alors

$$D = {}^tOSO \iff S = OD^tO$$

Puisque S est positive, les valeurs propres de S sont positives et les éléments diagonaux de D sont positifs ou nuls. Il existe une matrice diagonale Δ telle que $\Delta^2 = D$. Alors,

$$\text{si } A = O\Delta^tO \quad \text{alors} \quad {}^tA = O\Delta^tO = A \quad \text{et} \quad {}^tAA = D\Delta^2{}^tO = S$$

Si S est définie positive, on obtient le même résultat en disant que la forme bilinéaire associée admet une base orthonormale et donc que S est congruente à la matrice I .

- (3) Le cas hermitien se traite de façon symétrique...

Chapitre 7

Groupes classiques

7.1 LES GROUPES LINÉAIRES ET SPÉCIAL LINÉAIRES

Quand on étudie un groupe, il est très utile de connaître des générateurs qui soient le plus simple possible. Ainsi, les transpositions sont les générateurs privilégiés du groupe symétrique. Pour l'étude du groupe alterné, on utilise les 3-cycles. Une idée qui va nous guider : les transpositions sont, parmi les permutations différentes de l'identité, celles qui ont le plus de points fixes.

7.1.1. Les dilatations et les transvections

Soit E un K -espace vectoriel de dimension finie n , et H un hyperplan. On s'intéresse aux endomorphismes de E qui fixent tous les éléments de H ; il y en a de deux sortes, comme le précise le théorème suivant.

Théorème 7.1. *H est un hyperplan d'un K -espace vectoriel, et u un endomorphisme différent de l'identité qui fixe les vecteurs de H . Alors*

- *soit u est diagonalisable, de valeurs propres 1 et $\lambda \neq 1$, on dit que u est une **dilatation** de rapport λ .*
- *soit u est non diagonalisable et admet 1 pour seule valeur propre ; on dit que f est une **transvection**.*

Démonstration. Il suffit d'observer que, puisque H est formé de vecteurs invariants, le sous-espace propre E_1 est au moins de dimension $n - 1$ et $(X - 1)^{n-1}$ est facteur du polynôme caractéristique. On a donc

$$\chi_u(X) = (-1)^n (X - 1)^{n-1} (X - \lambda)$$

Alors, si λ n'est pas égal à 1, il existe une droite propre E_λ et u est diagonalisable avec $E = H \oplus E_\lambda$. Si λ est égal à 1 et si $E_1 = E$, u est l'identité (donc diagonalisable). Il reste le cas où $\lambda = 1$ et $E_1 = H$, u est non diagonalisable. $\square$

On peut être plus précis, et caractériser géométriquement les dilatations et les transvections. On en profite pour donner une forme canonique à la matrice de ces transformations.

- Si u est une dilatation autre que l'identité, on note $(e_1, \dots, e_{n-1})$ une base de H et e_n un vecteur propre associé à la valeur propre λ . On dit alors que u est la dilatation d'hyperplan H , de direction $\text{vect}(e_n)$ et de rapport λ . Dans la base $(e_1, e_2, \dots, e_n)$, sa matrice prend la forme bloc :

$$\begin{pmatrix} I_{n-1} & \vdots & 0 \\ \cdots & & \cdots \\ 0 & \vdots & \lambda \end{pmatrix}$$

où I_{n-1} est la matrice de l'identité dans un espace vectoriel de dimension $n - 1$.

- Si u est une transvection, $H = E_1$ est de dimension $n - 1$ donc $\text{Im}(u - \text{id})$ est de dimension un, par le théorème du rang. Notons a un vecteur de base de $\text{Im}(u - \text{id})$. Alors $a = u(b) - b$ où b n'est pas dans H (sinon a serait nul) et

$$u(a) = u(u(b)) - u(b) = (u - \text{id})(u(b)) \in \text{Im}(u - \text{id})$$

Comme $\text{Im}(u - \text{id})$ est de dimension un, il existe un scalaire λ tel que $u(a) = \lambda a$. Mais nécessairement $\lambda = 1$, car 1 est la seule valeur propre de u . En conclusion, a est dans H , l'image $\text{Im}(u - \text{id})$ est incluse dans $\text{Ker}(u - \text{id})$. Prenons $a = e_{n-1}$, $b = e_n$ et $(e_1, \dots, e_{n-2})$, de sorte que $(e_1, \dots, e_{n-1})$ soit une base de H . La matrice de u dans cette base prend la forme :

$$\begin{pmatrix} I_{n-2} & \vdots & 0 \\ \cdots & & \cdots \\ 0 & \vdots & \begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix} \end{pmatrix}$$

Toutes les formes linéaires dont le noyau est H sont proportionnelles, on en choisit une, ϕ , celle pour laquelle $\phi(b) = 1$. Alors, tout vecteur x de E se décompose dans la somme directe $H \oplus \text{vect}(b)$, par $x = x_1 + kb$. Comme $\phi(x) = \phi(x_1) + k\phi(b)$, on a $k = \phi(x)$ et $x = x_1 + \phi(x)b$. Et donc :

$$u(x) = u(x_1) + \phi(x)u(b) = x_1 + \phi(x)(a + b) = x + \phi(x)a$$

Réciproquement, si ϕ est une forme linéaire nulle pour le vecteur a non nul, la transformation définie par

$$\tau_{\phi,a}(x) = x + \phi(x)a$$

est une transvection d'hyperplan le noyau H de ϕ . Il n'y a pas unicité de cette écriture, mais

$$(\tau_{\phi,a} = \tau_{\psi,a'}) \iff \left(\exists k \in K^*, \psi = k\phi \text{ et } a' = \frac{1}{k}a \right)$$

7.1.2. Le groupe spécial linéaire

On appelle **groupe spécial linéaire** le sous-groupe de $\mathbf{GL}(E)$ formé des endomorphismes de déterminant 1. Plus précisément

Proposition 7.2. *L'ensemble des automorphismes de E de déterminant 1 est un sous-groupe distingué de $\mathbf{GL}(E)$ noté $\mathbf{SL}(E)$. Le quotient est isomorphe à $(K^*, \times)$.*

$$\mathbf{GL}(E)/\mathbf{SL}(E) \simeq K^*$$

Démonstration. Il suffit en effet d'appliquer le théorème d'isomorphisme au morphisme « déterminant », qui est surjectif (prendre une matrice diagonale avec tous les coefficients diagonaux égaux à 1, sauf un). Et le noyau est le groupe spécial linéaire. $\square$

Arrivons maintenant au théorème principal.

Théorème 7.3. *Le groupe spécial linéaire est engendré par les transvections, le groupe linéaire est engendré par les transvections et les dilatations.*

Démonstration. Commençons par donner deux propriétés géométriques des transvections :

- *Étant donnés deux vecteurs distincts non nuls u et v , il existe une transvection, ou un produit de deux transvections qui transforment u en v .* En effet, si u et v sont indépendants, il suffit de prendre une transvection définie par $\tau(x) = x + \phi(x)(v - u)$ où ϕ est une forme linéaire nulle pour $v - u$ et valant 1 pour u . On a alors en effet

$$\tau(u) = u + (v - u) = v$$

Si u et v sont dépendants et distincts, il ne peut y avoir une seule transvection (car l'image de u serait v colinéaire à u et distinct de u , or une transvection n'a pas de valeur propre différente de 1). On choisit alors un troisième vecteur w indépendant des deux premiers, et on compose deux transvections, transformant u en w puis w en v .

- *Étant donnés u non nul et H, H' deux hyperplans distincts ne contenant pas u , il existe une transvection qui fixe u et qui transforme H en H' .* En effet, les deux hyperplans ont pour intersection un sous-espace F de dimension $n - 2$; une transvection d'hyperplan $H'' = F \oplus Ku$ va fixer le vecteur u . Soit maintenant $y \in H' \setminus F$.

On peut décomposer y en $x + z$ où x est dans H et z dans H'' (puisque la somme $H + H''$ est l'espace tout entier). Comme précédemment, il existe donc une transvection d'hyperplan H'' qui transforme x en y : on prend $\tau(x) = x + \phi(x)z$, où ϕ est une forme linéaire nulle sur H'' et valant 1 en x . Comme $H = (H \cap H'') \oplus Kx$ et $H' = (H' \cap H'') \oplus Ky$ cette transvection transforme H en H' .

Venons en à la démonstration du théorème. Soit f un élément du groupe linéaire et $\mathcal{B} = (e_1, \dots, e_n)$. Si $f(e_1) \neq e_1$, il existe une transvection t (ou deux transvections t et t' telles que $g = t \circ f$ (ou $g = t \circ t' \circ f$) fixe le vecteur e_1 . Si $H = g(\text{vect}(e_2, e_3, \dots, e_n))$ est distinct de $H' = \text{vect}(e_2, e_3, \dots, e_n)$, il existe une transvection τ qui fixe e_1 et qui envoie H sur H' . Et donc $h = \tau \circ g$ laisse e_1 fixe et stabilise H' . On continue le processus avec la restriction de h à H' , les transvections de H' se prolongent à des transvections de E qui laissent fixe le vecteur e_1 . On arrivera en définitive à une application dont la matrice dans la base $\mathcal{B}$ est de la forme $\text{diag}(1, 1, \dots, 1, \lambda)$. Comme les transvections ont pour déterminant 1, λ est le déterminant de f . Conclusion : cette matrice est celle de l'identité si on est dans $\mathbf{SL}(n, K)$, ou la matrice d'une dilatation sinon. Comme l'application réciproque d'une transvection est une transvection, on a bien démontré le théorème. $\square$

Cette démonstration « géométrique » peut aussi être faite matriciellement : la première étape décrite peut s'interpréter ainsi : si M est une matrice inversible, on peut la multiplier par des matrices de transvections pour obtenir une matrice N de sorte que l'élément n_{11} soit égal à 1, tous les autres de la première ligne et de la première colonne étant nuls. On peut également n'utiliser que des matrices de transvection de la forme $I + \lambda E_{ij}$, et on aboutit donc au même résultat par une méthode de pivot. Donnons maintenant quelques propriétés supplémentaires des transvections.

Proposition 7.4. *Le conjugué d'une transvection de E par un automorphisme de E est une transvection, et toutes les transvections sont conjuguées.*

Démonstration. Soit $\tau_{\phi,a}$ la transvection d'hyperplan $H = \text{Ker } \phi$ et de vecteur $a \in H$. Alors, si g est un automorphisme de E ,

$$g \circ \tau_{\phi,a} \circ g^{-1}(x) = g(g^{-1}(x) + \phi(g^{-1}(x))a) = x + \phi \circ g^{-1}(x)g(a) = \tau_{\psi,g(a)}(x)$$

où $\psi = \phi \circ g^{-1}$, ce qui définit une transvection de plan $g(H)$ et de vecteur $g(a)$, puisque $\text{Ker}(\phi \circ g^{-1}) = g(\text{Ker } \phi)$ comme on le vérifie aisément. On déduit également de ce calcul que deux transvections sont toujours conjuguées dans $\mathbf{GL}(E)$, puisque, si on se donne H et H' d'équations respectives ϕ et ψ , a dans H et b dans H' , on peut construire g de sorte que $g(a) = b$ et $g(H) = H'$, en complétant a et b en une base de H (resp. de H'), et en complétant la base de H (resp. de H') par un vecteur e_n tel que $\phi(e_n) = 1$ (resp. e'_n tel que $\psi(e'_n) = 1$) en une base de E . On aura alors $\phi \circ g^{-1} = \psi$, puisque ces applications coïncident pour les vecteurs de H et pour e_n . $\square$

1. $e_1 = g(e_1)$ n'est pas dans H car le rang de g est n .

Une petite application (déjà obtenue dans l'exercice 4.1) :

Proposition 7.5. *Le centre de $\mathbf{GL}(E)$ est formé des homothéties.*

Démonstration. Si g est dans le centre de $\mathbf{GL}(E)$, g doit commuter à toutes les transvections, d'après le calcul précédent, on doit avoir $g(a)$ colinéaire à a pour tout a , ce qui impose que g est une homothétie. Réciproquement, toute homothétie commute aux éléments de $\mathbf{GL}(E)$. $\square$

Mais la seconde partie du théorème est plus intéressante : on en déduit que si un sous-groupe distingué de $\mathbf{GL}(E)$ contient **une** transvection, il les contient toutes, donc il contient tout le groupe spécial linéaire.

Exercice 7.1. Montrer que l'ensemble des transvections ayant un hyperplan donné H (auxquelles on joint l'identité) est un groupe G pour la composition, isomorphe au groupe $(H, +)$. Montrer que la réunion de ce groupe avec l'ensemble des dilatations ayant le même hyperplan H est encore un groupe G' , et montrer que G est distingué dans G' .

Exercice 7.2. Soit $K = \mathbb{F}_q$ un corps fini ayant q éléments. Déterminer le nombre d'éléments de $\mathbf{GL}(n, K)$, de $\mathbf{SL}(n, K)$: on remarquera qu'il suffit, pour se donner un élément de $\mathbf{GL}(E)$, de se donner l'image d'une base. Montrer que le groupe $\mathbf{SL}(n, \mathbb{F}_2)$ est isomorphe à $\mathcal{S}_3$.

7.2 LE GROUPE ORTHOGONAL

7.2.1. Généralités

Soit E un espace vectoriel euclidien, de dimension n . L'objectif de ce paragraphe est d'étudier le groupe orthogonal $\mathbf{O}(E)$ des automorphismes de E qui conservent le produit scalaire. On peut commencer par rappeler

Théorème 7.6. *Un élément f de $\mathbf{GL}(E)$ est une **transformation orthogonale** (ou une **isométrie**) si et seulement si l'une des conditions suivantes est réalisée :*

- $\forall (x, y) \in E^2, \quad \langle f(x), f(y) \rangle = \langle x, y \rangle \quad (\text{c'est la définition.})$
- $\forall x \in E, \quad \|f(x)\| = \|x\|$
- *L'image d'une base orthonormale est une base orthonormale.*

L'ensemble des transformations orthogonales de E est un sous-groupe de $\mathbf{GL}(E)$. On le note $\mathbf{O}(E)$, c'est le groupe orthogonal de E .

Démonstration. Si f conserve le produit scalaire, f conserve la norme. Réciproquement, la relation de polarisation montre qu'une application linéaire qui conserve la norme conserve le produit scalaire :

$$\begin{aligned}\langle f(x), f(y) \rangle &= \frac{1}{2} (\|f(x) + f(y)\|^2 - \|f(x) - f(y)\|^2) \\ &= \frac{1}{2} (\|f(x+y)\|^2 - \|f(x-y)\|^2) \\ &= \frac{1}{2} (\|x+y\|^2 - \|x-y\|^2) \\ &= \langle x, y \rangle\end{aligned}$$

Par ailleurs, si f conserve norme et produit scalaire, alors f transforme toute base orthonormale en une base orthonormale. Réciproquement, si (e_i) est une base orthonormale et si f est une application linéaire telle que $(f(e_i))$ soit aussi orthonormale, alors pour tout x de E :

$$\|f(x)\|^2 = \left\| f \left(\sum_i x_i e_i \right) \right\|^2 = \left\| \sum_i x_i f(e_i) \right\|^2 = \sum_i x_i^2 = \|x\|^2$$

De plus, on a immédiatement, pour f et g dans $\mathbf{O}(E)$

$$\|f \circ g(x)\| = \|g(x)\| = \|x\| \quad \text{et} \quad \|f^{-1}(x)\| = \|f \circ f^{-1}(x)\| = \|x\|.$$

□

Si $\mathcal{B}$ est une base orthonormale, la matrice O d'une transformation orthogonale est... orthogonale, c'est-à-dire qu'elle vérifie ${}^tOO = I$. En particulier,

$$\det({}^tOO) = \det({}^tO) \det(O) = \det(O)^2 = 1$$

Donc le déterminant d'une transformation orthogonale est ± 1 .

Définition 7.7. L'ensemble des transformations orthogonales de déterminant $+1$ est un sous-groupe distingué de $\mathbf{O}(E)$; c'est le groupe spécial orthogonal, noté $\mathbf{SO}(E)$. Ses éléments sont aussi appelées transformations orthogonales directes. [note : les autres transformations orthogonales sont dites indirectes]

La démonstration est analogue celle vue pour le groupe spécial linéaire : le groupe spécial orthogonal est le noyau du morphisme déterminant, restreint au groupe orthogonal. Remarquons que le groupe spécial orthogonal contient au moins l'identité. Nous verrons dans un paragraphe suivant que $\mathbf{O}(E) \setminus \mathbf{SO}(E) = \mathbf{O}^-(E)$ est non vide : le groupe spécial orthogonal est donc d'indice 2.

Terminons ce paragraphe par l'exemple immédiat de la dimension 1 : le groupe orthogonal ne contient que $\pm \text{id}_E$, la seule transformation directe est id_E .

7.2.2. Images de sous-espaces

Parmi les propriétés des transformations orthogonales, l'une est particulièrement utile, car elle permet souvent de décrire une transformation en se restreignant à des sous-espaces.

Proposition 7.8. *Soit f une transformation orthogonale de E . Alors, si F est stable par f , son supplémentaire orthogonal $F^\perp$ est aussi stable par f .*

Démonstration.

$$\forall x \in F, \forall y \in F^\perp, \quad \langle f(y), f(x) \rangle = \langle x, y \rangle = 0$$

et, par hypothèse $f(x)$ parcourt $F = f(F)$, donc l'image de $F^\perp$ est incluse dans $F^\perp$, il y a égalité puisque f est bijective. On peut aussi se rappeler des résultats concernant l'adjoint d'un endomorphisme. $\square$

Parmi les sous-espaces stables d'un endomorphisme, il y a les sous-espaces propres. Pour une transformation orthogonale f , les seules valeurs propres possibles sont 1 et -1 , puisque $\|f(x)\| = \|x\|$. On a alors :

Proposition 7.9. *Si f est une transformation orthogonale,*

$$\text{Ker}(f - \text{id}) \oplus \text{Im}(f - \text{id}) = E$$

et cette somme est orthogonale.

Démonstration. Soit $x \in \text{Ker}(f - \text{id})$ et $y \in \text{Im}(f - \text{id})$. Il existe $z \in E$ tel que $y = f(z) - z$ et :

$$\langle x, y \rangle = \langle x, f(z) - z \rangle = \langle f(x), f(z) \rangle - \langle x, z \rangle = \langle x, z \rangle - \langle x, z \rangle = 0$$

Les sous-espaces sont donc orthogonaux, le théorème du rang permet d'affirmer qu'ils sont supplémentaires orthogonaux. $\square$

7.2.3. Les réflexions

Comme dans tout groupe, il est intéressant d'étudier les éléments dont l'ordre est petit ; on sait déjà que les éléments d'ordre deux du groupe linéaire sont les symétries vectorielles. Dans le cas euclidien, on peut dire :

Proposition 7.10.

- *Un automorphisme involutif s de E est une transformation orthogonale si et seulement si c'est une symétrie par rapport à un sous-espace vectoriel F , suivant son orthogonal $F^\perp$. Dans le cas où F est un hyperplan, on dit que s est une **réflexion**, dans le cas où F est de codimension 2, on dit que s est un **retournement**.*
- *Les symétries sont dans le groupe spécial orthogonal lorsque la codimension de F est paire. En particulier, les réflexions sont toujours dans $\mathbf{O}^-(E)$, les retournements dans $\mathbf{SO}(E)$*

Démonstration. On se donne $E = F \oplus F'$, s est la symétrie par rapport à F suivant F' . Alors, si $x = x_1 + x_2$ est l'écriture d'un vecteur x quelconque dans cette somme directe,

$$\|s(x)\|^2 - \|x\|^2 = \|x_1 - x_2\|^2 - \|x_1 + x_2\|^2 = -4\langle x_1, x_2 \rangle$$

et s sera une transformation orthogonale si et seulement si tout vecteur de F est orthogonal à tout vecteur de F' . Enfin, l'écriture dans une base de diagonalisation (que l'on peut prendre d'ailleurs orthonormale), prouve que le déterminant de s est $(-1)^{\dim F^\perp}$. $\square$

Muni de ce vocabulaire, on peut donner maintenant des générateurs du groupe orthogonal et du groupe spécial orthogonal.

Théorème 7.11. *Le groupe $O(E)$ est engendré par les réflexions, et toute isométrie peut s'écrire comme composée de k réflexions où $k = n - \dim(\text{Ker}(f - \text{id}_E))$. Le groupe $SO(E)$ est engendré par les retournements, quand la dimension de E est supérieure ou égale à trois.*

Démonstration. Commençons par noter l'analogie avec le groupe S_n qui est engendré par les transpositions, tandis que le groupe $\mathcal{A}_n$ est engendré par les 3-cycles ($n \geq 3$).

Cas du groupe orthogonal. On procède par récurrence sur la dimension, le résultat étant immédiat en dimension 1 puisque la seule réflexion est $s = -\text{id}$ et $s \circ s = \text{id}$. Supposons le théorème vrai pour la dimension n et soit f une transformation orthogonale de E de dimension $n + 1$.

- Si f admet un vecteur invariant non nul x , l'hyperplan $H = \{x\}^\perp$ est stable par f , et la restriction f' de f à cet hyperplan est composée de k réflexions s_i . On peut prolonger ces réflexions de l'hyperplan H en des réflexions σ_i de l'espace tout entier, il suffit de poser $\sigma_i(x) = x$ et $\sigma_i|_H = s_i$. Ainsi f est composé de k réflexions. De plus,

$$\begin{aligned} k = n - \dim(\text{Ker}(f' - \text{id}_H)) &= (n + 1) - (\dim(\text{Ker}(f' - \text{id}_H)) + 1) \\ &= (n + 1) - \dim(\text{Ker}(f - \text{id}_E)) \end{aligned}$$

puisque $\text{Ker}(f - \text{id}_E) = \text{Ker}(f' - \text{id}_H) \oplus \text{vect}(x)$

- Si f n'admet aucun vecteur invariant non nul, soit x quelconque non nul et $y = f(x)$. Alors il existe une réflexion s telle que $s(y) = x$; il suffit de prendre la réflexion d'hyperplan $\{y - x\}^\perp$: en effet, cet hyperplan contient alors $x + y$ puisque :

$$\langle x - y, x + y \rangle = \langle x - f(x), x + f(x) \rangle = \langle x, x \rangle - \langle f(x), f(x) \rangle = 0$$

d'où

$$x = \frac{1}{2}(x + y) + \frac{1}{2}(x - y) \text{ donc } s(x) = \frac{1}{2}(x + y) - \frac{1}{2}(x - y) = y$$

On considère alors la transformation orthogonale $g = s \circ f$. Elle admet x comme vecteur invariant, donc est composée de réflexions. De plus, $\text{Ker}(g - \text{id}_E)$ est de

dimension 1, car si g admettait au moins un plan de vecteurs invariants, il y aurait un vecteur invariant non nul dans l'hyperplan de s , et ce vecteur serait invariant par g et par s donc par f . En utilisant la première partie de la démonstration, g est composé de n réflexions, donc $f = s \circ g$ est composé de $n + 1$ réflexions.

Cas du groupe spécial orthogonal. Tout élément de $\mathbf{SO}(E)$ est composé d'un nombre pair de réflexions. Nous allons montrer que le composé de deux réflexions distinctes est égal au composé de deux retournements. En dimension 3, si s_H est la réflexion de plan H , alors $-s_H$ est le retournement d'axe $H^\perp$: en effet, le rôle des valeurs propres 1 et -1 est échangé. On a donc $s_H \circ s_{H'} = (-s_H) \circ (-s_{H'})$ et le résultat est démontré. En dimension supérieure, on se ramène à ce cas : si $H = \{x\}^\perp$ et $H' = \{y\}^\perp$ sont deux hyperplans distincts, et z un vecteur non nul de leur intersection, alors $L = \text{vect}(x, y, z)$ est un sous-espace de dimension 3, stable par les réflexions. Le composé de ces restrictions est égal au composé des deux retournements d'axes $\text{vect}(x)$ et $\text{vect}(y)$. En prolongeant ses retournements par id de $L^\perp$, on obtient deux retournements de l'espace tout entier, dont le composé est égal au composé des réflexions.

Bien entendu, le résultat est faux en dimension deux (il n'y a qu'un seul retournement, $-\text{id}$, qui n'engendre pas le groupe spécial orthogonal). $\square$

Exercice 7.3. Soit E un espace vectoriel euclidien et u un vecteur non nul. Démontrer que l'application s définie par

$$\forall x \in E \quad s(x) = x - 2 \frac{\langle u, x \rangle}{\langle u, u \rangle} u$$

est la réflexion d'hyperplan $\{u\}^\perp$. Application : en dimension 3, écrire la forme générale de la matrice d'une réflexion, en base orthonormale.

Exercice 7.4. Soit $E = \mathbb{R}$ muni de sa structure euclidienne canonique. Trouver une application u de E dans E qui conserve la norme mais qui n'est pas une transformation orthogonale.

7.3 LA DIMENSION DEUX

7.3.1. Rotations et réflexions planes

On a vu que le groupe orthogonal d'une droite euclidienne est simplissime puisqu'il se réduit à $\pm \text{id}_E$. Le groupe orthogonal d'un plan vectoriel euclidien est plus riche, mais on peut le décrire explicitement assez facilement.

Théorème 7.12. Si E est de dimension 2, le groupe $\mathbf{SO}(E)$ est commutatif. Ses éléments sont appelés **rotations**. L'ensemble des transformations orthogonales indirectes est formé de toutes les réflexions.

Démonstration. Le théorème général prouve que, puisque la dimension est deux, toute transformation orthogonale est composée de une ou de deux réflexions. Le composé de deux réflexions est forcément un élément du groupe spécial orthogonal, d'où la seconde affirmation du théorème. Soit maintenant $r \in \mathbf{SO}(E)$, et s une réflexion. Alors, $s \circ r \circ s = r^{-1}$. En effet, $s \circ r$ est un élément de $\mathbf{O}(E) \setminus \mathbf{SO}(E)$, donc est une réflexion, donc est involutive :

$$s \circ r = (s \circ r)^{-1} = r^{-1} \circ s \text{ d'où } s \circ r \circ s = r^{-1}$$

Si maintenant $r' = s' \circ s$ est une rotation quelconque,

$$r' \circ r \circ r'^{-1} = s' \circ (s \circ r \circ s^{-1}) \circ s' = s' \circ r^{-1} \circ s' = r$$

et donc r et r' commutent. $\square$

Remarque : Toute rotation est composée de deux réflexions. Il peut être utile de remarquer que l'une des deux peut être choisie arbitrairement. En effet, si on se donne r rotation et s réflexion, alors $r \circ s$ est une isométrie indirecte donc une réflexion s' , on a alors $r \circ s = s'$ d'où $r = s' \circ s$. On peut faire le même calcul en échangeant les rôles de s et de s' .

7.3.2. Matrices orthogonales en dimension deux

Dans le cas du plan on peut donner la forme explicite des matrices orthogonales.

Proposition 7.13. *Les éléments de $\mathbf{O}(2, \mathbb{R})$ sont de la forme*

$$M = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

où θ est un réel. De même, les éléments de $\mathbf{O}(2, \mathbb{R}) \setminus \mathbf{SO}(2, \mathbb{R})$ sont de la forme

$$N = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}, \quad \theta \in \mathbb{R}$$

Démonstration. Soit $\mathcal{B} = (e_1, e_2)$ une base orthonormale. Si f est un élément de $\mathbf{O}(E)$, $r(e_1)$ est un vecteur unitaire, dont les coordonnées (a, b) dans la base $\mathcal{B}$ vérifient $a^2 + b^2 = 1$. Il existe donc un réel θ tel que $a = \cos \theta$ et $b = \sin \theta$. Comme $f(e_2)$ est orthogonal à $f(e_1)$, il a pour coordonnées (a', b') tels que $aa' + bb' = 0$. Donc (a', b') est proportionnel à $(-b, a)$. Comme de plus (a', b') doit être unitaire, il y a deux possibilités seulement, $(a', b') = (-b, a) = (-\sin \theta, \cos \theta)$ ou $(a', b') = (b, -a) = (\sin \theta, -\cos \theta)$. Dans le premier cas, la matrice est de déterminant 1, c'est la matrice d'un élément de $\mathbf{SO}(E)$, dans le second cas, c'est la matrice d'une réflexion. $\square$

Les deux ensembles de matrices ainsi décrits se notent $\mathbf{SO}(2, \mathbb{R})$ et $\mathbf{O}^-(2, \mathbb{R})$, groupe des matrices orthogonales directes, ensemble des matrices orthogonales indirectes respectivement.

7.3.3. Orientation du plan

Examinons plus précisément la matrice d'une rotation donnée, en base orthonormale. Si M est la matrice de r dans une base $\mathcal{B}$ et M' sa matrice dans une autre base $\mathcal{C}$, également orthonormale, alors la matrice de passage P est orthogonale et $M' = P^{-1}MP$. Mais si P est orthogonale directe, les matrices P et M commutent et $M' = M$. Si par contre P est indirecte, le calcul fait plus haut montre que $M' = M^{-1} = {}^tM$. On décide donc la convention suivante.

Définition 7.14. On dit que l'on a **orienté** le plan vectoriel euclidien si on a choisi une base orthonormale. Les bases qui s'en déduisent par un élément de $\text{SO}(E)$ seront appelées **directes**, les autres seront **indirectes**.

On peut alors attacher à chaque rotation une unique matrice orthogonale, de la forme $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ celle qui la représente dans toutes les bases directes.

Remarque : La situation est très différente pour les réflexions : si s est une réflexion, sa matrice en base orthonormale est orthogonale indirecte, mais change lorsqu'on change de base orthonormale. En particulier, il existe toujours une base orthonormale telle que la matrice de s prenne la forme $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, c'est une base de diagonalisation de s .

Rappelons maintenant le résultat d'Analyse suivant : l'application $\theta \mapsto e^{i\theta}$ est un morphisme surjectif du groupe $(\mathbb{R}, +)$ sur le groupe multiplicatif $\mathbb{U}$ des nombres complexes de module 1. Son noyau est $2\pi\mathbb{Z}$. On a donc :

$$(\mathbb{R}/2\pi\mathbb{Z}, +) \simeq (\mathbb{U}, \times)$$

On utilise ce résultat pour obtenir la proposition

Proposition 7.15. L'application de $\mathbb{R}$ dans $\text{SO}(E)$ donnée par $\theta \mapsto \text{rot}(\theta)$ où $\text{rot}(\theta)$ est la rotation dont la matrice dans une base orthonormale directe est $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ est un morphisme surjectif de groupes, de noyau $2\pi\mathbb{Z}$. On dit que θ (ou sa classe modulo 2π .) est l'**angle de la rotation**, et on a

$$\text{rot}(\theta) \circ \text{rot}(\theta') = \text{rot}(\theta + \theta')$$

Démonstration. Il suffit d'observer que le groupe $\text{SO}(2, \mathbb{R})$ est isomorphe à $\mathbb{U}$. Cela résulte de ce que l'ensemble des matrices de la forme

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = aI + bJ, \quad a, b \in \mathbb{R}$$

est un corps isomorphe au corps $\mathbb{C}$ des complexes, en associant à une telle matrice le complexe $a + bi$, car la matrice J a pour carré $-I$. Tout nombre complexe de la forme $e^{i\theta}$ correspond à la matrice $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$. □

7.3.4. Les angles de vecteurs en dimension deux

La notion d'angle est fondamentale en géométrie. Dans ce paragraphe nous allons nous contenter d'une introduction, en restant dans le cadre d'un espace vectoriel euclidien de dimension 2 ou 3. On peut s'aider d'une analogie, en regardant le cours de géométrie, entre le cas des rotations et des vecteurs unitaires d'un côté, des translations et des points d'un autre côté.

Théorème 7.16. *Le groupe des rotations agit simplement transitivement sur l'ensemble des vecteurs unitaires du plan vectoriel.*

Démonstration. Soit u et v deux vecteurs unitaires du plan vectoriel. On peut compléter u par u' , de sorte que (u, u') soit une base orthonormale. Il est alors possible de trouver une rotation unique qui transforme u en v , celle dont la matrice dans la base (u, v) a pour première colonne les coordonnées de v dans la base (u, v) . Il y a alors unicité, car la seconde colonne est imposée. Remarquons qu'il y a aussi unicité de la réflexion qui transforme u en v . $\square$

Si u et v sont deux vecteurs unitaires du plan, il existe donc une seule rotation r telle que $v = r(u)$. Dans le cas où le plan est orienté, on peut donc énoncer :

Définition 7.17. *Soit u et v deux vecteurs unitaires. Si la rotation qui transforme u en v est d'angle θ , on dit que θ est l'angle du couple (u, v) et on écrit :*

$$\theta = \widehat{(u, v)}$$

*L'ensemble $\mathbb{R}/2\pi\mathbb{Z}$ est noté $\mathcal{A}$, c'est le **groupe des angles**.*

On notera 0 (angle nul), l'élément neutre de $\mathcal{A}$, c'est l'angle de l'identité, et donc $\widehat{(u, u)} = 0$ pour tout u . Dans le même ordre d'idées, l'angle plat est l'angle de la rotation $-\text{id}_E = \text{rot}(\pi)$. La proposition suivante donne des propriétés importantes des angles.

Proposition 7.18.

$$\forall u, v, w \in E, \widehat{(u, v)} + \widehat{(v, w)} = \widehat{(u, w)}$$

$$\forall u, v \in E, \widehat{(u, v)} = -\widehat{(v, u)}$$

$$\forall u, v, u', v' \in E, \widehat{(u, v)} = \widehat{(u', v')} \iff \widehat{(u, u')} = \widehat{(v, v')}$$

Démonstration. La première propriété est connue sous le nom de « relation de Chasles ». Soit r telle que $v = r(u)$ et r' telle que $w = r'(v)$. On a donc $w = r'(r(u))$ et il suffit d'appliquer la définition ci-dessus concernant la somme des angles.

La seconde propriété résulte de l'équivalence $v = r(u) \iff u = r^{-1}(v)$, les angles $\widehat{(u, v)}$ et $\widehat{(v, u)}$ correspondent donc à des angles opposés.

Montrons la troisième propriété. Soit r et r' les rotations telles que $v = r(u)$ et $u' = r'(u)$. Alors, si $\widehat{(u, v)} = \widehat{(u', v')}$, on a

$$v' = r(u') = r \circ r'(u) = r' \circ r(u) = r'(v)$$

et on a donc bien $\widehat{(u, u')} = \widehat{(v, v')}$. On peut remarquer le rôle décisif de la commutativité du groupe des rotations. $\square$

Les transformations orthogonales conservent les produits scalaires : quelle est leur action sur les angles ?

Proposition 7.19. *Les rotations conservent les angles de vecteurs, les réflexions les opposent.*

Démonstration. Soient $\widehat{(u, v)}$ un angle de vecteurs unitaires, r la rotation telle que $v = r(u)$ et s une réflexion. Alors, comme nous l'avons vu dans la démonstration du théorème 7.12, $s \circ r \circ s = r^{-1}$ d'où $s(v) = s(r(u)) = r^{-1}(s(u))$ et donc $\widehat{(s(u), s(v))} = -\widehat{(u, v)}$. Comme une rotation est composée des deux réflexions, en appliquant deux fois de suite ce résultat, on obtient qu'une rotation conserve les angles. $\square$

Terminons en remarquant que pour les angles, on confond souvent un élément de $\mathbb{R}/2\pi\mathbb{Z}$ avec un de ses représentants, que l'on appelle une de ses mesures. Ainsi l'angle nul est l'angle dont une mesure est de la forme $2k\pi$, avec $k \in \mathbb{Z}$, ou dont les mesures sont congrues à 0 modulo 2π . Le représentant qui est dans l'intervalle $] -\pi, \pi]$ s'appelle souvent mesure principale de l'angle.

7.3.5. Autres angles

La notion d'angle peut être généralisée, mais il n'est pas toujours possible de faire de l'ensemble des angles un groupe comme nous l'avons fait pour l'ensemble des angles orientés de vecteurs unitaires. Donnons seulement des pistes qui seront prolongées dans les chapitres de géométrie.

- Angle orienté de vecteurs non nuls u, v du plan. C'est par définition l'angle (orienté) des vecteurs unitaires $\frac{u}{\|u\|}$ et $\frac{v}{\|v\|}$. Comme pour les vecteurs unitaires, un tel angle (confondu avec une mesure θ) est caractérisé par le cosinus et le sinus :

$$\widehat{(u, v)} = \theta \iff \begin{cases} \cos \theta &= \frac{\langle u, v \rangle}{\|u\| \|v\|} \\ \sin \theta &= \frac{\det_{\mathcal{B}}(u, v)}{\|u\| \|v\|} \end{cases}$$

où $\mathcal{B}$ est orthonormale directe. En particulier, le déterminant des deux vecteurs est indépendant de la base orthonormale directe.

- Angle orienté de droites vectorielles D et D' du plan. Si u et v sont des vecteurs directeurs respectifs de D et de D' , l'angle (D, D') est la réunion des angles $\widehat{(u, v)}$ et $\widehat{(u, -v)}$. Ces angles ont des mesures qui diffèrent de π , et le groupe $\mathcal{A}'$ des angles de droites est isomorphe au groupe $\mathbb{R}/\pi\mathbb{Z}$ muni de l'addition. On vérifie facilement qu'un tel angle orienté est caractérisé par sa tangente.
- Angle non orienté de vecteurs non nuls u et v d'un espace vectoriel euclidien (de dimension quelconque). C'est le réel

$$\widehat{u, v} = \arccos \left(\frac{\langle u, v \rangle}{\|u\| \|v\|} \right)$$

qui est donc dans l'intervalle $[0, \pi]$. Si u et v sont non colinéaires et si le plan qu'ils engendrent est orienté, il existe une mesure de l'angle orienté dans $] -\pi, \pi]$, et l'angle non orienté est la valeur absolue de cette mesure. On a ainsi $\widehat{u, v} = \widehat{v, u}$, ce qui justifie le vocabulaire « angle non orienté ». L'ensemble des angles non orientés ne peut être muni d'une structure de groupe de sorte que la relation de Chasles soit satisfaite.

Terminons en remarquant que l'on peut orienter un espace vectoriel euclidien de dimension quelconque comme nous l'avons fait en dimension 2, en choisissant une base orthonormale. Mais cette orientation n'induit pas une orientation des sous-espaces. Même si un espace de dimension 3 a été orienté, il n'en résulte pas d'orientation canonique des plans, on ne peut alors définir l'angle orienté d'un couple de vecteurs de l'espace.

Exercice 7.5. Soit $D = \text{vect}(u)$ une droite vectorielle du plan vectoriel euclidien orienté. On suppose que u a pour coordonnées $(\cos \phi, \sin \phi)$ dans une base orthonormale directe $\mathcal{B}$. Donner la matrice de la réflexion s_D d'axe D .

Avec les hypothèses précédentes, on suppose que D' est dirigée par u' de coordonnées $(\cos \phi', \sin \phi')$, déterminer la rotation $s_{D'} \circ s_D$.

Exercice 7.6. Montrer que l'application $\theta \mapsto \theta + \theta$ est un isomorphisme entre le groupe des angles orientés de droites et le groupe des angles orientés de vecteurs. Lien avec l'exercice précédent ?

Exercice 7.7. Déterminer, dans le groupe des angles, les éléments d'ordre 2 et les éléments d'ordre 4.

7.4 DÉCOMPOSITION DES TRANSFORMATIONS ORTHOGONALES

7.4.1. La dimension trois

Nous allons maintenant examiner le cas des transformations orthogonales d'un espace vectoriel euclidien de dimension trois. Moins simple qu'en dimension deux, cette étude est néanmoins très importante pour les applications à la géométrie. Le théorème fondamental sur les transformations orthogonales se précise :

Théorème 7.20. *Soit E un espace vectoriel euclidien de dimension 3. Si $f \neq \text{id}$ est un élément de $\text{SO}(E)$, l'ensemble des ses vecteurs invariants est une droite vectorielle D , et la restriction de f à $D^\perp$ est une rotation d'angle non nul. On dit que f est une rotation d'axe D . Si f est un élément de $\text{O}(E) \setminus \text{SO}(E)$, soit c 'est une réflexion, soit c 'est le composé commutatif d'une rotation d'axe D et de la réflexion par rapport à $D^\perp$.*

Démonstration. On sait qu'une transformation orthogonale f est la composée d'une, de deux ou de trois réflexions : si elle est directe, c'est la composée de deux réflexions. Examinons ce cas. Si les réflexions sont les mêmes, c'est l'identité. Sinon, $f = s_H \circ s_{H'}$ et $D = H \cap H'$ est formée de vecteurs invariants. De plus, $H'' = D^\perp$ est un plan, stable par f . La restriction de f à ce plan est composée des deux réflexions par rapport aux droites $H \cap H''$ et $H' \cap H''$, c'est donc une rotation. On peut observer que toute isométrie admettant pour ensemble de vecteurs invariants une droite D est réciproquement composée de deux réflexions, car la restriction à $H'' = D^\perp$ est une transformation orthogonale n'ayant pas de vecteurs invariants, donc composée de deux réflexions de ce plan : il est alors aisé de reconstruire les plans H et H' . De même, une transformation orthogonale admettant pour ensemble de vecteurs invariants un plan H est une réflexion, car la restriction à $H^\perp$ ne peut être que $-\text{id}$.

Reste à examiner le cas où f est composée de trois réflexions, et dans ce cas f ne peut avoir aucun vecteur invariant autre que 0. Comme on est en dimension 3, le polynôme caractéristique admet au moins une valeur propre réelle, qui ne peut être que -1 . Soit D la droite engendrée par un vecteur propre. Alors la restriction de f à $D^\perp$ est une transformation orthogonale du plan sans vecteur invariant, c'est une rotation. Si on note s la réflexion de plan $D^\perp$ et r la rotation de l'espace ayant pour axe D et même restriction que f à $D^\perp$, il est facile de vérifier que $f = r \circ s = s \circ r$. Sauf dans le cas de $-\text{id}$, il y a unicité de r et de s .

□

Remarque : Il est intéressant de remarquer que l'application $f \mapsto -f$ est une bijection de $\text{SO}(E)$ sur $\text{O}(E) \setminus \text{SO}(E)$ (cela est vrai plus généralement en dimension impaire).

Autre remarque : supposons que f soit une rotation d'axe D et que l'espace E soit orienté ; alors on peut orienter D (i.e. choisir une base orthonormale $\mathcal{B}_1 = (k)$) et décider d'orienter $D^\perp$ en choisissant une base orthonormale $\mathcal{B}_2 = (i, j)$ de sorte que $\mathcal{B} = (i, j, k)$ soit directe. Il est alors possible de mesurer l'angle de la rotation f en disant que c'est l'angle (orienté) de sa restriction dans le plan orienté $D^\perp$.

7.4.2. Cas général

On peut donner des résultats qui généralisent les études en dimension 2 et 3.

Proposition 7.21. *Soit f une transformation orthogonale de l'espace vectoriel euclidien E . Il existe des sous-espaces $E_1, E_{-1}, P_1, \dots, P_k$ tels que :*

- (1) *Les P_i sont de dimension 2.*
- (2) *E est somme directe orthogonale de tous ces sous-espaces.*
- (3) *Ces sous-espaces sont stables par f et : la restriction de f à E_1 est l'identité, la restriction de f à E_{-1} est $-\text{id}$, la restriction de f à P_k est une rotation d'angle ni nul, ni plat.*

Démonstration. On procède par récurrence sur la dimension. Les sous-espaces E_1 et E_{-1} sont les sous-espaces propres associés à la valeur propre 1 (resp. -1), ou bien sont réduits au vecteur nul. Si donc on admet que la proposition est vraie en dimension strictement inférieur à n , on a :

- Soit f admet la valeur propre réelle 1 et $E = E_1 \oplus E_1^\perp$, on applique l'hypothèse de récurrence à la restriction de f à $E_1^\perp$.
- Même démarche si f a la valeur propre réelle -1 .
- Reste le cas où f n'a pas de valeur propre réelle. Montrons qu'il existe un plan stable : on sait que le polynôme minimal de f peut s'écrire

$$\mu_f(X) = (X^2 + aX + b)R(X) \text{ où } X^2 + aX + b \text{ est sans racine réelle.}$$

Soit alors $x \in \text{Ker}(f^2 + af + b\text{id}_E)$, x non nul. Alors x et $f(x)$ engendrent un plan P puisque x n'est pas vecteur propre, et $f^2(x) = -af(x) - bx$ est dans P . Donc l'image de tout vecteur combinaison de x et de $f(x)$ est dans P . Ce plan est stable, la restriction de f à P est une rotation (il n'y a pas de vecteur propre), différente de $\pm \text{id}$. On applique ensuite l'hypothèse de récurrence au supplémentaire orthogonal de P .

Une remarque : il n'y a pas unicité en général de la suite des plans P_i , mais les $\cos \theta_i$ sont déterminés.

□

EXERCICES

Exercice 7.8. Soit E un espace vectoriel euclidien. Soit f une transformation antisymétrique, c'est-à-dire telle que $f^* = -f$.

- (1) Montrer que $f + \text{id}$ et $f - \text{id}$ sont bijectives.
- (2) On pose $u = (\text{id} - f) \circ (\text{id} + f)^{-1}$. Montrer que u est dans $\text{SO}(E)$. Vérifier que -1 n'est pas valeur propre de u .
- (3) Préciser u en fonction de la matrice de f dans une base orthonormale, dans les cas de la dimension 2 et de la dimension 3.
- (4) On se donne v de $\text{SO}(E)$ n'admettant pas la valeur propre -1 . Trouver f antisymétrique telle que $v = (\text{id} - f) \circ (\text{id} + f)^{-1}$.

Exercice 7.9.

- (1) Soit $r \in \text{SO}(E)$ où E est euclidien de dimension 3. Montrer qu'il existe une base orthonormale $\mathcal{B}$ telle que la matrice de r dans cette base soit de la forme

$$M_{\mathcal{B}}(r) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$$

- (2) Soit A une matrice orthogonale directe en dimension 3. Montrer que c'est la matrice en base orthonormale d'un retournement (on dit aussi demi-tour) si et seulement si sa trace est -1 .
- (3) Déterminer la nature de l'isométrie dont la matrice dans une base orthonormale est :

$$A = \frac{1}{4} \begin{pmatrix} 2 & -\sqrt{6} & -\sqrt{6} \\ \sqrt{6} & 3 & -1 \\ \sqrt{6} & -1 & 3 \end{pmatrix}$$

Exercice 7.10. E est un espace vectoriel euclidien de dimension 3. Si s et s' sont des réflexions de plans H et H' , montrer que $s \circ s' = s' \circ s$ si, et seulement si, H et H' sont confondus ou perpendiculaires².

Exercice 7.11. Soit $\mathcal{B} = (e_1, e_2, e_3)$ une base orthonormale d'un espace vectoriel euclidien de dimension 3. On considère l'ensemble G des transformations orthogonales qui laissent stable l'ensemble $\{\pm e_1, \pm e_2, \pm e_3\}$. Combien G a-t-il d'éléments ? Montrer que G est un groupe et décrire les trois cas suivants :

- $g(e_1) = e_2, g(e_2) = e_1, g(e_3) = e_3$.
- $h(e_1) = e_2, h(e_2) = e_3, h(e_3) = e_1$.
- $k(e_1) = -e_1, k(e_2) = e_3, k(e_3) = e_2$.

2. Deux plans sont dits perpendiculaires lorsque l'orthogonal de l'un est inclus dans l'autre.

PROBLÈME

Le corrigé de ce problème est disponible sur le site de Dunod : www.dunod.com.

7.1. DÉCOMPOSITIONS MATRICIELLES

(1) La décomposition LU.

On suppose que A est une matrice carrée inversible, et que chaque étape de la méthode de Gauss donne un pivot non nul. Montrer qu'il existe deux matrices triangulaires L (triangulaire inférieure) et U (triangulaire supérieure) telles que $A = LU$. On utilisera la méthode du pivot de Gauss, en observant que chaque opération élémentaire de la forme $L_i \mapsto L_i + \lambda L_j$ où $j < i$, revient à multiplier A par une matrice de transvection, triangulaire inversible, et que le résultat du pivot est une matrice triangulaire supérieure inversible. On parle de décomposition LU (lower, upper).

Traiter l'exemple de A est la matrice
$$\begin{pmatrix} 5 & 2 & 1 \\ 5 & -6 & 2 \\ -4 & 2 & 1 \end{pmatrix},$$

(2) Montrer qu'une matrice A a une décomposition LU si et seulement si ses mineurs principaux sont non nuls. On rappelle que les mineurs principaux de A sont les n déterminants $\det(a_{i,j})_{1 \leq i,j \leq k}$ pour $k = 1$ à n .

(3) La décomposition polaire.

Il s'agit de décomposer un endomorphisme dans un espace vectoriel réel E , muni d'une structure euclidienne, ou dans un espace vectoriel complexe E , muni d'une structure hermitienne. Montrer que tout élément u de $\mathbf{GL}(E)$ s'écrit de façon unique comme composé d'une transformation orthogonale o et d'un endomorphisme symétrique défini positif s (resp. d'une transformation unitaire et d'un endomorphisme autoadjoint défini positif). Matriciellement, on aura $M = OS$ ou $M = S'O'$. On commencera par trouver S en calculant tMM (resp. M^*M) Examiner le cas de la dimension 1 (dans le cadre réel ou dans le cadre complexe).

(4) Montrer qu'un endomorphisme non inversible admet aussi une décomposition polaire, mais on perd l'unicité de O , et S n'est pas définie positive.

(5) La décomposition QR.

L'idée est encore de faire apparaître une matrice triangulaire, en multipliant M par des matrices convenables. Montrer que toute matrice inversible peut s'écrire $M = QR$ où Q est (orthogonale)(unitaire) et R est triangulaire supérieure.

Il y a unicité si les coefficients diagonaux de R sont positifs. On emploiera une méthode de « pivot », avec l'idée que si a est un vecteur non nul, il est facile de trouver une réflexion qui transforme a en un vecteur colinéaire à e_1 ; il suffit de déterminer un vecteur normal, $a - \|a\|e_1$ convient. Multipliant M de première colonne a par la matrice de la réflexion correspondante, on obtient une première colonne colinéaire à e_1 et on poursuit.

(6) La décomposition d'Iwasawa.

Montrer que toute matrice inversible s'écrit comme produit d'une matrice orthogonale, d'une matrice diagonale à coefficients strictement positifs, d'une matrice unipotente triangulaire supérieure³. On pourra s'inspirer du procédé dit de Gram-Schmidt qui permet d'obtenir une base orthonormale à partir d'une base quelconque.

(7) La décomposition de Cartan.

Montrer que toute matrice de $\mathbf{SL}(E)$ s'écrit $KA K'$ où K et K' sont dans $\mathbf{SO}(E)$ et A est diagonale à coefficients positifs.

SOLUTIONS DES EXERCICES

Solution 7.1. Soit ϕ une forme linéaire de noyau H . Alors toute transvection τ d'hyperplan H a une expression de la forme : $\tau(x) = x + \phi(x)a$ où $a \in H$. Si on autorise $a = 0$, on a aussi l'identité. On peut alors composer τ et τ' :

$$\tau' \circ \tau(x) = \tau'(x + \phi(x)a) = x + \phi(x)a + \phi(x + \phi(x)a)a' = x + \phi(x)(a + a')$$

puisque $\phi(a) = 0$. Cette relation prouve que $\tau \mapsto a$ est un morphisme de l'ensemble des transvections d'hyperplan H (auquel on ajoute l'identité) avec le groupe additif $(H, +)$. Il est vite vérifié que c'est un isomorphisme de cet ensemble G avec le groupe H . Si on ajoute à cet ensemble de transvections les dilatations d'hyperplan H , on obtient nécessairement un groupe G' , qui est le stabilisateur de H dans l'action du groupe linéaire. Attention, le composé de deux dilatations de même hyperplan H est tantôt une dilatation, tantôt une transvection (lorsque le produit des rapports est égal à 1). Il suffit, pour s'en convaincre, d'utiliser une représentation matricielle. Par ailleurs, en considérant le morphisme déterminant, G est distingué dans G' .

Solution 7.2. Soit K un corps fini ayant q éléments, et E un espace vectoriel de dimension n sur K . On va dénombrer les bases de E : ayant choisi un vecteur e_1 non nul parmi $q^n - 1$ (on a enlevé le vecteur nul), il faut choisir un vecteur e_2 qui n'est pas dans les q vecteurs colinéaires à e_1 , cela donne $q^n - q$ choix, puis on doit choisir un vecteur qui n'est pas dans le plan engendré par les deux vecteurs précédemment choisis, cela laisse $q^n - q^2$ choix, et donc le cardinal de l'ensemble des bases de E est :

$$B = (q^n - 1)(q^n - q)(q^n - q^2) \dots (q^n - q^{n-1})$$

Un élément de $\mathbf{GL}(n, \mathbb{F}_q)$ est déterminé par l'image d'une base, donc le cardinal de ce groupe est le nombre B écrit plus haut.

Pour dénombrer l'ensemble $\mathbf{SL}(n, \mathbb{F}_q)$, il suffit d'utiliser le théorème d'isomorphisme : comme le morphisme déterminant de $\mathbf{GL}(n, \mathbb{F}_q)$ dans $\mathbb{F}_q^\times$ est surjectif et que

3. C'est une matrice triangulaire supérieure dont tous les éléments diagonaux sont égaux à 1.

son noyau est $\mathbf{SL}(n, \mathbb{F}_q)$, le cardinal de ce groupe est

$$\frac{(q^n - 1)(q^n - q)(q^n - q^2) \dots (q^n - q^{n-1})}{q - 1}$$

En particulier $\mathbf{SL}(2, \mathbb{F}_2)$ a pour cardinal 6. Ce groupe se confond avec $GL(2, \mathbb{F}_2)$, puisque la seule valeur possible d'un déterminant non nul est 1 ($\mathbb{F}_2 = \mathbb{Z}/2\mathbb{Z}$). C'est un groupe non commutatif à six éléments, il est donc nécessairement isomorphe à $\mathcal{S}_3$, et on peut réaliser cet isomorphisme en regardant l'action sur les trois droites $\text{vect}(e_1)$, $\text{vect}(e_2)$ et $\text{vect}(e_1 + e_2)$. À noter, le groupe $\mathbf{SL}(2, \mathbb{F}_3)$ est de cardinal 24 mais il n'est pas isomorphe à $\mathcal{S}_4$. Voir par exemple [6].

Solution 7.3. Commençons par déterminer la projection orthogonale $p(x)$ de x sur la droite $\text{vect}(u)$. C'est un vecteur de la forme λu , et $x - \lambda u$ doit être orthogonal à u :

$$\langle x - \lambda u, u \rangle = 0 \iff \lambda = \frac{\langle u, x \rangle}{\langle u, u \rangle}$$

Maintenant, si on considère la somme directe $\text{vect}(u) \oplus \{u\}^\perp$, on a $x = p(x) + x_2$ et $s(x) = -p(x) + x_2$ donc $s(x) = x - 2p(x)$ d'où le résultat.

Dans le cas de la dimension 3, prenons pour simplifier les écritures u de coordonnées a, b, c telles que $a^2 + b^2 + c^2 = 1$. On a alors :

$$\begin{aligned} s(x) &= \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} - 2(ax_1 + bx_2 + cx_3) \begin{pmatrix} a \\ b \\ c \end{pmatrix} \\ &= \begin{pmatrix} 1 - 2a^2 & -2ab & -2ac \\ -2ab & 1 - 2b^2 & -2bc \\ -2ac & -2bc & 1 - 2c^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \end{aligned}$$

et l'on a la forme générale des matrices de réflexion en dimension 3.

Solution 7.4. La norme « canonique » sur $\mathbb{R}$ est la valeur absolue. Prenons alors tout simplement $f(x) = |x|$. Alors $|f(x)| = |x|$, et f n'est pas linéaire.

Solution 7.5. On peut utiliser la méthode de l'exercice 7.3, en remarquant qu'un vecteur normal à D est $(-\sin \phi, \cos \phi)$.

$$\begin{aligned} s_D(x) &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - 2(-x_1 \sin \phi + x_2 \cos \phi) \begin{pmatrix} -\sin \phi \\ \cos \phi \end{pmatrix} \\ &= \begin{pmatrix} 1 - 2 \sin^2 \phi & 2 \sin \phi \cos \phi \\ 2 \sin \phi \cos \phi & 1 - 2 \cos^2 \phi \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \end{aligned}$$

Cette matrice peut aussi s'écrire

$$M_{\mathcal{B}}(s_D) = \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix}$$

qui est bien orthogonale indirecte. Si maintenant on fait le produit des matrices de $s_{D'}$ et de s_D , on obtient :

$$M_{\mathcal{B}}(s_{D'} \circ s_D) = \begin{pmatrix} \cos 2(\phi' - \phi) & -\sin 2(\phi' - \phi) \\ \sin 2(\phi' - \phi) & \cos 2(\phi' - \phi) \end{pmatrix}$$

c'est la matrice de la rotation d'angle

$$2(\phi' - \phi) = 2(\widehat{(e_1, u')} - \widehat{(e_1, u)}) = 2\widehat{(u, u')}$$

Solution 7.6. Soit $\mathcal{A}'$ le groupe des angles de droites, identifié au groupe quotient $\mathbb{R}/\pi\mathbb{Z}$. Comme $2\pi\mathbb{Z}$ est inclus dans $\pi\mathbb{Z}$, l'application $\theta \mapsto 2\theta$ est un morphisme. Ainsi, le double d'un angle de droites est un angle de vecteurs. Plus géométriquement, à un angle de droites correspond deux angles de vecteurs selon le choix qu'on fait des vecteurs directeurs, et ces deux angles diffèrent de π donc ils ont même double. Pour l'exercice précédent, on peut dire que la rotation $s_{D'} \circ s_D$ a pour angle le « double » de l'angle (D, D') .

Solution 7.7. Dans le groupe des angles $\mathcal{A}$, un angle α est d'ordre 2 s'il est non nul et si $\alpha + \alpha = 0$. C'est l'angle d'une rotation involutive. Si on utilise les mesures, la seule solution est π modulo 2π , mais on peut aussi utiliser les rotations, on trouve la rotation $-\text{id}_E$. Cet angle représente les couples $(u, -u)$, c'est l'angle plat. De même, les angles d'ordre 4 sont les deux angles droits, de mesure $\frac{\pi}{2}$ et $-\frac{\pi}{2}$, et plus généralement les angles d'ordre n sont les arguments des racines primitives n -ièmes de l'unité.

Solution 7.8.

(1) Si x est vecteur propre de f , associé à la valeur propre λ , on a :

$$\langle f(x), x \rangle = \langle x, f^*(x) \rangle \quad \text{donc} \quad \lambda \|x\|^2 = -\langle f(x), x \rangle = -\lambda \|x\|^2$$

et la seule valeur propre possible d'un endomorphisme antisymétrique est 0. On en déduit que $\text{id} + f$ et $\text{id} - f$ sont bijectives.

(2) Calculons u^* :

$$u^* = (\text{id} + f)^{* -1} \circ (\text{id} - f)^* = (\text{id} - f)^{-1} \circ (\text{id} + f) = u^{-1}$$

On en déduit que u est orthogonale. De plus, le déterminant d'un endomorphisme et de son transposé étant égaux,

$$\det(u) = \det(\text{id} - f) \det(\text{id} + f)^{-1} = \det(\text{id} - f) \det(\text{id} - f)^{-1} = 1$$

puisque $(\text{id} + f)^* = \text{id} - f$; u est donc dans $\text{SO}(E)$. Par ailleurs, si on calcule $(\text{id} + u) \circ (\text{id} + f)$, on trouve :

$$(\text{id} + u) \circ (\text{id} + f) = (\text{id} - f) \circ (\text{id} + f)^{-1} \circ (\text{id} + f) + \text{id} + f = 2\text{id}$$

et donc $\text{id} + u$ est inversible, -1 n'est pas valeur propre de u .

- (3) En dimension 2, la matrice de f dans une base orthonormale prend la forme $\begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix}$ et on trouve que la matrice de u se présente sous la forme

$$\begin{pmatrix} \frac{1-a^2}{1+a^2} & -\frac{2a}{1+a^2} \\ \frac{2a}{1+a^2} & \frac{1-a^2}{1+a^2} \end{pmatrix}$$

qui est bien la matrice d'une rotation (penser à la tangente de l'arc moitié), l'angle plat (et donc la valeur propre -1) étant exclus. Dans le cas de la dimension 3, on part d'une matrice antisymétrique de la forme

$$\begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix}$$

et on obtient

$$\frac{1}{1+a^2+b^2+c^2} \begin{pmatrix} 1+c^2-a^2-b^2 & -2bc-2a & -2b+2ac \\ 2a-2bc & 1+b^2-a^2-c^2 & -2c-2ab \\ 2b+2ac & 2c-2ab & 1+a^2-b^2-c^2 \end{pmatrix}$$

Le lecteur courageux pourra vérifier qu'il s'agit bien d'une matrice orthogonale directe.

- (4) Il suffit de « calculer » f en fonction de u . On trouve que nécessairement,

$$f = (\text{id} - u) \circ (\text{id} + u)^{-1}$$

qui existe car -1 n'est pas valeur propre de u . Des calculs immédiats montrent alors que f est antisymétrique.

Solution 7.9.

- (1) Soit r un élément de $\text{SO}(E)$, où E est euclidien de dimension 3. La transformation r est donc une rotation, il existe une droite de vecteurs invariants. Si e_1 est un vecteur unitaire qui dirige cette droite, on peut le compléter en (e_1, e_2, e_3) base orthonormale. Mais alors, $\text{vect}(e_2, e_3)$ est stable et la restriction v de r à ce plan est orthogonale directe, puisque $\det r = 1 = \det v$. On en déduit qu'il existe un réel θ tel que la matrice de u dans la base $\mathcal{B} = (e_1, e_2, e_3)$ soit

$$M_{\mathcal{B}}(r) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}$$

- (2) La trace de la matrice précédente est $1 + 2 \cos \theta$. Or r est un retournement si et seulement si sa restriction au plan orthogonal à l'axe est la rotation d'angle π . Puisque la trace est un invariant de similitude, on en déduit que les demi-tours sont les rotations dont la matrice a pour trace -1 .
- (3) On vérifie rapidement que ${}^tAA = I$, il s'agit bien d'une matrice orthogonale. L'ensemble des vecteurs invariant est $\text{vect}(e_2 - e_3)$, et la trace de la matrice est 2. On en déduit qu'il s'agit d'une rotation d'angle θ tel que $2 \cos \theta + 1 = 2$ soit $\cos \theta = \frac{1}{2}$, soit un angle géométrique de mesure $\frac{\pi}{3}$.

Solution 7.10. Soient u et u' des vecteurs unitaires directeurs de $H^\perp$ et $H'^\perp$ respectivement. Alors si $s' \circ s = s \circ s'$, $s' \circ s(u) = -s'(u) = s(s'(u))$ prouve que $s'(u) \in \text{vect}(u)$, puisque c'est un vecteur propre pour la valeur -1 . Comme s est une isométrie, on aura

- soit $s'(u) = u$, et l'orthogonal de H est inclus dans H' , les deux plans sont perpendiculaires ;
- soit $s'(u) = -u$, H et H' ont même orthogonal, ils coïncident.

Solution 7.11. Soit u un endomorphisme qui transforme la base (e_1, e_2, e_3) en $(\pm e_i, \pm e_j, \pm e_k)$. Ce sera une base (orthonormale) si les (i, j, k) sont distincts, donc forment une permutation de $(1, 2, 3)$. Il y a donc $2^3 \times 6 = 48$ possibilités. G est un groupe par sa définition même, c'est le stabilisateur de l'ensemble des 6 vecteurs $(e_1, e_2, e_3, -e_1, -e_2, -e_3)$. Les géomètres verront également que c'est le groupe des isométries qui conserve le cube dont les huit sommets sont les points de coordonnées $(\pm 1, \pm 1, \pm 1)$: voir le chapitre 10. Examinons les trois cas proposés.

- La matrice de g a un déterminant négatif, g est une transformation orthogonale indirecte. De plus, $e_1 + e_2$ et e_3 sont invariants, il s'agit de la réflexion de plan $\text{vect}(e_1 + e_2, e_3)$.
- C'est cette fois une transformation orthogonale directe puisque le déterminant est 1. De plus $h^3 = \text{id}_E$, donc h est une rotation d'angle géométrique de mesure $\frac{2\pi}{3}$. La recherche des vecteurs invariants montre que l'axe est $\text{vect}(e_1 + e_2 + e_3)$.
- On a $\det(k) = 1$ et $k^2 = \text{id}$. C'est un demi-tour, d'axe $\text{vect}(e_2 + e_3)$.

Chapitre 8

Espaces affines euclidiens

Ce chapitre propose un modèle mathématique de l'espace physique qui nous est familier : **l'espace affine euclidien de dimension 3**. La première section se développe dans un cadre purement affine. Les notions métriques (distance, orthogonalité, angles, etc.) sont ensuite introduites.

8.1 NOTIONS AFFINES

8.1.1. Espaces affines

La terminologie de l'algèbre linéaire a une connotation géométrique (espace vectoriel, dimension, droite vectorielle, plan vectoriel, orthogonalité, etc.). Cependant, même si l'on développe en l'étudiant des images mentales géométriques, il s'agit plus d'algèbre, c'est-à-dire de calcul, que de géométrie.

L'espace vectoriel $\mathbb{R}^3$ rapporté à sa base canonique $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ n'est pas un bon modèle de l'espace physique, dans lequel il n'y a pas d'addition naturelle : qu'est-ce que la somme de deux points ? L'espace physique n'est pas muni d'un point particulier jouant le rôle du vecteur nul $(0, 0, 0)$. On va dégager la notion d'**espace affine**, formé de **points** et non de vecteurs, où tous les points jouent le même rôle.

Définition 8.1. On appelle *espace affine* la donnée

- d'un ensemble $\mathcal{E}$ formé de **points**,
- d'un $\mathbb{R}$ -espace vectoriel E formé de **vecteurs**,
- d'une action simplement transitive du groupe additif $(E, +)$ sur $\mathcal{E}$,

$$\begin{aligned} E \times \mathcal{E} &\longrightarrow \mathcal{E} \\ (\vec{v}, M) &\longmapsto M + \vec{v} \end{aligned}$$

Dans ce triplet de données, l'objet géométrique est l'ensemble $\mathcal{E}$ des points. Contrairement aux premiers chapitres de ce livre, les vecteurs sont notés par des lettres surmontées d'une flèche, le vecteur nul étant noté $\vec{0}$. On dit que $\mathcal{E}$ est l'**espace affine** et que E est son **espace vectoriel associé**. Le signe $+$ est donc employé dans deux contextes :

- pour additionner deux vecteurs de E ,
- pour additionner un vecteur de E à un point de $\mathcal{E}$.

Le fait qu'il s'agit d'une action se traduit par

$$(1) \quad \forall M \in \mathcal{E}, \forall (\vec{u}, \vec{v}) \in E \times E, \quad M + (\vec{u} + \vec{v}) = (M + \vec{u}) + \vec{v}$$

$$(2) \quad \forall M \in \mathcal{E}, \quad M + \vec{0} = M$$

Pour tout $\vec{v} \in E$, l'application $\mathcal{E} \longrightarrow \mathcal{E}, M \longmapsto M + \vec{v}$ s'appelle la **translation de vecteur** $\vec{v}$, elle est notée $T_{\vec{v}}$.

Proposition 8.2. Les translations sont des bijections de $\mathcal{E}$ sur lui-même et forment un groupe commutatif $\mathcal{T}(\mathcal{E})$ pour la loi $\circ$ de composition des applications. L'application $\vec{v} \mapsto T_{\vec{v}}$ est un isomorphisme du groupe additif $(E, +)$ sur le groupe des translations $(\mathcal{T}(\mathcal{E}), \circ)$.

Démonstration. Exprimons les conditions (1) et (2) :

$$(1) \text{ pour deux vecteurs quelconques } \vec{u}, \vec{v}, \text{ on a } T_{\vec{v} + \vec{u}} = T_{\vec{u}} \circ T_{\vec{v}}.$$

$$(2) \text{ la translation de vecteur nul est l'identité : } T_{\vec{0}} = I_{\mathcal{E}}$$

Pour tout $\vec{v} \in E$, on a $T_{\vec{v}} \circ T_{-\vec{v}} = T_{-\vec{v}} \circ T_{\vec{v}} = T_{\vec{0}} = I_{\mathcal{E}}$, donc toute translation $T_{\vec{v}}$ est bijective et $T_{-\vec{v}} = T_{\vec{v}}^{-1}$.

D'après la condition (1), l'ensemble $\mathcal{T}(\mathcal{E})$ des translations est un sous-groupe commutatif du groupe des bijections de $\mathcal{E}$ sur lui-même et l'application $\vec{v} \mapsto T_{\vec{v}}$ est un morphisme de groupes de $(E, +)$ vers $\mathcal{T}(\mathcal{E})$, surjectif par définition de $\mathcal{T}(\mathcal{E})$.

Montrons que le noyau se réduit à $\{\vec{0}\}$, ce qui donnera l'injectivité, donc la bijectivité. L'action étant simplement transitive, pour tout couple de points (M, N) , il existe $\vec{v} \in E$ unique tel que $N = M + \vec{v} = T_{\vec{v}}(M)$. En particulier, pour tout M , l'unique $\vec{v}$ tel que $M = T_{\vec{v}}(M)$ est le vecteur nul. Donc $\vec{0}$ est l'unique vecteur dont la translation associée est l'identité. $\square$

Pour tout couple de points (M, N) , on note $\overrightarrow{MN}$ le vecteur de l'unique translation envoyant M en N . La simple transitivité de l'action se traduit ainsi : pour tout point O , on a des bijections inverses l'une de l'autre

$$\begin{array}{ccc} E & \longrightarrow & \mathcal{E} \\ \overrightarrow{v} & \longmapsto & O + \overrightarrow{v} = T_{\overrightarrow{v}}(O) \end{array} \quad \text{et} \quad \begin{array}{ccc} \mathcal{E} & \longrightarrow & E \\ M & \longmapsto & \overrightarrow{OM} \end{array}$$

En particulier, pour tout point O , on a $\overrightarrow{OO} = \overrightarrow{0}$.

Contrairement à l'espace $\mathcal{E}$ des points, l'espace vectoriel E donne prise au calcul. Pour étudier une configuration dans $\mathcal{E}$ où un point O joue un rôle particulier (par exemple l'isobarycentre d'un système de points), on peut **vectorialiser** $\mathcal{E}$ en O : on transpose dans E le problème posé par la bijection $\mathcal{E} \rightarrow E, M \mapsto \overrightarrow{OM}$ pour tenter une solution par calcul vectoriel ; on retraduit ensuite dans $\mathcal{E}$ les résultats obtenus par la bijection inverse $\overrightarrow{v} \mapsto O + \overrightarrow{v}$.

Proposition 8.3. Relation de Chasles. *Étant donné trois points A, B, C de $\mathcal{E}$, on a la relation vectorielle $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$.*

Démonstration. En effet

$$T_{\overrightarrow{AC}}(A) = C = T_{\overrightarrow{BC}}(B) = (T_{\overrightarrow{BC}} \circ T_{\overrightarrow{AB}})(A) = T_{\overrightarrow{BC} + \overrightarrow{AB}}(A)$$

Comme il existe une seule translation transformant A en C , on a $T_{\overrightarrow{AC}} = T_{\overrightarrow{BC} + \overrightarrow{AB}}$, donc $\overrightarrow{AC} = \overrightarrow{BC} + \overrightarrow{AB} = \overrightarrow{AB} + \overrightarrow{BC}$. $\square$

En particulier, $\overrightarrow{0} = \overrightarrow{AA} = \overrightarrow{AB} + \overrightarrow{BA}$, donc $\overrightarrow{BA} = -\overrightarrow{AB}$.

Proposition 8.4. Règle du parallélogramme. *Soit A, A', B, B' quatre points de $\mathcal{E}$. On a l'équivalence $(\overrightarrow{AB} = \overrightarrow{A'B'}) \iff (\overrightarrow{AA'} = \overrightarrow{BB'})$.*

S'il en est ainsi, on dit que A, A', B', B forment un **parallélogramme**.

Démonstration. L'égalité $\overrightarrow{AB'} = \overrightarrow{AB} + \overrightarrow{BB'} = \overrightarrow{AA'} + \overrightarrow{A'B'}$ est toujours vraie (relation de Chasles). On en déduit l'équivalence. $\square$

Définition 8.5. *Étant donné deux points P, Q , on appelle **milieu** de (P, Q) le point I défini par $\overrightarrow{PI} = \frac{1}{2}\overrightarrow{PQ}$.*

On a alors $\overrightarrow{QI} = \overrightarrow{QP} + \overrightarrow{PI} = \overrightarrow{QP} + \frac{1}{2}\overrightarrow{PQ} = \overrightarrow{QP} - \frac{1}{2}\overrightarrow{QP} = \frac{1}{2}\overrightarrow{QP}$ donc I est aussi milieu de (Q, P) . On écrira $I = \frac{P+Q}{2}$ (notation expliquée au chapitre 9).

Proposition 8.6. *Pour que quatre points A, A', B', B forment un parallélogramme, il faut et il suffit que (A, B') et (A', B) aient même milieu.*

Démonstration. Soit $I = \frac{A+B'}{2}$. On a $\overrightarrow{AB'} = 2\overrightarrow{AI}$, d'où

$$\overrightarrow{A'B} - 2\overrightarrow{A'I} = (\overrightarrow{A'A} + \overrightarrow{AB}) - 2(\overrightarrow{A'A} + \overrightarrow{AI}) = \overrightarrow{AB} - \overrightarrow{A'A} - \overrightarrow{AB'} = \overrightarrow{AB} - \overrightarrow{A'B'}$$

Pour que A, B', A', B soit un parallélogramme, il faut et il suffit que ce vecteur soit nul, donc que I soit aussi milieu de (A', B) . $\square$

Définition 8.7. On appelle dimension de l'espace affine $\mathcal{E}$ la dimension de l'espace vectoriel E et on écrit $\dim \mathcal{E} = \dim E$.

Le plus souvent dans la suite, la dimension sera égale à 2 en géométrie plane, 3 en géométrie dans l'espace.

Définition 8.8. Pour $\dim \mathcal{E} = n$, on appelle repère de $\mathcal{E}$ la donnée d'un point $O \in \mathcal{E}$ dit origine et d'une base $\mathcal{B} = (\vec{e}_1, \dots, \vec{e}_n)$ de E . Le repère $R = (O, \mathcal{B}) = (O, \vec{e}_1, \dots, \vec{e}_n)$ détermine la bijection

$$\mathbb{R}^n \longrightarrow \mathcal{E}, \quad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \longmapsto M = O + \sum_{i=1}^n x_i \vec{e}_i$$

On dit que $x_1, \dots, x_n$ sont les **coordonnées** de M dans le repère R .

En choisissant un repère, on transpose un problème géométrique dans $\mathbb{R}^n$ pour le résoudre par calcul numérique. Le choix du repère s'appuie souvent sur des considérations géométriques pour que les calculs soient simples. La vectorialisation en un point et le calcul vectoriel évitent parfois la lourdeur de calculs numériques dans un repère.

8.1.2. Sous-espaces affines

Soit $\mathcal{F}$ une partie non vide de $\mathcal{E}$. Pour tout $A \in \mathcal{F}$, on note $\overline{\mathcal{F}}_A$ l'ensemble des vecteurs $\overrightarrow{AM}$ où M décrit $\mathcal{F}$.

Proposition 8.9. On a équivalence entre

- (i) il existe $A \in \mathcal{F}$ tel que $\overline{\mathcal{F}}_A$ soit un sous-espace vectoriel de E ,
- (ii) pour tout $A \in \mathcal{F}$, $\overline{\mathcal{F}}_A$ est un sous-espace vectoriel de E .

S'il en est ainsi, les sous-espaces vectoriels $\overline{\mathcal{F}}_A$ sont tous égaux à un même sous-espace noté $\overline{\mathcal{F}} \subset E$. On dit alors que $\mathcal{F}$ est un sous-espace affine de $\mathcal{E}$ et que le **sous-espace vectoriel** $\overline{\mathcal{F}} \subset E$ est la **direction** de $\mathcal{F}$.

Démonstration. L'implication (ii) $\Rightarrow$ (i) est claire. Inversement, supposons que $\overline{\mathcal{F}}_A$ est un sous-espace vectoriel de E pour un certain point $A \in \mathcal{F}$. Soit B un point quelconque de $\mathcal{F}$. Il suffit de montrer que $\overline{\mathcal{F}}_A = \overline{\mathcal{F}}_B$. Pour tout $\vec{v}$, on a $B + \vec{v} = A + \overrightarrow{AB} + \vec{v}$, d'où les équivalences

$$(\vec{v} \in \overline{\mathcal{F}}_B) \Leftrightarrow (B + \vec{v} \in \mathcal{F}) \Leftrightarrow (A + \overrightarrow{AB} + \vec{v} \in \mathcal{F}) \Leftrightarrow (\overrightarrow{AB} + \vec{v} \in \overline{\mathcal{F}}_A)$$

Comme $B \in \mathcal{F}$, $\overrightarrow{AB} \in \overline{\mathcal{F}}_A$ et puisque $\overline{\mathcal{F}}_A$ est un sous-espace vectoriel de E , on peut écrire $(\overrightarrow{AB} + \overrightarrow{v} \in \overline{\mathcal{F}}_A) \Leftrightarrow (\overrightarrow{v} \in \overline{\mathcal{F}}_A)$, d'où l'équivalence :

$$(\overrightarrow{v} \in \overline{\mathcal{F}}_B) \Leftrightarrow (\overrightarrow{AB} + \overrightarrow{v} \in \overline{\mathcal{F}}_A) \Leftrightarrow (\overrightarrow{v} \in \overline{\mathcal{F}}_A)$$

On en déduit l'égalité cherchée $\overline{\mathcal{F}}_B = \overline{\mathcal{F}}_A$. $\square$

Inversement, étant donné un sous-espace vectoriel $F \subset E$ et un point $A \in \mathcal{E}$, l'ensemble

$$A + F = \{M \in \mathcal{E} \mid \exists \overrightarrow{v} \in F, M = A + \overrightarrow{v}\} = \{M \in \mathcal{E} \mid \overrightarrow{AM} \in F\}$$

est l'**unique sous-espace affine passant par A de direction F**.

Étant donné un sous-espace affine $\mathcal{F} \subset \mathcal{E}$ de direction $\overline{\mathcal{F}} \subset E$, la restriction à $\overline{\mathcal{F}} \times \mathcal{F} \subset E \times \mathcal{E}$ de l'action $(\overrightarrow{v}, M) \mapsto M + \overrightarrow{v}$ définit une action $\overline{\mathcal{F}} \times \mathcal{F} \rightarrow \mathcal{F}$ faisant de $\mathcal{F}$ un espace affine d'espace vectoriel associé $\overline{\mathcal{F}}$.

Exemples :

- (1) Pour tout $A \in \mathcal{E}$, $\mathcal{E} = A + E$, donc $\mathcal{E}$ est le seul sous-espace affine de direction E .
- (2) Pour tout A , l'ensemble $\{A\}$ réduit au point A est le sous-espace affine passant par A de direction $\{\overrightarrow{0}\}$. Sa dimension est 0.
- (3) On appelle **droite affine** tout sous-espace affine de dimension 1, sa direction est une droite vectorielle.

Pour $A \in \mathcal{E}$ et $\overrightarrow{v} \neq \overrightarrow{0}$, on appelle **droite passant par A de vecteur directeur** $\overrightarrow{v}$ l'ensemble $A + \mathbb{R}\overrightarrow{v} = \{M \mid \exists \lambda \in \mathbb{R}, M = A + \lambda \overrightarrow{v}\}$. L'application $\lambda \mapsto A + \lambda \overrightarrow{v}$ est une bijection de $\mathbb{R}$ sur cette droite. On dit que c'est une **représentation paramétrique** de cette droite.

Étant donné deux points distincts A et B , l'unique droite affine passant par A et B , notée (AB) , est $A + \mathbb{R}\overrightarrow{AB}$.

- (4) On appelle **plan affine** tout sous-espace affine de dimension 2. Par trois points non alignés A, B, C passe un unique plan affine notée (ABC) : celui passant par A (ou B , ou C) de direction le plan vectoriel

$$\text{Vect}(\overrightarrow{AB}, \overrightarrow{AC}) = \text{Vect}(\overrightarrow{BC}, \overrightarrow{BA}) = \text{Vect}(\overrightarrow{CA}, \overrightarrow{CB})$$

- (5) Si $\dim \mathcal{E} = n$, on appelle **hyperplan affine** tout sous-espace affine de dimension $n - 1$, c'est-à-dire du type $\mathcal{H} = A + H$ où $H = \overline{\mathcal{H}}$ est un hyperplan vectoriel. En géométrie plane (resp. dans l'espace), les hyperplans sont les droites du plan (resp. les plans de l'espace).

Dans un repère $(0, \overrightarrow{e}_1, \dots, \overrightarrow{e}_n)$, donnons A par ses coordonnées $(a_1, \dots, a_n)$ et H comme noyau de la forme linéaire de matrice $(u_1, \dots, u_n)$. Pour qu'un point M de coordonnées $x_1, \dots, x_n$ soit dans $\mathcal{H}$, il faut et il suffit que $\overrightarrow{AM} \in H$, donc que

$$\sum_{i=1}^n u_i(x_i - a_i) = \sum_{i=1}^n u_i x_i + h = 0 \text{ où } h = - \sum_{i=1}^n a_i$$

Ainsi, $u_1x_1 + \dots + u_nx_n + h = 0$ est une **équation** de $\mathcal{H}$ dans le repère. Inversement, toute équation de ce type où $(u_1, \dots, u_n)$ n'est pas le n -uplet nul, est l'équation d'un hyperplan. Un hyperplan a une infinité d'équations, toutes proportionnelles.

Définition 8.10. Deux sous-espaces affines sont dits **parallèles** s'ils ont même direction $F \subset E$.

Ils sont de la forme $A + F$ et $B + F$ où A et B sont des points de $\mathcal{E}$. Ce sont les orbites de A et B pour l'action du groupe additif $(F, +)$ sur $\mathcal{E}$:

$$F \times \mathcal{E} \longrightarrow \mathcal{E}, \quad (\vec{v}, M) \longmapsto M + \vec{v}$$

Ils sont donc, ou confondus, ou d'intersection vide.

Proposition 8.11. Droites parallèles d'un plan. Soit $\mathcal{P}$ un plan affine de plan vectoriel associé P . Pour que deux droites distinctes de $\mathcal{P}$ soient parallèles, il faut et il suffit que leur intersection soit vide.

Démonstration. Soit deux droites distinctes $\mathcal{A} = A + \mathbb{R}\vec{u}$ et $\mathcal{B} = B + \mathbb{R}\vec{v}$. Par ce qui précède, si elles sont parallèles, leur intersection est vide.

Supposons-les non parallèles. Alors $(\vec{u}, \vec{v})$ est base de P . On cherche un point C commun à $\mathcal{A}$ et $\mathcal{B}$; alors $C = A + x\vec{u} = B + y\vec{v}$. Ceci revient à chercher $(x, y) \in \mathbb{R} \times \mathbb{R}$ tel que $\vec{AB} = x\vec{u} - y\vec{v}$. Comme $(\vec{u}, \vec{v})$ est base de P , il existe un unique tel couple (x, y) , d'où un unique point commun. $\square$

8.1.3. Applications affines

Soit $\mathcal{E}$ et $\mathcal{E}'$ deux espaces affines, $f: \mathcal{E} \rightarrow \mathcal{E}'$ une application. Soit $A \in \mathcal{E}$. On définit $\overline{f}_A: E \rightarrow E'$ par $\overline{f}_A(\vec{AM}) = \overrightarrow{f(A)f(M)}$ pour tout $M \in \mathcal{E}$.

Proposition 8.12. On a équivalence entre

- (i) il existe $A \in \mathcal{E}$ tel que $\overline{f}_A$ soit une application linéaire,
- (ii) pour tout $A \in \mathcal{E}$, $\overline{f}_A$ est une application linéaire.

S'il en est ainsi, les applications linéaires $\overline{f}_A$ sont toutes égales à une même application linéaire notée $\overline{f}: E \rightarrow E'$. On dit alors que f est une **application affine** de $\mathcal{E}$ vers $\mathcal{E}'$ et que $\overline{f}$ est l'**application linéaire (ou vectorielle) associée** à f .

Démonstration. Supposons $\overline{f}_A$ linéaire pour un certain $A \in \mathcal{E}$. Soit B un autre point. Montrons que $\overline{f}_A = \overline{f}_B$. Pour $M \in \mathcal{E}$, on a $\vec{BM} = \vec{AM} - \vec{AB}$, donc

$$\begin{aligned} \overline{f}_B(\vec{BM}) &= \overrightarrow{f(B)f(M)} = \overrightarrow{f(A)f(M)} - \overrightarrow{f(A)f(B)} \\ &= \overline{f}_A(\vec{AM}) - \overline{f}_A(\vec{AB}) = \overline{f}_A(\vec{AM} - \vec{AB}) = \overline{f}_A(\vec{BM}) \end{aligned}$$

Ceci montre que $\overline{f}_B = \overline{f}_A$, d'où la linéarité de $\overline{f}_B$. $\square$

Ainsi, une application affine $f: \mathcal{E} \rightarrow \mathcal{E}'$ est déterminée par la donnée

- de l'image $f(A) \in \mathcal{E}'$ d'un point A fixé dans $\mathcal{E}$,
- de l'application linéaire associée $\overline{f}: E \rightarrow E'$.

L'image d'un point $M \in \mathcal{E}$ est alors $f(M) = f(A) + \overline{f}(\overrightarrow{AM})$. Autrement dit, pour tout $A \in \mathcal{E}$ et tout $\overrightarrow{u} \in E$, on a

$$f(A + \overrightarrow{u}) = f(A) + \overline{f}(\overrightarrow{u}).$$

Proposition 8.13. *Soit une application affine $f: \mathcal{E} \rightarrow \mathcal{E}'$.*

- Soit $A \in \mathcal{E}$. L'image de f est $\text{Im } f = f(A) + \text{Im } \overline{f}$, sous-espace affine de $\mathcal{E}'$. L'application f est surjective si et seulement si $\overline{f}$ l'est.*
- Soit $M \in \mathcal{E}$, l'ensemble des $N \in \mathcal{E}$ tels que $f(N) = f(M)$ est $M + \text{Ker } \overline{f}$, sous-espace affine de $\mathcal{E}$. L'application f est injective si et seulement si $\overline{f}$ l'est.*

Démonstration.

- L'image de f est formée des $f(A + \overrightarrow{v}) = f(A) + \overline{f}(\overrightarrow{v})$, où $\overrightarrow{v} \in E$. Si $\overrightarrow{v}$ décrit E , $\overline{f}(\overrightarrow{v})$ décrit $\text{Im } \overline{f}$, donc $f(A + \overrightarrow{v}) = f(A) + \overline{f}(\overrightarrow{v})$ décrit $\text{Im } f = f(A) + \text{Im } \overline{f}$, sous-espace affine de $\mathcal{E}'$.

Pour que f soit surjective, il faut et il suffit que $\text{Im } f = f(A) + \text{Im } \overline{f} = \mathcal{E}'$. Comme le seul sous-espace affine de $\mathcal{E}'$ de direction E' est $\mathcal{E}'$ lui-même, il faut et il suffit que $E' = \text{Im } \overline{f}$, donc que $\overline{f}$ soit surjective.

- Pour tout $N \in \mathcal{E}$, on a $f(N) = f(M) + \overline{f}(\overrightarrow{MN})$. On a $f(N) = f(M)$ si et seulement si $\overline{f}(\overrightarrow{MN}) = \overrightarrow{0}$, donc si et seulement si $\overrightarrow{MN} \in \text{Ker } \overline{f}$.

Pour que f soit injective, il faut et il suffit que $M + \text{Ker } \overline{f}$ soit réduit au singleton $\{M\}$ pour tout M , donc que $\text{Ker } \overline{f} = \{\overrightarrow{0}\}$, i.e. que $\overline{f}$ soit injective.

Le lecteur montrera de même que pour un sous-espace vectoriel $F \subset E$ et un point $A \in \mathcal{E}$, l'image par f du sous-espace affine $\mathcal{F} = A + F$ est le sous-espace affine $f(\mathcal{F}) = f(A) + \overline{f}(F) \subset \mathcal{E}'$. $\square$

Proposition 8.14. Translations. *Pour qu'une application affine $T: \mathcal{E} \rightarrow \mathcal{E}$ ait pour application linéaire associée l'identité de E , il faut et il suffit que T soit une translation.*

Démonstration. Soit $T: \mathcal{E} \rightarrow \mathcal{E}$ affine, $O \in \mathcal{E}$, $O' = T(O)$. Pour tout $M \in \mathcal{E}$, on note M' l'image de M par T . On a $\overrightarrow{O'M'} = \overline{T}(\overrightarrow{OM})$. Pour que $\overline{T} = I_E$, il faut et il suffit que pour tout $M \in \mathcal{E}$, on ait $\overrightarrow{O'M'} = \overrightarrow{OM}$, soit $\overrightarrow{MM'} = \overrightarrow{OO'}$ (règle du parallélogramme, p. 205). Ainsi, la translation $T_{\overrightarrow{OO'}}$ est l'application affine $\mathcal{E} \rightarrow \mathcal{E}$ transformant O en O' d'application linéaire associée I_E . $\square$

Définition 8.15. Soit $0 \in \mathcal{E}$ et $\lambda \in \mathbb{R}$ non nul. On appelle **homothétie** de centre 0 et de rapport λ l'application

$$\begin{aligned} H_{0,\lambda}: \mathcal{E} &\longrightarrow \mathcal{E} \\ M &\longmapsto 0 + \lambda \overrightarrow{OM} \end{aligned}$$

C'est l'application affine transformant 0 en 0 d'application linéaire associée l'homothétie vectorielle λI_E de rapport λ .

Proposition 8.16. Homothéties. Étant donné $\lambda \in \mathbb{R}$ distinct de 0 et 1 , pour qu'une application affine h soit une homothétie de rapport λ , il faut et il suffit que $\bar{h}$ soit l'homothétie vectorielle λI_E .

Démonstration. Par définition des homothéties, la condition est nécessaire. Inversement, soit $h: \mathcal{E} \rightarrow \mathcal{E}$ affine telle que $\bar{h} = \lambda I_E$. La seule chose à montrer est que h a un point fixe 0 . Soit M et $M' = h(M)$ tel que $M' \neq M$. Cherchons 0 tel que $\overrightarrow{M'O} = \lambda \overrightarrow{MO}$. Comme $\overrightarrow{M'O} = \overrightarrow{M'M} + \overrightarrow{MO}$, le seul point 0 possible est donné par $\overrightarrow{MO} = \frac{1}{\lambda-1} \overrightarrow{M'M}$. On vérifie que ce point là convient. $\square$

Définition 8.17. Soit $E = F \oplus G$ une décomposition de E en somme directe de deux sous-espaces supplémentaires. Tout $\overrightarrow{u} \in E$ se décompose de façon unique en $\overrightarrow{u} = \overrightarrow{v} + \overrightarrow{w}$, $\overrightarrow{v} \in F$, $\overrightarrow{w} \in G$. L'application $\pi: \overrightarrow{u} \mapsto \overrightarrow{v}$ est la **projection** de E sur F parallèlement à G , elle est linéaire. Soit $\mathcal{F}$ un sous-espace affine de direction F et $0 \in \mathcal{F}$. L'application affine p définie par $p(0) = 0$ et $\bar{p} = \pi$ est appelée la **projection** de $\mathcal{E}$ sur $\mathcal{F}$ parallèlement à G .

(i) On a $\text{Im } p = p(0) + \text{Im } \pi = 0 + F = \mathcal{F}$ (8.13, p. 209). Pour tout $M \in \mathcal{F}$, $\overrightarrow{OM} \in F$, donc $p(M) = p(0) + \pi(\overrightarrow{OM}) = 0 + \overrightarrow{OM} = M$, donc la restriction de p à $\mathcal{F}$ est l'identité de $\mathcal{F}$. Tout point de $\mathcal{F}$ joue le même rôle que 0 .

(ii) Soit $M \in \mathcal{E}$. L'ensemble des N tels que $p(N) = p(M)$ est $\mathcal{G}_M = M + \text{Ker } \pi = M + G$. D'après (i), $p(p(M)) = p(M)$, donc $p(M) \in \mathcal{G}_M \cap \mathcal{F}$.

Inversement, si $M' \in \mathcal{G}_M \cap \mathcal{F}$, $p(M') = M'$ car $M' \in \mathcal{F}$, $p(M') = p(M)$ car $M' \in \mathcal{G}_M$. Donc $M' = p(M)$ est l'unique point commun à $\mathcal{G}_M$ et à $\mathcal{F}$.

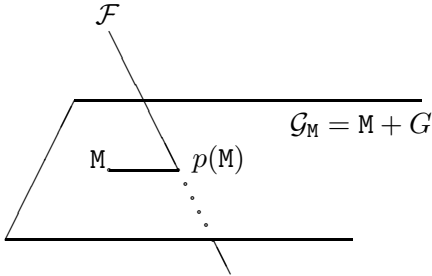
Les figures page suivante illustrent cette construction de $p(M)$ comme intersection de $\mathcal{G}_M$ et de $\mathcal{F}$.

Pour tout $0 \in \mathcal{E}$, notons $\text{Fix}(0)$ l'ensemble des $f: \mathcal{E} \rightarrow \mathcal{E}$ tels que $f(0) = 0$.

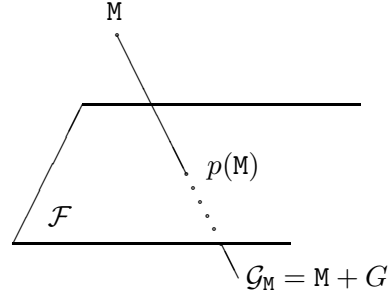
Proposition 8.18. a) La composée $h = g \circ f$ de deux applications affines $\mathcal{E} \rightarrow \mathcal{E}$ est une application affine d'application linéaire associée $\bar{h} = \bar{g} \circ \bar{f}$.

b) Pour tout $0 \in \mathcal{E}$, la restriction à $\text{Fix}(0)$ de l'application $f \mapsto \bar{f}$ est une bijection de $\text{Fix}(0)$ sur $\mathcal{L}(E)$.

$$\dim \mathcal{E} = 3, \dim \mathcal{F} = 1, \dim G = 2$$



$$\dim \mathcal{E} = 3, \dim \mathcal{F} = 2, \dim G = 1$$



Démonstration. a) Soit $A \in \mathcal{E}$ et $B = f(A)$. Pour $M \in \mathcal{E}$, soit $N = f(M)$, alors $\overrightarrow{BN} = \overrightarrow{f(AM)}$. Soit $C = g(B) = h(A)$ et $P = g(N) = h(M)$, alors

$$\overrightarrow{CP} = \overrightarrow{g(BN)} = (\overrightarrow{g} \circ \overrightarrow{f})(\overrightarrow{AM})$$

Comme $\overrightarrow{g} \circ \overrightarrow{f}$ est linéaire, $h = g \circ f$ est l'application affine transformant A en C d'application linéaire associée $\overrightarrow{g} \circ \overrightarrow{f}$.

b) Soit $\varphi \in \mathcal{L}(E)$ et $O \in \mathcal{E}$. La donnée de $f(O) = O$ et de $\overrightarrow{f} = \varphi$ détermine l'unique antécédent $f \in \text{Fix}(O)$ de l'application linéaire φ . $\square$

L'application $f \mapsto \overrightarrow{f}$ est donc surjective car pour tout $\varphi \in \mathcal{L}(E)$ et tout $O \in \mathcal{E}$, φ a un antécédent dans $\text{Fix}(O)$. Mais il se peut que φ ait des antécédents sans point fixe. Par exemple, toute translation est un antécédent de l'identité I_E . On verra (pb.4) que φ a des antécédents sans point fixe si et seulement si 1 est valeur propre de φ .

8.1.4. Expression analytique d'une application affine

Soit f une application affine $\mathcal{E} \rightarrow \mathcal{F}$ où les dimensions de $\mathcal{E}$ et $\mathcal{F}$ sont n et p . Rapportons $\mathcal{E}$ et $\mathcal{F}$ aux repères $(O, \overrightarrow{e_1}, \dots, \overrightarrow{e_n})$ et $(\Omega, \overrightarrow{\varepsilon_1}, \dots, \overrightarrow{\varepsilon_p})$. Alors f est déterminée par l'image $f(O)$ de coordonnées $(b_1, \dots, b_p)$ et la matrice $(a_{i,j})$ de $\overrightarrow{f}$. Il est facile de voir que f opère comme il est indiqué ci-dessous :

$$M \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto f(M) \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{p,1} & \cdots & a_{p,n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix}$$

formule s'écrivant $Y = AX + B$ avec les notations ci-dessous :

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix}, A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{p,1} & \cdots & a_{p,n} \end{pmatrix}, B = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix}$$

Proposition 8.19. *Étant donné un espace affine $\mathcal{E}$ d'espace vectoriel associé E , l'ensemble $\mathbf{GA}(\mathcal{E})$ des endomorphismes affines bijectifs est un groupe. L'application $f \mapsto \bar{f}$ est un morphisme surjectif de groupes $\mathbf{GA}(\mathcal{E}) \rightarrow \mathbf{GL}(E)$ dont le noyau est le sous-groupe $\mathcal{T}(\mathcal{E})$ des translations.*

Cela résulte des propositions 8.13, 8.14, 8.18 (p. 209 à 210).

8.1.5. Théorème de conjugaison

Comme on l'a vu sur les exemples (translations, homothéties, projections, etc.), les applications affines ne sont pas toutes de même nature géométrique. Le théorème de conjugaison précise cette notion. Étant donné $f: \mathcal{E} \rightarrow \mathcal{E}$ et $g \in \mathbf{GA}(\mathcal{E})$, rappelons que la **conjuguée** de f par g est l'application $g \circ f \circ g^{-1}$.

Théorème 8.20. *théorème de conjugaison.* *Étant donné un repère R de $\mathcal{E}$, la conjuguée $f' = g \circ f \circ g^{-1}$ a même expression analytique dans le repère $R' = g(R)$ que f dans le repère R .*

Démonstration. Soit O l'origine et $\mathcal{B} = (\vec{e}_1, \dots, \vec{e}_n)$ la base de R . À tout $M \in \mathcal{E}$ de coordonnées $x_1, \dots, x_n$ correspond $f(M) = N$ de coordonnées $y_1, \dots, y_n$. On a la relation matricielle

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

où $A = (a_{i,j})$ est la matrice de $\bar{f}$ dans la base $\mathcal{B}$, $b_1, \dots, b_n$ les coordonnées de $f(O)$. Soit le repère $R' = g(R)$ d'origine $O' = g(O)$, de base $\mathcal{B}' = (\vec{g}(\vec{e}_1), \dots, \vec{g}(\vec{e}_n))$, M de coordonnées $x_1, \dots, x_n$ et $M' = g(M)$. On a

$$\vec{OM} = \sum_i x_i \vec{e}_i \text{ donc } \vec{O'M'} = \vec{g}(\vec{OM}) = \sum_i x_i \vec{g}(\vec{e}_i)$$

Donc $M' = g(M)$ a mêmes coordonnées dans R' que M dans R . De plus,

$$f'(M') = (g \circ f \circ g^{-1})(g(M)) = (g \circ f)(M) = g(N)$$

d'où $f'(M') = g(N)$ a mêmes coordonnées $y_1, \dots, y_n$ dans R' que $f(M) = N$ dans R . $\square$

Ainsi, $f' = g \circ f \circ g^{-1}$ opère relativement au repère $R' = g(R)$ de la même façon que f relativement à R . Ceci justifie et précise l'assertion : f et $f' = g \circ f \circ g^{-1}$ ont même nature géométrique. La relation « f et f' sont conjuguées » est une relation d'équivalence sur l'ensemble des applications affines de $\mathcal{E}$ dans lui-même. Les classes d'équivalence s'appellent les **classes de conjugaison**. Elles rassemblent les applications affines de mêmes propriétés géométriques.

- **Conjuguée d'une translation** La conjugée par g de la translation de vecteur $\vec{v}$ est la translation de vecteur $\bar{g}(\vec{v}) : g \circ T_{\vec{v}} \circ g^{-1} = T_{\bar{g}(\vec{v})}$. Les translations de vecteur non nul constituent donc une classe de conjugaison.

En effet, soit M un point, $N = g^{-1}(M)$. On a

$$\begin{aligned} (g \circ T_{\vec{v}} \circ g^{-1})(M) &= g(T_{\vec{v}}(N)) = g(N + \vec{v}) = g(N) + \bar{g}(\vec{v}) \\ &= M + \bar{g}(\vec{v}) = T_{\bar{g}(\vec{v})}(M) \end{aligned}$$

- **Conjuguée d'une homothétie** La conjugée par g de l'homothétie de centre O et de rapport λ est l'homothétie $g \circ H_{O,\lambda} \circ g^{-1} = H_{g(O),\lambda}$ de centre $O' = g(O)$ et de même rapport λ . Les homothéties de même rapport λ forment donc une classe de conjugaison.

En effet, posons $h = (g \circ H_{O,\lambda} \circ g^{-1})$. On a $h(O') = O'$ car

$$h(O') = (g \circ H_{O,\lambda} \circ g^{-1})(g(O)) = (g \circ H_{O,\lambda})(O) = g(O) = O'$$

L'application linéaire associée à h est $\bar{h} = \bar{g} \circ \lambda I_E \circ \bar{g}^{-1}$. L'homothétie vectorielle λI_E commutant avec toute application linéaire, $\bar{h} = \lambda I_E \circ \bar{g} \circ \bar{g}^{-1} = \lambda I_E$ est l'homothétie vectorielle de rapport λ . D'où $h = H_{g(O),\lambda}$ (8.16, p. 210).

- **Conjuguée d'une projection** Le lecteur vérifiera que la conjugée par g de la projection $\mathcal{E} \rightarrow \mathcal{F} \subset \mathcal{E}$ parallèlement à un supplémentaire G de $\mathcal{F}$ est la projection $\mathcal{E} \rightarrow g(\mathcal{F})$ parallèlement à $\bar{g}(G)$.

8.1.6. Groupe des homothéties-translations

Le groupe des homothéties vectorielles λI_E , $\lambda \neq 0$, est sous-groupe distingué de $\mathbf{GL}(E)$, isomorphe au groupe multiplicatif $\mathbb{R}^\times$ des réels non nuls. Son image réciproque par le morphisme $\mathbf{GA}(\mathcal{E}) \rightarrow \mathbf{GL}(E)$ (8.19, p. 212) est le sous-groupe distingué de $\mathbf{GA}(\mathcal{E})$ formé des translations et des homothéties affines (8.14 p. 209 et 8.16, p. 210), noté $\mathcal{H}(\mathcal{E})$. Pour tout $h \in \mathcal{H}(\mathcal{E})$, le rapport $\lambda(h)$ de l'homothétie vectorielle $\bar{h}$ est appelé **rapport** de h . L'application $h \mapsto \lambda(h)$ est un morphisme de groupes $\mathcal{H}(\mathcal{E}) \rightarrow \mathbb{R}^\times$ de noyau le sous-groupe $\mathcal{T}(\mathcal{E})$ des translations.

Pour tout $h \in \mathcal{H}(\mathcal{E})$ distinct de I_E , soit $\Delta(h)$ l'ensemble des droites stables par h . On vérifie facilement que :

- si $h = T_{\vec{v}}$, $\vec{v} \neq \vec{0}$, $\Delta(T_{\vec{v}})$ est l'ensemble des droites de vecteur directeur $\vec{v}$,
- si $h = H_{O,\lambda}$, $\lambda \neq 1$, $\Delta(H_{O,\lambda})$ est l'ensemble des droites passant par O .

Proposition 8.21. *a) Soit les homothéties, distinctes de l'identité, de centres distincts $f = H_{A,\alpha}$ et $g = H_{B,\beta}$. Les composées $g \circ f$ et $f \circ g$ sont alors :*

- *pour $\alpha\beta \neq 1$, des homothéties de rapports $\alpha\beta$ et de centres distincts situés sur la droite (AB) ,*
- *pour $\alpha\beta = 1$, des translations de vecteurs distincts colinéaires à $\overrightarrow{AB}$.*

b) Soit une homothétie $f = H_{A,\alpha}$ et une translation $g = T_{\vec{v}}$, distinctes de l'identité. Les composées $g \circ f$ et $f \circ g$ sont alors des homothéties de rapport α et de centres distincts situés sur la droite $A + \mathbb{R}\vec{v}$.

Démonstration. Soit f et g dans $\mathcal{H}(\mathcal{E})$ avec les rapports α et β , $h_1 = f \circ g$ et $h_2 = g \circ f$. Les homothéties vectorielles $\bar{f} = \alpha I_E$ et $\bar{g} = \beta I_E$ commutent, d'où $h_1 = h_2 = \alpha\beta I_E$. Donc h_1 et h_2 sont dans $\mathcal{H}(\mathcal{E})$ et ont même rapport $\alpha\beta$. Ce sont des homothéties si $\alpha\beta \neq 1$, des translations si $\alpha\beta = 1$.

On suppose que $f = H_{A,\alpha}$ est une homothétie de centre A et de rapport $\alpha \neq 1$. Montrons que si $g(A) \neq A$, i.e. si g n'est pas une homothétie de centre A , alors $h_1 = f \circ g$ et $h_2 = g \circ f$ sont distinctes.

Par l'absurde, supposons $h_1 = f \circ g = g \circ f = h_2$, alors $f = g \circ f \circ g^{-1}$. Or $g \circ f \circ g^{-1}$ admet $g(A) \neq A$ pour point fixe. Comme A est l'unique point fixe de f , on a $g \circ f \circ g^{-1} \neq f$, donc $h_2 \neq h_1$.

a) Supposons que $f = H_{A,\alpha}$ et $g = H_{B,\beta}$ sont des homothéties de centres distincts A et B . La droite $\mathcal{D} = (AB)$ est commune à $\Delta(f)$ et $\Delta(g)$. C'est donc une droite stable pour $h_1 = f \circ g$ et $h_2 = g \circ f$.

- Si $\alpha\beta \neq 1$, h_1 et h_2 sont des homothéties de même rapport $\alpha\beta$ ayant (AB) pour droite stable. Leurs centres sont donc situés sur (AB) . Ils sont distincts car h_2 et h_1 sont distinctes, de même rapport.
- Si $\alpha\beta = 1$, h_1 et h_2 sont des translations ayant (AB) pour droite stable, donc de vecteurs colinéaires à $\overrightarrow{AB}$. Ces vecteurs sont distincts car h_1 et h_2 sont des translations distinctes.

b) Supposons que $f = H_{A,\alpha}$ est une homothétie et $g = T_{\vec{v}}$ est une translation. Alors h_1 et h_2 sont des homothéties distinctes de même rapport α . La droite $\mathcal{D} = A + \mathbb{R}\vec{v}$ appartient à $\Delta(f) \cap \Delta(g)$, donc est stable pour h_1 et h_2 . Les centres de h_1 et h_2 sont sur $\mathcal{D}$, distincts car h_1 et h_2 sont des homothéties distinctes de même rapport α . $\square$

L'énoncé ne précise pas la localisation exacte sur $\mathcal{D}$ des centres de $g \circ f$ et $f \circ g$ si ce sont des homothéties. Si ce sont des translations, leurs vecteurs colinéaires à $\mathcal{D}$ ne sont pas non plus précisés. Ces renseignements complémentaires ne sont pas inaccessibles, mais sont d'intérêt secondaire.

Remarque : Si $f \in \mathbf{GA}(\mathcal{E})$ admet une droite stable $\mathcal{D}$, l'application $\mathcal{D} \rightarrow \mathcal{D}$ induite par f est une homothétie ou une translation.

En effet, l'application induite par f appartient au groupe $\mathbf{GA}(\mathcal{D})$, réduit aux homothéties et translation car $\mathbf{GL}(\overline{\mathcal{D}})$ se réduit aux homothéties vectorielles.

Si f a deux points fixes A, B distincts, tout point de (AB) est fixe.

8.1.7. Rapports de mesures algébriques

Définition 8.22. Soit $\mathcal{D}$ une droite affine et $\vec{u} \in \overline{\mathcal{D}}$ non nul. Si P, Q sont deux points de $\mathcal{D}$, la **mesure algébrique** $\overline{PQ}$ est le réel défini par $\overrightarrow{PQ} = \overline{PQ} \vec{u}$. Ceci dépend du choix de $\vec{u}$. En revanche, les rapports de mesures algébriques sont **intrinsèques**, c'est-à-dire indépendants du choix de $\vec{u}$.

Ainsi A, B, C, D étant quatre points de $\mathcal{D}$, (C, D distincts), le rapport $\frac{\overline{AB}}{\overline{CD}}$ ne dépend pas du vecteur directeur $\vec{u}$. Plus généralement, on définit le rapport $\frac{\overline{AB}}{\overline{CD}}$ si (AB) et (CD) sont parallèles (en convenant de les munir du même vecteur directeur). Par exemple, si A, B sont les transformés de C, D par l'homothétie $H_{O, \lambda}$, on aura $\frac{\overline{AB}}{\overline{CD}} = \lambda$.

Étant donné A, B distincts sur une droite $\mathcal{D}$, l'application

$$\begin{aligned} \mathcal{D} \setminus \{B\} &\rightarrow \mathbb{R} \setminus \{1\} \\ M &\mapsto \frac{\overline{MA}}{\overline{MB}} \end{aligned}$$

est bijective : au rapport $\lambda \neq 1$ correspond le point $M = A + \frac{\lambda}{\lambda-1} \overrightarrow{AB}$.

Théorème 8.23. Étant donné deux espaces affines $\mathcal{E}$ et $\mathcal{E}'$ d'espaces vectoriels associés E et E' , une application $f: \mathcal{E} \rightarrow \mathcal{E}'$ est affine si et seulement si elle conserve l'alignement et les rapports de mesures algébriques.

Démonstration. Précisons la propriété « f conserve l'alignement et les rapports de mesures algébriques ». Soit A, B, C trois points alignés distincts de $\mathcal{E}$ d'images A', B', C' . Alors A', B', C' sont alignés et

ou bien A', B', C' sont distincts avec $\frac{\overline{A'C'}}{\overline{A'B'}} = \frac{\overline{AC}}{\overline{AB}}$,

ou bien A', B', C' sont confondus.

L'unique condition $\overrightarrow{A'C'} = \frac{\overline{AC}}{\overline{AB}} \overrightarrow{A'B'}$ englobe ces deux cas.

Si f est affine, la propriété est facile à vérifier. Inversement, supposons que f a cette propriété. Soit $O \in \mathcal{E}$ et $O' = f(O)$. Il suffit de montrer que l'application $\overline{f}_O: E \rightarrow E'$ est linéaire.

Étant donné $\lambda \in \mathbb{R}$, $\vec{v} \in E$ non nul, $M = O + \vec{v}$, $N = O + \lambda \vec{v}$, O', M', N' les images de O, M, N . La propriété de f donne

$$\lambda = \frac{\overline{ON}}{\overline{OM}} = \frac{\overline{O'N'}}{\overline{O'M'}} \quad \text{soit} \quad \overrightarrow{O'N'} = \lambda \overrightarrow{O'M'} \quad \text{donc} \quad \overline{f}_O(\lambda \vec{v}) = \lambda \overline{f}_O(\vec{v})$$

Montrons que pour $\vec{u}, \vec{v}$ dans E , $\overline{f}_O(\vec{u} + \vec{v}) = \overline{f}_O(\vec{u}) + \overline{f}_O(\vec{v})$. On pose $P = O + \vec{u}$, $Q = O + \vec{v}$, $R = O + \vec{u} + \vec{v}$. Alors O, P, R, Q est un parallélogramme, donc (O, R) et (P, Q) ont même milieu I . Notant O', P', Q', R' les images de O, P, Q, R , la propriété de f donne que (O', R') et (P', Q') ont même milieu $f(I)$, donc que

O', P', R', Q' est un parallélogramme, donc que

$$\overline{f_0}(\vec{u} + \vec{v}) = \overline{f_0}(\overrightarrow{OR}) = \overrightarrow{O'R'} = \overrightarrow{O'P'} + \overrightarrow{O'Q'} = \overline{f_0}(\overrightarrow{OP}) + \overline{f_0}(\overrightarrow{OQ}) = \overline{f_0}(\vec{u}) + \overline{f_0}(\vec{v})$$

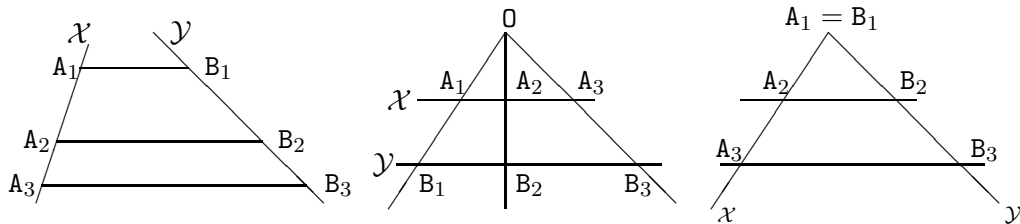
□

Proposition 8.24. de Thalès. Soit $\mathcal{P}$ un plan affine de plan vectoriel associé P , $\mathcal{X}$ et $\mathcal{Y}$ deux droites de $\mathcal{P}$, δ une droite vectorielle distincte de $\mathcal{X}$ et $\mathcal{Y}$, A_1, A_2, A_3 des points de $\mathcal{X}$, B_1, B_2, B_3 les points où les droites passant par A_1, A_2, A_3 de direction δ coupent $\mathcal{Y}$. Alors on a

$$\frac{\overline{A_1A_3}}{\overline{A_1A_2}} = \frac{\overline{B_1B_3}}{\overline{B_1B_2}}$$

Démonstration. En effet, la projection de $\mathcal{P}$ sur $\mathcal{Y}$ parallèlement à δ transforme A_1, A_2, A_3 en B_1, B_2, B_3 et conserve les rapports des mesures algébriques. □

Inversement, cette égalité de rapports n'implique pas le parallélisme des droites (A_iB_i) car une application affine $\mathcal{X} \rightarrow \mathcal{Y}$ conserve les rapports des mesures algébriques, mais peut ne pas être restriction à $\mathcal{X}$ d'une projection $\mathcal{P} \rightarrow \mathcal{Y}$. Pour un contre-exemple, prendre $\mathcal{X}, \mathcal{Y}$ parallèles, O un point non sur les droites $\mathcal{X}, \mathcal{Y}$, B_i l'intersection de (OA_i) avec $\mathcal{Y}$. L'homothétie de centre O et de rapport $\frac{\overline{OB_1}}{\overline{OA_1}}$ transforme A_i en B_i et conserve les rapports algébriques. On a égalité des rapports et pourtant les droites (A_iB_i) se coupent en O .



En revanche, si on a égalité des rapports et si $A_1 = B_1$, alors (A_2B_2) et (A_3B_3) sont parallèles. En effet, la projection p parallèlement à (A_2B_2) sur $\mathcal{Y}$ transforme A_1, A_2 en B_1, B_2 . Comme p conserve les rapports des mesures algébriques, $p(A_3) = B_3$, d'où le parallélisme de (A_2B_2) et de (A_3B_3) .

8.1.8. Divisions harmoniques

Définition 8.25. Étant donné deux points A, B distincts d'une droite $\mathcal{D}$, on dit que les points distincts M et N de $\mathcal{D}$, distincts de A et B , sont **harmoniquement conjugués** par rapport à A et B si on a la relation

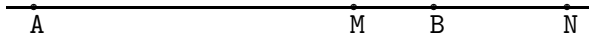
$$\frac{\overline{MA}}{\overline{MB}} = -\frac{\overline{NA}}{\overline{NB}}$$

L'application $N \mapsto \frac{\overline{NA}}{\overline{NB}}$ étant une bijection de $\mathcal{D} \setminus \{B\}$ sur $\mathbb{R} \setminus \{1\}$, on voit que *tout point* $M \in \mathcal{D}$, *distinct de* A, B *et du milieu de* (A, B) (pour $-\frac{\overline{MA}}{\overline{MB}} \neq 1$) *admet un unique conjugué harmonique* $N \in \mathcal{D}$.

Cette relation est clairement équivalente aux relations

$$\frac{\overline{NA}}{\overline{NB}} = -\frac{\overline{MA}}{\overline{MB}} \text{ ou } \frac{\overline{AM}}{\overline{AN}} = -\frac{\overline{BM}}{\overline{BN}} \text{ ou } \frac{\overline{AN}}{\overline{AM}} = -\frac{\overline{BN}}{\overline{BM}}$$

Autrement dit, A et B jouent le même rôle, de même que M et N . De plus, M et N sont conjugués harmoniques par rapport à A et B si et seulement si A et B sont conjugués harmoniques par rapport à M et N . Il s'agit donc d'une relation où les paires de points $\{A, B\}$ et $\{M, N\}$ jouent le même rôle l'une par rapport à l'autre, les deux points d'une même paire jouant aussi le même rôle et pouvant être intervertis. On dit alors que les quatre points alignés A, B, M, N forment une **division harmonique**. C'est une configuration qu'on rencontre souvent en géométrie.



On a alors la relation $\overline{MA} \cdot \overline{NB} + \overline{MB} \cdot \overline{NA} = 0$ qui n'a de sens que si on a choisi un vecteur directeur de $\mathcal{D}$ pour définir les mesures algébriques.

Étant donné une origine $O \in \mathcal{D}$ et un vecteur directeur $\vec{v}$, posons $\overline{OA} = \alpha$, $\overline{OB} = \beta$, $\overline{OM} = x$ et $\overline{ON} = y$. La relation précédente devient

$$2(\alpha\beta + xy) - (\alpha + \beta)(x + y) = 0$$

Cette relation garde un sens si trois de ces points sont confondus, par exemple M et N avec A ou B , ce qui permet de dire qu'*étant donné deux points* A, B , *le conjugué harmonique de* A *relativement à* A *et* B *est* A *lui-même*.

Enfin, il est immédiat (8.23, p. 215) que *les applications affines conservent les divisions harmoniques*.

Proposition 8.26. Relation de Newton. *Soit* A, B, M, N *quatre points alignés, I et J les milieux de* (A, B) *et de* (M, N) . *On a équivalence entre*

- (i) *Les quatre points* A, B, M, N *sont en division harmonique.*
- (ii) $\overline{IA}^2 = \overline{IB}^2 = \overline{IM} \times \overline{IN}$,
- (iii) $\overline{JM}^2 = \overline{JN}^2 = \overline{JA} \times \overline{JB}$.

Démonstration. Il suffit de placer l'origine O en I ou J . □

Plus généralement, on définit le **birapport** de quatre points alignés distincts A, B, M, N comme étant

$$[A, B, M, N] = \frac{\overline{AM}}{\overline{AN}} : \frac{\overline{BM}}{\overline{BN}} = \frac{\overline{MA}}{\overline{MB}} : \frac{\overline{NA}}{\overline{NB}} = \frac{\overline{AM}}{\overline{BM}} \times \frac{\overline{BN}}{\overline{AN}}$$

où le signe $:$ signifie **divisé par**. Ainsi, A, B, M, N sont en division harmonique si et seulement si $[A, B, M, N] = -1$. On a $[A, B, M, N] = [B, A, N, M] = [M, N, A, B] = [N, M, B, A]$.

En revanche, pour d'autres ordonnancements, le birapport change. Par exemple, $[A, B, N, M] = \frac{1}{[A, B, M, N]}$.

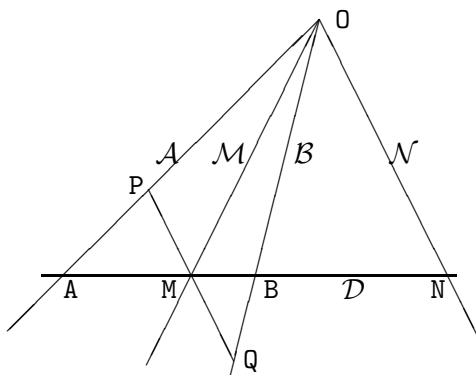
Proposition 8.27. Faisceaux harmoniques. Dans un plan $\mathcal{P}$, soit quatre droites distinctes, concourantes ou parallèles, $\mathcal{A}, \mathcal{B}, \mathcal{M}, \mathcal{N}$. S'il existe une sécante $\mathcal{D}$ coupant $\mathcal{A}, \mathcal{B}, \mathcal{M}, \mathcal{N}$ en quatre points distincts A, B, M, N formant une division harmonique, il en est de même pour toute autre sécante $\mathcal{D}'$ coupant ces droites en quatre points distincts.

On dit alors que $\mathcal{A}, \mathcal{B}, \mathcal{M}, \mathcal{N}$ forment un **faisceau harmonique** ce qu'on indique par la notation $[\mathcal{A}, \mathcal{B}, \mathcal{M}, \mathcal{N}] = -1$.

Démonstration. a) Si $\mathcal{A}, \mathcal{B}, \mathcal{M}, \mathcal{N}$ ont même direction δ , on utilise la conservation des rapports des mesures algébriques par projection suivant δ d'une sécante $\mathcal{D}$ sur l'autre $\mathcal{D}'$.

b) Supposons $\mathcal{A}, \mathcal{B}, \mathcal{M}, \mathcal{N}$ sécantes en O et les sécantes $\mathcal{D}$ et $\mathcal{D}'$ parallèles. On conclut en utilisant la conservation des rapports des mesures algébriques par l'homothétie de centre O transformant $\mathcal{D}$ en $\mathcal{D}'$.

c) Supposons $\mathcal{A}, \mathcal{B}, \mathcal{M}, \mathcal{N}$ sécantes en O et $\mathcal{D}, \mathcal{D}'$ de directions distinctes coupant $\mathcal{A}, \mathcal{B}, \mathcal{M}, \mathcal{N}$ en respectivement A, B, M, N et A', B', M', N' . On doit prouver l'équivalence $([A, B, M, N] = -1) \Leftrightarrow ([A', B', M', N'] = -1)$.



Soit $\mathcal{D}'_1$ la parallèle à $\mathcal{D}'$ passant par M coupant $\mathcal{A}, \mathcal{B}, \mathcal{M}, \mathcal{N}$ en respectivement $A'_1, B'_1, M'_1 = M, N'_1$. D'après b), $[A', B', M', N'] = -1$ est équivalent à $[A'_1, B'_1, M'_1, N'_1] = -1$. On est donc ramené à prouver l'équivalence

$$([A, B, M, N] = -1) \Leftrightarrow ([A'_1, B'_1, M'_1, N'_1] = -1)$$

Soit la parallèle à $\mathcal{N}$ passant par M coupant $\mathcal{A}, \mathcal{B}$ en P, Q . Considérant les homothéties de centres A et B transformant (PQ) en $\mathcal{N}$, on a les égalités :

$$\frac{\overline{AM}}{\overline{AN}} = \frac{\overline{MP}}{\overline{NO}}, \quad \frac{\overline{BM}}{\overline{BN}} = \frac{\overline{MQ}}{\overline{NO}}$$

On a $[A, B, M, N] = -1$ si et seulement si ces rapports sont opposés, i.e. si M est milieu de (P, Q) . Le même raisonnement appliqué à $\mathcal{D}'_1$ permet de conclure. $\square$

La preuve précédente permet de compléter ainsi cet énoncé : *quatre droites $\mathcal{A}, \mathcal{B}, \mathcal{M}, \mathcal{N}$ concourantes forment un faisceau harmonique si et seulement si une parallèle à $\mathcal{N}$ coupe $\mathcal{A}, \mathcal{B}, \mathcal{M}$ en P, Q, M de sorte que M soit milieu de (P, Q) .*

Exemple : On verra (8.40, p. 227) qu'étant donné deux droites $\mathcal{A}$ et $\mathcal{B}$ d'un plan affine euclidien sécantes en O , si $\mathcal{M}$ et $\mathcal{N}$ sont les bissectrices de $(\widehat{\mathcal{A}, \mathcal{B}})$, alors $[\mathcal{A}, \mathcal{B}, \mathcal{M}, \mathcal{N}] = -1$.

8.1.9. Théorèmes de Ménélaüs et de Ceva

Ce sont des théorèmes de géométrie plane. La scène se passe dans un plan affine $\mathcal{P}$ de plan vectoriel associé P .

Proposition 8.28. Soit un triangle ABC , P, Q, R des points de (BC) , (CA) , (AB) distincts des sommets.

(i) Les points P, Q, R sont alignés si et seulement si

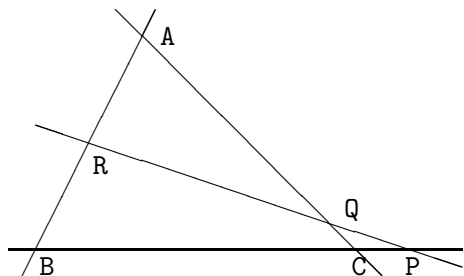
$$\frac{\overline{PB}}{\overline{PC}} \times \frac{\overline{QC}}{\overline{QA}} \times \frac{\overline{RA}}{\overline{RB}} = +1$$

(ii) Les droites (AP) , (BQ) , (CR) sont concourantes ou parallèles si et seulement si

$$\frac{\overline{PB}}{\overline{PC}} \times \frac{\overline{QC}}{\overline{QA}} \times \frac{\overline{RA}}{\overline{RB}} = -1$$

Démonstration. Les propriétés (i) et (ii) sont les théorèmes de, respectivement, Ménélaüs et Ceva.

(i) **Théorème de Ménélaüs.** Supposons P, Q, R alignés. Soit les homothéties H_P, H_Q de centres P, Q de rapports $\frac{\overline{PC}}{\overline{PB}}$ et $\frac{\overline{QA}}{\overline{QC}}$. Elles transforment respectivement B en C et C en A .



Si $H_Q \circ H_P$ était une translation, ce serait $T_{\overrightarrow{BA}}$; c'est impossible car $\overrightarrow{BA}$ et $\overrightarrow{PQ}$ sont non colinéaires. C'est donc une homothétie de centre sur (PQ) et aussi sur (AB) car transformant B en A . Son centre est donc R et son rapport est

$$\frac{\overline{RA}}{\overline{RB}} = \frac{\overline{PC}}{\overline{PB}} \times \frac{\overline{QA}}{\overline{QC}}$$

d'où l'on déduit la relation.

Inversement, supposons la relation vérifiée. Comme $\overrightarrow{RA} \neq \overrightarrow{RB}$,

$$\frac{\overline{PC}}{\overline{PB}} \times \frac{\overline{QA}}{\overline{QC}} = \frac{\overline{RA}}{\overline{RB}} \neq 1 \text{ donc } \frac{\overline{PC}}{\overline{PB}} \neq \frac{\overline{QC}}{\overline{QA}}$$

La droite (PQ) n'est donc pas parallèle à (AB) et la coupe en un point R' qui d'après la partie directe vérifie

$$\frac{\overline{R'A}}{\overline{R'B}} = \frac{\overline{PC}}{\overline{PB}} \times \frac{\overline{QA}}{\overline{QC}}$$

La relation donnée implique donc $\frac{\overline{R'A}}{\overline{R'B}} = \frac{\overline{RA}}{\overline{RB}}$, d'où $R = R'$. Les points P, Q, R sont donc alignés.

(ii) **Théorème de Ceva.** a) Supposons (AP), (BQ), (CR) concourantes en I. L'application du théorème de Ménélaüs aux triangles ABP et APC munis respectivement des sécantes (CR) et (BQ) donne :

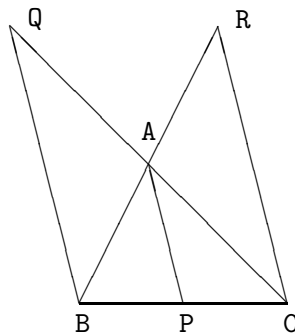
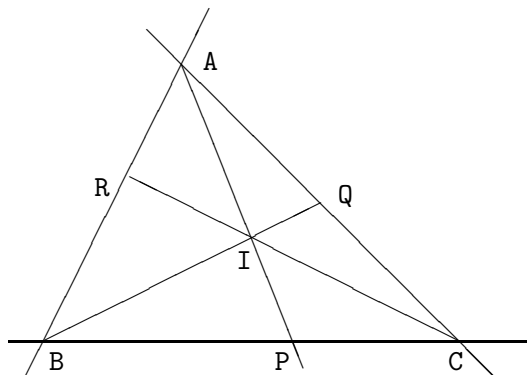
$$\frac{\overline{IA}}{\overline{IP}} \times \frac{\overline{RB}}{\overline{RA}} \times \frac{\overline{CP}}{\overline{CB}} = 1, \quad \frac{\overline{IP}}{\overline{IA}} \times \frac{\overline{BC}}{\overline{BP}} \times \frac{\overline{QA}}{\overline{QC}} = 1$$

On obtient le résultat annoncé en multipliant ces deux relations.

b) Supposons (AP), (BQ), (CR) parallèles. Par projection sur (BC), on a

$$\frac{\overline{QA}}{\overline{QC}} = \frac{\overline{BP}}{\overline{BC}}, \quad \frac{\overline{RB}}{\overline{RA}} = \frac{\overline{CB}}{\overline{CP}}$$

d'où le résultat en multipliant ces égalités.



Réciproquement, supposons la relation satisfaite.

c) Supposons (BQ) et (CR) sécantes en I. Montrons que (AI) et (BC) ne peuvent pas être parallèles.

Supposons qu'elles le soient. Soit les homothéties H_Q et H_R de centres Q et R transformant C en A et A en B. On aurait $H_Q(B) = I$ et $H_R(I) = C$. Alors $H_R \circ H_Q$ échangerait B et C et serait une symétrie centrale, donc une homothétie de rapport -1 . On aurait donc

$$-\frac{\overline{PB}}{\overline{PC}} = \frac{\overline{QA}}{\overline{QC}} \times \frac{\overline{RB}}{\overline{RA}} = -1$$

ce qui est impossible car $B \neq C$.

Donc (AI) coupe (BC) en un point P' . Par la partie directe, on a

$$\frac{\overline{P'B}}{\overline{P'C}} \times \frac{\overline{QC}}{\overline{QA}} \times \frac{\overline{RA}}{\overline{RB}} = -1 = \frac{\overline{PB}}{\overline{PC}} \times \frac{\overline{QC}}{\overline{QA}} \times \frac{\overline{RA}}{\overline{RB}}$$

On en déduit $\frac{\overline{PB}}{\overline{PC}} = \frac{\overline{P'B}}{\overline{P'C}}$, d'où $P = P'$.

d) Supposons enfin (BQ) et (CR) de même direction δ . La droite passant par A de direction δ coupe (BC) en P' . On voit de même que $P = P'$. $\square$

8.1.10. Géométrie sur une droite affine réelle

Soit $\mathcal{D}$ une droite affine de direction D , trois points distincts A, B, C de $\mathcal{D}$, $\vec{u}$ un vecteur directeur de $\mathcal{D}$. On dit que C est **entre** A et B si les vecteurs colinéaires $\overrightarrow{AC}$ et $\overrightarrow{CB}$ sont de même sens : les mesures algébriques $\overline{AC}$ et $\overline{CB}$ ont même signe, soit $\frac{\overline{AC}}{\overline{CB}} > 0$. Le point C est entre A et B si et seulement s'il est entre B et A. En effet, les rapports $\frac{\overline{AC}}{\overline{CB}}$ et $\frac{\overline{BC}}{\overline{CA}}$ sont inverses, donc de même signe.

Proposition 8.29. *Étant donné trois points distincts A, B, C de $\mathcal{D}$, il en est un et un seul qui est entre les deux autres.*

Démonstration. On a

$$\overrightarrow{AC} + \overrightarrow{CB} + \overrightarrow{BA} = (\overline{AC} + \overline{CB} + \overline{BA})\vec{u} = \vec{0}$$

d'où $\overline{AC} + \overline{CB} + \overline{BA} = 0$. Les réels $\overline{AC}$, $\overline{CB}$, $\overline{BA}$ ne sont donc pas de même signe. Si $\overline{BA}$ est d'un signe et $\overline{AC}$, $\overline{CB}$ de l'autre signe, alors C est entre A et B. $\square$

Définition 8.30. *Étant donné deux points A, B de $\mathcal{D}$, on appelle **segment d'extrémités** A et B l'ensemble, noté $[A, B]$ ou $[B, A]$, formé de A, de B et des points $M \in \mathcal{D}$ qui sont entre A et B.*

*Un point A $\in \mathcal{D}$ détermine deux **demi-droites** : étant donné un vecteur directeur $\vec{u}$, les demi-droites sont*

- la demi-droite du sens de $\vec{u}$ formée des $M \in \mathcal{D}$ tels que $\overrightarrow{AM}$ et $\vec{u}$ soient de même sens,
- la demi-droite du sens opposé à $\vec{u}$ formée des $M \in \mathcal{D}$ tels que $\overrightarrow{AM}$ et $\vec{u}$ soient de sens opposés.

La bijection, représentation paramétrique de $\mathcal{D}$,

$$\begin{aligned} \mathbb{R} &\longrightarrow \mathcal{D} \\ \lambda &\longmapsto M = A + \lambda \overrightarrow{AB} \end{aligned}$$

envoie $[0, 1] \subset \mathbb{R}$ sur le segment $[A, B]$ de $\mathcal{D}$, car $\overrightarrow{AM} = \lambda \overrightarrow{AB}$ et $\overrightarrow{MB} = (1 - \lambda) \overrightarrow{AB}$ ont même sens si et seulement si λ et $1 - \lambda$ ont même signe, i.e. $\lambda \in [0, 1]$.

De même, elle envoie $] - \infty, 0]$ et $[1, +\infty[$ sur les demi-droites d'origine A ne contenant pas B et d'origine B ne contenant pas A.

Exercice 8.1. Montrer que deux droites d'un espace de dimension 3 peuvent avoir une intersection vide sans être parallèles

Exercice 8.2. Étant donné deux sous-espaces affines $\mathcal{F}$ et $\mathcal{G}$ de dimensions p et q avec $p \leq q$. On dit que $\mathcal{F}$ et $\mathcal{G}$ sont **parallèles** si $\overline{\mathcal{F}} \subset \overline{\mathcal{G}}$.

1) Soit F et G deux sous-espaces vectoriels de E , A et B deux points de $\mathcal{E}$. Montrer que $A + F$ et $B + G$ sont d'intersection non vide si et seulement si $\overrightarrow{AB} \in F + G$.

2) En déduire que deux sous-espaces affines de dimensions p et q comprises entre 1 et $n - 2$ peuvent être d'intersection vide sans être parallèles.

3) Montrer que si un sous-espace affine $\mathcal{F}$ et un hyperplan affine $\mathcal{G}$ sont d'intersection vide, alors ils sont parallèles.

Exercice 8.3. Dans $\mathcal{E}$ affine de dimension 3, quelles sont les configurations possibles pour que trois droites distinctes $\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3$ soient deux à deux coplanaires?

Exercice 8.4. Soit un triangle ABC , une droite $\mathcal{D}$ coupant (BC) , (CA) , (AB) en A', B', C' distincts des sommets, les milieux A'', B'', C'' de (A, A') , (B, B') , (C, C') . Montrer que A'', B'', C'' sont alignés (utiliser la commutativité des homothéties h et h' de centre C' transformant respectivement B en A et B' en A' et considérer les symétriques P et Q de C par rapport à A'' et B'').

8.2 GÉOMÉTRIE EUCLIDIENNE

8.2.1. Distance et produit scalaire

Un **espace affine euclidien** de dimension n est la donnée d'un espace affine $\mathcal{E}$ de dimension n dont l'espace vectoriel associé E est muni d'un **produit scalaire euclidien** (et donc d'une norme euclidienne)

$$E \times E \rightarrow \mathbb{R}, \quad (\vec{v}, \vec{w}) \mapsto \vec{v} \cdot \vec{w}$$

L'espace affine euclidien de dimension 3 est un modèle mathématique de l'espace physique ambiant. Dans la suite, l'espace euclidien est supposé **orienté**. Pour l'espace physique ambiant, la convention d'orientation communément employée est celle du « bonhomme d'Ampère » ou des « trois doigts de la main droite » : pouce, index, majeur, pour une base directe.

De la norme euclidienne de E , on déduit la **distance** sur $\mathcal{E}$

$$\mathcal{E} \times \mathcal{E} \rightarrow \mathbb{R}^+, (M, N) \mapsto MN = \|\overrightarrow{MN}\|$$

L'**inégalité triangulaire** prend alors la forme suivante : étant donné trois points A, B, C, on a $AC \leq AB + BC$, car, par la relation de Chasles, on a

$$AC = \|\vec{AC}\| = \|\vec{AB} + \vec{BC}\| \leq \|\vec{AB}\| + \|\vec{BC}\| = AB + BC$$

Cette inégalité large est une égalité si et seulement si $\vec{AB}$ et $\vec{BC}$ sont :

- colinéaires, c'est-à-dire A, B, C alignés,
- de même sens, auquel cas B appartient au segment $[A, C]$.

Proposition 8.31. Étant donné trois points A, B, C, on a les inégalités

$$|AB - BC| \leq AC \leq AB + BC$$

Démonstration. Il suffit de prouver l'inégalité de gauche. On a $AB \leq AC + BC$, donc $AB - BC \leq AC$. De même $BC - AB \leq AC$, d'où l'inégalité.

Inversement, étant donné a, b, c positifs vérifiant $|c - a| \leq b \leq c + a$, existe-t-il des points A, B, C tels que $a = BC$, $b = CA$, $c = AB$? La réponse est oui, mais la démonstration n'est pas évidente (Corollaire 11.6, p. 286) $\square$

Définition 8.32. La mesure de l'angle non orienté $\widehat{\vec{v}, \vec{w}}$ de deux vecteurs non nuls $\vec{v}$ et $\vec{w}$ est l'unique nombre de $[0, \pi]$ tel que

$$\cos \widehat{\vec{v}, \vec{w}} = \frac{\vec{v} \cdot \vec{w}}{\|\vec{v}\| \times \|\vec{w}\|}$$

Dans l'espace physique, les notions d'angle et d'orthogonalité sont intrinsèques et ne résultent pas d'une convention alors que la distance dépend de l'**unité de longueur** choisie. L'espace vectoriel E est donc muni d'une **famille de formes quadratiques définies positives proportionnelles**. Les produits scalaires associés définissent les mêmes notions d'orthogonalité et d'angle. Le choix de l'un d'eux correspond au choix d'une unité de longueur. Dans la suite, ce choix est supposé fait une fois pour toutes : E est donc muni d'une norme euclidienne et du produit scalaire associé.

8.2.2. Orthocentre, droite d'Euler, cercle des neuf points

Proposition 8.33. Identité du produit scalaire. Soit A, B, C, D quatre points. On a $\vec{DA} \cdot \vec{BC} + \vec{DB} \cdot \vec{CA} + \vec{DC} \cdot \vec{AB} = 0$

Démonstration. En développant l'expression X suivante, on trouve 0 :

$$\begin{aligned} X &= \vec{DA} \cdot \vec{BC} + \vec{DB} \cdot \vec{CA} + \vec{DC} \cdot \vec{AB} \\ &= \vec{DA} \cdot (\vec{DC} - \vec{DB}) + \vec{DB} \cdot (\vec{DA} - \vec{DC}) + \vec{DC} \cdot (\vec{DB} - \vec{DA}) \end{aligned}$$

$\square$

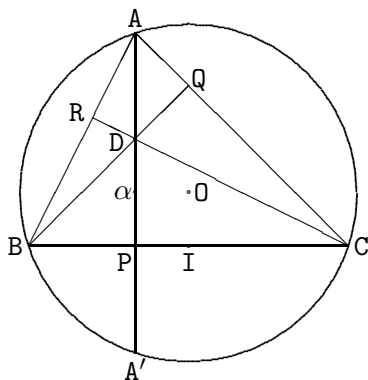
Soit un triangle ABC, P, Q, R les projections orthogonales de A, B, C sur les côtés (BC), (CA), (AB), I, J, K les milieux de (B, C), (C, A), (A, B), Γ le cercle circonscrit à ABC de centre O, point de concours des médiatrices de (B, C), (C, A), (A, B).

Proposition 8.34. (i) Les hauteurs (AP), (BQ) et (CR) concourent en un point D appelé orthocentre.

(ii) On a $\vec{OD} = \vec{OA} + \vec{OB} + \vec{OC}$.

(iii) Les symétriques A', B', C' de A, B, C relativement aux côtés (BC), (CA), (AB) sont sur le cercle circonscrit.

Démonstration. (i) Soit P, Q, R les pieds des hauteurs issues de A, B, C. Les droites (AP) et (BQ) ne sont pas parallèles car elles sont orthogonales à (BC) et (CA) qui ne sont pas parallèles. Soit D le point où elles se coupent. Il suffit de prouver que $\vec{DC} \cdot \vec{AB} = 0$: cela vient de l'identité du produit scalaire et de ce que $\vec{DA} \cdot \vec{BC} = \vec{DB} \cdot \vec{CA} = 0$.



(ii) Soit H le point défini par $\vec{OH} = \vec{OA} + \vec{OB} + \vec{OC}$. Montrons que $H = D$. On a $\vec{AH} = \vec{OH} - \vec{OA} = \vec{OB} + \vec{OC} = 2\vec{OI}$ où I est milieu de (B, C). Comme O est sur la médiatrice de (B, C), $\vec{AH} = 2\vec{OI}$ est orthogonal à (BC), donc H est sur la hauteur (AP). Il est de même sur (BQ) et (CR), donc $H = D$.

(iii) Soit p la projection orthogonale du plan sur la hauteur (AP). Il suffit de montrer que O est sur la médiatrice de (A, A'), donc que $p(O) = \alpha$ est milieu de (A, A'). L'égalité $\vec{OD} = \vec{OA} + \vec{OB} + \vec{OC}$ devient (car P est milieu de (D, A'))

$$\vec{\alpha D} = \vec{\alpha A} + 2\vec{\alpha P} = \vec{\alpha A} + \vec{\alpha D} + \vec{\alpha A'} \quad \text{d'où} \quad \vec{\alpha A} + \vec{\alpha A'} = \vec{0}$$

□

Proposition 8.35. (i) Il existe une homothétie de rapport $-\frac{1}{2}$ transformant A, B, C, D en I, J, K, O. Soit G son centre.

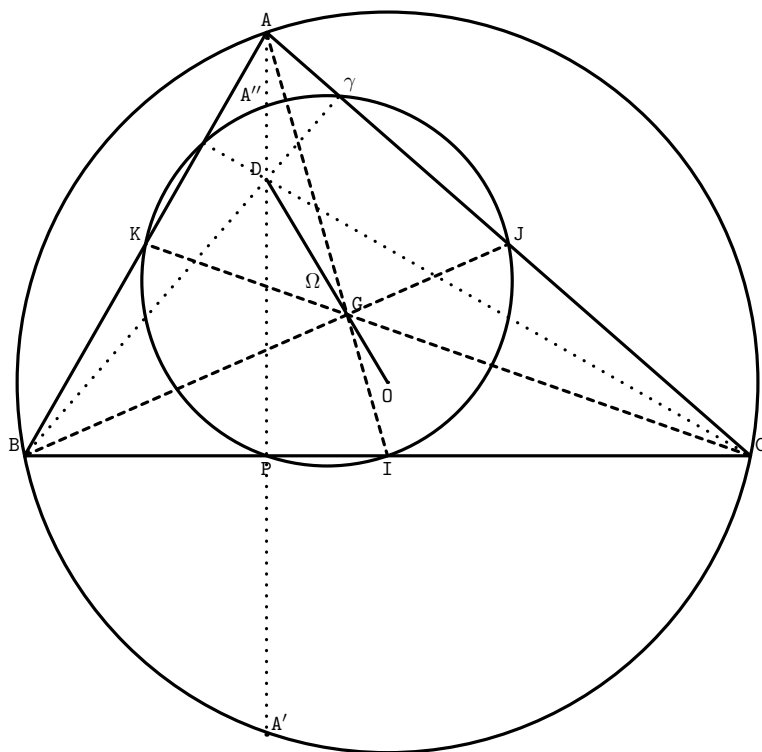
(ii) Les homothéties $h' = H_{G, -\frac{1}{2}}$ et $h'' = H_{D, \frac{1}{2}}$ transforment le cercle Γ circonscrit à ABC en un cercle γ passant par I, J, K, P, Q, R, les milieux A'', B'', C'' de (D, A), (D, B), (D, C).

Démonstration. (i) L'homothétie $H_{A, \frac{1}{2}}$ transforme B et C en K et J, donc $\vec{KJ} = \frac{1}{2}\vec{BC} = -\frac{1}{2}\vec{CB}$. On a donc une homothétie h' de rapport $-\frac{1}{2}$ transformant B, C en J, K (8.6, p. 233). Son centre G est commun aux médianes (BJ) et (CK). On a de même $\vec{KI} = -\frac{1}{2}\vec{CA}$ et $\vec{IJ} = -\frac{1}{2}\vec{AB}$, donc $h'(A) = I$. Les médianes (AI), (BJ), (CK) joignent chaque sommet à son image par h' , donc concourent en le centre G de h' . On dit que G est **centre de gravité** ou **isobarycentre** (voir chap.9) du triangle.

La hauteur (AP) se transforme par h' en la parallèle passant par I, qui est donc la médiatrice de (B, C). Les deux autres hauteurs se transforment de même en les deux autres médiatrices. Donc $h'(D) = O$, point de concours des médiatrices.

Le centre de gravité G, l'orthocentre D et le centre du cercle circonscrit O sont alignés car $h'(D) = O$. La droite portant ces points est la **droite d'Euler** du triangle.

(ii) L'homothétie h' transforme Γ de rayon R et de centre O en le cercle circonscrit γ à IJK de rayon $\frac{1}{2}R$ et de centre Ω , milieu de (D, O). Il en est de même de $h'' = H_{D, \frac{1}{2}}$. Les symétriques A', B', C' de D relativement aux côtés sont sur Γ (8.34) donc $h''(A') = P$, $h''(B') = Q$, $h''(C') = R$ sont sur $\gamma = h''(\Gamma)$. Enfin γ passe par $A'' = h''(A)$, $B'' = h''(B)$, $C'' = h''(C)$, milieux de (D, A), (D, B), (D, C). Le cercle γ s'appelle le **cercle des neuf points** du triangle ABC. $\square$



Définition 8.36. Soit un triangle ABC non rectangle d'orthocentre D. On a quatre points et quatre triangles, chaque point étant orthocentre du triangle formé par les trois autres. On a six droites : les côtés et les hauteurs d'un quelconque des quatre triangles. On dit que A, B, C, D sont en **situation orthocentrique**. Les pieds des hauteurs P, Q, R sont les mêmes pour les quatre triangles et sont sommets du **triangle orthique** de la configuration orthocentrique. Le cercle des neuf points est le même pour les quatre triangles.

8.2.3. Isométries

Définition 8.37. Une application affine $f: \mathcal{E} \rightarrow \mathcal{E}$ est une **isométrie** si l'application linéaire associée $\overline{f}: E \rightarrow E$ est une isométrie vectorielle. Autrement dit, l'ensemble des isométries de $\mathcal{E}$ est image réciproque du sous-groupe $\mathbf{O}(E) \subset \mathbf{GL}(E)$ par le morphisme $f \mapsto \overline{f}$, $\mathbf{GA}(\mathcal{E}) \rightarrow \mathbf{GL}(E)$. C'est un sous-groupe $\mathbf{O}(\mathcal{E}) \subset \mathbf{GA}(\mathcal{E})$. Une isométrie f est un **déplacement** (resp. un **antidéplacement**) si $\det \overline{f} = +1$ (resp. $\det \overline{f} = -1$). Les déplacements forment un sous-groupe distingué $\mathbf{SO}(\mathcal{E})$ d'indice 2 de $\mathbf{O}(\mathcal{E})$.

Une isométrie f conserve les distances : pour M et N d'images $f(M) = M'$ et $f(N) = N'$, on a $MN = \|\overrightarrow{MN}\| = \|\overline{f}(\overrightarrow{MN})\| = \|\overrightarrow{M'N'}\| = M'N'$. On montre plus généralement que toute application $\mathcal{E} \rightarrow \mathcal{E}$ (non supposée a priori affine) conservant les distances est une application affine, donc est une isométrie au sens précédent.

Proposition 8.38. Réflexions. Étant donné un hyperplan affine $\mathcal{H}$, il existe une unique isométrie, notée $\sigma_{\mathcal{H}}$, dont l'ensemble des points fixes est $\mathcal{H}$. Son application linéaire associée est la réflexion vectorielle autour de l'hyperplan vectoriel $\overline{\mathcal{H}}$. On dit que $\sigma_{\mathcal{H}}$ est la **réflexion** d'hyperplan $\mathcal{H}$.

Démonstration. Supposons qu'une telle isométrie σ existe. Alors $\overline{\sigma}$ laisse fixes les vecteurs de $\overline{\mathcal{H}}$ et n'est pas I_E car σ n'est pas une translation. Donc $\overline{\sigma}$ est nécessairement la réflexion vectorielle $s_{\overline{\mathcal{H}}}$. L'application linéaire associée $\overline{\sigma} = s_{\overline{\mathcal{H}}}$ et l'image de tout point de $\mathcal{H}$ étant imposés, σ est unique.

Inversement, soit un point O et un hyperplan vectoriel H . L'application affine d'application linéaire associée s_H et conservant O est une isométrie dont l'ensemble des points fixes est l'hyperplan affine $\mathcal{H} = O + H$. $\square$

La droite vectorielle $\delta = \overline{\mathcal{H}}^\perp$ est propre pour $s_{\overline{\mathcal{H}}}$ pour la valeur propre -1 . Soit un point $M \notin \mathcal{H}$ et H la projection orthogonale de M sur $\mathcal{H}$, intersection de la droite affine $M + \delta$ avec $\mathcal{H}$. L'image $M' = \sigma_{\mathcal{H}}(M)$ est donnée par $\overrightarrow{HM'} = s_{\overline{\mathcal{H}}}(\overrightarrow{HM}) = -\overrightarrow{HM}$. Ceci décrit l'action de la réflexion $\sigma_{\mathcal{H}}$.

Proposition 8.39. Les réflexions affines engendrent le groupe des isométries affines $\mathbf{O}(\mathcal{E})$.

Démonstration. Montrons que toute isométrie f est composée de réflexions.

1) Cas où f a un point fixe O . On vectorialise en O . Pour tout M , $\overrightarrow{Of(M)} = \overline{f}(\overrightarrow{OM})$. On sait qu'il existe des réflexions vectorielles $s_1, \dots, s_p$ d'hyperplans vectoriels $H_1, \dots, H_p$ telles que $\overline{f} = s_1 \circ \dots \circ s_p$. Pour tout i , soit σ_i la réflexion affine d'hyperplan $O + H_i$. Alors $\sigma_1 \circ \dots \circ \sigma_p = f$ car c'est l'application affine laissant O fixe d'application linéaire associée $\overline{\sigma_1} \circ \dots \circ \overline{\sigma_p} = \overline{f}$.

2) Cas où f n'a pas de point fixe. Soit $O \in \mathcal{E}$, $O' = f(O)$ et $\mathcal{H}$ l'**hyperplan médiateur** de (O, O') : hyperplan orthogonal à (OO') passant par le milieu de (O, O') . Alors $g = \sigma_{\mathcal{H}} \circ f$ est une isométrie laissant O fixe. Il suffit d'appliquer ce qui précède à g puis d'utiliser la relation $f = \sigma_{\mathcal{H}} \circ g$. $\square$

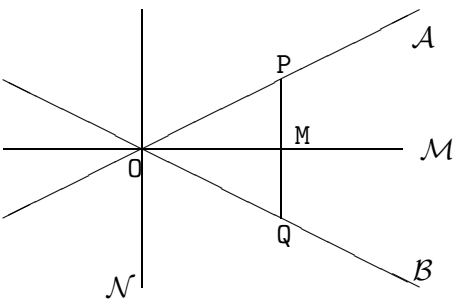
Comme en algèbre linéaire, cette décomposition d'une isométrie en composée de réflexions $f = \sigma_1 \circ \dots \circ \sigma_p$ n'est pas unique, mais la parité de p est un invariant. En effet, $\bar{f} = \bar{\sigma}_1 \circ \dots \circ \bar{\sigma}_p$, $\det(\bar{f}) = (-1)^p$, d'où p est pair pour un déplacement, impair pour un antidéplacement.

Proposition 8.40. Bissectrices. $\mathcal{A}, \mathcal{B}$ sont deux droites d'un plan, sécantes en O . Il existe deux réflexions qui les échangent, et leurs axes $\mathcal{M}, \mathcal{N}$ sont caractérisées par les propriétés :

- (i) $\mathcal{M}$ et $\mathcal{N}$ sont orthogonales,
- (ii) le faisceau $\mathcal{A}, \mathcal{B}, \mathcal{M}, \mathcal{N}$ est harmonique.

Démonstration. Supposons que $\mathcal{X}$ est axe d'une réflexion σ échangeant $\mathcal{A}$ et $\mathcal{B}$. Comme $\{O\} = \mathcal{A} \cap \mathcal{B}$, $\{\sigma(O)\} = \sigma(\mathcal{A}) \cap \sigma(\mathcal{B}) = \mathcal{B} \cap \mathcal{A} = \{O\}$ ce qui implique $O \in \mathcal{X}$. Soit $\vec{u}$ et $\vec{v}$ des vecteurs unitaires de $A = \overline{\mathcal{A}}$ et $B = \overline{\mathcal{B}}$. Nécessairement $\bar{\sigma}(\vec{u}) = \pm \vec{v}$, et donc deux possibilités pour $\bar{\sigma}$:

- ou bien $\bar{\sigma}$ échange $\vec{u}$ et $\vec{v}$, $\bar{\sigma}(\vec{u} + \vec{v}) = \vec{u} + \vec{v}$ et $\bar{\sigma}(\vec{u} - \vec{v}) = -(\vec{u} - \vec{v})$, alors $\mathcal{X}$ est la droite $\mathcal{M}$ passant par O de vecteur directeur $\vec{u} + \vec{v}$, non nul car $A \neq B$,
- ou bien $\bar{\sigma}$ échange $\vec{u}$ et $-\vec{v}$, $\bar{\sigma}(\vec{u} - \vec{v}) = \vec{u} - \vec{v}$ et $\bar{\sigma}(\vec{u} + \vec{v}) = -(\vec{u} + \vec{v})$, alors $\mathcal{X}$ est la droite $\mathcal{N}$ passant par O de vecteur directeur $\vec{u} - \vec{v}$, non nul car $A \neq B$.



Inversement, les réflexions d'axes $\mathcal{M}$ et $\mathcal{N}$ conservent O et échangent les droites vectorielles A et B , donc échangent les droites affines $\mathcal{A}$ et $\mathcal{B}$. Ce sont les deux réflexions annoncées.

Comme $(\vec{u} + \vec{v}) \cdot (\vec{u} - \vec{v}) = \|\vec{u}\|^2 - \|\vec{v}\|^2 = 0$, $\mathcal{M}$ et $\mathcal{N}$ sont orthogonales.

Soit $P \in \mathcal{A}$, $Q = \sigma_{\mathcal{M}}(P)$. Alors (PQ) est parallèle à $\mathcal{N}$ et coupe $\mathcal{M}$ en M milieu de (P, Q) . On en déduit que $[\mathcal{A}, \mathcal{B}, \mathcal{M}, \mathcal{N}] = -1$ (8.27, p. 218)

Réciproquement, soit $\mathcal{M}, \mathcal{N}$ orthogonales en O telles que $[\mathcal{A}, \mathcal{B}, \mathcal{M}, \mathcal{N}] = -1$. Pour tout $P \in \mathcal{A}$, la parallèle à $\mathcal{N}$ passant par P coupe $\mathcal{B}$ en Q et $\mathcal{M}$ en M tels que M soit milieu de (P, Q) . Donc $Q = \sigma_{\mathcal{M}}(P)$ et $\sigma_{\mathcal{M}}$ échange bien $\mathcal{A}$ et $\mathcal{B}$.

On dit que $\mathcal{M}, \mathcal{N}$ sont les **bissectrices** de $(\widehat{\mathcal{A}, \mathcal{B}})$. $\square$

8.2.4. Classification des isométries

À une isométrie affine f , on associe le sous-espace $\text{Fix}(\overline{f}) = E_1$ des vecteurs fixes par $\overline{f}$, sous-espace propre de $\overline{f}$ pour la valeur 1 (réduit à $\{\vec{0}\}$ si 1 non valeur propre). Deux cas se présentent :

- ou bien f est à points fixes : l'ensemble $\text{Fix}(f)$ des points fixes par f est non vide,
- ou bien f est sans points fixes : aucun point de $\mathcal{E}$ n'est fixe par f (cas des translations par exemple).

Le cas où f a un point fixe O est plus facile à imaginer. Par vectorialisation en O , f opère comme $\overline{f}$: pour tout M , $\overrightarrow{Of(M)} = \overline{f}(\overrightarrow{OM})$. Le b) de la proposition suivante montre comment f opère si elle est sans point fixe :

Proposition 8.41. a) Si f est à points fixes, l'ensemble $\text{Fix}(f)$ est un sous-espace affine de direction E_1 et tout sous-espace de direction $E_1^\perp$ est stable.

b) Si f est sans point fixe, il existe un unique vecteur $\vec{v}$ et une unique isométrie affine à points fixes f_1 tels que $f = T_{\vec{v}} \circ f_1 = f_1 \circ T_{\vec{v}}$; de plus, $\vec{v} \in E_1$ et la restriction de f à $\text{Fix}(f_1)$ est la translation $T_{\vec{v}}$.

c) L'entier $n - \dim E_1$ est pair ou impair selon que f est un déplacement ou un antidéplacement.

Démonstration. On sait (voir la proposition 7.9, p. 185) que $E_1 = \text{Ker}(I_E - \overline{f})$ et $\text{Im}(I_E - \overline{f})$ sont supplémentaires orthogonaux.

a) Supposons f à points fixes. Soit $O \in \text{Fix}(f)$. Pour tout M , $\overrightarrow{Of(M)} = \overline{f}(\overrightarrow{OM})$, donc $f(M) = M$ si et seulement si $\overline{f}(\overrightarrow{OM}) = \overrightarrow{OM}$, i.e. $\overrightarrow{OM} \in E_1$, i.e. $M \in O + E_1$. Ainsi, $\text{Fix}(f) = O + E_1$ est sous-espace affine de direction E_1 .

Un sous-espace $\mathcal{F}$ de direction $E_1^\perp$ coupe $\text{Fix}(f)$ en un point H fixe (p. 210). On a $f(\mathcal{F}) = f(H) + \overline{f}(E_1^\perp) = \mathcal{F}$ car $f(H) = H$ et $\overline{f}(E_1^\perp) = E_1^\perp$.

b) Supposons f sans point fixe. On est ramené à chercher $\vec{u} = -\vec{v}$ vérifiant deux conditions :

- (1) d'une part $T_{\vec{u}}$ doit commuter avec f ,
- (2) d'autre part $f_1 = T_{\vec{u}} \circ f$ doit être à points fixes.

1) La translation $T_{\vec{u}}$ commute avec f si et seulement si $T_{\vec{u}} \circ f = f \circ T_{\vec{u}}$, donc si $T_{\vec{u}} = f \circ T_{\vec{u}} \circ f^{-1}$. La conjuguée de $T_{\vec{u}}$ par f étant $f \circ T_{\vec{u}} \circ f^{-1} = T_{\overline{f}(\vec{u})}$ (p. 213), il faut et il suffit que $\vec{u} = \overline{f}(\vec{u})$, donc que $\vec{u} \in E_1$.

2) Soit A fixé dans $\mathcal{E}$. Pour tout point $O \in \mathcal{E}$, on a

$$(T_{\vec{u}} \circ f)(O) = f(O) + \vec{u} = f(A) + \overline{f}(\overrightarrow{AO}) + \vec{u} = A + \overrightarrow{Af(A)} + \overline{f}(\overrightarrow{AO}) + \vec{u}$$

Pour que O soit fixe par $T_{\vec{u}} \circ f$, il faut et il suffit que $O = A + \overrightarrow{Af(A)} + \overline{f}(\overrightarrow{AO}) + \vec{u}$, donc que $\overrightarrow{Af(A)} = (I_E - \overline{f})(\overrightarrow{AO}) - \vec{u}$. Donc $T_{\vec{u}} \circ f$ aura un point fixe si et seulement si $\overrightarrow{Af(A)}$ est somme de $\vec{v} = -\vec{u}$ appartenant à E_1 et d'un vecteur du type $(I_E - \overline{f})(\overrightarrow{AO})$ de $\text{Im}(I_E - \overline{f}) = E_1^\perp$.

En utilisant la décomposition en somme directe orthogonale $E = E_1 \oplus {}^\perp E_1$, on voit que $\overrightarrow{Af(A)}$ se décompose de façon unique en somme d'un vecteur $-\vec{u} \in E_1$ et d'un vecteur de $E_1^\perp = \text{Im}(I_E - \bar{f})$.

c) Le polynôme caractéristique de $\bar{f}$ est $\chi_{\bar{f}}(X) = (X - 1)^p(X + 1)^qQ(X)$ où $p = \dim E_1$ et Q n'a pas de racine réelle, donc est de degré pair. Alors $q = n - p - \deg Q$ a la même parité que $n - p$. $\square$

Les isométries sans point fixe, du type $f = T_{\vec{v}} \circ f_1 = f_1 \circ T_{\vec{v}}$ ($\vec{v} \neq \vec{0}$, $\text{Fix}(f_1) \neq \phi$) sont les **isométries glissées**, $\vec{v} \in E_1$ est le **vecteur de glissement**. Les isométries à points fixes sont à glissement nul. On classe les isométries affines dans un tableau selon qu'elles sont ou non glissées (2 cas) et selon la valeur de $\dim E_1 \in \{0, 1, \dots, n\}$ ($n + 1$ cas). Voici quelques cas particuliers :

- **Cas où 1 n'est pas valeur propre.** En ce cas $\dim E_1 = 0$ et $I_E - \bar{f}$ est bijective de E sur lui-même. Soit A un point de $\mathcal{E}$. Pour que $0 \in \mathcal{E}$ soit fixe, il faut et il suffit que $\overrightarrow{Af(A)} = (I_E - \bar{f})(\overrightarrow{AO})$. Comme $I_E - \bar{f}$ est bijective, il existe un unique point fixe $0 = A + (I_E - \bar{f})^{-1}(\overrightarrow{Af(A)})$. Un cas sur les $2n + 2$ recensés ne se produit pas : il n'existe pas d'isométrie glissée si $\dim E_1 = 0$.
- **Cas où $\dim E_1 = n$.** On a $\bar{f} = I_E$,
 - ou bien $f = I_{\mathcal{E}}$ est l'identité,
 - ou bien $f = T_{\vec{v}}$ est une translation.
- **Cas où $\dim E_1 = n - 1$.** Alors E_1 est un hyperplan et $\bar{f}$ est la réflexion autour de l'hyperplan E_1 .
 - ou bien f est une réflexion d'hyperplan affine $\mathcal{H}$ de direction E_1 ,
 - ou bien f est composée commutative d'une réflexion f_1 d'hyperplan $\mathcal{H}$ de direction $\mathcal{H} = E_1$ et d'une translation $T_{\vec{v}}$ où $\vec{v} \in E_1$. On dit que f est la **réflexion glissée** d'hyperplan $\mathcal{H}$ et de vecteur $\vec{v}$.

En dimension 2, on a cinq types d'isométries :

- $\dim E_1 = 2$, les déplacements identité et translations,
- $\dim E_1 = 1$, les antidéplacements réflexions et réflexions glissées autour d'une droite,
- $\dim E_1 = 0$, les déplacements rotations dotées d'un centre et d'un angle.

Rappelons qu'étant donné deux droites $\mathcal{A}$ et $\mathcal{B}$ d'un plan, la composée de réflexions $\sigma_{\mathcal{B}} \circ \sigma_{\mathcal{A}}$ est le déplacement :

- $T_{\vec{v}}$ si $\mathcal{A}$ et $\mathcal{B}$ sont parallèles, $\mathcal{B}$ étant image de $\mathcal{A}$ par la translation de vecteur $\frac{1}{2}\vec{v}$ orthogonal à ces droites,
- $\text{rot}(0, \theta)$ si $\mathcal{A}$ et $\mathcal{B}$ sont sécantes en O , $\mathcal{B}$ étant image de $\mathcal{A}$ par la rotation de centre O et d'angle $\frac{1}{2}\theta$.

8.2.5. Isométries en dimension 3

On a sept types d'isométries :

- $\dim E_1 = 3$, les déplacements identité et translations,
- $\dim E_1 = 2$, les antidéplacements réflexions et réflexions glissées autour d'un plan,
- $\dim E_1 = 1$, les déplacements rotations et **vissages** (rotations glissées) autour d'un axe,
- $\dim E_1 = 0$, les antidéplacements isométries à point fixe unique, dont les symétries centrales sont un cas particulier.

L'espace E de dimension 3 étant orienté, les données nécessaires pour définir une rotation f sont :

- un axe $\mathcal{D}$ muni d'une **orientation** donnée par un de ses deux vecteurs unitaires directeurs $\vec{w}$,
- un angle θ défini modulo 2π .

La donnée du vecteur $\vec{w}$ déterminant l'orientation de $\mathcal{D}$ permet de définir une orientation sur le plan vectoriel orthogonal $\overline{\mathcal{D}}^\perp = P$: une base $(\vec{u}, \vec{v})$ orthonormale de P sera dite **directe** si la base orthonormale $(\vec{u}, \vec{v}, \vec{w})$ de E est directe. Pour tout $\Omega \in \mathcal{D}$, le plan $\mathcal{P}_\Omega = \Omega + P$ est stable et f induit sur $\mathcal{P}_\Omega$ la rotation de centre Ω et d'angle θ .

Si on remplace $\vec{w}$ par $-\vec{w}$ et θ par $-\theta$, on obtient la même rotation¹.

Un **vissage** d'axe $\mathcal{D}$ est composé d'une rotation d'axe $\mathcal{D}$ et d'une translation $T_{\vec{v}}$ où $\vec{v} \in \overline{\mathcal{D}}$. Le vecteur de glissement peut servir à orienter $\mathcal{D}$.

Définition 8.42. Soit une droite $\mathcal{D}$. La rotation d'axe $\mathcal{D}$ et d'angle π est le **retournement** (ou symétrie axiale) d'axe $\mathcal{D}$, noté $\rho_{\mathcal{D}}$. Comme π et $-\pi$ sont congrus modulo 2π , il n'est pas besoin en ce cas d'une orientation sur $\mathcal{D}$.

Proposition 8.43. En dimension 3, tout déplacement f est un composé de deux retournements.

Démonstration. Soit f un déplacement. Alors $\overline{f}$ est une rotation vectorielle d'axe une droite vectorielle D orientée et d'angle θ . Soit D_1 une droite vectorielle orthogonale à D , D_2 l'image de D_1 par la rotation vectorielle d'axe D et d'angle $\frac{\theta}{2}$, r_1 et r_2 les retournements vectoriels d'axes D_1 et D_2 , on sait alors que $\overline{f} = r_2 \circ r_1$.

a) Cas où $f = T_{\vec{v}}$ est une translation. Alors $\overline{f} = I_E$. Ce qui précède est encore valable en prenant $D = \text{Vect}(\vec{v})$, $D_1 = D_2 = \delta$ orthogonale à D . Soit O_1 un point, $O_2 = O_1 + \frac{1}{2}\vec{v}$, $\mathcal{D}_1 = O_1 + \delta$, $\mathcal{D}_2 = O_2 + \delta$, ρ_1 et ρ_2 les retournements affines

1. C'est une difficulté de ne pas pouvoir choisir **canoniquement** l'orientation de $\mathcal{D}$. Alors qu'en géométrie plane, une rotation est définie de façon unique par son centre et son angle, il n'en est pas ainsi en dimension 3.

d'axes $\mathcal{D}_1$ et $\mathcal{D}_2$. Alors $\rho_2 \circ \rho_1$ a $r_2 \circ r_1 = I_E$ pour application linéaire associée, donc est une translation. L'image de O_1 est $\rho_2(O_1) = O_1 + 2\overrightarrow{O_1O_2} = T_{\vec{v}}(O_1)$, donc $\rho_2 \circ \rho_1 = T_{\vec{v}} = f$.

b) Cas où f est une rotation d'axe $\mathcal{D}$. Soit un point $O \in \mathcal{D}$, les droites affines $\mathcal{D}_1 = O + D_1$ et $\mathcal{D}_2 = O + D_2$, ρ_1 et ρ_2 les retournements affines autour de $\mathcal{D}_1$ et $\mathcal{D}_2$. Alors $\overline{\rho_2} = r_2$, $\overline{\rho_1} = r_1$, donc $\rho_2 \circ \rho_1$ laisse $O = f(O)$ fixe et admet $r_2 \circ r_1 = \overline{f}$ comme application linéaire associée. On a donc $f = \rho_2 \circ \rho_1$ (p. 209).

c) Cas où f est un vissage : $f = T_{\vec{v}} \circ g = g \circ T_{\vec{v}}$ où g est une rotation d'axe $\mathcal{D}$ ayant $\vec{v}$ pour vecteur directeur. Soit $\mathcal{D}_3$ une droite coupant $\mathcal{D}$ orthogonalement et ρ_3 le retournement d'axe $\mathcal{D}_3$.

Par b), on peut écrire $g = \rho_3 \circ \rho_1$ où ρ_1 est le retournement dont l'axe $\mathcal{D}_1$ se déduit de $\mathcal{D}_3$ par la rotation d'axe $\mathcal{D}$ et d'angle $-\frac{\theta}{2}$.

Par a), on peut écrire $T_{\vec{v}} = \rho_2 \circ \rho_3$ où ρ_2 est le retournement d'axe $\mathcal{D}_2$ où $\mathcal{D}_2$ se déduit de $\mathcal{D}_3$ par la translation $T_{\frac{1}{2}\vec{v}}$.

On en déduit $f = T_{\vec{v}} \circ g = (\rho_2 \circ \rho_3) \circ (\rho_3 \circ \rho_1) = \rho_2 \circ \rho_1$. $\square$

Les isométries ayant un unique point fixe O sont des antidéplacements. La symétrie centrale S_O de centre O (homothétie $H_{O,-1}$) est un exemple. Soit f une isométrie distincte de S_O de seul point fixe O . Alors $g = S_O \circ f = f \circ S_O$ est un déplacement ayant O pour point fixe ; g est distinct de l'identité car f est différent de S_O . Par la classification des déplacements, g est une rotation d'axe $\mathcal{D}$ passant par O . Ainsi f se décompose de façon unique en $f = S_O \circ g = g \circ S_O$ où g est une rotation d'axe passant par O .

8.2.6. Similitudes en dimension 2

Définition 8.44. Soit $\mathcal{P}$ un plan affine euclidien orienté de plan vectoriel associé P . Étant donné $k > 0$, on appelle **similitude de rapport k** toute application affine f ayant la propriété :

pour tout couple (M, N) de points de transformés $M' = f(M)$ et $N' = f(N)$, on a $M'N' = k \times MN$. Autrement dit, f multiplie les longueurs par k .

- Les isométries sont les similitudes de rapport 1.
- Une homothétie de rapport $\pm k$ est une similitude de rapport k .

Lemme 8.45. Soit f une similitude de rapport k . Alors f est composée (d'une infinité de façons) d'une homothétie h de rapport $\pm k$ et d'une isométrie u . On a $\det \overline{f} = \pm k^2$.

Démonstration. Soit h une homothétie de rapport $\pm k$. Alors $f \circ h^{-1} = u$ est une isométrie et $f = u \circ h$. On a

$$\overline{f} = \overline{u} \circ \overline{h} = \overline{u} \circ (\pm k I_P),$$

d'où $\det \overline{f} = \det \overline{u} \times \det(\pm k I_P) = \det \overline{u} \times k^2$. Comme $\det \overline{u} = \pm 1$ selon que u est directe ou indirecte, $\det \overline{f} = \pm k^2$. $\square$

Définition 8.46. On dit que f est une **similitude directe** si $\det \overline{f} > 0$, **indirecte** si $\det \overline{f} < 0$.

- Une isométrie est une similitude directe ou indirecte selon qu'elle est directe ou indirecte comme isométrie.
- Les homothéties sont des similitudes directes.

L'énoncé suivant est immédiat : Les similitudes forment un groupe $\text{Sim}(\mathcal{P})$. L'application associant à une similitude son rapport $f \mapsto k(f)$ est un morphisme $\text{Sim}(\mathcal{P}) \rightarrow \mathbb{R}^{\times+}$ dont le noyau est le sous-groupe des isométries $\mathbf{O}(\mathcal{P})$. Les similitudes directes forment un sous-groupe $\text{Sim}^+(\mathcal{P})$ d'indice 2.

Proposition 8.47. Soit une similitude f de rapport $k \neq 1$. Alors :

- f a un unique point fixe $\mathbf{O}$,
- il existe un unique couple (h, u) d'une homothétie h de rapport positif et d'une isométrie u tels que $f = h \circ u = u \circ h$.
L'homothétie h est de rapport k et de centre $\mathbf{O}$. L'isométrie est
 - une rotation de centre $\mathbf{O}$ si f est directe,
 - une réflexion d'axe passant par $\mathbf{O}$ si f est indirecte.
- f conserve les angles ou les change en leurs opposés selon que f est directe ou indirecte.

Démonstration. a) Soit $A \in \mathcal{P}$. Pour tout $M \in \mathcal{P}$, on a

$$\begin{aligned} \overrightarrow{Mf(M)} &= \overrightarrow{Mf(A)} + \overrightarrow{f(A)f(M)} = \overrightarrow{Mf(A)} + \overline{f}(\overrightarrow{AM}) \\ &= \overrightarrow{MA} + \overrightarrow{Af(A)} - \overline{f}(\overrightarrow{MA}) = \overrightarrow{Af(A)} - (\mathbf{I}_P - \overline{f})(\overrightarrow{AM}) \end{aligned}$$

Un point $\mathbf{O}$ est fixe si et seulement si $\overrightarrow{Af(A)} = (\mathbf{I}_P - \overline{f})(\overrightarrow{AO})$. Comme $k \neq 1$, 1 n'est pas valeur propre de $\overline{f}$, donc $\mathbf{I}_P - \overline{f}$ est bijective. L'unique point fixe $\mathbf{O}$ est donc défini par $\overrightarrow{AO} = (\mathbf{I}_P - \overline{f})^{-1}(\overrightarrow{Af(A)})$.

b) Supposons qu'il existe un couple (h, u) d'une homothétie positive h et d'une rotation u tel que $f = h \circ u = u \circ h$. Alors $k(f) = k(h)k(u) = k(h)$, donc l'homothétie positive h est nécessairement de rapport k . Soit Ω le centre de h . Alors $h = u \circ h \circ u^{-1}$ est de centre $\Omega = u(\Omega)$ (p. 213). Donc Ω est fixe par h et u , donc par f , d'où $\Omega = \mathbf{O}$ est l'unique point fixe de f . Nécessairement, $h = H_{\mathbf{O},k}$.

Inversement, $f \circ H_{\mathbf{O},k} \circ f^{-1}$ est l'homothétie de même rapport et de centre $f(\mathbf{O}) = \mathbf{O}$. Donc $H_{\mathbf{O},k}$ commute avec f . L'application $u = f \circ H_{\mathbf{O},k}^{-1} = H_{\mathbf{O},k}^{-1} \circ f$ est une isométrie commutant avec $H_{\mathbf{O},k}$.

On a $\det \bar{f} = k^2 \det \bar{u}$, donc f est directe ou indirecte selon que u est un déplacement ou un antidéplacement. Comme $u(0) = 0$, u est une rotation de centre 0 si f est directe, une réflexion d'axe passant par 0 si f est indirecte.

c) Comme les homothéties conservent les angles, $f = H_{0,k} \circ u$ et u ont même action sur les angles.

Si f est directe, on dit que f est la **similitude directe de centre 0, de rapport k et d'angle θ** , elle est notée $\text{sim}(0, k, \theta)$. $\square$

EXERCICES

Exercice 8.5. Soit dans un plan affine $\mathcal{P}$ trois droites distinctes $\mathcal{A}, \mathcal{B}, \mathcal{C}$ et deux droites vectorielles F et G distinctes des directions de $\mathcal{A}, \mathcal{B}, \mathcal{C}$. Au point $M \in \mathcal{A}$, on associe le point $N = f(M)$ où la droite $M + F$ coupe $\mathcal{B}$ et le point $P = g(N) = h(M)$ où la droite $N + G$ coupe $\mathcal{C}$. Les droites (MN) et (MP) sont donc de directions fixes F et G si M décrit $\mathcal{A}$. La question est : la direction de la droite (MP) reste-t-elle fixe ?

1) Montrer que c'est le cas si $\mathcal{A}, \mathcal{B}, \mathcal{C}$ sont concourantes ou parallèles.

2) On suppose que les droites $\mathcal{A} = (BC)$, $\mathcal{B} = (CA)$ et $\mathcal{C} = (AB)$ forment un véritable triangle de sommets ABC . Montrer que (MP) a une direction fixe si et seulement si $F = G$.

Exercice 8.6. Dans un plan affine, soit A_1, B_1 distincts, $\lambda \in \mathbb{R}$, distinct de 0, 1, -1, A_2, B_2 tels que $\overrightarrow{A_2 B_2} = \lambda \overrightarrow{A_1 B_1}$.

1) Montrer qu'il existe exactement deux homothéties H' et H'' transformant la paire de points $\{A_1, B_1\}$ en la paire de points $\{A_2, B_2\}$.

2) Soit O' et O'' les centres de H' et H'' , I_1 et I_2 les milieux de (A_1, B_1) et (A_2, B_2) . Montrer que O', O'', I_1, I_2 sont alignés et que $[O', O'', I_1, I_2] = -1$.

Exercice 8.7. Soit un triangle ABC , P, Q, R les milieux de (B, C) , (C, A) , (A, B) , A', B', C' des points de (BC) , (CA) , (AB) , A'', B'', C'' leurs symétriques respectifs par rapport à P, Q, R .

Montrer que $s(AA'), (BB'), (CC')$ sont concourantes ou parallèles si et seulement s'il en est ainsi pour $(AA''), (BB''), (CC'')$.

Montrer que A', B', C' sont alignés si et seulement s'il en est ainsi pour A'', B'', C'' .

Exercice 8.8. Soit ABC un triangle, A', B', C' des points de (BC) , (CA) , (AB) . Les parallèles à (AC) passant par A' , (BC) passant par B' , (CA) passant par C' coupent respectivement (CA) en B'' , (AB) en C'' , (BC) en A'' .

Montrer que (AA') , (BB') , (CC') sont concourantes ou parallèles si et seulement s'il en est ainsi pour (AA'') , (BB'') , (CC'') .

Montrer que A', B', C' sont alignés si et seulement s'il en est ainsi pour A'', B'', C'' .

Exercice 8.9. Soit un triangle ABC , une droite ne passant ni par les sommets, ni par les milieux des côtés coupant (BC) , (CA) , (AB) en A', B', C' . Soient A'', B'', C'' les conjugués harmoniques de A', B', C' par rapport à (B, C) , (C, A) , (A, B) . Donner les relations d'alignement entre $A', B', C', A'', B'', C''$ et de concours ou parallélisme entre (AA') , (BB') , (CC') , (AA'') , (BB'') , (CC'') .

Exercice 8.10. Quadrilatère complet. Soit un triangle ABC , A', B', C' sur (BC) , (CA) , (AB) alignés. On suppose que les droites (AA') , (BB') (resp. (BB') , (CC') , resp. (CC') , (AA')) sont sécantes en R (resp. P , resp. Q). Montrer que

$$[A, A', Q, R] = [B, B', R, P] = [C, C', P, Q] = -1$$

Exercice 8.11. Soit A, B deux points distincts, I le milieu de (A, B) et $\mathcal{D}$ la droite (AB) . À tout $M \in \mathcal{D}$, on associe son conjugué harmonique N relativement à A et B .

1) Montrer que pour que M soit entre A et B , il faut et il suffit que N ne soit pas entre A et B .

2) Montrer que l'application $M \mapsto N$ induit une bijection entre le segment $[I, B]$ privé de I et la demi-droite limitée par B du sens de $\overrightarrow{AB}$.

Exercice 8.12. Soit $\mathcal{P}$ un plan euclidien orienté, O un point de $\mathcal{P}$, θ un angle. Montrer que la conjuguée de la rotation $f = \text{rot}(O, \theta)$ de centre O et d'angle θ par une isométrie g est la rotation de centre $g(O)$ et d'angle θ ou $-\theta$ selon que g est un déplacement ou un antidéplacement.

Exercice 8.13. Soit une rotation f d'axe une droite $\mathcal{D}$ orientée par un vecteur unitaire $\vec{k}$ et d'angle θ . Déterminer l'ensemble des isométries commutant avec f .

Exercice 8.14. Soit dans un plan euclidien $\mathcal{P}$ des points A, B, A', B' tels que $A \neq B$ et $A' \neq B'$. Montrer qu'il existe une unique similitude directe s telle que $s(A) = A'$ et $s(B) = B'$.

PROBLÈMES

Les corrigés de ces problèmes sont disponibles sur le site de Dunod : www.dunod.com.

8.1. PROBLÈME

Soit un espace affine $\mathcal{E}$. On donne trois vecteurs colinéaires distincts de même sens $\vec{v}_1, \vec{v}_2, \vec{v}_3$, trois points A_1, A_2, A_3 de $\mathcal{E}$ et les trois points B_1, B_2, B_3 où $B_i = A_i + \vec{v}_i$. Pour deux indices distincts i, j de $\{1, 2, 3\}$, on définit $\lambda_{j,i} > 0$ par $\vec{v}_j = \lambda_{j,i} \vec{v}_i$. Soit $H'_{j,i}$ (resp. $H''_{j,i}$) l'homothétie de rapport $\lambda_{j,i}$ (resp. $-\lambda_{j,i}$) de centre $O'_{j,i}$ (resp. $O''_{j,i}$) transformant A_i, B_i en A_j, B_j (resp. B_j, A_j) (8.6, p. 233).

1) Montrer que $O'_{i,j} = O'_{j,i}$ et $O''_{i,j} = O''_{j,i}$. On notera ces points O'_k et O''_k où k est le troisième indice de $\{1, 2, 3\}$ distinct de i et j .

2) Quels alignements a-t-on entre les six points $O'_1, O'_2, O'_3, O''_1, O''_2, O''_3$?

3) Faire une figure où seront représentés les six points $O'_1, O'_2, O'_3, O''_1, O''_2, O''_3$ et les milieux I_1 de (A_1, B_1) , I_2 de (A_2, B_2) , I_3 de (A_3, B_3)

8.2. PROBLÈME

Soit $\mathcal{E}$ affine de dimension au moins 2, $f: \mathcal{E} \rightarrow \mathcal{E}, M \mapsto f(M) = M_1$, une application (non supposée a priori affine) telle que, pour tout couple de points M, N , $\overrightarrow{MN}$ et $\overrightarrow{M_1N_1}$ soient colinéaires. Il s'agit de montrer que si f n'est pas constante, alors c'est une homothétie ou une translation.

1) Montrer que si f n'est pas une application constante, alors elle est injective, ce qu'on supposera dans la suite.

2) Montrer que si f a un point fixe O , c'est une homothétie de centre O .

3) Montrer que si f n'a pas de point fixe, c'est une translation.

8.3. PROBLÈME

Dans un plan affine $\mathcal{P}$ de plan vectoriel associé P , on donne une droite affine $\mathcal{D}$ de direction $\overline{\mathcal{D}} = D$. Soit $G_{\mathcal{D}}$ le groupe des $f \in \mathbf{GA}(\mathcal{E})$ induisant l'identité sur $\mathcal{D}$.

1) Soit $O \in \mathcal{D}$. Montrer que pour qu'une application affine f appartienne à $G_{\mathcal{D}}$, il faut et il suffit que $f(O) = O$ et que l'application linéaire $\overline{f}$ soit une dilatation ou une transvection d'axe D .

2) On dit alors que l'application affine f est une *dilatation* d'axe $\mathcal{D}$ si l'application linéaire $\overline{f}$ est une dilatation. Décrire alors l'action de f .

3) On dit que f est une *transvection* affine si l'application linéaire $\overline{f}$ est une transvection. Montrer que les droites parallèles à $\mathcal{D}$ sont stables et que la restriction de f à chacune de ces droites est une translation.

4) Montrer que l'ensemble des transvections de $G_{\mathcal{D}}$ auquel on adjoint l'identité est un sous-groupe commutatif distingué de $G_{\mathcal{D}}$.

5) Une application affine f dont l'application linéaire associée est une transvection est-elle toujours une transvection affine ?

8.4. PROBLÈME

Soit f une application affine d'un espace affine $\mathcal{E}$ dans lui-même.

1) On forme les sous-espaces vectoriels $E_1 = \text{Ker}(I_E - \bar{f})$ et $F_1 = \text{Im}(I_E - \bar{f})$. Soit $A \in \mathcal{E}$. Montrer que $\text{Fix}(f)$ est non vide si et seulement si $\overrightarrow{Af(A)} \in F_1$. En déduire que l'ensemble des points fixes $\text{Fix}(f)$ est réduit à un unique point si et seulement si 1 n'est pas valeur propre de $\bar{f}$.

2) On dira que f **admet une décomposition canonique** si et seulement s'il existe un unique couple $(\vec{v}, f_1)$ où f_1 est affine avec $\text{Fix}(f_1)$ non vide tel que $f = T_{\vec{v}} \circ f_1 = f_1 \circ T_{\vec{v}}$. Montrer qu'il en est ainsi si et seulement si E_1 et F_1 sont en somme directe.

3) Soit $\mathcal{E}$ un plan affine, $\mathcal{D}$ une droite, τ une transvection d'axe $\mathcal{D}$ (pb. 8.3), $\vec{v} \in E$. La composée $f = T_{\vec{v}} \circ \tau$ a-t-elle une décomposition canonique ?

8.5. PROBLÈME (groupes diédraux et polygones réguliers)

Soit $\mathcal{P}$ un plan affine euclidien de plan vectoriel associé P .

1) Soit G un groupe d'isométries ne contenant pas de translation, $G^+ = G \cap \text{SO}(\mathcal{P})$ le sous-groupe des déplacements de G . Montrer qu'il existe un point $0 \in \mathcal{P}$ tel que, pour tout $f \in G$, on ait $f(0) = 0$ (pour f, g dans G^+ , considérer $f \circ g \circ f^{-1} \circ g^{-1}$).

2) Désormais, on suppose G de cardinal fini non réduit à G^+ . Montrer que $\text{Card}(G) = 2n$ est pair et que G^+ est cyclique de cardinal n .

3) Montrer que l'ensemble des antidéplacements de G est formé de n réflexions σ_p d'axes $\mathcal{D}_p$ passant par 0 ($0 \leq p \leq n-1$), avec $(\widehat{\mathcal{D}_0}, \widehat{\mathcal{D}_p}) = \frac{p\pi}{n}$.

4) Soit M distinct de 0 . Montrer que l'orbite $G(M)$ de M par G est formée

- de $2n$ points distincts si M n'est sur aucune des droites $\mathcal{D}_p$,
- de n points distincts, coïncidant avec l'orbite $G^+(M)$ de M pour G^+ si M est sur l'une des droites $\mathcal{D}_p$.

5) Montrer que le centre de G est $\{I_{\mathcal{P}}\}$ si n impair, $\{I_{\mathcal{P}}, S_0\}$ si n pair.

6) On fait opérer G sur l'ensemble Σ des réflexions de G par conjugaison : à $(f, \sigma) \in G \times \Sigma$, on associe $f \circ \sigma \circ f^{-1}$. Montrer que les actions de G et G^+ définissent les mêmes orbites.

En conclure que les réflexions de G se groupent en deux classes de conjugaison si n est pair, une seule si n est impair. Proposer une interprétation géométrique pour les polygones réguliers d'ordre n .

7) Montrer que si n est impair, tout sous-groupe distingué non trivial de G est contenu dans G^+ . Qu'en est-il si n est pair ?

SOLUTIONS DES EXERCICES

Solution 8.1. On suppose $\dim \mathcal{E} = 3$. Soit $\vec{v}_1$ et $\vec{v}_2$ non colinéaires, $\mathcal{D}_1$ une droite de vecteur directeur $\vec{v}_1$, $A_1 \in \mathcal{D}_1$. Si $A_2 \notin A_1 + \text{Vect}(\vec{v}_1, \vec{v}_2)$, la droite $\mathcal{D}_2 = A_2 + \mathbb{R}\vec{v}_2$ ne rencontre pas $\mathcal{D}_1$, mais ne lui est pas parallèle.

Solution 8.2. 1) Supposons $(A + F) \cap (B + G)$ non vide. Soit M un point de l'intersection. Alors $\vec{AM} \in F$ et $\vec{BM} \in G$, d'où $\vec{AB} = \vec{AM} - \vec{BM} \in F + G$. Inversement, supposons $\vec{AB} \in F + G$. Il existe $\vec{u} \in F$ et $\vec{v} \in G$ tels que $\vec{AB} = \vec{u} + \vec{v}$. Alors $M = B - \vec{v} \in B + G$. On a

$$\vec{AM} = \vec{AB} + \vec{BM} = (\vec{u} + \vec{v}) - \vec{v} = \vec{u} \in F$$

donc $M \in A + F$, d'où $(A + F) \cap (B + G) \neq \emptyset$.

2) Soit $G \subset E$ de dimension q , $H \subset G$ de dimension $p-1$, $\vec{\varepsilon}$ un vecteur non dans G et $F = H \oplus \mathbb{R}\vec{\varepsilon}$. Alors $F + G = G \oplus \mathbb{R}\vec{\varepsilon}$, de dimension $q+1 \leq n-1$ est strictement contenu dans E . Prenons A, B dans $\mathcal{E}$ tel que $\vec{AB} \notin G + \mathbb{R}\vec{\varepsilon}$. Alors $\mathcal{F} = A + F$ et $\mathcal{G} = B + G$ ne sont pas parallèles car $F \not\subset G$ et $\mathcal{F} \cap \mathcal{G} = \emptyset$ puisque $\vec{AB} \notin F + G$.

3) Supposons $\mathcal{F} = A + F$ et $\mathcal{G} = B + G$ d'intersection vide, $\dim \mathcal{G} = n-1$. Alors $\vec{AB} \notin F + G$ d'après a). Comme $\dim G = n-1$, ceci impose $G = F + G$, donc $F \subset G$. D'où $\mathcal{F}$ et $\mathcal{G}$ sont bien parallèles.

Solution 8.3. Supposons ces droites deux à deux coplanaires et montrons que nécessairement, $\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3$ ou bien sont dans un même plan, ou bien passent par un même point, ou bien ont même direction. Soit les plans $\mathcal{P}_1 = (\mathcal{D}_2, \mathcal{D}_3)$, $\mathcal{P}_2 = (\mathcal{D}_3, \mathcal{D}_1)$, $\mathcal{P}_3 = (\mathcal{D}_1, \mathcal{D}_2)$.

Cas où les plans ne sont pas distincts. Supposons $\mathcal{P}_1 = \mathcal{P}_2$. Ce plan contient les deux droites de $\mathcal{P}_3$, donc est aussi confondu avec $\mathcal{P}_3$. Les trois droites sont bien dans ce même plan.

Cas où les plans sont distincts. On a $\mathcal{P}_i \cap \mathcal{P}_j = \mathcal{D}_k$, i, j, k étant distincts dans $\{1, 2, 3\}$.

- Si $\mathcal{D}_2 \cap \mathcal{D}_3 = \{M\}$, alors $M \in \mathcal{P}_2 \cap \mathcal{P}_3 = \mathcal{D}_1$, les trois droites passent par M .
- Si $\mathcal{D}_2$ et $\mathcal{D}_3$ sont de direction δ , droite vectorielle contenue dans chacun des $\overline{\mathcal{P}}_i$, $\delta = \overline{\mathcal{P}}_2 \cap \overline{\mathcal{P}}_3 = \overline{\mathcal{D}}_1$. Les trois droites sont parallèles.

Solution 8.4. En considérant des parallélogrammes évidents, les droites $(B'Q)$, (BCA') , (AP) sont parallèles, de même que les droites (BQ) , (ACB') , $(A'P)$. L'homothétie $h \circ h' = h' \circ h$ transforme $(B'Q)$ en (AP) et (BQ) en $(A'P)$. L'intersection Q de $(B'Q)$ et (BQ) est transformée en l'intersection P de $(A'P)$ et (AP) . Cette homothétie composée étant de centre C' , on a prouvé l'alignement de P, Q, C' . Faisant opérer $H_{C', \frac{1}{2}}$, on a l'alignement de A'', B'', C'' .

Solution 8.5. Soit $M_0 \in \mathcal{A}$ tel que les points $M_0, N_0 = f(M_0), P_0 = g(N_0)$ soient distincts. La question est : les droites (MP) et (M_0P_0) sont-elles parallèles ?

- Supposons $\mathcal{A}, \mathcal{B}, \mathcal{C}$ parallèles. La translation de vecteur $\overrightarrow{M_0N_0}$ transforme $\mathcal{A}$ en $\mathcal{B}$ et M en N . On a donc $\overrightarrow{MN} = \overrightarrow{M_0N_0} \in F$ et de même $\overrightarrow{NP} = \overrightarrow{N_0P_0} \in G$. On en déduit

$$\overrightarrow{MP} = \overrightarrow{MN} + \overrightarrow{NP} = \overrightarrow{M_0N_0} + \overrightarrow{N_0P_0} = \overrightarrow{M_0P_0}$$

La droite (MP) est donc de direction fixe.

- Supposons $\mathcal{A}, \mathcal{B}, \mathcal{C}$ sécantes en O . Par le théorème de Thalès, on a

$$\frac{\overline{OM}}{\overline{OM_0}} = \frac{\overline{ON}}{\overline{ON_0}} = \frac{\overline{OP}}{\overline{OP_0}}$$

Par la remarque sur la réciproque du théorème de Thalès, on en déduit que (M_0P_0) et (MP) sont parallèles.

- Soit φ (resp. ψ) la projection de $\mathcal{P}$ sur $\mathcal{B}$ (resp. $\mathcal{C}$) parallèlement à F (resp. G). Alors f (resp. g) est la restriction de φ à $\mathcal{A}$ (resp. de ψ à $\mathcal{B}$). La question est : $h = g \circ f$ est-elle la restriction à $\mathcal{A}$ d'une projection de $\mathcal{P}$ sur $\mathcal{C}$? S'il en est ainsi, $B = \mathcal{A} \cap \mathcal{C}$ est fixe par h . La droite $B + F$ coupe $\mathcal{B}$ en $f(B) = B'$ et la droite $B' + G$ coupe la droite $\mathcal{C}$ en B . La droite (BB') a donc F et G pour direction. On a donc $F = G$. Inversement, si $F = G$, le résultat est clair.

Solution 8.6. 1) Soit H' l'application affine transformant A_1 en A_2 , d'application linéaire associée l'homothétie vectorielle de rapport λ . C'est une homothétie de rapport λ , $H'(B_1) = H'(A_1) + \lambda \overrightarrow{A_1B_1} = A_2 + \overrightarrow{A_2B_2} = B_2$. Son centre est l'intersection O' des droites (A_1A_2) et (B_1B_2) .

Inversement, une homothétie transformant A_1 en A_2 et B_1 en B_2 est nécessairement $H' = H_{O', \lambda}$.

De même, l'application affine H'' transformant A_1 en B_2 , d'application linéaire associée l'homothétie vectorielle de rapport $-\lambda$, est une homothétie transformant B_1 en $H''(A_1) - \lambda \overrightarrow{A_1B_1} = B_2 + \overrightarrow{B_2A_2} = A_2$. Son centre est l'intersection O'' des droites (A_1B_2) et (B_1A_2) .

Inversement, une homothétie transformant A_1 en B_2 et B_1 en A_2 est nécessairement $H'' = H_{O'', -\lambda}$.

2) Comme H', H'' respectent les rapports de mesures algébriques, elles transforment toutes deux I_1 en I_2 . On a donc l'alignement de O', O'', I_1, I_2 et

$$\frac{\overline{O'I_2}}{\overline{O'I_1}} = \lambda = -\frac{\overline{O''I_2}}{\overline{O''I_1}}$$

Solution 8.7. En faisant opérer les symétries centrales de centres P, Q, R, on obtient les égalités :

$$\begin{aligned} \frac{\overline{BA'}}{\overline{CA'}} &= \frac{\overline{CA''}}{\overline{BA''}}, & \frac{\overline{CB'}}{\overline{AB'}} &= \frac{\overline{AB''}}{\overline{CB''}}, & \frac{\overline{AC'}}{\overline{BC'}} &= \frac{\overline{BC''}}{\overline{AC''}} \\ \frac{\overline{BA'}}{\overline{CA'}} \times \frac{\overline{CB'}}{\overline{AB'}} \times \frac{\overline{AC'}}{\overline{BC'}} &= \frac{\overline{CA''}}{\overline{BA''}} \times \frac{\overline{AB''}}{\overline{CB''}} \times \frac{\overline{BC''}}{\overline{AC''}} \end{aligned}$$

Les théorèmes de Ménélaüs et Céva donnent les résultats demandés.

Solution 8.8. Le théorème de Thalès donne

$$\frac{\overline{B''A}}{\overline{B''C}} = \frac{\overline{A'B}}{\overline{A'C}}, \quad \frac{\overline{C''B}}{\overline{C''A}} = \frac{\overline{B'C}}{\overline{B'A}}, \quad \frac{\overline{A''C}}{\overline{A''B}} = \frac{\overline{C'A}}{\overline{C'B}},$$

donc

$$\frac{\overline{B''A}}{\overline{B''C}} \times \frac{\overline{C''B}}{\overline{C''A}} \times \frac{\overline{A''C}}{\overline{A''B}} = \frac{\overline{A'B}}{\overline{A'C}} \times \frac{\overline{B'C}}{\overline{B'A}} \times \frac{\overline{C'A}}{\overline{C'B}}$$

On applique alors les théorèmes de Ménélaüs et Céva.

Solution 8.9. En considérant les égalités de conjugaison harmonique et l'égalité de Ménélaüs

$$\frac{\overline{A'B}}{\overline{A'C}} = -\frac{\overline{A''B}}{\overline{A''C}}, \quad \frac{\overline{B'C}}{\overline{B'A}} = -\frac{\overline{B''C}}{\overline{B''A}}, \quad \frac{\overline{C'A}}{\overline{C'B}} = -\frac{\overline{C''A}}{\overline{C''B}}, \quad \frac{\overline{A'B}}{\overline{A'C}} \times \frac{\overline{B'C}}{\overline{B'A}} \times \frac{\overline{C'A}}{\overline{C'B}} = 1$$

on obtient l'alignement de (A', B'', C'') , (A'', B', C') , (A'', B'', C') de même que le concours ou le parallélisme de chacun des triplets de droites $((AA''), (BB'), (CC'))$, $((AA'), (BB''), (CC'))$, $((AA'), (BB'), (CC''))$ et $((AA''), (BB''), (CC''))$.

Solution 8.10. On applique le théorème de Ménélaüs aux triangles CAA' et BAA' coupés respectivement par les droites (RBB') et (QCC') . Ceci donne

$$\frac{\overline{RA}}{\overline{RA'}} \times \frac{\overline{B'C}}{\overline{B'A}} \times \frac{\overline{BA'}}{\overline{BC}} = 1 = \frac{\overline{QA'}}{\overline{QA}} \times \frac{\overline{CB}}{\overline{CA'}} \times \frac{\overline{C'A}}{\overline{C'B}}$$

D'autre-part, (toujours d'après Ménélaüs)

$$\frac{\overline{A'B}}{\overline{A'C}} \times \frac{\overline{B'C}}{\overline{B'A}} \times \frac{\overline{C'A}}{\overline{C'B}} = 1, \quad \frac{\overline{CB}}{\overline{BC}} = -1$$

En multipliant les deux premières relations, on obtient le résultat

$$\frac{\overline{QA'}}{\overline{QA}} \times \frac{\overline{RA}}{\overline{RA'}} = -1 \quad \text{soit} \quad \frac{\overline{QA'}}{\overline{QA}} = -\frac{\overline{RA'}}{\overline{RA}}$$

Solution 8.11. 1) L'égalité $\frac{\overline{AM}}{\overline{BM}} = -\frac{\overline{AN}}{\overline{BN}}$ montre que $\overrightarrow{AM}$ et $\overrightarrow{MB} = -\overrightarrow{BM}$ sont de même sens si et seulement si $\overrightarrow{AN}$ et $\overrightarrow{NB} = -\overrightarrow{BN}$ sont de sens opposés.

2) Prenons $\overrightarrow{IB}$ pour vecteur unitaire. On pose $x = \overline{IM}$ et $y = \overline{BN}$. La relation de Newton donne

$$1 = \overline{IB}^2 = \overline{IM} \times \overline{IN} = \overline{IM} \times (\overline{IB} + \overline{BN}) = x(1 + y)$$

L'application $[0, +\infty[\rightarrow]0, 1], y \mapsto x = \frac{1}{1+y}$ est bijective, d'où le résultat.

Solution 8.12. On a $(g \circ \text{rot}(0, \theta) \circ g^{-1})(g(0)) = (g \circ \text{rot}(0, \theta))(0) = g(0)$, donc $g(0)$ est fixe par $g \circ \text{rot}(0, \theta) \circ g^{-1}$.

Soit $\vec{e}_1, \vec{e}_2$ une base orthonormale directe. La base $\vec{\varepsilon}_1 = \overline{g}(\vec{e}_1), \vec{\varepsilon}_2 = \overline{g}(\vec{e}_2)$ est orthonormale, directe ou indirecte selon que $\overline{g}$ est une isométrie directe ou indirecte.

La matrice la rotation vectorielle $\text{rot}(\theta)$ dans la base $\vec{e}_1, \vec{e}_2$ est $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$.

C'est aussi la matrice de $\overline{g} \circ \text{rot}(\theta) \circ \overline{g}^{-1}$ dans la base orthonormale $\vec{\varepsilon}_1, \vec{\varepsilon}_2$ (p. 212). Donc $\overline{g} \circ \text{rot}(\theta) \circ \overline{g}^{-1}$ est la rotation vectorielle d'angle θ ou $-\theta$ selon que la base orthonormale $\vec{\varepsilon}_1, \vec{\varepsilon}_2$ est directe ou indirecte, c'est-à-dire selon que g est un déplacement ou un antidéplacement.

Solution 8.13. Soit $\mathcal{R}(f)$ l'ensemble des repères orthonormés tels que f soit d'expression analytique

$$f: \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto f(M) \begin{pmatrix} x \cos \theta - y \sin \theta \\ x \sin \theta + y \cos \theta \\ z \end{pmatrix}$$

Si f n'est pas un retournement, les repères orthonormés directs de $\mathcal{R}(f)$ sont ceux de la forme $(0, \vec{i}, \vec{j}, \vec{k})$ où $0 \in \mathcal{D}$. L'ensemble des déplacements qui commutent avec f est formé des vissages d'axes $\mathcal{D}$.

Si f est un retournement, il faut ajouter à cet ensemble les repères orthonormés directs $(0, \vec{j}, \vec{i}, -\vec{k})$. L'ensemble des déplacements qui commutent avec f est formé des vissages d'axes $\mathcal{D}$ et des retournements d'axes rencontrant $\mathcal{D}$ orthogonalement.

Si f n'est pas un retournement, les repères orthonormés indirects de $\mathcal{R}(f)$ sont ceux de la forme $(0, \vec{i}, \vec{j}, -\vec{k})$. L'ensemble des antidéplacements qui commutent avec f est formé des symétries centrales autour de points de $\mathcal{D}$.

Si f est un retournement, il faut ajouter à l'ensemble précédent les repères orthonormés indirects $(0, \vec{i}, \vec{j}, \vec{k})$ où $0 \in \mathcal{D}$.

Il faut ajouter à cet ensemble d'antidéplacements les réflexions glissées par un vecteur de $\overline{\mathcal{D}}$ autour d'un plan passant par $\mathcal{D}$.

Solution 8.14. L'unique similitude vectorielle directe σ transformant $\overrightarrow{AB}$ en $\overrightarrow{A'B'}$ est de rapport $k = \frac{A'B'}{AB}$ et d'angle $\theta = \widehat{(\overrightarrow{AB}, \overrightarrow{A'B'})}$. L'application affine s définie par $s(A) = A'$ et $\overline{s} = \sigma$ est l'unique similitude directe transformant (A, B) en (A', B') .

Chapitre 9

Calculs barycentriques

L'introduction dans la section 1 de l'espace vectoriel $\tilde{\mathcal{E}}$ des points pondérés d'un espace affine $\mathcal{E}$ permet l'application à la géométrie affine du calcul de l'algèbre linéaire. Des applications en géométrie plane sont présentées : aire algébrique dans un plan euclidien (9.2.1), problèmes de régionnement (9.2.2), géométrie du triangle (9.2.3). Les exercices proposés ont pour but de convaincre le lecteur de l'efficacité des méthodes calculatoires de l'algèbre linéaire. Et par exemple pour les problèmes de régionnement, l'étude du signe de certaines entités permet d'éviter de fastidieuses discussions.

9.1 ESPACE VECTORIEL DES POINTS PONDÉRÉS

9.1.1. Barycentres

Soit un espace affine $\mathcal{E}$ de dimension n d'espace vectoriel associé E . Soit $(\alpha_i, M_i)_i$ une famille finie de couples d'un coefficient $\alpha_i \in \mathbb{R}$ et d'un point $M_i \in \mathcal{E}$. Deux cas sont à distinguer pour l'application

$$\begin{aligned}\mathcal{E} &\longrightarrow E \\ P &\longmapsto \sum_i \alpha_i \overrightarrow{PM_i}\end{aligned}$$

Proposition 9.1. (i) Si $\sum_i \alpha_i = 0$, cette application est constante et le vecteur $\sum_i \alpha_i \overrightarrow{PM_i}$ ne dépend pas du point $P \in \mathcal{E}$.

(ii) Si $\sum_i \alpha_i \neq 0$, cette application est une bijection de $\mathcal{E}$ sur lui-même.

Démonstration. En effet, étant donné deux points P et O, on a

$$\sum_i \alpha_i \overrightarrow{PM_i} - \sum_i \alpha_i \overrightarrow{OM_i} = \left(\sum_i \alpha_i \right) \overrightarrow{PO}$$

(i) Si $\sum \alpha_i = 0$, $\sum_i \alpha_i \overrightarrow{PM_i} = \sum_i \alpha_i \overrightarrow{OM_i}$ est indépendant de P (et de O).

(ii) Supposons $\sum_i \alpha_i \neq 0$. Soit $\vec{v} \in E$. Pour tout P $\in \mathcal{E}$, on a

$$\begin{aligned} \sum_i \alpha_i \overrightarrow{PM_i} &= \sum_i \alpha_i \overrightarrow{OM_i} - \left(\sum_i \alpha_i \right) \overrightarrow{OP} \\ \text{donc} \quad \left(\sum_i \alpha_i \overrightarrow{PM_i} = \vec{v} \right) &\Leftrightarrow \left(\overrightarrow{OP} = \frac{\sum_i \alpha_i \overrightarrow{OM_i} - \vec{v}}{\sum_i \alpha_i} \right) \end{aligned}$$

Il existe bien un unique point P $\in \mathcal{E}$ tel que $\sum_i \alpha_i \overrightarrow{PM_i} = \vec{v}$. □

Définition 9.2. Si $\sum \alpha_i \neq 0$, l'unique point M tel que $\sum_i \alpha_i \overrightarrow{MM_i} = \vec{0}$ est appelé **barycentre** des M_i affectés des α_i . Par ce qui précède, O étant donné quelconque, le barycentre M est caractérisé par la relation

$$\left(\sum_i \alpha_i \right) \overrightarrow{OM} = \sum_i \alpha_i \overrightarrow{OM_i}$$

9.1.2. L'espace vectoriel $\tilde{\mathcal{E}}$

Soit l'ensemble produit $\mathbb{R}^\times \times \mathcal{E}$ des couples (λ, M) où $\lambda \in \mathbb{R}^\times$, ensemble des réels non nuls, et $M \in \mathcal{E}$. Considérant $\mathbb{R}^\times \times \mathcal{E}$ et E comme disjoints, soit la réunion $\tilde{\mathcal{E}}$:

$$\tilde{\mathcal{E}} = (\mathbb{R}^\times \times \mathcal{E}) \sqcup E$$

le signe $\sqcup$ (au lieu de $\cup$) rappelant que ces ensembles sont **disjoints**. On va munir $\tilde{\mathcal{E}}$ d'une structure d'espace vectoriel de dimension $n + 1$ sur $\mathbb{R}$.

• **Addition** Soit x, y deux éléments de $\tilde{\mathcal{E}}$. On définit $x + y$ ainsi :

(1) Si $x = (\alpha, A)$, $y = (\beta, B)$ où α, β sont non nuls et vérifient $\alpha + \beta \neq 0$,

$$(\alpha, A) + (\beta, B) = (\alpha + \beta, C)$$

avec C barycentre de α, A et β, B .

(2) Si $x = (\alpha, A)$, $y = (-\alpha, B)$ où $\alpha \neq 0$,

$$(\alpha, A) + (-\alpha, B) = \alpha \overrightarrow{BA}$$

(3) Si $x = (\alpha, A)$ où $\alpha \neq 0$, $y = \vec{v}$,

$$(\alpha, A) + \vec{v} = \vec{v} + (\alpha, A) = (\alpha, A')$$

où $A' = A + \frac{\vec{v}}{\alpha}$ est le translaté de A par la translation de vecteur $\frac{\vec{v}}{\alpha}$.

(4) Si $x = \vec{u}$, $y = \vec{v}$ sont deux vecteurs, la somme $x + y$ est la somme $\vec{u} + \vec{v}$ au sens de l'addition de E .

• **Multiplication scalaire** Soit $x \in \tilde{\mathcal{E}}$ et $\lambda \in \mathbb{R}$. On définit λx ainsi :

- (1) Si $\lambda = 0$, pour tout $x \in \tilde{\mathcal{E}}$, $0x = \vec{0}$
- (2) Si $x = (\alpha, A)$ où $\alpha \neq 0$ et si $\lambda \neq 0$, $\lambda(\alpha, A) = (\lambda\alpha, A)$
- (3) Si $x = \vec{v} \in E$, $\lambda\vec{v}$ est à entendre au sens de la structure d'espace vectoriel de E .

Théorème 9.3. *L'ensemble $\tilde{\mathcal{E}}$ muni de ces opérations est un espace vectoriel sur $\mathbb{R}$ de dimension $n + 1$.*

Démonstration. Cet énoncé remplace les énoncés habituels sur l'associativité, la commutativité de l'opération barycentre. La démonstration est facile.

Pour l'associativité de l'addition α, β, γ étant dans $\mathbb{R}^\times$ et A, B, C dans $\mathcal{E}$, on doit distinguer plusieurs cas pour prouver la formule

$$((\alpha, A) + (\beta, B)) + (\gamma, C) = (\alpha, A) + ((\beta, B) + (\gamma, C))$$

$$\begin{array}{lll} \alpha + \beta + \gamma \neq 0, & \alpha + \beta \neq 0, & \beta + \gamma \neq 0 \\ \alpha + \beta + \gamma = 0, & \alpha + \beta \neq 0, & \beta + \gamma \neq 0 \\ \alpha + \beta + \gamma \neq 0, & \alpha + \beta = 0, & \beta + \gamma \neq 0 \\ \alpha + \beta + \gamma \neq 0, & \alpha + \beta \neq 0, & \beta + \gamma = 0 \end{array}$$

Traisons par exemple le cas où $\alpha + \beta + \gamma \neq 0$, $\alpha + \beta \neq 0$, $\beta + \gamma \neq 0$. On définit P, Q, R, S par les formules :

$$\begin{array}{ll} (\alpha, A) + (\beta, B) = (\alpha + \beta, P) & (\alpha + \beta, P) + (\gamma, C) = (\alpha + \beta + \gamma, R) \\ (\beta, B) + (\gamma, C) = (\beta + \gamma, Q) & (\alpha, A) + (\beta + \gamma, Q) = (\alpha + \beta + \gamma, S) \end{array}$$

Il s'agit de prouver que $R = S$. Soit M un point quelconque. On a

$$\begin{array}{ll} (\alpha + \beta + \gamma)\vec{MR} &= (\alpha + \beta)\vec{MP} + \gamma\vec{MC} = \alpha\vec{MA} + \beta\vec{MB} + \gamma\vec{MC} \\ (\alpha + \beta + \gamma)\vec{MS} &= \alpha\vec{MA} + (\beta + \gamma)\vec{MQ} = \alpha\vec{MA} + \beta\vec{MB} + \gamma\vec{MC} \end{array}$$

Comme $\alpha + \beta + \gamma \neq 0$, on a $R = S$, barycentre de $(\alpha, A), (\beta, B), (\gamma, C)$.

Le lecteur est invité à faire d'autres vérifications pour se familiariser avec l'espace vectoriel un peu inhabituel des points pondérés.

Le couple (λ, M) est le point M **pondéré** par le coefficient λ appelé **masse** de (λ, M) . On vérifie immédiatement que l'application masse

$$\begin{array}{ll} \mu: \tilde{\mathcal{E}} & \longrightarrow \mathbb{R} \\ (\lambda, M) & \longmapsto \lambda \\ \vec{v} & \longmapsto 0 \end{array}$$

est une forme linéaire. Le noyau E est donc un hyperplan vectoriel de $\tilde{\mathcal{E}}$, d'où $\dim \tilde{\mathcal{E}} = \dim E + 1 = n + 1$. $\square$

9.1.3. Nouvelles notations

L'espace affine $\mathcal{E}$ s'injecte canoniquement dans $\tilde{\mathcal{E}}$ par l'injection $M \mapsto (1, M)$. Identifiant $\mathcal{E}$ à son image, $\mathcal{E}$ est l'ensemble des éléments de $\tilde{\mathcal{E}}$ de masse 1, on note λM au lieu de (λ, M) le point M pondéré par la masse λ . Explicitons ces nouvelles notations :

- Pour $M \in \mathcal{E}$ et $\vec{v} \in E$, $M + \vec{v}$ est le translaté de M par la translation $T_{\vec{v}}$, ce qui éclaire la notation introduite au chapitre 8 (p. 204)
- Si A, B sont deux points, on peut écrire $\overrightarrow{AB} = B - A$
- Si $(\alpha_i M_i)_i$ est une famille de points pondérés telle que $\sum \alpha_i \neq 0$, de barycentre M , on écrit :

$$\sum \alpha_i M_i = (\sum \alpha_i) M \text{ ou } M = \frac{\sum \alpha_i M_i}{\sum \alpha_i}, \text{ ainsi } \frac{A+B}{2} \text{ est milieu de } (A, B)$$

- Si $(\alpha_i M_i)_i$ est une famille de points pondérés telle que $\sum \alpha_i = 0$, le vecteur $\sum \alpha_i \overrightarrow{PM_i}$, indépendant de P , est noté $\sum \alpha_i M_i$.

9.1.4. Sous-espaces vectoriels de $\tilde{\mathcal{E}}$ et sous-espaces affines de $\mathcal{E}$

Soit $\mathcal{F}$ un sous-espace affine de $\mathcal{E}$ de direction $\overline{\mathcal{F}}$. On voit immédiatement que $\tilde{\mathcal{F}} = (\mathbb{R}^\times \times \mathcal{F}) \sqcup \overline{\mathcal{F}}$ est sous-espace vectoriel de $\tilde{\mathcal{E}}$.

Proposition 9.4. *L'application $\mathcal{F} \mapsto \tilde{\mathcal{F}}$ est une bijection de l'ensemble des sous-espaces affines de $\mathcal{E}$ sur l'ensemble des sous-espaces vectoriels de $\tilde{\mathcal{E}}$ non contenus dans E .*

Démonstration. Montrons qu'un sous-espace vectoriel $Z \not\subset E$ de dimension d est du type $\tilde{\mathcal{F}}$. Posons $F = Z \cap E$, sous-espace de dimension $d-1$ de E . Il existe $z \in Z$ tel que $\mu(z) \neq 0$. Remplaçant z par $\frac{z}{\mu(z)}$, on se ramène à $z \in \mathcal{E}$. Soit $\mathcal{F} = z + F$. On a $\tilde{\mathcal{F}} \subset Z$, $\overline{\mathcal{F}} = F$, $\dim \tilde{\mathcal{F}} = d = \dim Z$, d'où $\tilde{\mathcal{F}} = Z$. $\square$

On a donc deux sortes de sous-espaces vectoriels dans $\tilde{\mathcal{E}}$:

- les sous-espaces vectoriels non contenus dans l'hyperplan E , du type $\tilde{\mathcal{F}}$ où $\mathcal{F}$ est sous-espace affine de $\mathcal{E}$,
- les sous-espaces vectoriels contenus dans l'hyperplan E de $\tilde{\mathcal{E}}$.

Les phénomènes d'incidence (alignement, coplanarité, etc.) dans $\mathcal{E}$ sont des phénomènes de dépendance linéaire dans $\tilde{\mathcal{E}}$ et peuvent être traités par des techniques d'algèbre linéaire (nullité de déterminants, etc.). Par exemple, trois points A, B, C d'un espace affine $\mathcal{E}$ sont alignés si et seulement si A, B, C sont liés en tant que vecteurs de $\tilde{\mathcal{E}}$.

Définition 9.5. On appelle **isobarycentre** d'un système de points $M_1, \dots, M_k$ le point $M = \frac{1}{k}(M_1 + \dots + M_k)$. Pour $k = 2$, le milieu de M_1, M_2 est $\frac{1}{2}(M_1 + M_2)$.

9.1.5. Exemple

Soit un triangle ABC , A' , B' , C' les milieux de (B, C) , (C, A) , (A, B) , G l'isobarycentre ou **centre de gravité** de A, B, C . On a

$$G = \frac{1}{3}(A + B + C) = \frac{1}{3}(A + 2A') = \frac{1}{3}(B + 2B') = \frac{1}{3}(C + 2C')$$

Ainsi, A, A', G est lié dans $\tilde{\mathcal{E}}$, donc G est sur la médiane (AA') . Il est de même sur (BB') et (CC') . Donc les médianes concourent en G .

On a aussi $\overrightarrow{GA} + 2\overrightarrow{GA'} = \vec{0}$ et deux autres relations analogues, d'où

$$\frac{\overrightarrow{GA}}{\overrightarrow{GA'}} = \frac{\overrightarrow{GB}}{\overrightarrow{GB'}} = \frac{\overrightarrow{GC}}{\overrightarrow{GC'}} = -2$$

9.1.6. Repères barycentriques

Soit un repère affine d'origine O et de base $(\vec{e}_1, \dots, \vec{e}_n)$. À tout $M \in \mathcal{E}$ correspond son système $(x_1, \dots, x_n)$ de coordonnées avec

$$\overrightarrow{OM} = \sum_{i=1}^n x_i \vec{e}_i \text{ ou } M = O + \sum_{i=1}^n x_i \vec{e}_i$$

Comme $(\vec{e}_1, \dots, \vec{e}_n)$ est base de E et $O \notin E$, $(O, \vec{e}_1, \dots, \vec{e}_n)$ est base de $\tilde{\mathcal{E}}$ et $1, x_1, \dots, x_n$ sont les composantes de $M \in \mathcal{E} \subset \tilde{\mathcal{E}}$ dans cette base.

Définition 9.6. Un **repère barycentrique** ou **simplexe** est la donnée de $n + 1$ points $(A_0, A_1, \dots, A_n)$ non dans un même hyperplan affine de $\mathcal{E}$, donc non dans un même hyperplan vectoriel de $\tilde{\mathcal{E}}$ (par exemple, trois points non alignés d'un plan, ou quatre points non coplanaires d'un espace de dimension 3). C'est donc une base de $\tilde{\mathcal{E}}$.

Un point $M \in \mathcal{E}$ étant considéré comme vecteur de $\tilde{\mathcal{E}}$, il existe un unique système de $n + 1$ scalaires $\xi_0, \dots, \xi_n$ de $\mathbb{R}$ tel que $M = \sum_0^n \xi_i A_i$.

Autrement dit M est barycentre des A_i affectés des ξ_i , avec la relation $\sum_i \xi_i = 1$, obtenue en écrivant l'égalité des masses des deux membres :

$$1 = \mu(M) = \mu \left(\sum_0^n \xi_i A_i \right) = \sum_0^n \xi_i$$

Pour tout $\lambda \neq 0$, $\lambda M = \sum_i (\lambda \xi_i) A_i$. On dit que $(\lambda \xi_0, \lambda \xi_1, \dots, \lambda \xi_n)$ est un **système de coordonnées barycentriques** de M dans ce repère. Le point M a une infinité de systèmes de coordonnées barycentriques tous proportionnels. Parmi eux, $(\xi_0, \dots, \xi_n)$ est le seul dont la somme des coordonnées est 1

Si $\vec{v} \in E$, on pourra écrire, de façon unique $\vec{v} = \sum_i \lambda_i \mathbf{A}_i$ et on a

$$0 = \mu(\vec{v}) = \mu\left(\sum_0^n \lambda_i \mathbf{A}_i\right) = \sum_0^n \lambda_i$$

9.1.7. Applications affines et applications linéaires

Soit $\mathcal{E}, \mathcal{E}'$ deux espaces affines, μ, μ' les formes linéaires masses. Étant donné une application affine $f: \mathcal{E} \rightarrow \mathcal{E}'$, on définit l'application $\tilde{f}: \tilde{\mathcal{E}} \rightarrow \tilde{\mathcal{E}}'$ ainsi :

$$\forall \mathbf{M} \in \mathcal{E}, \forall \lambda \in \mathbb{R}, \lambda \neq 0, \tilde{f}(\lambda \mathbf{M}) = \lambda f(\mathbf{M}) ; \forall \vec{v} \in E, \tilde{f}(\vec{v}) = \vec{f}(\vec{v})$$

Proposition 9.7. Pour toute $f: \mathcal{E} \rightarrow \mathcal{E}'$ affine, $\tilde{f}$ est linéaire. L'application $f \mapsto \tilde{f}$ est une bijection de l'ensemble des applications affines $\mathcal{E} \rightarrow \mathcal{E}'$ sur l'ensemble des applications linéaires $\tilde{\mathcal{E}} \rightarrow \tilde{\mathcal{E}}'$ telles que $\mu' \circ \varphi = \mu$.

Démonstration. Montrons la linéarité de $\tilde{f}$. Par définition, on a $\tilde{f}(\lambda x) = \lambda \tilde{f}(x)$ pour tout $\lambda \in \mathbb{R}$ et tout $x \in \tilde{\mathcal{E}}$.

Montrons que $\tilde{f}(x + y) = \tilde{f}(x) + \tilde{f}(y)$ pour tout $(x, y) \in \tilde{\mathcal{E}} \times \tilde{\mathcal{E}}$. Comme pour le théorème 9.3, p. 245, on distingue plusieurs cas. On se bornera au cas où $x = \alpha \mathbf{A}$, $y = \beta \mathbf{B}$, $\mathbf{A}, \mathbf{B}$ étant dans $\mathcal{E}$, α, β des réels non nuls tels que $\alpha + \beta \neq 0$, les autres cas se vérifiant facilement. Soit $\mathbf{P}$ le barycentre de $\mathbf{A}, \alpha$ et $\mathbf{B}, \beta$. On a $\alpha \vec{\mathbf{P}\mathbf{A}} + \beta \vec{\mathbf{P}\mathbf{B}} = \vec{0}$. Soit $\mathbf{P}' = f(\mathbf{P})$, $\mathbf{A}' = f(\mathbf{A})$ et $\mathbf{B}' = f(\mathbf{B})$. On est ramené à prouver que $\mathbf{P}'$ est barycentre de $\mathbf{A}', \alpha$ et $\mathbf{B}', \beta$. Comme f est affine d'application linéaire associée $\vec{f}$, on a

$$\alpha \vec{\mathbf{P}'\mathbf{A}'} + \beta \vec{\mathbf{P}'\mathbf{B}'} = \alpha \vec{f}(\vec{\mathbf{P}\mathbf{A}}) + \beta \vec{f}(\vec{\mathbf{P}\mathbf{B}}) = \vec{f}(\alpha \vec{\mathbf{P}\mathbf{A}} + \beta \vec{\mathbf{P}\mathbf{B}}) = \vec{f}(\vec{0}) = \vec{0}$$

Soit $x \in \tilde{\mathcal{E}}$. Montrons que $\mu(x) = \mu'(\tilde{f}(x))$.

Si $x = \alpha \mathbf{A} \in \mathbb{R}^\times \times \mathcal{E}$, on a $\tilde{f}(x) = \alpha f(\mathbf{A})$ et $\alpha = \mu(x) = \mu'(\tilde{f}(x))$.

Si $x = \vec{v} \in E$, alors $\tilde{f}(x) = \vec{f}(\vec{v}) \in E'$, on a $0 = \mu(x) = \mu'(\tilde{f}(x))$.

Inversement, soit $\varphi: \tilde{\mathcal{E}} \rightarrow \tilde{\mathcal{E}}'$ linéaire telle que $\mu' \circ \varphi = \mu$. Comme $\mathcal{E}$ (resp. $\mathcal{E}'$) est l'ensemble des éléments de $\tilde{\mathcal{E}}$ (resp. $\tilde{\mathcal{E}}'$) de masse 1, on a $\varphi(\mathcal{E}) \subset \mathcal{E}'$. On vérifie alors, en utilisant la linéarité de φ , que l'application $f: \mathcal{E} \rightarrow \mathcal{E}'$ induite par φ est affine et que $\varphi = \tilde{f}$. $\square$

9.2 APPLICATION EN GÉOMÉTRIE PLANE

Dans cette section $\mathcal{P}$ est un plan affine euclidien orienté de plan vectoriel associé P . On sait qu'étant donné un système de deux vecteurs $\vec{v}, \vec{w}$ et une base orthonormale directe $\mathcal{B} = (\vec{i}, \vec{j})$, la valeur du déterminant $\det_{\mathcal{B}}(\vec{v}, \vec{w})$ ne dépend pas de la base orthonormale directe $\mathcal{B}$. (Voir 7.3.5., page 191). On peut donc ne pas la mentionner. La notation $\det(\vec{v}, \vec{w})$ désignera le déterminant de ces vecteurs relativement à n'importe quelle base orthonormale directe.

9.2.1. Aires algébriques

Lemme 9.8. *Étant donné trois points A, B, C de $\mathcal{P}$, on a les égalités de déterminants $\det(\vec{AB}, \vec{AC}) = \det(\vec{BC}, \vec{BA}) = \det(\vec{CA}, \vec{CB})$. Leur valeur commune sera notée $[A, B, C]$.*

Démonstration. Il suffit de prouver une des égalités, les autres se montrant de la même façon. On a

$$\det(\vec{BC}, \vec{BA}) = \det(\vec{AC} - \vec{AB}, -\vec{AB}) = \det(\vec{AC}, -\vec{AB}) = \det(\vec{AB}, \vec{AC})$$

□

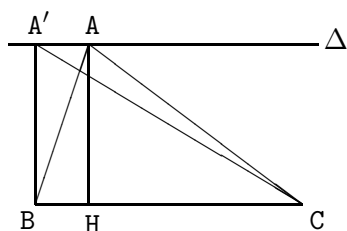
Le lemme montre que $[A, B, C]$ ne varie pas si on fait une permutation circulaire sur A, B, C et se change en son opposé si on transpose deux des lettres A, B, C : $[A, B, C] = [B, C, A] = [C, A, B] = -[A, C, B] = -[C, B, A] = -[B, A, C]$.

Définition 9.9. On appelle **aire algébrique du triangle orienté** ABC la quantité

$$\frac{1}{2}[A, B, C] = \frac{1}{2}\det(\vec{AB}, \vec{AC}) = \frac{1}{2}\det(\vec{BC}, \vec{BA}) = \frac{1}{2}\det(\vec{CA}, \vec{CB})$$

ABC étant **direct** si $[A, B, C]$ est positif, **inverse** sinon. S'il en est ainsi, les triangles ABC, BCA, CAB sont directs, les triangles ACB, CBA, BAC étant inverses.

Voyons en quoi la notion précédente se raccroche à la notion d'aire d'un triangle des classes élémentaires. Considérons un triangle direct ABC.



Soit Δ la parallèle à (BC) passant par A, H la projection orthogonale de A sur (BC), A' la projection orthogonale de B sur Δ . Selon la notion élémentaire, l'aire du triangle ABC est

$$S = \frac{1}{2}BC \times AH = \frac{1}{2}BC \times BA'$$

D'autre part, en prenant les déterminants relativement à la base orthonormale directe $(\vec{i}, \vec{j})$ où $\vec{i}$ et $\vec{j}$ sont colinéaires de même sens à $\vec{BC}$ et $\vec{BA}'$, on a

$$\begin{aligned} [A, B, C] &= [B, C, A] = \det(\vec{BC}, \vec{BA}) = \det(\vec{BC}, \vec{BA}' + \vec{A'A}) \\ &= \det(\vec{BC}, \vec{BA}') = BC \times BA' = 2S \end{aligned}$$

Théorème 9.10. Soit A, B, C un repère barycentrique de $\mathcal{P}$ et M un point.

Alors $[M, B, C], [A, M, C], [A, B, M]$ est l'unique système de coordonnées barycentriques de M vérifiant

$$[M, B, C] + [A, M, C] + [A, B, M] = [A, B, C]$$

Démonstration. Soit x, y, z le système de coordonnées barycentriques de M tel que $x + y + z = 1$. On a $M = xA + yB + zC$, d'où $\vec{AM} = y\vec{AB} + z\vec{AC}$ et

$$[A, M, C] = \det(\vec{AM}, \vec{AC}) = y \det(\vec{AB}, \vec{AC}) = y[A, B, C]$$

On a de même $[M, B, C] = x[A, B, C]$ et $[A, B, M] = z[A, B, C]$. Ainsi, $[M, B, C]$, $[A, M, C]$, $[A, B, M]$ sont proportionnels à x, y, z et

$$[M, B, C] + [A, M, C] + [A, B, M] = (x + y + z)[A, B, C] = [A, B, C]$$

□

9.2.2. Aires algébriques et déterminants dans l'espace $\tilde{\mathcal{P}}$

Soit $\mathcal{B} = (\vec{i}, \vec{j})$ une base orthonormale directe et O un point origine. Alors $(O, \mathcal{B}) = (O, \vec{i}, \vec{j})$ est base de l'espace vectoriel $\tilde{\mathcal{P}}$ de dimension 3 des points pondérés.

Proposition 9.11. Pour tout triangle ABC , on a dans l'espace vectoriel $\tilde{\mathcal{P}}$

$$[A, B, C] = \det_{O, \mathcal{B}}(A, B, C)$$

Démonstration. Posons $\vec{OA} = \alpha\vec{i} + \beta\vec{j}$, $\vec{AB} = x\vec{i} + y\vec{j}$ et $\vec{AC} = z\vec{i} + t\vec{j}$. Considérons A, B, C comme vecteurs de $\tilde{\mathcal{P}}$, alors

$$A = O + \vec{OA} = O + \alpha\vec{i} + \beta\vec{j}$$

$$B = O + \vec{OA} + \vec{AB} = O + (\alpha + x)\vec{i} + (\beta + y)\vec{j}$$

$$C = O + \vec{OA} + \vec{AC} = O + (\alpha + z)\vec{i} + (\beta + t)\vec{j}$$

$$\begin{aligned} \det_{O, \mathcal{B}}(A, B, C) &= \begin{vmatrix} 1 & 1 & 1 \\ \alpha & \alpha + x & \alpha + z \\ \beta & \beta + y & \beta + t \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ \alpha & x & z \\ \beta & y & t \end{vmatrix} \\ &= \begin{vmatrix} x & z \\ y & t \end{vmatrix} = \det_{\mathcal{B}}(\vec{AB}, \vec{AC}) = [A, B, C] \end{aligned}$$

On a retranché la première colonne des deux autres dans le premier déterminant, puis développé par rapport à la première ligne. □

9.2.3. Régionnement

Définition 9.12. Étant donné un plan affine euclidien $\mathcal{P}$ de plan vectoriel associé P , une **fonction affine** est une application affine non constante $f: \mathcal{P} \rightarrow \mathbb{R}$. Autrement dit, f est déterminée par (p. 209)

1) l'image $f(O) \in \mathbb{R}$ d'un point O de $\mathcal{P}$,

2) la forme linéaire associée $\vec{f}: P \rightarrow \mathbb{R}$.

Pour tout point M , on a alors $f(M) = f(O) + \vec{f}(\vec{OM})$.

Voici trois aspects sous lesquels se présente souvent une fonction affine :

- (1) **Analytiquement.** Soit $(O, \vec{i}, \vec{j})$ un repère. Supposons $f(O) = c$ et $\text{Mat}_{\vec{i}, \vec{j}}(\vec{f}) = \begin{pmatrix} a & b \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \end{pmatrix}$, alors

$$f: M \begin{pmatrix} x \\ y \end{pmatrix} \mapsto f(M) = ax + by + c$$

- (2) Comme **produit scalaire.** Soit un point O et un vecteur non nul $\vec{v}$, l'application $M \mapsto \vec{OM} \cdot \vec{v}$ est une fonction affine.
- (3) Comme **aire algébrique.** Soit deux points distincts A, B , l'application $M \mapsto [A, B, M]$ est une fonction affine.

Définition 9.13. Si f est une fonction affine, les ensembles

$$\Delta_\lambda = \{M \in \mathcal{P} \mid f(M) = \lambda\} \text{ où } \lambda \in \mathbb{R}$$

forment une famille de droites affines parallèles de direction $\text{Ker } \vec{f}$. La donnée de f permet de partager le plan en

- le **demi-plan ouvert supérieur** $\mathcal{P}_f^+$ des M tels que $f(M) > 0$,
- le **demi-plan ouvert inférieur** $\mathcal{P}_f^-$ des M tels que $f(M) < 0$,
- la droite $V(f)$ des M tels que $f(M) = 0$.

On définit les demi-plans **fermés**

$$\overline{\mathcal{P}}_f^+ = \{M \mid f(M) \geq 0\}, \quad \overline{\mathcal{P}}_f^- = \{M \mid f(M) \leq 0\}$$

Proposition 9.14. a) La donnée d'une droite Δ détermine une unique partition de $\mathcal{P}$ en la droite Δ et deux demi-plans ouverts.

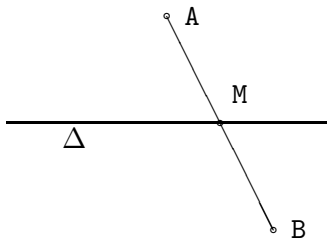
b) Deux points A et B non sur Δ sont dans le même demi-plan si et seulement si l'intersection du segment $[A, B]$ avec Δ est vide.

Démonstration. a) Soit des fonctions affines f, g déterminant la même droite $\Delta = V(f) = V(g)$. Les formes linéaires $\vec{f}$ et $\vec{g}$ sont proportionnelles car $\text{Ker } \vec{f} = \text{Ker } \vec{g}$. Il existe $\lambda \neq 0$ tel que $\vec{g} = \lambda \vec{f}$. Soit $O \in \Delta$. Pour tout M , on a $g(M) = \vec{g}(\vec{OM}) = \lambda \vec{f}(\vec{OM}) = \lambda f(M)$.

Si $\lambda > 0$, $\mathcal{P}_g^+ = \mathcal{P}_f^+$ et $\mathcal{P}_g^- = \mathcal{P}_f^-$.

Si $\lambda < 0$, $\mathcal{P}_g^- = \mathcal{P}_f^+$ et $\mathcal{P}_g^+ = \mathcal{P}_f^-$.

Les notions de demi-plan supérieur et demi-plan inférieur ne dépendent pas seulement de la droite Δ , mais aussi des fonctions affines déterminant Δ . En revanche, la partition du complémentaire de Δ en deux demi-plans ouverts est intrinsèque, ne dépendant que de la droite Δ .



b) Soit f fonction affine nulle sur Δ et A, B distincts non sur Δ . On a vu (chap. 8, p. 221) que l'application

$$t \mapsto A + t\vec{AB} = (1-t)A + tB$$

envoie $\mathbb{R}$ sur la droite (AB) et l'intervalle fermé $[0, 1]$ sur le segment $[A, B]$. Pour

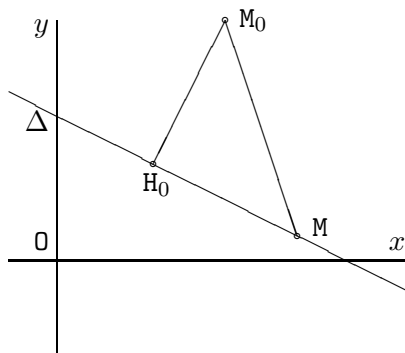
$M = tA + (1-t)B \in (A, B)$, on a (9.7, p. 248)

$$f(M) = f(tA + (1-t)B) = tf(A) + (1-t)f(B)$$

Si t varie de 0 à 1, M décrit le segment $[A, B]$, $f(M)$ décrit l'intervalle $[f(A), f(B)]$. Donc $[A, B]$ rencontre Δ si et seulement si $0 \in [f(A), f(B)]$, c'est-à-dire si et seulement si $f(A)$ et $f(B)$ sont de signes contraires, donc si et seulement si A et B ne sont pas dans le même demi-plan ouvert défini par Δ . $\square$

Proposition 9.15. Soit $(O, \vec{i}, \vec{j})$ un repère orthonormé, Δ une droite d'équation $ax + by + c = 0$. La distance d'un point $M_0(x_0, y_0)$ à Δ est

$$d(M_0, \Delta) = \frac{|ax_0 + by_0 + c|}{\sqrt{a^2 + b^2}}$$



Démonstration. La droite vectorielle $\overline{\Delta}$ est l'ensemble des $\vec{v} = \xi \vec{i} + \eta \vec{j}$ tels que $a\xi + b\eta = 0$. C'est donc la droite vectorielle orthogonale à $\vec{N} = a\vec{i} + b\vec{j}$. Soit $M(x, y) \in \Delta$ et H_0 la projection orthogonale de M_0 sur Δ . On a $ax + by = -c$ car $M(x, y) \in \Delta$. La formule se déduit alors des égalités :

$$\begin{aligned} \overrightarrow{MM_0} \cdot \vec{N} &= a(x_0 - x) + b(y_0 - y) = ax_0 + by_0 + c \\ |\overrightarrow{MM_0} \cdot \vec{N}| &= |\overrightarrow{MH_0} \cdot \vec{N}| = MH_0 \times \|\vec{N}\| = d(M_0, \Delta) \times \sqrt{a^2 + b^2} \end{aligned}$$

$\square$

La quantité $\frac{ax_0 + by_0 + c}{\sqrt{a^2 + b^2}}$ est la **distance algébrique** de M_0 à Δ . Son signe dépend du sens du vecteur $\vec{N}$ définissant l'orientation de $\overline{\Delta}^\perp$. Toutes les fonctions affines f telles que $V(f) = \Delta$ lui sont proportionnelles.

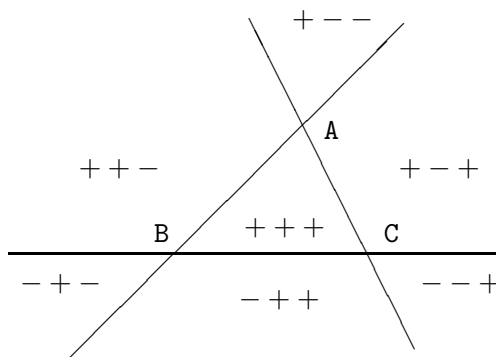
9.2.4. Régionnement relatif à un triangle

Soit un triangle direct ABC. Pour tout M, notons $x(M)$, $y(M)$, $z(M)$ les coordonnées barycentriques de M telles que $x(M) + y(M) + z(M) = 1$. On a (9.10, p. 249)

$$x(M) = \frac{[M, B, C]}{[A, B, C]}, \quad y(M) = \frac{[A, M, C]}{[A, B, C]}, \quad z(M) = \frac{[A, B, M]}{[A, B, C]}$$

Les trois droites (AB), (BC), (CA) déterminent 6 demi-plans ouverts :

$$\begin{aligned} \mathcal{A}^+ &= \{M \mid [M, B, C] > 0\} & \mathcal{A}^- &= \{M \mid [M, B, C] < 0\} \\ \mathcal{B}^+ &= \{M \mid [A, M, C] > 0\} & \mathcal{B}^- &= \{M \mid [A, M, C] < 0\} \\ \mathcal{C}^+ &= \{M \mid [A, B, M] > 0\} & \mathcal{C}^- &= \{M \mid [A, B, M] < 0\} \end{aligned}$$



On a donc 8 régions qui sont :

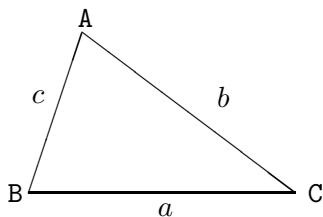
$$\begin{aligned} &\mathcal{A}^+ \cap \mathcal{B}^+ \cap \mathcal{C}^+ & \mathcal{A}^- \cap \mathcal{B}^+ \cap \mathcal{C}^+ & \mathcal{A}^+ \cap \mathcal{B}^- \cap \mathcal{C}^+ & \mathcal{A}^+ \cap \mathcal{B}^+ \cap \mathcal{C}^- \\ &\mathcal{A}^+ \cap \mathcal{B}^- \cap \mathcal{C}^- & \mathcal{A}^- \cap \mathcal{B}^+ \cap \mathcal{C}^- & \mathcal{A}^- \cap \mathcal{B}^- \cap \mathcal{C}^+ & \mathcal{A}^- \cap \mathcal{B}^- \cap \mathcal{C}^- \end{aligned}$$

La relation $[M, B, C] + [A, M, C] + [A, B, M] = [A, B, C] > 0$ (9.10, p. 249) impose que la région $\mathcal{A}^- \cap \mathcal{B}^- \cap \mathcal{C}^-$ est vide.

Les autres régions ne sont pas vides. Par exemple, $A + B - C \in \mathcal{A}^+ \cap \mathcal{B}^+ \cap \mathcal{C}^-$ et $3A - B - C \in \mathcal{A}^+ \cap \mathcal{B}^- \cap \mathcal{C}^-$. La région $\mathcal{A}^+ \cap \mathcal{B}^+ \cap \mathcal{C}^+$ s'appelle l'**intérieur** du triangle. L'isobarycentre $G = \frac{1}{3}(A + B + C)$ en fait partie.

L'interprétation géométrique du théorème 9.10 (p. 249) est que si M est intérieur au triangle, l'aire de ABC est somme des aires des triangles MBC, AMC, ABM, ce qu'on voit sur la figure. Si M n'est pas à l'intérieur, le théorème 9.10 énonce que c'est encore vrai à condition de considérer les aires **algébriques**.

9.3 APPLICATION EN GÉOMÉTRIE DU TRIANGLE



On donne un triangle non aplati ABC. Les longueurs des côtés sont notées $a = BC$, $b = CA$, $c = AB$. On note respectivement $\widehat{A}$, $\widehat{B}$, $\widehat{C}$ les mesures comprises entre $-\pi$ et $+\pi$ des angles orientés $(\overrightarrow{AB}, \overrightarrow{AC})$, $(\overrightarrow{BC}, \overrightarrow{BA})$, $(\overrightarrow{CA}, \overrightarrow{CB})$.

Proposition 9.16. a) Les mesures des trois angles $\widehat{A}$, $\widehat{B}$, $\widehat{C}$ sont de même signe, positives si le triangle est direct, négatives dans le cas contraire.
 b) On a la formule $[A, B, C] = bc \sin \widehat{A} = ca \sin \widehat{B} = ab \sin \widehat{C}$.
 c) La somme $S = \widehat{A} + \widehat{B} + \widehat{C}$ est égale à π ou à $-\pi$ selon que le triangle ABC est direct ou inverse.

Démonstration. Comme les signes de $\widehat{A}$, $\widehat{B}$, $\widehat{C}$ sont ceux de leur sinus, il suffit de prouver b) pour obtenir a).

On a $[A, B, C] = \det(\overrightarrow{AB}, \overrightarrow{AC})$. Il est clair que b) est conséquence du lemme suivant :

Lemme 9.17. Expression du déterminant.

Soit $\vec{v}$ et $\vec{w}$ deux vecteurs non nuls. On a

$$\det(\vec{v}, \vec{w}) = \|\vec{v}\| \times \|\vec{w}\| \times \sin(\widehat{\vec{v}, \vec{w}}).$$

Soit $\vec{v}$ le vecteur unitaire colinéaire de même sens que $\vec{v}$ et $\vec{j}$ le vecteur unitaire directement orthogonal. On a

$$\vec{v} = \|\vec{v}\| \vec{v}, \vec{w} = \|\vec{w}\| (\cos(\widehat{\vec{v}, \vec{w}}) \vec{v} + \sin(\widehat{\vec{v}, \vec{w}}) \vec{j})$$

$$\det(\vec{v}, \vec{w}) = \det \begin{pmatrix} \|\vec{v}\| & \|\vec{w}\| \cos(\widehat{\vec{v}, \vec{w}}) \\ 0 & \|\vec{w}\| \sin(\widehat{\vec{v}, \vec{w}}) \end{pmatrix} = \|\vec{v}\| \times \|\vec{w}\| \times \sin(\widehat{\vec{v}, \vec{w}})$$

c) Par la relation de Chasles, on a les congruences d'angles modulo 2π

$$\begin{aligned} S &\equiv (\widehat{\overrightarrow{AB}, \overrightarrow{AC}}) + (\widehat{\overrightarrow{BC}, \overrightarrow{BA}}) + (\widehat{\overrightarrow{CA}, \overrightarrow{CB}}) \\ &\equiv (\widehat{\overrightarrow{AB}, \overrightarrow{AC}}) + (\widehat{\overrightarrow{AC}, \overrightarrow{BC}}) + (\widehat{\overrightarrow{BC}, \overrightarrow{BA}}) \equiv (\widehat{\overrightarrow{AB}, \overrightarrow{BA}}) \equiv \pi \end{aligned}$$

Donc S est de la forme $S = (2k + 1)\pi$ où $k \in \mathbb{Z}$.

Si le triangle est direct, $\widehat{A}$, $\widehat{B}$, $\widehat{C}$ sont dans l'intervalle ouvert $]0, \pi[$, donc $S = (2k + 1)\pi$ est dans l'intervalle ouvert $]0, 3\pi[$. La seule valeur possible est $k = 0$ d'où $S = \pi$. Le cas où le triangle est inverse se traite de même. $\square$

Dans la suite, ABC est direct, p est le demi-périmètre, ($2p = a + b + c$), O est centre du cercle circonscrit et R le rayon. Nous utiliserons le résultat suivant, conséquence du **théorème de l'angle inscrit** (11.15, p. 292) :

$$(\widehat{OB, OC}) \equiv 2\widehat{A}, (\widehat{OC, OA}) \equiv 2\widehat{B}, (\widehat{OA, OB}) \equiv 2\widehat{C} \text{ modulo } 2\pi$$

Proposition 9.18. Avec les notations précédentes, on a

$$\begin{aligned} a^2 &= b^2 + c^2 - 2bc \cos \widehat{A}, & a &= 2R \sin \widehat{A} \\ b^2 &= c^2 + a^2 - 2ca \cos \widehat{B}, & b &= 2R \sin \widehat{B} \\ c^2 &= a^2 + b^2 - 2ab \cos \widehat{C}, & c &= 2R \sin \widehat{C} \end{aligned}$$

Démonstration. Pour commencer

$$a^2 = \|\vec{BC}\|^2 = \|\vec{AC} - \vec{AB}\|^2 = AC^2 - 2\vec{AC} \cdot \vec{AB} + AB^2 = b^2 + c^2 - 2bc \cos \widehat{A}$$

On a de même

$$a^2 = \|\vec{BC}\|^2 = \|\vec{OC} - \vec{OB}\|^2 = OC^2 - 2\vec{OC} \cdot \vec{OB} + OB^2 = 2R^2 - 2R^2 \cos 2\widehat{A} = 4R^2 \sin^2 \widehat{A}$$

d'où l'on déduit $a = 2R \sin \widehat{A}$. Les autres formules s'obtiennent de la même façon. $\square$

9.4 FONCTION DE LEIBNITZ

Proposition 9.19. Fonction de Leibnitz. Soit (α_i, A_i) une famille de points pondérés. On appelle fonction de Leibnitz l'application

$$\varphi: \mathcal{E} \longrightarrow \mathbb{R}, \quad M \longmapsto \sum_i \alpha_i MA_i^2$$

(i) Si $\sum \alpha_i \neq 0$, pour tout point M , on a $\varphi(M) = (\sum \alpha_i)MA^2 + \varphi(A)$ où A est barycentre des (α_i, A_i) .

(ii) Si $\sum_i \alpha_i = 0$ et $\vec{v} = \sum \alpha_i A_i \neq \vec{0}$, pour tout couple de points M, M_0 dans $\mathcal{P}$, on a $\varphi(M) = \varphi(M_0) + \vec{MM_0} \cdot \vec{v}$.

(iii) Si $\sum_i \alpha_i = 0$ et $\vec{v} = \sum \alpha_i A_i = \vec{0}$, la fonction φ est constante.

Démonstration. (i) Dans le cas $\sum \alpha_i \neq 0$, le point A vérifie $(\sum \alpha_i)A = \sum \alpha_i A_i$, et pour tout M ,

$$\begin{aligned} \varphi(M) &= \sum_i \alpha_i \|\vec{MA_i}\|^2 = \sum_i \alpha_i \|\vec{MA} + \vec{AA_i}\|^2 = \sum_i \alpha_i (MA^2 + 2\vec{MA} \cdot \vec{AA_i} + AA_i^2) \\ &= \left(\sum_i \alpha_i \right) MA^2 + 2\vec{MA} \cdot \left(\sum_i \alpha_i \vec{AA_i} \right) + \sum_i \alpha_i AA_i^2 = \left(\sum_i \alpha_i \right) MA^2 + \varphi(A) \end{aligned}$$

La fonction φ a donc un maximum ou minimum en A selon le signe de $\sum \alpha_i$. L'ensemble des M où $\varphi(M) = \lambda$ est un cercle de centre A , ou est vide selon la place de λ par rapport à $\varphi(A)$.

(ii) et (iii) Dans le cas où $\sum \alpha_i = 0$, $\sum \alpha_i \mathbf{A}_i$ est un vecteur constant $\vec{v}$. Soit M et M_0 deux points, alors

$$\begin{aligned}\varphi(M) &= \sum_i \alpha_i \|\overrightarrow{MA_i}\|^2 = \sum_i \alpha_i \|\overrightarrow{MM_0} + \overrightarrow{M_0A_i}\|^2 = \sum_i \alpha_i \left(MM_0^2 + 2\overrightarrow{MM_0} \cdot \overrightarrow{M_0A_i} + M_0A_i^2 \right) \\ &= \left(\sum_i \alpha_i \right) MM_0^2 + 2\overrightarrow{MM_0} \cdot \left(\sum_i \alpha_i \overrightarrow{M_0A_i} \right) + \sum_i \left(\alpha_i M_0A_i^2 \right) = 2\overrightarrow{MM_0} \cdot \vec{v} + \varphi(M_0)\end{aligned}$$

(ii) Si $\sum \alpha_i \mathbf{A}_i = \vec{v} \neq \vec{0}$, φ prend toute valeur de $\mathbb{R}$, l'ensemble des M où $\varphi(M) = \lambda$ est une droite orthogonale à $\vec{v}$.

(iii) Si $\sum \alpha_i \mathbf{A}_i = \vec{0}$, la fonction φ est constante. $\square$

EXERCICES

Exercice 9.1. Théorème de Gergonne Soit un triangle ABC, un point M tel que (MA), (MB), (MC) coupent (BC), (CA), (AB) en A' , B' , C' , x, y, z le système de coordonnées barycentriques de M tel que $x + y + z = 1$. Montrer que

$$\frac{\overline{A'M}}{\overline{A'A}} + \frac{\overline{B'M}}{\overline{B'B}} + \frac{\overline{C'M}}{\overline{C'C}} = 1$$

Exercice 9.2. Soit ABC un triangle, M non sur les côtés du triangle de coordonnées barycentriques x, y, z telles que $x + y + z = 1$, A', B', C' les points où (MA), (MB), (MC) coupent (BC), (CA), (AB), A'', B'', C'' les milieux de (B', C') , (C', A') , (A', B') . Montrer que (AA'') , (BB'') , (CC'') sont concourantes ou parallèles.

Exercice 9.3. Soit A, B deux points distincts d'une droite $\mathcal{D}$, x, y deux scalaires non nuls distincts et non opposés. Montrer que les points $P = \frac{x\mathbf{A} + y\mathbf{B}}{x+y}$ et $Q = \frac{x\mathbf{A} - y\mathbf{B}}{x-y}$ sont conjugués harmoniques relativement à A et B.

Exercice 9.4. Soit (f_i) une famille finie d'applications affines de $\mathcal{E}$ dans lui-même et (λ_i) une famille de réels tels que $\sum \lambda_i = 1$.

1) Montrer que l'application $f = \sum \lambda_i f_i: \mathcal{E} \rightarrow \mathcal{E}, M \mapsto \sum \lambda_i f_i(M)$ est affine.

2) Soit quatre points A, B, C, D d'un plan $\mathcal{P}$ tels qu'il n'y ait pas alignement de trois d'entre eux. À tout $P \in (AB)$, on associe $Q \in (BC)$ tel que (PQ) soit parallèle à (AC) et $S \in (AD)$ tel que (PS) soit parallèle à (BD) . On définit R par $\overrightarrow{PR} = \overrightarrow{PQ} + \overrightarrow{PS}$. Quel est le lieu géométrique de R ?

Exercice 9.5. Soit un triangle ABC , $S = [A, B, C]$, des points P, Q, R sur (BC) , (CA) , (AB) , A'', B'', C'' les milieux de (A, P) , (B, Q) , (C, R) , A', B', C' les points donnés par $\overrightarrow{AA'} = \overrightarrow{AQ} + \overrightarrow{AR}$, $\overrightarrow{BB'} = \overrightarrow{BR} + \overrightarrow{BP}$, $\overrightarrow{CC'} = \overrightarrow{CP} + \overrightarrow{CQ}$.

Comparer $[P, Q, R]$, $[A', B', C']$ et $[A'', B'', C'']$.

Exercice 9.6. Dans un plan affine orienté, soit un triangle ABC direct, Δ une droite coupant (BC) , (CA) , (AB) en respectivement P, Q, R .

1) Montrer que $[A, Q, R] + [P, B, R] + [P, Q, C] = -[A, B, C]$.

2 Montrer que parmi les quantités $[A, Q, R]$, $[P, B, R]$, $[P, Q, C]$, deux sont négatives et une est positive.

Exercice 9.7. On appelle quadrilatère une suite de quatre points d'un plan. Soit Q un quadrilatère P, Q, R, S tel que trois quelconques de ces points ne sont pas alignés.

Montrer que $[P, Q, R] + [Q, S, R] + [P, R, S] + [S, Q, P] = 0$.

En déduire que pour la convexité, il existe deux cas de figure pour le quadrilatère Q .

Exercice 9.8. Étant donné un triangle direct ABC , pour tout M non sur les côtés, on note $\alpha(M)$, $\beta(M)$, $\gamma(M)$ les mesures des angles $(\widehat{MB, MC})$, $(\widehat{MC, MA})$, $(\widehat{MA, MB})$ comprises entre $-\pi$ et $+\pi$. Montrer que

(i) si M est intérieur au triangle, on a $\alpha(M) + \beta(M) + \gamma(M) = 2\pi$,

(ii) si M est extérieur au triangle, on a $\alpha(M) + \beta(M) + \gamma(M) = 0$.

Exercice 9.9. Soit un triangle ABC .

Montrer que $(\sin 2\hat{A}, \sin 2\hat{B}, \sin 2\hat{C})$ est un système de coordonnées barycentriques du centre O du cercle circonscrit au triangle. À quelle condition O est intérieur au triangle ? Quelle expression de la surface peut-on déduire ?

Exercice 9.10. Soit un triangle ABC , I le centre du cercle inscrit et r son rayon, J, K, L les centres et r_A, r_B, r_C les rayons des cercles exinscrits relatifs aux sommets A, B, C . Montrer que (a, b, c) , $(-a, b, c)$, $(a, -b, c)$, $(a, b, -c)$ sont des systèmes de coordonnées barycentriques pour I, J, K, L . En déduire

$$\frac{r}{r_A} + \frac{r}{r_B} + \frac{r}{r_C} = 1$$

Exercice 9.11. Soit un triangle ABC de cercle circonscrit de centre O et de rayon R . Montrer que le cercle circonscrit est l'ensemble des points M tels que

$$MA^2 \sin 2\hat{A} + MB^2 \sin 2\hat{B} + MC^2 \sin 2\hat{C} = 4R^2 \sin \hat{A} \sin \hat{B} \sin \hat{C} = \frac{abc}{2R}$$

Exercice 9.12. Dans un plan euclidien $\mathcal{P}$, soit quatre points distincts A, B, C, D en situation orthocentrique, PQR le triangle orthique (8.36, p. 225).

1) Montrer l'égalité $\vec{AB} \cdot \vec{AC} = \vec{AC} \cdot \vec{AD} = \vec{AD} \cdot \vec{AB}$, la valeur commune de ces produits scalaires est notée α . On définit de même β, γ, δ .

2) Montrer que $\alpha + \beta = AB^2$ et cinq autres égalités analogues.

3) Montrer qu'un système de coordonnées barycentriques de D relativement à A, B, C est $(\frac{1}{\alpha}, \frac{1}{\beta}, \frac{1}{\gamma})$.

4) Montrer dans $\mathbb{R}$ et $\tilde{\mathcal{P}}$ les égalités

$$\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} + \frac{1}{\delta} = 0, \quad \frac{1}{\alpha}A + \frac{1}{\beta}B + \frac{1}{\gamma}C + \frac{1}{\delta}D = \vec{0}$$

5) Que peut-on en déduire quant à la convexité ?

Exercice 9.13. Soit un triangle ABC non rectangle et non équilatéral, D l'orthocentre, O le centre du cercle circonscrit, G le centre de gravité, Ω le centre du cercle des neuf points, R le rayon du cercle circonscrit, a, b, c les longueurs des côtés.

1) Montrer que Ω est isobarycentre de A, B, C, D.

2) Montrer que la droite d'Euler de ABC est l'ensemble des M tels que

$$(b^2 - c^2)MA^2 + (c^2 - a^2)MB^2 + (a^2 - b^2)MC^2 = 0$$

3) Montrer que $\Omega A^2 + \Omega B^2 + \Omega C^2 + \Omega D^2 = 3R^2$.

4) En déduire que le cercle des neuf points est l'ensemble des M tels que

$$MA^2 + MB^2 + MC^2 + MD^2 = 4R^2$$

PROBLÈME

Le corrigé de ce problème est disponible sur le site de Dunod : www.dunod.com.

9.1. CONVEXES

Soit $\mathcal{P}$ un plan euclidien. Une partie $C \subset \mathcal{P}$ est dite **convexe** si, pour tout couple (P, Q) de points de C , le segment $[P, Q]$ est contenu dans C .

1) Montrer qu'une partie C de $\mathcal{P}$ est convexe si et seulement si, pour toute famille de points M_i de C et toute famille de coefficients $\lambda_i \geq 0$ telle que $\sum \lambda_i = 1$, le barycentre $M = \sum \lambda_i M_i$ appartient à C .

2) Montrer qu'un demi-plan est convexe.

3) Montrer que l'intérieur d'un cercle est convexe.

4) Montrer que l'adhérence d'un convexe est convexe.

Dans la suite, C désigne un convexe fermé.

5) Soit un point $A \notin C$. Montrer qu'il existe un unique point $A' \in C$ tel que, pour tout $M \in C$ distinct de A' , on ait $AA' < AM$ (commencer par le cas où C est compact). On dit que A' est la **projection** de A sur C .

6) Dans ces conditions, soit Δ la droite orthogonale en A' à (AA') . Montrer que C est situé dans le demi-plan fermé limité par Δ ne contenant pas A .

7) En déduire que les convexes fermés distincts de $\mathcal{P}$ sont les intersections de familles de demi-plans fermés.

8) On considère l'application projection $A \mapsto A'$ du complémentaire de C dans C . Montrer que cette application est contractante.

SOLUTIONS DES EXERCICES

Solution 9.1. On a $M = xA + yB + zC$. Les plans vectoriels $\text{Vect}(A, M)$ et $\text{Vect}(B, C)$ se coupent suivant la droite vectorielle $\text{Vect}(A')$. On a donc $M - xA = yB + zC = (y + z)A'$, d'où

$$\overrightarrow{A'M} - x\overrightarrow{A'A} = \overrightarrow{0} \text{ soit } \frac{\overrightarrow{A'M}}{\overrightarrow{A'A}} = x$$

et deux autres formules analogues. Il suffit d'additionner ces trois formules.

Solution 9.2. On a $M - xA = yB + zC = (y + z)A'$ et de même $zC + xA = (z + x)B'$. Donc

$$\begin{aligned} 2(y + z)(z + x)C'' &= (z + x)(yB + zC) + (y + z)(zC + xA) \\ &= (y + z)xA + (z + x)yB + (x + y + 2z)zC \end{aligned}$$

$2(y+z)(z+x)\mathbf{C}'' - 2z^2\mathbf{C} = (y+z)x\mathbf{A} + (z+x)y\mathbf{B} + (x+y)z\mathbf{C}$
 et deux autres formules analogues. Les droites (AA'') , (BB'') , (CC'') ,

- ou bien concourent au point de coordonnées barycentriques

$$(yx + zx, zy + xy, xz + yz)$$

si la somme de ces coefficients est non nulle,

- ou bien ont la direction du vecteur $\vec{v}$ ayant ces nombres pour coordonnées barycentriques si la somme en est nulle. Comme M n'est pas sur les côtés (BC), (CA), (AB), x, y, z sont non nuls. Comme $x + y + z = 1$, les trois quantités $yx + zx, zy + xy, xz + yz$ ne sont pas toutes nulles et $\vec{v} \neq \vec{0}$.

Solution 9.3. Comme $x + y$ et $x - y$ sont non nuls, ces points P, Q existent. On a $x\vec{PA} + y\vec{PB} = \vec{0}$ et $x\vec{QA} - y\vec{QB} = \vec{0}$. On en déduit $[A, B, P, Q] = -1$:

$$\frac{\overline{PA}}{\overline{PB}} = -\frac{y}{x} = -\frac{\overline{QA}}{\overline{QB}}$$

Solution 9.4. 1) Soit l'application $\varphi = \sum \lambda_i \tilde{f}_i$ linéaire de $\tilde{\mathcal{E}}$ dans lui-même. Comme $\sum \lambda_i = 1$, on a $\mu \circ \varphi = \varphi$. Donc f est affine et $\varphi = \tilde{f}$ (9.7, p. 248).

2) On a $R = Q + S - P$. Soit u (resp. v) la projection de $\mathcal{P}$ sur (BC) parallèlement à (AC) (resp. (AD) parallèlement à (BD)). L'application $u + v - I$ envoie P en R. Le lieu de R est l'image de (AB), donc est une droite.

Si $P = B$, alors $Q = B$ et $S = D$, donc $R = D$.

Si $P = A$, alors $Q = C$ et $S = A$, donc $R = C$.

Le lieu de R est donc la droite (DC).

Solution 9.5. Soit les coordonnées barycentriques de P, Q, R par rapport à (A, B, C)

$$P \begin{pmatrix} 0 \\ \lambda \\ 1 - \lambda \end{pmatrix}, \quad Q \begin{pmatrix} 1 - \mu \\ 0 \\ \mu \end{pmatrix}, \quad R \begin{pmatrix} \nu \\ 1 - \nu \\ 0 \end{pmatrix}$$

$$[P, Q, R] = \det \begin{pmatrix} 0 & 1 - \mu & \nu \\ \lambda & 0 & 1 - \nu \\ 1 - \lambda & \mu & 0 \end{pmatrix} S = (1 - \lambda - \mu - \nu + \mu\nu + \nu\lambda + \lambda\mu)S$$

a) On a $A' = Q + R - A$, $B' = R + P - B$, $C' = P + Q - C$, donc A', B', C' ont pour coordonnées barycentriques

$$A' \begin{pmatrix} \nu - \mu \\ 1 - \nu \\ \mu \end{pmatrix}, \quad B' \begin{pmatrix} \nu \\ \lambda - \nu \\ 1 - \lambda \end{pmatrix}, \quad C' \begin{pmatrix} 1 - \mu \\ \lambda \\ \mu - \lambda \end{pmatrix}$$

$$[A', B', C'] = \det \begin{pmatrix} \nu - \mu & \nu & 1 - \mu \\ 1 - \nu & \lambda - \nu & \lambda \\ \mu & 1 - \lambda & \mu - \lambda \end{pmatrix} S = (1 - \lambda - \mu - \nu + \mu\nu + \nu\lambda + \lambda\mu)S$$

Les aires de ces triangles sont donc égales. En particulier, si P, Q, R sont alignés, il en est de même de A', B', C'.

b) On a $2A'' = A + P$, $2B'' = B + Q$, $2C'' = C + R$, donc A'', B'', C'' ont pour coordonnées barycentriques

$$2A'' \begin{pmatrix} 1 \\ \lambda \\ 1 - \lambda \end{pmatrix}, \quad 2B'' \begin{pmatrix} 1 - \mu \\ 1 \\ \mu \end{pmatrix}, \quad 2C'' \begin{pmatrix} \nu \\ 1 - \nu \\ 1 \end{pmatrix}$$

$$[A'', B'', C''] = \frac{1}{8} \det \begin{pmatrix} 1 & 1 - \mu & \nu \\ \lambda & 1 & 1 - \nu \\ 1 - \lambda & \mu & 1 \end{pmatrix} S = \frac{1}{4} (1 - \lambda - \mu - \nu + \mu\nu + \nu\lambda + \lambda\mu) S$$

En particulier, si P, Q, R sont alignés, il en est de même de A'', B'', C''.

Solution 9.6. a) On définit α, β, γ par $P = \alpha B + (1 - \alpha)C$, $Q = \beta C + (1 - \beta)A$, $R = \gamma A + (1 - \gamma)B$. On a

$$[A, Q, R] = \det \begin{pmatrix} 1 & 1 - \beta & \gamma \\ 0 & 0 & 1 - \gamma \\ 0 & \beta & 0 \end{pmatrix} [A, B, C] = \beta(\gamma - 1)[A, B, C]$$

et deux autres formules analogues. Le théorème de Ménélaüs s'écrit (p. 219)

$$\frac{\alpha}{\alpha - 1} \times \frac{\beta}{\beta - 1} \times \frac{\gamma}{\gamma - 1} = +1 \quad \text{soit} \quad \alpha + \beta + \gamma - \beta\gamma - \gamma\alpha - \alpha\beta = 1$$

En additionnant les expressions de $[A, Q, R]$, $[P, B, R]$, $[P, Q, C]$, on obtient le résultat.

b) En multipliant ces aires algébriques, on obtient

$$\begin{aligned} [A, Q, R] \times [P, B, R] \times [P, Q, C] &= \alpha\beta\gamma(\alpha - 1)(\beta - 1)(\gamma - 1)[A, B, C]^3 \\ &= \frac{\alpha}{\alpha - 1} \times \frac{\beta}{\beta - 1} \times \frac{\gamma}{\gamma - 1} (\alpha - 1)^2(\beta - 1)^2(\gamma - 1)^2 [A, B, C]^3 \\ &= (\alpha - 1)^2(\beta - 1)^2(\gamma - 1)^2 [A, B, C]^3 > 0 \end{aligned}$$

Comme la somme est $-[A, B, C] < 0$, deux de ces quantités sont négatives et la troisième positive. Et cela quelle que soit la position de la droite Δ relativement au triangle. Le lecteur est invité à faire des figures.

Solution 9.7. Il existe $\binom{2}{4} = 6$ paires de points et 3 partitions de Q en 2 paires disjointes : $(\{P, Q\}, \{R, S\})$, $(\{P, R\}, \{Q, S\})$, $(\{P, S\}, \{Q, R\})$. Une structure de **quadrilatère** sur Q est la donnée de 2 paires disjointes appelées **diagonales** du quadrilatère, les 4 autres étant appelées **côtés**.

Considérant (P, Q, R) comme repère barycentrique, on a

$$[P, Q, R] = [S, Q, R] + [P, S, R] + [P, Q, S]$$

d'où la formule : *Pour tout sommet $X \in \{P, Q, R, S\}$, on forme $[Y, Z, T]$ où Y, Z, T sont les autres sommets rangés de sorte que les permutations X, Y, Z, T de Q aient toutes même signature. La somme des quatre aires algébriques obtenues est nulle.*

Pour la convexité, on a deux possibilités :

- (1) Le cas non convexe : parmi les quatre aires de somme nulle, trois $[Q, S, R]$, $[P, R, S]$, $[S, Q, P]$, sont d'un signe, $[P, Q, R]$ étant de l'autre signe. Le sommet S se singularise comme intérieur au triangle PQR . En revanche, les trois quadrilatères sont de même type avec une diagonale intérieure ayant S pour extrémité et l'autre extérieure.
- (2) Le cas convexe : Parmi ces aires de somme nulle, deux, $[P, Q, R]$ et $[Q, S, R]$, sont d'un signe, les deux autres, $[P, R, S]$ et $[S, Q, P]$, de l'autre signe. Comme $[P, Q, R]$ et $[S, Q, R]$ sont de signes contraires, P et S sont de part et d'autre de (QR) . De même Q et R sont de part et d'autre de (PS) . Les segments $[P, S]$ et $[Q, R]$ se coupent. Aucun des sommets ne se singularise, mais parmi les trois quadrilatères, il en est un **convexe** de diagonales sécantes. Les deux autres sont dits **croisés**.

Solution 9.8. Appliquons la relation de Chasles sur les angles orientés. Modulo 2π ,

$$\alpha(M) + \beta(M) + \gamma(M) \equiv (\widehat{MB, MC}) + (\widehat{MC, MA}) + (\widehat{MA, MB}) \equiv (\widehat{MB, MB}) \equiv 0$$

Donc il existe $k \in \mathbb{Z}$ tel que $\alpha(M) + \beta(M) + \gamma(M) = 2k\pi$.

Comme $\alpha(M), \beta(M), \gamma(M)$ sont dans l'intervalle ouvert $]-\pi, +\pi[$, $\alpha(M) + \beta(M) + \gamma(M)$ est dans $] - 3\pi, +3\pi[$. Donc les valeurs possibles de k sont $-2, 0, 2$.

On sait que $\alpha(M)$ est du signe de $[M, B, C] = MB \times MC \times \sin \alpha(M)$ (9.15, p. 252). De même, $\beta(M)$ et $\gamma(M)$ sont du signe de $[A, M, C]$ et $[A, B, M]$.

Par ce qu'on sait sur l'intérieur d'un triangle (p. 253),

(i) si M est intérieur, $\alpha(M), \beta(M), \gamma(M)$ sont tous trois positifs, donc la seule valeur possible pour k est 2,

(ii) si M est extérieur, $\alpha(M), \beta(M), \gamma(M)$ ne sont pas tous de même signe. Supposons $\alpha(M) > 0$ et $\beta(M) < 0$, alors

$$-2\pi < \beta(M) + \gamma(M) < \alpha(M) + \beta(M) + \gamma(M) < \alpha(M) + \gamma(M) < 2\pi$$

donc la seule valeur possible de k est 0.

Solution 9.9. Par le théorème de l'angle inscrit, $\widehat{(\overrightarrow{OB}, \overrightarrow{OC})} = 2\widehat{A}$. Le résultat se déduit de

$$[O, B, C] = R^2 \sin 2\widehat{A}, \quad [O, C, A] = R^2 \sin 2\widehat{B}, \quad [O, A, B] = R^2 \sin 2\widehat{C}$$

Le point O est intérieur si $\sin 2\widehat{A}, \sin 2\widehat{B}, \sin 2\widehat{C}$ sont positifs, c'est-à-dire si les 3 angles sont aigus (p. 253).

On obtient une expression de l'aire (9.10, p. 249)

$$\frac{1}{2}[A, B, C] = \frac{1}{2}R^2(\sin 2\widehat{A} + \sin 2\widehat{B} + \sin 2\widehat{C}).$$

Solution 9.10. Sachant que l'aire d'un triangle vaut, en valeur absolue, la moitié du produit d'une hauteur par la base associée, on a

$$\begin{aligned} |[I, B, C]| &= ar, & |[I, C, A]| &= br, & |[I, A, B]| &= cr \\ |[J, B, C]| &= ar_A, & |[J, C, A]| &= br_A, & |[J, A, B]| &= cr_A \\ |[K, B, C]| &= ar_B, & |[K, C, A]| &= br_B, & |[K, A, B]| &= cr_B \\ |[L, B, C]| &= ar_C, & |[L, C, A]| &= br_C, & |[L, A, B]| &= cr_C \end{aligned}$$

On en déduit que des coordonnées barycentriques sont proportionnelles à des triplets de la forme $(\pm a, \pm b, \pm c)$.

Soit les points I', J', K', L' de coordonnées barycentriques (a, b, c) , $(-a, b, c)$, $(a, -b, c)$, $(a, b, -c)$. Les ensembles de points $\{I, J, K, L\}$ et $\{I', J', K', L'\}$ sont nécessairement identiques. Comme les sommes ci-dessous sont positives (inégalité triangulaire),

$$\begin{aligned} 2p &= a + b + c, & 2(p - a) &= b + c - a, \\ 2(p - b) &= c + a - b, & 2(p - c) &= a + b - c \end{aligned}$$

I' est intérieur au triangle, J', K', L' sont dans les régions occupées respectivement par J, K, L (p. 253). On a donc bien $I = I', J = J', K = K', L = L'$. On a de plus

$$[A, B, C] = 2pr = 2(p - a)r_A = 2(p - b)r_B = 2(p - c)r_C$$

dont on déduit

$$\frac{r}{r_A} + \frac{r}{r_B} + \frac{r}{r_C} = \frac{(p - a) + (p - b) + (p - c)}{p} = \frac{3p - 2p}{p} = 1$$

Solution 9.11. Le centre du cercle circonscrit a pour coordonnées barycentriques $\sin 2\hat{A}$, $\sin 2\hat{B}$, $\sin 2\hat{C}$ (9.9). Posons $\varphi(M) = MA^2 \sin 2\hat{A} + MB^2 \sin 2\hat{B} + MC^2 \sin 2\hat{C}$. Le cercle circonscrit est donc le lieu des points M tels que

$$\varphi(M) = \varphi(A) = c^2 \sin 2\hat{B} + b^2 \sin 2\hat{C}.$$

Comme $b = 2R \sin \hat{B}$ et $c = 2R \sin \hat{C}$, on a dans un premier temps

$$\varphi(A) = 4R^2 \sin \hat{B} \sin \hat{C} (\sin \hat{C} \cos \hat{B} + \sin \hat{B} \cos \hat{C}),$$

soit

$$\varphi(A) = 4R^2 \sin \hat{B} \sin \hat{C} \sin(\hat{B} + \hat{C}) = 4R^2 \sin \hat{B} \sin \hat{C} \sin \hat{A} = \frac{abc}{2R}$$

Solution 9.12. 1) On a $\vec{AB} \cdot \vec{AC} = \vec{AB} \cdot (\vec{AD} + \vec{DC}) = \vec{AB} \cdot \vec{AD}$ car $\vec{AB}$ et $\vec{DC}$ sont orthogonaux. Les autres égalités s'obtiennent par le même calcul.

2) De $\alpha = \vec{AB} \cdot \vec{AC}$, $\beta = \vec{BA} \cdot \vec{BC}$, on déduit $\alpha + \beta = \vec{AB} \cdot (\vec{AC} - \vec{BC}) = AB^2$ et de même cinq autres égalités analogues.

3) Soit (u, v, w) un système de coordonnées barycentriques de D. Multipliant scalairement $u\vec{AD} + v\vec{BD} + w\vec{CD} = \vec{0}$ par $\vec{AB}$, on a $u\alpha = v\beta$. Donc $u\alpha = v\beta = w\gamma$, d'où le résultat.

4) Soit O un point quelconque du plan, on a

$$\left(\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma}\right) \vec{OD} = \frac{1}{\alpha} \vec{OA} + \frac{1}{\beta} \vec{OB} + \frac{1}{\gamma} \vec{OC}$$

$$\left(\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} + \frac{1}{\delta}\right) \vec{OD} = \frac{1}{\alpha} \vec{OA} + \frac{1}{\beta} \vec{OB} + \frac{1}{\gamma} \vec{OC} + \frac{1}{\delta} \vec{OD}$$

Faisant jouer à A, B, C le rôle de D, on a trois égalités analogues. Les vecteurs $\vec{OA}$, $\vec{OB}$, $\vec{OC}$, $\vec{OD}$ étant distincts, la seule possibilité est

$$\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} + \frac{1}{\delta} = 0, \quad \frac{1}{\alpha} \vec{OA} + \frac{1}{\beta} \vec{OB} + \frac{1}{\gamma} \vec{OC} + \frac{1}{\delta} \vec{OD} = \vec{0}$$

5) De 2) et 4), on déduit que parmi $\alpha, \beta, \gamma, \delta$, l'un est négatif, les trois autres positifs. Donc $\{A, B, C, D\}$ n'est pas convexe. Parmi les quatre triangles extraits de A, B, C, D, un seul a ses trois angles aigus le quatrième point étant intérieur à ce triangle. Dans la configuration orthocentrique, ce point joue un rôle spécial pour la convexité.

Solution 9.13. 1) La disposition de D, G, Ω sur la droite d'Euler (8.35, p. 224) montre que $\Omega = 3G + D$. Comme $G = A + B + C$, on a le résultat.

2) L'ensemble Δ des M tels que $(b^2 - c^2)MA^2 + (c^2 - a^2)MB^2 + (a^2 - b^2)MC^2 = 0$ contient le centre du cercle circonscrit O. Puisque $3\vec{AG} = \vec{AC} + \vec{AB}$, $\vec{BC} = \vec{AC} - \vec{AB}$,

$$9GA^2 = b^2 + c^2 + 2\vec{AC} \cdot \vec{AB}, \quad a^2 = b^2 + c^2 - 2\vec{AC} \cdot \vec{AB}$$

d'où $9GA^2 = 2b^2 + 2c^2 - a^2$. En calculant de même GB^2 et GC^2 , on voit que $G \in \Delta$. Δ est la droite d'Euler (OG).

3) On a $\vec{\Omega A} = \vec{\Omega O} + \vec{OA}$, donc $\Omega A^2 = \Omega O^2 + OA^2 + 2\vec{\Omega O} \cdot \vec{OA}$ et

$$\Omega A^2 + \Omega B^2 + \Omega C^2 = 3\Omega O^2 + 3R^2 + 6\vec{\Omega O} \cdot \vec{OG}$$

La disposition de Ω, G, D, O sur la droite d'Euler donne $3\Omega O^2 + 6\vec{\Omega O} \cdot \vec{OG} = -\Omega D^2$, d'où le résultat.

4) Pour tout point M,

$$\varphi(M) = MA^2 + MB^2 + MC^2 + MD^2$$

donne

$$\varphi(M) = 4M\Omega^2 + \varphi(\Omega) = 4M\Omega^2 + 3R^2$$

Comme le cercle d'Euler est de rayon $\frac{1}{2}R$, on a bien le résultat annoncé.

Chapitre 10

Tétraèdres et parallélépipèdes

La géométrie du triangle a été développée aux chapitres 8 et 9. En dimension 3, tétraèdres et parallélépipèdes correspondent aux triangles et parallélogrammes. Mais la situation est plus compliquée. C'est ainsi que la notion d'orthocentre n'est pas toujours définie pour un tétraèdre quelconque.

Dans ce chapitre, $\mathcal{E}$ est un espace affine euclidien de dimension 3 d'espace vectoriel associé E , orienté.

10.1 PRODUIT MIXTE, PRODUIT VECTORIEL

10.1.1. Volumes algébriques

Définition 10.1. La formule de changement de base montre que le déterminant d'un système de trois vecteurs $\vec{V}_1, \vec{V}_2, \vec{V}_3$ relativement à une base orthonormale directe $\mathcal{B}$ ne dépend pas de $\mathcal{B}$. La quantité $\det_{\mathcal{B}}(\vec{V}_1, \vec{V}_2, \vec{V}_3)$ est appelée **produit mixte** de ce système, elle est notée $(\vec{V}_1, \vec{V}_2, \vec{V}_3)$.

Un **tétraèdre** $\mathcal{E}$ est la donnée de quatre points A_0, A_1, A_2, A_3 . Il est dit aplati si ces quatre points sont coplanaires.

Soit $\mathcal{B} = (\vec{i}, \vec{j}, \vec{k})$ une base orthonormale directe et O un point. Alors $(O, \mathcal{B})$ est base de l'espace vectoriel $\tilde{\mathcal{E}}$ de dimension 4 des points pondérés (9.3, p. 245 et p. 247).

Lemme 10.2. *Étant donné un tétraèdre A_0, A_1, A_2, A_3 , on a*

$$(\overrightarrow{A_0A_1}, \overrightarrow{A_0A_2}, \overrightarrow{A_0A_3}) = \det_{\mathcal{B}}(\overrightarrow{A_0A_1}, \overrightarrow{A_0A_2}, \overrightarrow{A_0A_3}) = \det_{0,\mathcal{B}}(A_0, A_1, A_2, A_3)$$

Démonstration. Ce lemme correspond à la proposition 9.11 (p. 250) relative aux triangles. Dans l'espace vectoriel $\tilde{\mathcal{E}}$ de dimension 4, on pose $V = \det_{0,\mathcal{B}}(A_0, A_1, A_2, A_3)$, $\overrightarrow{OA_0} = a\vec{i} + b\vec{j} + c\vec{k}$, $A_0 = 0 + a\vec{i} + b\vec{j} + c\vec{k}$ et pour $1 \leq \alpha \leq 3$

$$\overrightarrow{A_0A_\alpha} = x_\alpha\vec{i} + y_\alpha\vec{j} + z_\alpha\vec{k},$$

$$A_\alpha = 0 + \overrightarrow{OA_0} + \overrightarrow{A_0A_\alpha} = 0 + (a + x_\alpha)\vec{i} + (b + y_\alpha)\vec{j} + (c + z_\alpha)\vec{k}.$$

$$\begin{aligned} V &= \det \begin{pmatrix} 1 & 1 & 1 & 1 \\ a & a+x_1 & a+x_2 & a+x_3 \\ b & b+y_1 & b+y_2 & b+y_3 \\ c & c+z_1 & c+z_2 & c+z_3 \end{pmatrix} \\ &= \det \begin{pmatrix} 1 & 0 & 0 & 0 \\ a & x_1 & x_2 & x_3 \\ b & y_1 & y_2 & y_3 \\ c & z_1 & z_2 & z_3 \end{pmatrix} \\ &= \det \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{pmatrix} = (\overrightarrow{A_0A_1}, \overrightarrow{A_0A_2}, \overrightarrow{A_0A_3}) \end{aligned}$$

en retranchant la première colonne du premier déterminant aux trois autres, puis en développant par rapport à la première ligne. $\square$

Définition 10.3. On note $[A_0, A_1, A_2, A_3]$ cette quantité. On appelle **volume algébrique** du tétraèdre orienté $A = (A_0, A_1, A_2, A_3)$ la quantité $\frac{1}{6}[A_0, A_1, A_2, A_3]$. On dit que A est **direct** (resp. **inverse**) si $[A_0, A_1, A_2, A_3]$ est positif, (resp. négatif). Une permutation des sommets conserve le volume algébrique ou le change en son opposé selon qu'elle est de signature 1 ou -1 .

Proposition 10.4. Soit A_0, A_1, A_2, A_3 un repère barycentrique et M un point.

Alors $[M, A_1, A_2, A_3]$, $[A_0, M, A_2, A_3]$, $[A_0, A_1, M, A_3]$, $[A_0, A_1, A_2, M]$ est l'unique système de coordonnées barycentriques de M vérifiant

$$[M, A_1, A_2, A_3] + [A_0, M, A_2, A_3] + [A_0, A_1, M, A_3] + [A_0, A_1, A_2, M] = [A_0, A_1, A_2, A_3]$$

Démonstration. Cette proposition joue le rôle du théorème 9.10 (p. 249). Soit x_0, x_1, x_2, x_3 le système de coordonnées barycentriques de M tel que

$$x_0 + x_1 + x_2 + x_3 = 1$$

On a $M = x_0 A_0 + x_1 A_1 + x_2 A_2 + x_3 A_3$. Posons $V_0 = [M, A_1, A_2, A_3]$, $V_1 = [A_0, M, A_2, A_3]$, $V_2 = [A_0, A_1, M, A_3]$, $V_3 = [A_0, A_1, A_2, M]$ et $V = [A_0, A_1, A_2, A_3]$. Alors (lemme 10.2)

$$V_0 = [M, A_1, A_2, A_3] = \left[\sum_{i=0}^3 x_i A_i, A_1, A_2, A_3 \right] = \sum_{i=0}^3 x_i [A_i, A_1, A_2, A_3] = x_0 V$$

Et de même $V_1 = x_1 V$, $V_2 = x_2 V$, $V_3 = x_3 V$. Donc V_0, V_1, V_2, V_3 , proportionnel à x_0, x_1, x_2, x_3 , est bien un système de coordonnées barycentriques de M . Comme $\sum x_i = 1$, on a bien $\sum V_i = V$. $\square$

Remarque : Ces résultats se généralisent en toute dimension. Il en est de même de 9.2.2 pour le répartitionnement (fonctions affines, demi-espaces, etc.). Les exercices 10.1 et 10.2 sont en dimension n .

10.1.2. Produit vectoriel

Définition 10.5. Le produit scalaire étant une forme bilinéaire non dégénérée, l'application de E vers son dual E^* associant à un vecteur $\vec{W}$ la forme linéaire $\vec{X} \mapsto \vec{W} \cdot \vec{X}$ est bijective. Pour deux vecteurs $\vec{U}, \vec{V}$. L'application

$$E \longrightarrow \mathbb{R}, \quad \vec{X} \longmapsto (\vec{U}, \vec{V}, \vec{X})$$

est une forme linéaire. Donc il existe un unique vecteur $\vec{W} = \vec{U} \wedge \vec{V}$, appelé **produit vectoriel** de $\vec{U}$ et $\vec{V}$ (dans cet ordre), tel que

$$\forall \vec{X} \in E, \quad (\vec{U} \wedge \vec{V}) \cdot \vec{X} = \vec{W} \cdot \vec{X} = (\vec{U}, \vec{V}, \vec{X})$$

Proposition 10.6. Soit le produit vectoriel $\vec{W} = \vec{U} \wedge \vec{V}$ de $\vec{U}$ et $\vec{V}$.

- (i) $\vec{U} \wedge \vec{V} = \vec{0}$ si et seulement si $(\vec{U}, \vec{V})$ est lié,
- (ii) $(\vec{U}, \vec{V}, \vec{U} \wedge \vec{V})$ est une base directe de E si $(\vec{U}, \vec{V})$ est libre,
- (iii) $\vec{U} \wedge \vec{V}$ est orthogonal à $\vec{U}$ et $\vec{V}$,
- (iv) $\|\vec{U} \wedge \vec{V}\| = \|\vec{U}\| \times \|\vec{V}\| \times \sin \theta$ où $\theta \in [0, \pi]$ est l'angle (non orienté) de $\vec{U}$ et $\vec{V}$.

Démonstration. C'est une description complète du produit vectoriel.

(i) Le produit scalaire étant une forme bilinéaire non dégénérée, $\vec{U} \wedge \vec{V} = \vec{0}$ si et seulement si $(\vec{U} \wedge \vec{V}) \cdot \vec{X} = (\vec{U}, \vec{V}, \vec{X}) = 0$ pour tout $\vec{X}$, donc si $\vec{U}, \vec{V}$ ne fait partie d'aucune base, donc si $\vec{U}, \vec{V}$ sont colinéaires.

(ii) On a $(\vec{U} \wedge \vec{V}) \cdot \vec{U} = (\vec{U}, \vec{V}, \vec{U}) = 0$, donc $\vec{U} \wedge \vec{V}$ est orthogonal à $\vec{U}$ et de même à $\vec{V}$.

(iii) Supposons $\vec{U}, \vec{V}$ non colinéaires. On a

$$(\vec{U}, \vec{V}, \vec{U} \wedge \vec{V}) = (\vec{U} \wedge \vec{V}) \cdot (\vec{U} \wedge \vec{V}) = \|\vec{U} \wedge \vec{V}\|^2 > 0$$

donc $(\vec{U}, \vec{V}, \vec{U} \wedge \vec{V})$ est une base directe de E .

(iv) Si $\vec{U}, \vec{V}$ colinéaires, $\theta \in \{0, \pi\}$, $\sin \theta = 0$, le résultat est clair.

Supposons $\vec{U}, \vec{V}$ non colinéaires. Orientons $\Pi = \text{Vect}(\vec{U}, \vec{V})$ de sorte que $(\vec{U}, \vec{V})$ en soit une base directe. Soit la base orthonormale directe $\mathcal{B} = (\vec{u}, \vec{v}, \vec{w})$ où $\vec{u}$ est colinéaire de même sens à $\vec{U}$, $\vec{v}$ dans Π directement orthogonal à $\vec{v}$. On déduit le résultat de

$$\begin{aligned} \|\vec{U} \wedge \vec{V}\|^2 &= \det_{\mathcal{B}}(\vec{U}, \vec{V}, \vec{U} \wedge \vec{V}) = \begin{vmatrix} \|\vec{U}\| & \|\vec{V}\| \cos \theta & 0 \\ 0 & \|\vec{V}\| \sin \theta & 0 \\ 0 & 0 & \|\vec{U} \wedge \vec{V}\| \end{vmatrix} \\ &= \|\vec{U}\| \times \|\vec{V}\| \sin \theta \times \|\vec{U} \wedge \vec{V}\| \end{aligned}$$

□

Remarquons que $\|\vec{U} \wedge \vec{V}\|$ est l'aire du parallélogramme associé à $\vec{U}$ et $\vec{V}$.

Proposition 10.7. *L'application $E \times E \rightarrow E$, $(\vec{U}, \vec{V}) \mapsto \vec{U} \wedge \vec{V}$ est bilinéaire alternée.*

Démonstration. Prouvons qu'elle est alternée. Pour tout $\vec{U}, \vec{V}, \vec{X}$,

$$(\vec{V} \wedge \vec{U}) \cdot \vec{X} = (\vec{V}, \vec{U}, \vec{X}) = -(\vec{U}, \vec{V}, \vec{X}) = -(\vec{U} \wedge \vec{V}) \cdot \vec{X}$$

Donc $\vec{V} \wedge \vec{U} + \vec{U} \wedge \vec{V}$ est nul car orthogonal à tout $\vec{X}$.

Il suffit de prouver que l'application est linéaire relativement à la première variable.

Soit $\vec{U}_1, \vec{U}_2 \in E \times E$, $\lambda_1, \lambda_2 \in \mathbb{R} \times \mathbb{R}$, $\vec{V} \in E$. Pour tout $\vec{X}$, on a

$$\begin{aligned} ((\lambda_1 \vec{U}_1 + \lambda_2 \vec{U}_2) \wedge \vec{V}) \cdot \vec{X} &= (\lambda_1 \vec{U}_1 + \lambda_2 \vec{U}_2, \vec{V}, \vec{X}) \\ &= \lambda_1 (\vec{U}_1, \vec{V}, \vec{X}) + \lambda_2 (\vec{U}_2, \vec{V}, \vec{X}) \\ &= \lambda_1 (\vec{U}_1 \wedge \vec{V}) \cdot \vec{X} + \lambda_2 (\vec{U}_2 \wedge \vec{V}) \cdot \vec{X} \\ &= (\lambda_1 \vec{U}_1 \wedge \vec{V} + \lambda_2 \vec{U}_2 \wedge \vec{V}) \cdot \vec{X} \end{aligned}$$

d'où $(\lambda_1 \vec{U}_1 + \lambda_2 \vec{U}_2) \wedge \vec{V} - (\lambda_1 \vec{U}_1 \wedge \vec{V} + \lambda_2 \vec{U}_2 \wedge \vec{V}) = \vec{0}$ car ce vecteur est orthogonal à tout $\vec{X}$. □

Proposition 10.8. Double produit vectoriel. *Soit trois vecteurs $\vec{A}, \vec{B}, \vec{C}$. On a $(\vec{A} \wedge \vec{B}) \wedge \vec{C} = (\vec{A} \cdot \vec{C}) \vec{B} - (\vec{B} \cdot \vec{C}) \vec{A}$.*

Démonstration. Posons $\vec{P} = (\vec{A} \wedge \vec{B}) \wedge \vec{C}$. Si $(\vec{A}, \vec{B})$ est lié, $\vec{P} = \vec{0}$ et la formule est vraie. Supposons $\vec{A}, \vec{B}$ non colinéaires.

Alors $\vec{P}$, orthogonal à $\vec{A} \wedge \vec{B}$, est dans $\text{Vect}(\vec{A}, \vec{B})$, donc de la forme $\vec{P} = x\vec{A} + y\vec{B}$. Comme $\vec{P}$ est orthogonal à $\vec{C}$, $x\vec{A} \cdot \vec{C} + y\vec{B} \cdot \vec{C} = 0$, donc il existe λ tel que $x = -\lambda\vec{B} \cdot \vec{C}$ et $y = \lambda\vec{A} \cdot \vec{C}$, i.e. tel que

$$\vec{P} = \lambda((\vec{A} \cdot \vec{C})\vec{B} - (\vec{B} \cdot \vec{C})\vec{A})$$

Étant donné $\lambda \in \mathbb{R}$, soit l'application trilinéaire

$$\Phi_\lambda: (\vec{A}, \vec{B}, \vec{C}) \mapsto (\vec{A} \wedge \vec{B}) \wedge \vec{C} - \lambda((\vec{A} \cdot \vec{C})\vec{B} - (\vec{B} \cdot \vec{C})\vec{A})$$

Soit $(\vec{i}, \vec{j}, \vec{k})$ une base orthonormale. Alors, pour $\lambda = 1$, on vérifie que $\Phi_1(\vec{A}, \vec{B}, \vec{C}) = 0$ dès que $\vec{A}, \vec{B}, \vec{C}$ sont pris dans l'ensemble $\{\vec{i}, \vec{j}, \vec{k}\}$. Ceci montre que Φ_1 est identiquement nulle. $\square$

Exercice 10.1. En dimension n , soit un simplexe $(A_0, \dots, A_n)$, $\mathcal{F}_i$ la face opposée au sommet A_i , c'est-à-dire l'hyperplan engendré par les A_j pour $j \neq i$, F_i la direction de $\mathcal{F}_i$, δ une droite vectorielle contenue dans aucun des F_i . Pour tout i , la droite $A_i + \delta$ coupe $\mathcal{F}_i$ en A'_i . Comparer les volumes des simplexes $(A_0, \dots, A_n)$ et $(A'_0, \dots, A'_n)$.

Exercice 10.2. Les conditions sont les précédentes. Une droite affine Δ coupe l'hyperplan $\mathcal{F}_i$ en B_i . Soit C_i le milieu de (A_i, B_i) . Montrer que les C_i sont dans un même hyperplan (observer que la matrice des B_i est de rang 2, de trace nulle, de valeurs propres 1 et -1).

Exercice 10.3. Soit quatre vecteurs $\vec{A}, \vec{B}, \vec{C}, \vec{D}$. Montrer que

$$\begin{aligned} (\vec{A} \wedge \vec{B}) \wedge (\vec{C} \wedge \vec{D}) &= (\vec{A}, \vec{C}, \vec{D})\vec{B} - (\vec{B}, \vec{C}, \vec{D})\vec{A} \\ &= (\vec{A}, \vec{B}, \vec{D})\vec{C} - (\vec{A}, \vec{B}, \vec{C})\vec{D} \\ (\vec{A} \wedge \vec{B}) \cdot (\vec{C} \wedge \vec{D}) &= (\vec{A} \cdot \vec{C})(\vec{B} \cdot \vec{D}) - (\vec{A} \cdot \vec{D})(\vec{B} \cdot \vec{C}) \end{aligned}$$

Exercice 10.4. On donne un triangle $A_1A_2A_3$. Montrer que

$$\overrightarrow{A_1A_2} \wedge \overrightarrow{A_1A_3} = \overrightarrow{A_2A_3} \wedge \overrightarrow{A_2A_1} = \overrightarrow{A_3A_1} \wedge \overrightarrow{A_3A_2}$$

Ce vecteur, noté $\overrightarrow{[A_1, A_2, A_3]}$, est l'**aire vectorielle** du triangle orienté $A_1A_2A_3$.

Exercice 10.5. Soit $A_1A_2A_3A_4$ un tétraèdre. Étant donné une permutation paire i, j, k, l de 1, 2, 3, 4, on pose (exercice 10.4) $\vec{F}_i = \overrightarrow{[A_j, A_k, A_l]}$, et ce vecteur ne dépend que de l'indice i . Montrer que

$$(i) \vec{F}_1 + \vec{F}_2 + \vec{F}_3 + \vec{F}_4 = \vec{0},$$

$$(ii) \vec{F}_i \cdot \overrightarrow{A_i A_j} = [A_1, A_2, A_3, A_4],$$

$$(iii) \vec{F}_i \wedge \vec{F}_j = [A_1, A_2, A_3, A_4] \overrightarrow{A_k A_l},$$

$$(iv) (\vec{F}_i, \vec{F}_j, \vec{F}_k) = [A_1, A_2, A_3, A_4]^2.$$

10.2 APPLICATIONS À DES CONFIGURATIONS

10.2.1. Tétraèdres et parallélépipèdes

Proposition 10.9. Soit un tétraèdre $A = (A_1A_2A_3A_4)$ d'isobarycentre O , et soit $B = (B_1B_2B_3B_4)$ le tétraèdre symétrique relativement à O . Les huit points A_i et B_i forment un parallélépipède $A \cup B$ de centre O . Les six milieux $M_{i,j}$ des arêtes de A sont les centres des faces de $A \cup B$.

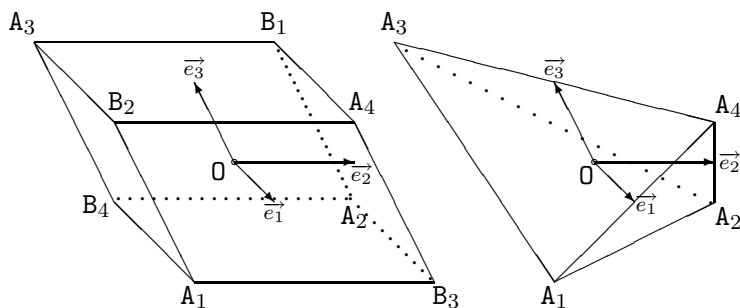
Démonstration. Les points A_i et B_i étant symétriques autour de O , on a $\overrightarrow{A_3B_1} = \overrightarrow{A_1B_3}$ et $\overrightarrow{B_2A_4} = \overrightarrow{B_4A_2}$. On déduit l'égalité de ces quatre vecteurs de

$$\overrightarrow{B_2A_4} - \overrightarrow{A_3B_1} = (\overrightarrow{OA_4} - \overrightarrow{OB_2}) - (\overrightarrow{OB_1} - \overrightarrow{OA_3}) = (\overrightarrow{OA_4} + \overrightarrow{OA_2}) - (-\overrightarrow{OA_1} - \overrightarrow{OA_3}) = \vec{0}$$

car O est isobarycentre du tétraèdre A . On prouve de même

$$\overrightarrow{A_3B_4} = \overrightarrow{A_4B_3} = \overrightarrow{B_1A_2} = \overrightarrow{B_2A_1}, \quad \overrightarrow{A_3B_2} = \overrightarrow{A_2B_3} = \overrightarrow{B_4A_1} = \overrightarrow{B_1A_4}$$

Il s'agit donc bien d'un parallélépipède.



Inversement, soit un repère $(O, \vec{e}_1, \vec{e}_2, \vec{e}_3)$ et le parallélépipède de sommets

$$A_1 \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} A_2 \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix} A_3 \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} A_4 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} B_1 \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} B_2 \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} B_3 \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} B_4 \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}$$

Alors les tétraèdres $A = A_1 A_2 A_3 A_4$ et $B = B_1 B_2 B_3 B_4$ sont bien symétriques autour de O et ont bien O pour isobarycentre. $\square$

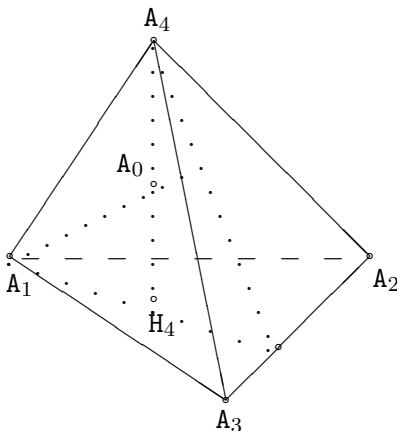
Proposition 10.10. Tétraèdre orthocentrique. Soit un tétraèdre $(A_1, A_2, A_3, A_4) = A$ d'isobarycentre O et $(B_1, B_2, B_3, B_4) = B$ le symétrique de A autour de O . Les conditions suivantes sont équivalentes :

- (i) il existe deux couples d'arêtes opposées formés d'arêtes orthogonales,
- (ii) les trois couples d'arêtes opposées sont formés d'arêtes orthogonales,
- (iii) il existe un sommet se projetant orthogonalement en l'orthocentre de la face opposée,
- (iv) tout sommet se projette orthogonalement en l'orthocentre de la face opposée,
- (v) les quatre hauteurs du tétraèdre A sont concourantes,
- (vi) $A_4 A_1^2 + A_2 A_3^2 = A_4 A_2^2 + A_3 A_1^2 = A_4 A_3^2 + A_1 A_2^2$,
- (vii) les faces du parallélépipède $A \cup B$ sont des losanges.

Démonstration. On dit que A est **orthocentrique** s'il vérifie ces conditions. Le point de concours A_0 des hauteurs est l'**orthocentre** de A .

On a les implications $(ii) \Rightarrow (i)$ et $(iv) \Rightarrow (iii)$. On obtient $(i) \Rightarrow (ii)$ par l'identité du produit scalaire (8.33, p. 223) :

$$\overrightarrow{A_4 A_1} \cdot \overrightarrow{A_2 A_3} + \overrightarrow{A_4 A_2} \cdot \overrightarrow{A_3 A_1} + \overrightarrow{A_4 A_3} \cdot \overrightarrow{A_1 A_2} = 0$$



Soit H_4 la projection orthogonale de A_4 sur $(A_1 A_2 A_3)$. On a $\overrightarrow{A_4 A_1} \cdot \overrightarrow{A_2 A_3} = \overrightarrow{H_4 A_1} \cdot \overrightarrow{A_2 A_3}$, donc $(A_4 A_1)$ et $(A_2 A_3)$ sont orthogonales si et seulement si H_4 est sur la hauteur du triangle $A_1 A_2 A_3$ issue de A_1 . D'où H_4 est orthocentre de $A_1 A_2 A_3$ si et seulement si les couples d'arêtes opposées $((A_4 A_1), (A_2 A_3))$, $((A_4 A_2), (A_3 A_1))$, $((A_4 A_3), (A_1 A_2))$ sont formés d'arêtes orthogonales. Ceci donne l'équivalence de (ii) , (iii) , (iv) , donc l'équivalence des quatre premières conditions.

Si elles sont remplies, les hauteurs issues de A_1 et A_4 sont dans le plan contenant $(A_1 A_4)$ et orthogonal à $(A_2 A_3)$. N'étant pas parallèles

car orthogonales à des plans de face non parallèles, elles se coupent. Les hauteurs du tétraèdre sont donc deux à deux concourantes. Étant trois à trois non coplanaires, les points communs à ces paires de hauteurs sont nécessairement confondus en un même point A_0 .

Inversement, si les quatre hauteurs de A concourent en A_0 , les droites (A_0A_1) et (A_0A_4) sont orthogonales respectivement aux plans $(A_2A_3A_4)$ et $(A_1A_2A_3)$, donc l'intersection (A_2A_3) de ces deux plans est orthogonale au plan $(A_0A_1A_4)$, d'où l'orthogonalité des arêtes opposées (A_1A_4) et (A_2A_3) . On fait de même pour les autres couples d'arêtes opposées.

On a donc l'équivalence des cinq premières conditions. Pour prouver l'équivalence avec (vi) , on utilisera l'identité vectorielle

$$\overrightarrow{A_1A_2} + \overrightarrow{A_3A_4} = \overrightarrow{A_1A_4} + \overrightarrow{A_3A_2}$$

obtenue ainsi : $\overrightarrow{A_1A_2} + \overrightarrow{A_3A_4} = \overrightarrow{A_1A_4} + \overrightarrow{A_4A_2} + \overrightarrow{A_3A_2} + \overrightarrow{A_2A_4} = \overrightarrow{A_1A_4} + \overrightarrow{A_3A_2}$. On a donc

$$\|\overrightarrow{A_1A_2} + \overrightarrow{A_3A_4}\|^2 = \|\overrightarrow{A_1A_4} + \overrightarrow{A_3A_2}\|^2,$$

d'où

$$A_1A_2^2 + A_3A_4^2 + 2\overrightarrow{A_1A_2} \cdot \overrightarrow{A_3A_4} = A_1A_4^2 + A_3A_2^2 + 2\overrightarrow{A_1A_4} \cdot \overrightarrow{A_3A_2}$$

Si (ii) est vérifié, on a $A_1A_2^2 + A_3A_4^2 = A_1A_4^2 + A_3A_2^2$ et deux autres égalités analogues aboutissant à (vi) .

Inversement, si (vi) est vérifié, on aura

$$\overrightarrow{A_1A_2} \cdot \overrightarrow{A_3A_4} = \overrightarrow{A_1A_4} \cdot \overrightarrow{A_3A_2}, \overrightarrow{A_2A_3} \cdot \overrightarrow{A_1A_4} = \overrightarrow{A_2A_4} \cdot \overrightarrow{A_1A_3}, \overrightarrow{A_3A_1} \cdot \overrightarrow{A_2A_4} = \overrightarrow{A_3A_4} \cdot \overrightarrow{A_2A_1}$$

Ces quantités sont nulles car deux à deux opposées, d'où (ii) .

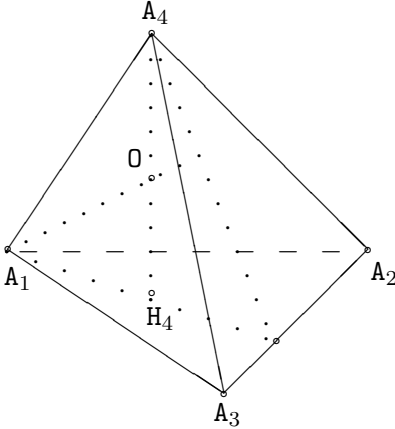
On a $\overrightarrow{A_1A_4} = \overrightarrow{B_4B_1}$, $\overrightarrow{A_2A_3} = \overrightarrow{B_3B_2}$, car (A_1A_4) et (B_3B_2) sont des diagonales de la face $A_1B_2A_4B_3$ du parallélépipède $A \cup B$ (10.9, p. 272). Elles sont orthogonales si et seulement si ce parallélogramme est un losange, d'où $(ii) \Leftrightarrow (vii)$. $\square$

Remarque : Un tétraèdre A n'a d'orthocentre que s'il est orthocentrique. En revanche, de même qu'un triangle a un cercle circonscrit, un tétraèdre $A = A_1A_2A_3A_4$ a une **sphère circonscrite**. En effet, soit Ω_4 le centre du cercle circonscrit au triangle $A_1A_2A_3$ et Δ_4 la droite orthogonale en Ω_4 au plan $(A_1A_2A_3)$. Pour tout $M \in \Delta_4$, les quantités $MA_i^2 = M\Omega_4^2 + \Omega_4A_i^2$ sont égales pour $i \in \{1, 2, 3\}$ (Pythagore). Le plan médiateur de (A_4, A_1) coupe Δ_4 en Ω équidistant des sommets de A .

10.2.2. Cube et tétraèdre régulier

Définition 10.11. Un tétraèdre A est dit **régulier** si les arêtes ont même longueur, c'est-à-dire si les faces sont des triangles équilatéraux. Un parallélépipède est un **cube** si les faces sont des carrés.

Lemme 10.12. *Un tétraèdre régulier $A = A_1A_2A_3A_4$ est orthocentrique et l'isobarycentre O est aussi orthocentre et centre de la sphère circonscrite.*



Démonstration. L'orthocentre H_4 du triangle équilatéral $A_1A_2A_3$ est en même temps l'isobarycentre et le centre du cercle circonscrit.

On a $O \in (A_4H_4)$ car

$$\overrightarrow{OH_4} = \frac{1}{3}(\overrightarrow{OA_1} + \overrightarrow{OA_2} + \overrightarrow{OA_3}) = -\frac{1}{3}\overrightarrow{OA_4}$$

Comme $A_4A_1 = A_4A_2 = A_4A_3$, les points A_1, A_2, A_3 sont sur une sphère de centre A_4 . La projection orthogonale de A_4 sur $(A_1A_2A_3)$ est donc le centre du cercle circonscrit à $A_1A_2A_3$, i.e. H_4 . Donc A est orthocentrique (10.10, (iii)) et O est sur la hauteur (A_4H_4) . Il est de même sur les autres hauteurs, donc c'est l'orthocentre de A .

On a $OA_1^2 = OH_4^2 + H_4A_1^2 = OH_4^2 + H_4A_2^2 = OA_2^2$ (Pythagore). D'où $OA_1 = OA_2$, tous les OA_i sont égaux d'où O est centre de la sphère circonscrite. $\square$

Théorème 10.13. *Le groupe des isométries (resp. déplacements) laissant stable un tétraèdre régulier laisse fixe l'isobarycentre. Il est isomorphe, par action sur les quatre sommets, au groupe symétrique S_4 (resp. alterné A_4).*

Démonstration. Prenons les notations du lemme. Soit G (resp. G^+) le groupe des isométries (resp. déplacements) de $\mathcal{E}$ laissant A stable. Pour tout $f \in G$, $f(O)$ est isobarycentre de $f(A) = A$, donc $f(O) = O$.

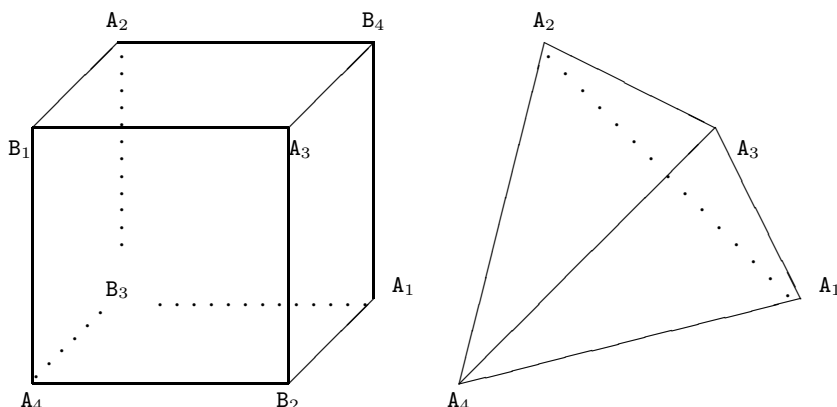
Soit f/A la restriction de $f \in G$ à l'ensemble des sommets de A . L'application $f \mapsto f/A$ est un morphisme φ de G dans le groupe des bijections de A sur lui-même que l'on identifie à S_4 .

Par le lemme, la projection orthogonale de (A_4A_1) sur le plan $(A_1A_2A_3)$ est (H_4A_1) , médiatrice de (A_2, A_3) , donc A_4 et A_1 sont dans le plan médiateur $\mathcal{P}_{2,3}$ de (A_2, A_3) . La réflexion $\sigma_{2,3}$ autour de $\mathcal{P}_{2,3}$ laisse fixe A_1, A_4 et échange A_2, A_3 . Donc $\sigma_{2,3} \in G$ et $\sigma_{2,3}/A$ est la transposition de S_4 échangeant A_2 et A_3 .

De même toute transposition de S_4 est dans l'image de φ . Les transpositions engendrant S_4 , φ est surjectif.

Le noyau $\text{Ker } \varphi$ est formé des isométries laissant fixes les quatre sommets. Comme ils forment un repère barycentrique, ce noyau se réduit à l'identité de $\mathcal{E}$. Donc φ est un isomorphisme.

L'application $\delta: f \mapsto \det \overline{f}$ est un morphisme $G \rightarrow \{-1, 1\}$ de noyau G^+ . Comme G contient des réflexions, δ est surjectif et G^+ est sous-groupe d'indice 2. Comme $\mathcal{A}_4$ est le seul sous-groupe d'indice 2 de S_4 , φ induit un isomorphisme de G^+ sur $\mathcal{A}_4$. $\square$

cube $C = A \cup B$ tétraèdre A

Lemme 10.14. Soit $A = A_1A_2A_3A_4$ un tétraèdre, O l'isobarycentre, $B = B_1B_2B_3B_4$ le symétrique de A par rapport à O . Le tétraèdre A est régulier si et seulement si le parallélépipède $A \cup B$ est un cube.

Démonstration. Supposons que $C = A \cup B$ est un cube dont les arêtes A_iB_j sont de longueur a . Les arêtes A_iA_j de A sont les diagonales des faces du cubes donc sont de même longueur $a\sqrt{2}$. Les faces de A sont des triangles équilatéraux, donc A est un tétraèdre régulier.

Supposons que A est un tétraèdre régulier. Il est orthocentrique (10.12). Les faces de $A \cup B$ sont des losanges (10.10). Pour voir que ce sont des carrés, il suffit de voir par exemple que $\overrightarrow{A_1B_2} = \overrightarrow{OB_2} - \overrightarrow{OA_1} = -\overrightarrow{OA_2} - \overrightarrow{OA_1}$ et $\overrightarrow{A_1B_3} = \overrightarrow{OB_3} - \overrightarrow{OA_1} = -\overrightarrow{OA_3} - \overrightarrow{OA_1}$ sont orthogonaux.

$$\begin{aligned} \overrightarrow{A_1B_2} \cdot \overrightarrow{A_1B_3} &= (-\overrightarrow{OA_2} - \overrightarrow{OA_1}) \cdot (-\overrightarrow{OA_3} - \overrightarrow{OA_1}) \\ &= \overrightarrow{OA_2} \cdot \overrightarrow{OA_3} + \overrightarrow{OA_1} \cdot \overrightarrow{OA_3} + \overrightarrow{OA_2} \cdot \overrightarrow{OA_1} + \overrightarrow{OA_1} \cdot \overrightarrow{OA_1} \\ &= \overrightarrow{OA_2} \cdot \overrightarrow{OA_3} + \overrightarrow{OA_1} \cdot (\overrightarrow{OA_3} + \overrightarrow{OA_2} + \overrightarrow{OA_1}) = \overrightarrow{OA_2} \cdot \overrightarrow{OA_3} - \overrightarrow{OA_1} \cdot \overrightarrow{OA_4} \end{aligned}$$

Soit l la longueur d'une arête de A et R le rayon de la sphère circonscrite. Pour tout couple d'indices (i, j) ,

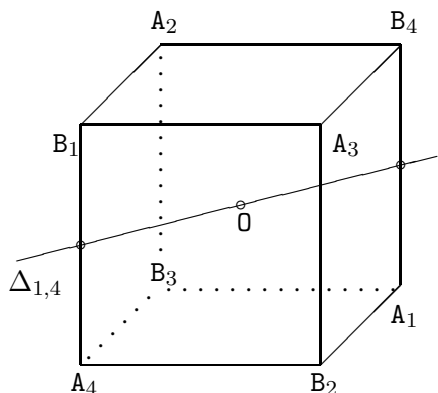
$$l^2 = A_iA_j^2 = \|\overrightarrow{OA_j} - \overrightarrow{OA_i}\|^2 = OA_j^2 + OA_i^2 - 2\overrightarrow{OA_i} \cdot \overrightarrow{OA_j} = 2R^2 - 2\overrightarrow{OA_i} \cdot \overrightarrow{OA_j}$$

Les produits scalaires $\overrightarrow{OA_i} \cdot \overrightarrow{OA_j}$ sont tous égaux, donc $\overrightarrow{A_1B_2} \cdot \overrightarrow{A_1B_3} = 0$. $\square$

Théorème 10.15. Le groupe des déplacements laissant stable un cube laisse fixe l'isobarycentre. Par son action sur l'ensemble des quatre diagonales, il est isomorphe au groupe symétrique S_4 .

Démonstration. Reprenons les mêmes notations. Soit G le groupe des déplacements laissant stable le cube $A \cup B$. Comme pour 10.13, on montre que $f(O) = O$ pour tout $f \in G$.

Le cube a quatre diagonales $D_i = [A_i, B_i]$, chacune de milieu 0. Comme 0 est laissé fixe, tout $f \in G$ transforme une diagonale en une diagonale. On note f/D l'action ainsi induite par $f \in G$ sur l'ensemble $D = \{D_1, D_2, D_3, D_4\}$ des quatre diagonales. On a ainsi un morphisme $\varphi: f \mapsto f/D$ de G dans le groupe identifié à S_4 des bijections de D sur lui-même.



Comme pour 10.13, on montre que φ est surjectif en montrant que toute transposition de S_4 est dans l'image. Soit $\Delta_{1,4}$ la droite joignant les milieux de (A_1, B_4) et (A_4, B_1) et $\rho_{1,4}$ le retournement d'axe $\Delta_{1,4}$. Il échange A_1, B_1 avec B_4, A_4 , donc échange les diagonales D_1 et D_4 . Les points A_2 et B_2 sont échangés, de même que A_3 et B_3 , donc les diagonales D_2 et D_3 sont stables. D'où $\rho_{1,4}/D$ est la transposition de D_1 et D_4 . De même, toute transposition est dans l'image de φ , d'où la surjectivité.

Soit $f \in G$ distincte de l'identité. Alors f est une rotation d'axe Δ . L'ensemble des droites

stables se réduit à $\{\Delta\}$ si f n'est pas un retournement, se compose de Δ et des droites perpendiculaires à Δ si f est un retournement. Si $f \in \text{Ker } \varphi$, f laisse stable les quatre diagonales. Comme trois quelconques d'entre elles n'ont pas de perpendiculaire commune, f est nécessairement l'identité, d'où l'injectivité. $\square$

Remarquons que G contient le sous-groupe d'indice 2 des déplacements laissant stable chacun des tétraèdres A et B , isomorphe à A_4 d'après 10.13.

EXERCICES

Exercice 10.6. Soit un tétraèdre $A = A_1A_2A_3A_4$, 0 l'isobarycentre, B le tétraèdre symétrique par rapport à 0. Montrer (utiliser l'exercice 10.5) l'équivalence des conditions suivantes :

- (i) les faces ont même aire,
- (ii) toute arête a même longueur que l'arête opposée,
- (iii) le parallélépipède $A \cup B$ est rectangle.

On dit alors que A est un tétraèdre **équifacial**. Que dire d'un tétraèdre équifacial et orthocentrique ?

Exercice 10.7. Soit $A = A_1A_2A_3A_4$ un tétraèdre régulier d'isobarycentre 0. Décrire le sous-groupe distingué de cardinal 4 du groupe isomorphe à S_4 des isométries laissant A stable.

Même question pour le sous-groupe distingué de cardinal 4 des déplacements laissant stable un cube.

Exercice 10.8. Soit un cube $C = A \cup B$ où A, B sont deux tétraèdres réguliers de même isobarycentre O , symétriques par rapport à O , G le groupe de toutes les isométries laissant C stable.

- 1) Montrer que G a trois sous-groupes d'indice 2 dont deux sont isomorphes à S_4 .
- 2) Montrer que G est isomorphe au produit direct $S_4 \times \{-1, +1\}$.
- 3) Déterminer un sous-groupe distingué de G cardinal 8 laissant C stable.

Exercice 10.9. Soit un tétraèdre $A = A_1A_2A_3A_4$, $M_{i,j}$ le milieu de (A_i, A_j) . Montrer que A est équi-facial si et seulement si les droites $(M_{1,2}M_{3,4})$, $(M_{1,3}M_{2,4})$, $(M_{1,4}M_{2,3})$ forment un trièdre trirectangle (montrer que A est équi-facial si et seulement si le groupe des isométries laissant A stable contient le groupe de Klein)

PROBLÈMES

Les corrigés de ces problèmes sont disponibles sur le site de Dunod : www.dunod.com.

10.1. PROBLÈME

Soit $A = (A_1A_2A_3A_4)$ un tétraèdre, O l'isobarycentre, Ω le centre de la sphère circonscrite. À toute arête $[A_i, A_j]$, on associe le plan $\mathcal{P}_{i,j}$ passant par le milieu de $[A_i, A_j]$ et orthogonal à l'arête opposée.

- 1) Montrer que les six plans $\mathcal{P}_{i,j}$ concourent en un même point M appelé **point de Monge** du tétraèdre.
- 2) Montrer que O est milieu de (M, Ω) (on pourra projeter orthogonalement sur les arêtes).
- 3) Que dire du point de Monge si A est orthocentrique ?
- 4) Montrer que A est équi-facial si et seulement si O, M, Ω sont confondus.

10.2. PROBLÈME

Soit $A = A_1A_2A_3A_4$ un tétraèdre, O l'isobarycentre, $M_{i,j}$ le milieu de (A_i, A_j) .

- 1) Montrer que A est orthocentrique si et seulement si les six points $M_{i,j}$ sont sur une sphère S_1 de centre O .

Dans la suite, on suppose A orthocentrique. Soit H l'orthocentre et S_2 le centre de la sphère circonscrite S_0 .

- 2) Montrer que S_1 coupe chaque face $(A_iA_jA_k)$ suivant les cercles des 9 points du triangle $A_iA_jA_k$.

- 3) Montrer que les isobarycentres et les orthocentres des faces sont sur une sphère S_2 (considérer les homothéties de centre O et H de rapport $-\frac{1}{3}$ et $\frac{1}{3}$ et utiliser la question 2) du problème 10.1).

SOLUTIONS DES EXERCICES

Solution 10.1. Soit $\mathcal{B}$ une base orthonormale de E et 0 une origine. Les déterminants de systèmes de vecteurs de $\tilde{\mathcal{E}}$ seront évalués relativement à la base $0, \mathcal{B}$ de $\tilde{\mathcal{E}}$. Pour tout i on a $A'_i = A_i + (A'_i - A_i)$, les vecteurs $A'_i - A_i = \overrightarrow{A_i A'_i}$ étant tous dans δ , donc colinéaires, on a donc

$$\begin{aligned} [A'_0, \dots, A'_n] &= \det(A_0 + (A'_0 - A_0), \dots, A_n + (A'_n - A_n)) \\ &= \det(A_0, \dots, A_n) + \sum_{i=0}^n D_i \end{aligned}$$

où D_i se déduit de $\det(A_0, \dots, A_i, \dots, A_n)$ en remplaçant A_i par $A'_i - A_i$. Comme A'_i et les $A_j, j \neq i$ sont dans le même hyperplan $\tilde{\mathcal{F}}_i$, l'expression déduite de $\det(A_0, \dots, A_i, \dots, A_n)$ en remplaçant A_i par A'_i est nulle. On a donc

$$D_i = -\det(A_0, \dots, A_n)$$

d'où $[A'_0, \dots, A'_n] = -n[A_0, \dots, A_n]$.

Solution 10.2. Dans $\tilde{\mathcal{E}}$, on écrit $B_j = \sum_i \beta_{i,j} A_i$, où $\beta_{j,j} = 0, \sum_i \beta_{i,j} = 1$. On a $2C_j = A_j + B_j$. Pour voir que ces éléments sont liés dans $\tilde{\mathcal{E}}$, vérifions que le déterminant de la matrice

$$\mathcal{C} = \begin{pmatrix} 1 & \beta_{0,1} & \cdots & \beta_{0,n} \\ \beta_{1,0} & 1 & \cdots & \beta_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n,0} & \beta_{n,1} & \cdots & 1 \end{pmatrix}$$

est nul. L'alignement des B_j se traduit par $\text{rg}(\mathcal{B}) = 2$ où

$$\mathcal{B} = \begin{pmatrix} 0 & \beta_{0,1} & \cdots & \beta_{0,n} \\ \beta_{1,0} & 0 & \cdots & \beta_{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n,0} & \beta_{n,1} & \cdots & 0 \end{pmatrix}$$

Donc 0 est valeur propre, racine du polynôme caractéristique à l'ordre au moins $n-1$. Le fait que $\sum_i \beta_{i,j} = 1$ pour tout j montre que 1 est valeur propre de $\mathcal{B}$. La dernière valeur propre de $\mathcal{B}$ est nécessairement -1 car $\text{Tr}(\mathcal{B}) = 0$, d'où $\det \mathcal{C} = 0$.

Solution 10.3. Par la formule du double produit vectoriel, on a

$$\begin{aligned} (\vec{A} \wedge \vec{B}) \wedge (\vec{C} \wedge \vec{D}) &= (\vec{A} \cdot (\vec{C} \wedge \vec{D})) \vec{B} - (\vec{B} \cdot (\vec{C} \wedge \vec{D})) \vec{A} \\ &= (\vec{A}, \vec{C}, \vec{D}) \vec{B} - (\vec{B}, \vec{C}, \vec{D}) \vec{A} \end{aligned}$$

La deuxième formule s'obtient de la même façon.

On a aussi

$$(\vec{A} \wedge \vec{B}) \cdot (\vec{C} \wedge \vec{D}) = (\vec{A} \wedge \vec{B}, \vec{C}, \vec{D}) = ((\vec{A} \wedge \vec{B}) \wedge \vec{C}) \cdot \vec{D}$$

puis on applique la formule du double produit vectoriel.

Solution 10.4. Il suffit d'écrire

$$\overrightarrow{A_2A_3} \wedge \overrightarrow{A_2A_1} = (\overrightarrow{A_1A_3} - \overrightarrow{A_1A_2}) \wedge (-\overrightarrow{A_1A_2}) = -\overrightarrow{A_1A_3} \wedge \overrightarrow{A_1A_2} = \overrightarrow{A_1A_2} \wedge \overrightarrow{A_1A_3}$$

Solution 10.5. (i) On a

$$\begin{aligned} \vec{F}_1 &= \overrightarrow{A_2A_3} \wedge \overrightarrow{A_2A_4} = (\overrightarrow{A_1A_3} - \overrightarrow{A_1A_2}) \wedge (\overrightarrow{A_1A_4} - \overrightarrow{A_1A_2}) \\ &= \overrightarrow{A_1A_3} \wedge \overrightarrow{A_1A_4} - \overrightarrow{A_1A_3} \wedge \overrightarrow{A_1A_2} - \overrightarrow{A_1A_2} \wedge \overrightarrow{A_1A_4} = -\vec{F}_2 - \vec{F}_4 - \vec{F}_3 \end{aligned}$$

(ii) On a

$$\begin{aligned} \vec{F}_i \cdot \overrightarrow{A_iA_j} &= (\overrightarrow{A_jA_k} \wedge \overrightarrow{A_jA_l}) \cdot \overrightarrow{A_iA_j} = -(\overrightarrow{A_jA_k}, \overrightarrow{A_jA_l}, \overrightarrow{A_jA_i}) \\ &= -[A_j, A_k, A_l, A_i] = [A_i, A_j, A_k, A_l] = [A_1, A_2, A_3, A_4] \end{aligned}$$

(iii) Par la formule du double produit vectoriel

$$\begin{aligned} \vec{F}_i \wedge \vec{F}_j &= \vec{F}_i \wedge (\overrightarrow{A_kA_l} \wedge \overrightarrow{A_kA_i}) = (\vec{F}_i \cdot \overrightarrow{A_kA_l}) \overrightarrow{A_kA_i} - (\vec{F}_i \cdot \overrightarrow{A_kA_i}) \overrightarrow{A_kA_l} \\ &= (\overrightarrow{A_kA_l}, \overrightarrow{A_kA_j}, \overrightarrow{A_kA_l}) \overrightarrow{A_kA_i} - (\overrightarrow{A_kA_l}, \overrightarrow{A_kA_j}, \overrightarrow{A_kA_i}) \overrightarrow{A_kA_l} \\ &= -[A_k, A_l, A_j, A_i] \overrightarrow{A_kA_l} = [A_1, A_2, A_3, A_4] \overrightarrow{A_kA_l} \end{aligned}$$

(iv) Par (iii), on a

$$\begin{aligned} (\vec{F}_i, \vec{F}_j, \vec{F}_k) &= (\vec{F}_i \wedge \vec{F}_j) \cdot \vec{F}_k = [A_1, A_2, A_3, A_4] \overrightarrow{A_kA_l} \cdot \vec{F}_k \\ &= [A_1, A_2, A_3, A_4] ((\overrightarrow{A_kA_l} \cdot (\overrightarrow{A_lA_i} \wedge \overrightarrow{A_lA_j}))) \\ &= [A_1, A_2, A_3, A_4] (\overrightarrow{A_kA_l}, \overrightarrow{A_lA_i}, \overrightarrow{A_lA_j}) = [A_1, A_2, A_3, A_4]^2 \end{aligned}$$

Solution 10.6. (ii) $\Rightarrow$ (i). On suppose que pour toute permutation i, j, k, l de $1, 2, 3, 4$, on a $A_iA_j = A_kA_l$. Les triangles $A_1A_2A_3$ et $A_1A_2A_4$ ont leurs côtés égaux deux à deux : $A_1A_3 = A_2A_4$, $A_2A_3 = A_1A_4$ et le côté A_1A_2 est commun. Ces triangles sont donc égaux, donc ont même aire.

(i) $\Rightarrow$ (ii). Avec les notations de l'exercice 10.5, les $\|\vec{F}_i\|$ sont égaux à une même quantité F . Soit $\theta_{i,j} \in [0, \pi]$ l'angle non orienté de $\vec{F}_i$ et $\vec{F}_j$. Pour i, j, k, l permutation de $1, 2, 3, 4$, $\vec{F}_i + \vec{F}_j = -\vec{F}_k - \vec{F}_l$. Prenant les carrés des normes de chaque membre, on obtient $\cos \theta_{i,j} = \cos \theta_{k,l}$, d'où $\theta_{i,j} = \theta_{k,l}$.

On a $\vec{F}_i \wedge \vec{F}_j = [A_1, A_2, A_3, A_4] \overrightarrow{A_k A_l}$, d'où

$$|[A_1, A_2, A_3, A_4] A_k A_l| = \|\vec{F}_i \wedge \vec{F}_j\| = F^2 \sin \theta_{i,j}$$

$$A_k A_l = \frac{F^2 \sin \theta_{i,j}}{|[A_1, A_2, A_3, A_4]|} = \frac{F^2 \sin \theta_{k,l}}{|[A_1, A_2, A_3, A_4]|} = A_i A_j$$

L'équivalence (ii) $\Leftrightarrow$ (iii) se déduit du fait qu'un parallélogramme est un rectangle si et seulement si ses diagonales ont même longueur.

Un tétraèdre est à la fois équi-facial et orthocentrique si et seulement si les faces du parallélépipède associé sont à la fois losanges et rectangles, donc carrées. Un tétraèdre équi-facial et orthocentrique est régulier.

Solution 10.7. Le sous-groupe distingué de cardinal 4 de S_4 est le groupe de Klein, formé de l'identité et des trois paires de transpositions disjointes. L'isométrie échangeant A_1 et A_2 , A_3 et A_4 est le retournement autour de la droite joignant les milieux $M_{1,2}$ de (A_1, A_2) et $M_{3,4}$ de (A_3, A_4) .

Le groupe de Klein est donc formé de l'identité et des trois retournements autour des axes $(M_{i,j}, M_{k,l})$, i, j, k, l étant une permutation de 1, 2, 3, 4. Le composé de deux de ces retournements étant le troisième, les axes $(M_{1,2}, M_{3,4})$, $(M_{1,3}, M_{2,4})$, $(M_{1,4}, M_{2,3})$ forment un système trirectangle de droites se rencontrant en l'isobarycentre O .

Soit B le tétraèdre symétrique de A autour de O et $C = A \cup B$ le cube associé. Les trois droites précédentes sont les axes des faces du cube.

Solution 10.8. Le groupe G^+ des déplacements laissant C stable est isomorphe à S_4 (10.15, p. 276). L'application $f \mapsto \det \bar{f}$ est un morphisme $G \rightarrow \{-1, 1\}$. Il est surjectif car la réflexion autour du plan médiateur d'une arête est un antidéplacement de G . Donc le noyau G^+ est d'indice 2.

Comme $G^+ \sim S_4$ est de cardinal 24, G est de cardinal 48.

Toute isométrie laissant stable A laisse fixe O donc laisse B stable : si $f(A_i) = A_j$, alors $f(B_i) = B_j$. On a donc un sous-groupe $\mathcal{A}$ de G isomorphe à S_4 (10.13, p. 275), donc de cardinal 24, donc d'indice 2.

L'intersection $\mathcal{A} \cap G^+$ est formée des déplacements laissant stables les tétraèdres. C'est un sous-groupe distingué d'indice 4 isomorphe à $\mathcal{A}_4$.

Le troisième sous-groupe d'indice 2 est formé des f qui

- ou bien sont dans $\mathcal{A} \cap G^+$, donc sont des déplacements laissant stables les tétraèdres,
- ou bien ne sont ni dans $\mathcal{A}$, ni dans G^+ , donc sont des antidéplacements échangeant les tétraèdres. Parmi ces isométries, la symétrie centrale S_0 .

On vérifie que l'application associant à $f \in G$ le couple

- $(1, f)$ si $f \in G^+$,
- $(-1, S_0 \circ f)$ si $f \notin G^+$,

est un isomorphisme $G \rightarrow \{-1, 1\} \times G^+ \simeq \{-1, 1\} \times S_4$. On a de même un isomorphisme $G \rightarrow \{-1, 1\} \times \mathcal{A} \simeq \{-1, 1\} \times S_4$.

Le sous-groupe distingué de cardinal 8 est formé de l'identité, de la symétrie S_0 , des trois retournements autour des axes des faces du cube, des composés de ces retournements avec S_0 qui sont les réflexions autour des plans médiateurs des arêtes du cube.

Solution 10.9. Soit O l'isobarycentre et B le symétrique de A autour de O . On a

$$O = \frac{1}{4}(A_1 + A_2 + A_3 + A_4) = \frac{1}{2}(M_{1,2} + M_{3,4})$$

donc O est milieu des trois segments $[M_{i,j}, M_{k,l}]$.

Les $M_{i,j}$ sont les isobarycentres des faces du parallélépipède $A \cup B = C$. Le tétraèdre A est équi facial si et seulement si C est rectangle, donc si et seulement si les trois droites $(M_{i,j}M_{k,l})$ sont deux à deux orthogonales.

Le groupe des isométries laissant A stable, identifié à un sous-groupe de S_4 , contient le groupe de Klein si seulement si A est équi facial.

Chapitre 11

Géométrie des cercles

D'importantes propriétés sont développées dans les quatre premières sections : positions relatives de droites et cercles (11.1.1), puissance d'un point (11.2), propriété angulaire (11.3), lieu des points dont le rapport des distances à deux points est constant (11.4). Ceci amène à une première approche de la notion de faisceau de cercles, à points de base dans la section 11.3, à points limites dans la section 11.4. La section 11.5 introduit une extension de la notion de cercle : on ajoute les droites. On y installe également un point à l'infini.

L'orthogonalité est définie à l'aide d'une forme quadratique de signature $(3, 1)$ dans l'espace vectoriel de dimension 4 des fonctions circulaires. Ceci permet une approche plus générale de la notion de faisceau de cercles.

11.1 POSITIONS RELATIVES DE CERCLES ET DE DROITES

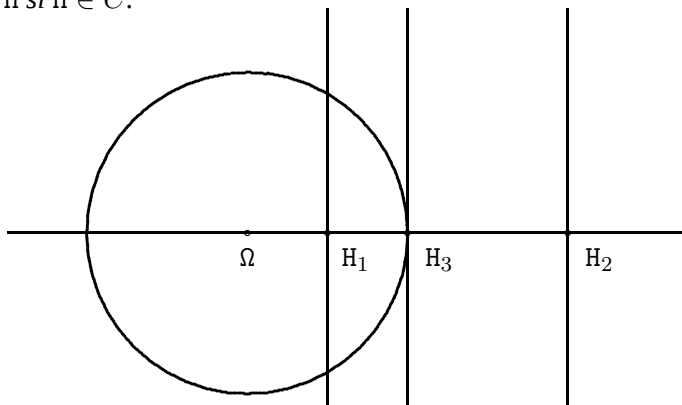
Commençons par reprendre des notions intuitives.

Définition 11.1. Soit C un cercle de centre Ω et de rayon $R > 0$, ensemble des points M du plan tels que $\Omega M = R$. Le complémentaire de C est formé de :

- 1) l'**extérieur** formé des M tels que $M\Omega > R$,
- 2) l'**intérieur** formé des M tels que $M\Omega < R$.

Proposition 11.2. Soit Δ une droite, H la projection de Ω sur Δ . L'intersection $\Delta \cap C$ est

- (i) formée de deux points distincts si H est intérieur à C ,
- (ii) vide si H est extérieur à C ,
- (iii) réduite au seul point H si $H \in C$.



Démonstration. Soit $\vec{u}$ un vecteur unitaire directeur de Δ et $M_t = H + t\vec{u}$ où $t \in \mathbb{R}$. On a

$$\|\vec{\Omega M_t}\|^2 = \|\vec{\Omega H}\|^2 + \|\vec{HM_t}\|^2 = \|\vec{\Omega H}\|^2 + t^2.$$

Donc $M_t \in C$ si et seulement si $t^2 = R^2 - \Omega H^2$. Une discussion immédiate aboutit au résultat. $\square$

Définition 11.3. On dit que Δ est **sécante** à C dans le premier cas, **extérieure** à C dans le deuxième cas, **tangente** à C dans le troisième cas.

Corollaire Soit $M \in C$. La perpendiculaire $\mathcal{T}$ en M à (ΩM) est tangente à C . Toute droite passant par M distincte de $\mathcal{T}$ est sécante.

Démonstration. Si $\mathcal{T}$ est la perpendiculaire en M à (ΩM) , alors M est projection orthogonale de Ω sur $\mathcal{T}$, donc $\mathcal{T} \cap C = \{M\}$.

Si Δ , distincte de $\mathcal{T}$, passe par M , la projection orthogonale H de Ω sur Δ est distincte de M , donc $\Omega H < \Omega M = R$ et Δ est sécante à C . $\square$

Proposition 11.4. Soit M un point. Alors

- (i) M est intérieur à C si et seulement si toute droite passant par M est sécante ;
- (ii) M est extérieur à C si et seulement si il passe par M (au moins) une droite extérieure.

Démonstration. a) Supposons M intérieur. Soit Δ passant par M et H la projection orthogonale de Ω sur Δ . On a $\Omega H \leq \Omega M < R$, donc H est intérieur à C et Δ est sécante.
 b) Supposons M extérieur. Si Δ est la perpendiculaire en M à (ΩM) , alors M est projection orthogonale de Ω sur Δ , donc Δ est extérieure à C .

c) Voyons la réciproque de a). Supposons que toute droite passant par M soit sécante. Alors M ne peut être extérieur d'après b). Et M ne peut être sur C car la perpendiculaire T en M à (ΩM) serait tangente. Donc M est intérieur. On voit de même la réciproque de b). $\square$

11.1.1. Positions relatives de deux cercles

Proposition 11.5. Soit deux cercles C_1, C_2 de centres Ω_1, Ω_2 , de rayons R_1, R_2 , $d = \Omega_1\Omega_2$.

(i) L'intersection $C_1 \cap C_2$ est non vide si et seulement si

$$|R_1 - R_2| \leq d \leq R_1 + R_2$$

(ii) si $|R_1 - R_2| < d < R_1 + R_2$, C_1 et C_2 sont **sécants** : ils ont deux points communs distincts ;

(iii) si $d = R_1 + R_2$, C_1 et C_2 sont **tangents extérieurement** en H : tout point de l'un distinct de H est extérieur à l'autre ;

(iv) si $d = R_1 - R_2$, C_1 et C_2 sont **tangents intérieurement** en H : tout point de C_2 distinct de H est intérieur à C_1 .

Démonstration. (i) Supposons que $C_1 \cap C_2$ contienne un point M. L'inégalité triangulaire appliquée au triangle $M\Omega_1\Omega_2$ donne

$$|R_1 - R_2| = |M\Omega_1 - M\Omega_2| \leq d = \Omega_1\Omega_2 \leq M\Omega_1 + M\Omega_2 = R_1 + R_2$$

Réciproquement, supposons $|R_1 - R_2| \leq d \leq R_1 + R_2$. Soit M quelconque.

$$\begin{aligned} (M \in C_1 \cap C_2) &\Leftrightarrow \left(\|\overrightarrow{\Omega_1 M}\|^2 = R_1^2 \text{ et } \|\overrightarrow{\Omega_2 M}\|^2 = R_2^2 \right) \\ &\Leftrightarrow \left(M \in C_1 \text{ et } \|\overrightarrow{\Omega_1 M}\|^2 - \|\overrightarrow{\Omega_2 M}\|^2 = R_1^2 - R_2^2 \right) \end{aligned}$$

Soit H la projection orthogonale de M sur $(\Omega_1\Omega_2)$ et I le milieu de (Ω_1, Ω_2) .

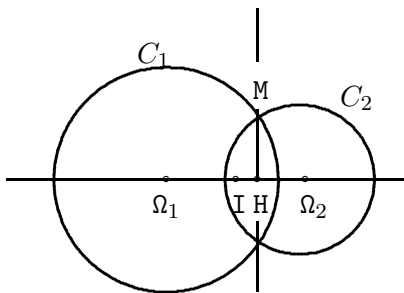
$$\|\overrightarrow{\Omega_1 M}\|^2 - \|\overrightarrow{\Omega_2 M}\|^2 = (\overrightarrow{\Omega_1 M} - \overrightarrow{\Omega_2 M}) \cdot (\overrightarrow{\Omega_1 M} + \overrightarrow{\Omega_2 M}) = 2\overrightarrow{\Omega_1\Omega_2} \cdot \overrightarrow{IM} = 2\overrightarrow{\Omega_1\Omega_2} \cdot \overrightarrow{IH}$$

Orientons la droite $(\Omega_1\Omega_2)$ par le vecteur $\overrightarrow{\Omega_1\Omega_2}$. On a l'équivalence

$$(M \in C_1 \cap C_2) \Leftrightarrow (M \in C_1 \text{ et } 2\overrightarrow{\Omega_1\Omega_2} \cdot \overrightarrow{IH} = R_1^2 - R_2^2) \Leftrightarrow (M \in C_1 \text{ et } \overrightarrow{IH} = \frac{R_1^2 - R_2^2}{2d})$$

Définissons H $\in (\Omega_1\Omega_2)$ par $\overrightarrow{IH} = \frac{R_1^2 - R_2^2}{2d}$. Existe-t-il M $\in C_1$ de projection orthogonale H ? Autrement dit, H est-il intérieur à C_1 ou sur C_1 ? On a

$$\overrightarrow{\Omega_1 H} = \overrightarrow{\Omega_1 I} + \overrightarrow{IH} = \frac{d}{2} + \frac{R_1^2 - R_2^2}{2d} = \frac{d^2 + R_1^2 - R_2^2}{2d}$$



$$\begin{aligned}
 R_1^2 - \Omega_1 H^2 &= (R_1 - \overline{\Omega_1 H})(R_1 + \overline{\Omega_1 H}) \\
 &= \frac{2dR_1 - d^2 - R_1^2 + R_2^2}{2d} \times \frac{2dR_1 + d^2 + R_1^2 - R_2^2}{2d} \\
 &= \frac{R_2^2 - (R_1 - d)^2}{2d} \times \frac{(R_1 + d)^2 - R_2^2}{2d} \\
 &= \frac{(R_2 - R_1 + d)(R_2 + R_1 - d)(R_1 + d - R_2)(R_1 + d + R_2)}{4d^2}
 \end{aligned}$$

Par hypothèse, ces quatre facteurs sont positifs ou nuls. Donc H est intérieur à C_1 ou sur C_1 et la perpendiculaire à $(\Omega_1 \Omega_2)$ en H coupe C_1 en deux points distincts ou lui est tangente en H.

(ii) Si $|R_1 - R_2| < d < R_1 + R_2$, H est intérieur aux cercles donc C_1 et C_2 ont deux points communs (11.2, p. 284).

(iii) Si $M \neq H$ est sur C_2 , on a : $M\Omega_1 + R_2 = M\Omega_1 + M\Omega_2 \geq \Omega_1 \Omega_2 = d = R_1 + R_2$, donc $M\Omega_1 \geq R_1$. Comme $M \in C_2$ et $M \neq H$, M n'est pas sur C_1 , donc $M\Omega_1 > R_1$.

(iv) On raisonne comme pour (iii). □

Corollaire 11.6. Soit trois longueurs a, b, c . Il existe un triangle non aplati ayant a, b, c pour longueurs des côtés si et seulement si $|a - b| < c < a + b$.

Démonstration. Par l'inégalité triangulaire, les côtés d'un triangle vérifient cette inégalité. Inversement, supposons cette inégalité et soit $[A, B]$ un segment de longueur c . Les cercles de centres A et B, de rayons b et a sont sécants. Si C est un point commun, le triangle ABC répond à la question. □

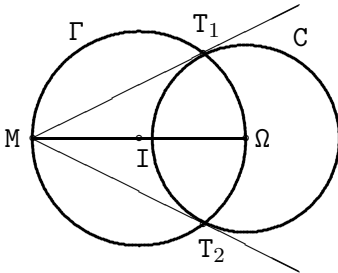
Proposition 11.7. Soit un cercle C de centre Ω , de rayon R , P, Q diamétralement opposés sur C . Un point M est sur C si et seulement si $\vec{MP} \cdot \vec{MQ} = 0$.

Démonstration. Comme $\vec{\Omega P} + \vec{\Omega Q} = \vec{0}$, on a $\vec{\Omega P} \cdot \vec{\Omega Q} = -R^2$ et

$$\vec{MP} \cdot \vec{MQ} = (\vec{M\Omega} + \vec{\Omega P}) \cdot (\vec{M\Omega} + \vec{\Omega Q}) = M\Omega^2 + \vec{M\Omega} \cdot (\vec{\Omega P} + \vec{\Omega Q}) + \vec{\Omega P} \cdot \vec{\Omega Q} = M\Omega^2 - R^2$$

D'où les équivalences $(M \in C) \Leftrightarrow (M\Omega^2 = R^2) \Leftrightarrow (\vec{MP} \cdot \vec{MQ} = 0)$ □

Proposition 11.8. Tangentes issues d'un point. Soit un cercle C de rayon R et de centre Ω . Par un point M extérieur passent deux tangentes.



Démonstration. Étant donné $T \in C$, (MT) est tangente à C si et seulement si (MT) et $(T\Omega)$ sont perpendiculaires, c'est-à-dire si $T \in \Gamma$, cercle de diamètre (Ω, M) . La question est donc : les cercles C et Γ sont-ils sécants ?

Le centre I de Γ est le milieu de (M, Ω) . La distance des centres est égale au rayon $\frac{1}{2}M\Omega$ de Γ .

Les inégalités $I\Omega - R < I\Omega < I\Omega + R$ sont claires. Reste à vérifier $R - I\Omega < I\Omega$, soit $R < 2I\Omega = M\Omega$. Ceci est vrai si et seulement si M est extérieur à C .

Remarquons que ceci donne une construction géométrique des tangentes à C issues de M : on prend les points T_1, T_2 communs à C et Γ . $\square$

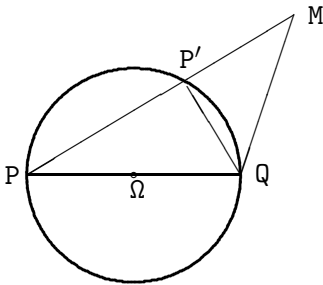
11.2 PUISSANCE D'UN POINT

La puissance d'un point par rapport à un cercle joue le rôle d'une distance.

11.2.1. Théorème fondamental

Théorème 11.9. Soit C un cercle de centre Ω et de rayon R et M un point du plan. Pour toute sécante passant par M coupant C en P et P' et pour tout diamètre (P, Q) de C , on a $\overrightarrow{MP} \cdot \overrightarrow{MQ} = \overrightarrow{MP} \times \overrightarrow{MP'} = M\Omega^2 - R^2$.

Cette quantité notée $C(M)$ est appelée **puissance**¹ de M par rapport à C .



Démonstration. Comme P, Q sont diamétralement opposés, $\overrightarrow{P'P}$ et $\overrightarrow{P'Q}$ sont orthogonaux (11.7, p. 286). On a :

$$\begin{aligned} \overrightarrow{MP} \times \overrightarrow{MP'} &= \overrightarrow{MP} \cdot (\overrightarrow{MQ} + \overrightarrow{QP'}) \\ &= \overrightarrow{MP} \cdot \overrightarrow{MQ} + \overrightarrow{MP} \cdot \overrightarrow{QP'} = \overrightarrow{MP} \cdot \overrightarrow{MQ} \\ &= (\overrightarrow{M\Omega} + \overrightarrow{\Omega P}) \cdot (\overrightarrow{M\Omega} + \overrightarrow{\Omega Q}) = M\Omega^2 - R^2 \end{aligned}$$

$\square$

Le signe de $C(M)$ détermine un régionnement du plan :

- si $C(M) < 0$, M est **intérieur**, le minimum de $C(M)$ étant atteint au centre avec $C(\Omega) = -R^2$,
- si $C(M) = 0$, M est sur C ,

1. Les valeurs algébriques $\overrightarrow{MP}$ et $\overrightarrow{MP'}$ dépendent du choix d'un des deux vecteurs unitaires opposés de la sécante, mais le produit $\overrightarrow{MP} \times \overrightarrow{MP'}$ n'en dépend pas.

- si $C(M) > 0$, M est **extérieur**. Si T est point de contact d'une tangente issue de M , on a (d'après le théorème de Pythagore) $C(M) = M\Omega^2 - T\Omega^2 = MT^2$.

Proposition 11.10. Soit Δ_1, Δ_2 , deux droites sécantes en O et M_1, N_1, M_2, N_2 des points de Δ_1, Δ_2 distincts de O . Pour qu'il existe un cercle ayant pour intersection M_1 et N_1 avec Δ_1 et M_2 et N_2 avec Δ_2 , il faut et il suffit que

$$\overline{OM_1} \times \overline{ON_1} = \overline{OM_2} \times \overline{ON_2}$$

Démonstration. La condition est nécessaire : s'il existe un tel cercle C , on a $C(O) = \overline{OM_1} \times \overline{ON_1} = \overline{OM_2} \times \overline{ON_2}$.

Inversement, supposons cette relation satisfaite.

Supposons M_1 et N_1 distincts. Le cercle $C = (M_1N_1M_2)$ recoupe Δ_2 en N'_2 éventuellement confondu avec M_2 si Δ_2 est tangente à C . On a

$$\overline{OM_2} \times \overline{ON_2} = \overline{OM_1} \times \overline{ON_1} = C(O) = \overline{OM_2} \times \overline{ON'_2}$$

d'où $N_2 = N'_2$ et les points sont cocycliques.

Si par contre M_1 et N_1 sont confondus ainsi que M_2 et N_2 , alors $OM_1 = OM_2$. Les droites orthogonales en M_1 et M_2 à Δ_1 et Δ_2 se coupent en Ω . On a

$$\Omega M_1^2 = O\Omega^2 - OM_1^2 = O\Omega^2 - OM_2^2 = \Omega M_2^2 = R^2$$

Le cercle de centre Ω et de rayon R est tangent en M_1 et M_2 à Δ_1 et Δ_2 . □

11.2.2. Expression analytique

Soit un repère orthonormé $(O, \vec{i}, \vec{j})$, Ω de coordonnées a, b , $R > 0$ et C le cercle de centre Ω et de rayon R . Alors C est l'ensemble des points $M(x, y)$ tels que $\|\vec{\Omega M}\|^2 = R^2$, c'est-à-dire tels que

$$(x - a)^2 + (y - b)^2 = R^2$$

C'est une équation du cercle C . Un cercle a une infinité d'équations, toutes proportionnelles : soit $\gamma \neq 0$, $\alpha = -\gamma a$, $\beta = -\gamma b$, $\delta = \gamma(a^2 + b^2 - R^2)$, alors

$$\gamma(x^2 + y^2) + 2\alpha x + 2\beta y + \delta = 0$$

est aussi une équation de C . Parmi ces équations, il en est une particulière : l'**équation normale** où le coefficient de $x^2 + y^2$ est 1.

Définition 11.11. Un **polynôme circulaire**, est un polynôme du type

$$P(X, Y) = \gamma(X^2 + Y^2) + 2\alpha X + 2\beta Y + \delta$$

Les polynômes circulaires forment un $\mathbb{R}$ -espace vectoriel de dimension 4.

Étant donné un polynôme circulaire non nul $P(X, Y)$, l'ensemble $V(P)$ des $M(x, y)$ tels que $P(x, y) = 0$ est-il un cercle ? Voici deux familles de cas :

- (1) **Cas où $\gamma \neq 0$** , i.e. P est effectivement de degré 2. Remplaçons P par son polynôme **normalisé**

$$\begin{aligned} P_0(X, Y) &= \frac{P(X, Y)}{\gamma} = X^2 + Y^2 + 2\frac{\alpha}{\gamma}X + 2\frac{\beta}{\gamma}Y + \frac{\delta}{\gamma} \\ &= \left(X - \frac{\alpha}{\gamma}\right)^2 + \left(Y - \frac{\beta}{\gamma}\right)^2 + \frac{\gamma\delta - \alpha^2 - \beta^2}{\gamma^2} \end{aligned}$$

Si $\alpha^2 + \beta^2 - \gamma\delta > 0$, $V(P)$ est le cercle de centre $\Omega(-\frac{\alpha}{\gamma}, -\frac{\beta}{\gamma})$ et de rayon $R = \frac{\sqrt{\alpha^2 + \beta^2 - \gamma\delta}}{|\gamma|}$.

Si $\alpha^2 + \beta^2 - \gamma\delta = 0$, $V(P)$ se réduit au point $\Omega(-\frac{\alpha}{\gamma}, -\frac{\beta}{\gamma})$. On dit que $V(P)$ est le **cercle-point** $\{\Omega\}$.

Si $\alpha^2 + \beta^2 - \gamma\delta < 0$, aucun point ne vérifie l'équation, $V(P)$ est vide.

- (2) **Cas où $\gamma = 0$** , c'est-à-dire que P est de degré 1. Ces polynômes forment un sous-espace vectoriel de dimension 3 de l'espace des polynômes circulaires.

Si $\alpha = \beta = 0$ et $\delta \neq 0$, aucun point ne vérifie l'équation, $V(P)$ est vide.

Si $(\alpha, \beta) \neq (0, 0)$, $V(P)$ est la droite d'équation $2\alpha x + 2\beta y + \delta = 0$.

Expression analytique de la puissance. Soit le polynôme **normal**

$$P(X, Y) = X^2 + Y^2 + 2\alpha X + 2\beta Y + \delta \text{ où } \alpha^2 + \beta^2 - \delta > 0$$

et le cercle $V(P) = C$ de centre $\Omega(-\alpha, -\beta)$ et de rayon $R = \sqrt{\alpha^2 + \beta^2 - \delta}$. Pour tout $M_0(x_0, y_0)$, on a

$$C(M_0) = \Omega M_0^2 - R^2 = (x_0 + \alpha)^2 + (y_0 + \beta)^2 - (\alpha^2 + \beta^2 - \delta) = P(x_0, y_0)$$

La puissance d'un point M_0 par rapport à un cercle C s'obtient en remplaçant x, y dans l'équation normale de C par les coordonnées x_0, y_0 de M_0 .

11.2.3. Espace vectoriel des fonctions circulaires

La discussion précédente est menée relativement à un repère orthonormé $(O, \vec{i}, \vec{j})$. Reprenons la « intrinsèquement », c'est-à-dire sans utiliser de repère.

Définition 11.12. Étant donné un point origine O , une **fonction circulaire** est une application non identiquement nulle

$$F: \mathcal{P} \rightarrow \mathbb{R}, M \mapsto F(M) = \gamma \|\vec{OM}\|^2 + 2\vec{V}_0 \cdot \vec{OM} + F(O) \text{ où } \vec{V}_0 \in \mathcal{P}$$

L'expression de F relativement à une autre origine O' est du même type avec la même constante γ ne dépendant que de F et notée $\gamma(F)$.

On a en effet :

$$\begin{aligned} F(\mathbf{M}) &= \gamma \|\vec{\mathbf{O}'\mathbf{M}} - \vec{\mathbf{O}'\mathbf{O}}\|^2 + 2\vec{V}_0 \cdot (\vec{\mathbf{O}'\mathbf{M}} - \vec{\mathbf{O}'\mathbf{O}}) + F(\mathbf{O}) \\ &= \gamma \|\vec{\mathbf{O}'\mathbf{M}}\|^2 + 2(\vec{V}_0 - \gamma \vec{\mathbf{O}'\mathbf{O}}) \cdot \vec{\mathbf{O}'\mathbf{M}} + F(\mathbf{O}') \\ &= \gamma \|\vec{\mathbf{O}'\mathbf{M}}\|^2 + 2\vec{V}_{0'} \cdot \vec{\mathbf{O}'\mathbf{M}} + F(\mathbf{O}') \end{aligned}$$

où on pose $\vec{V}_{0'} = \vec{V}_0 - \gamma \vec{\mathbf{O}'\mathbf{O}}$

Soit une fonction circulaire F non identiquement nulle, étudions l'ensemble $V(F) = \{\mathbf{M} \mid F(\mathbf{M}) = 0\}$

- **Cas où $\gamma(F) \neq 0$.** Soit le point $\Omega = \mathbf{O} - \frac{1}{\gamma} \vec{V}_0$. C'est le point tel que $\vec{V}_\Omega = \vec{0}$ appelé **centre** de F . On a donc

$$F(\mathbf{M}) = \gamma \|\vec{\Omega\mathbf{M}}\|^2 + F(\Omega) = \gamma \left(\|\vec{\Omega\mathbf{M}}\|^2 + \frac{F(\Omega)}{\gamma} \right)$$

Si $\frac{F(\Omega)}{\gamma} < 0$, $V(F)$ est le cercle de centre Ω et de rayon $\sqrt{-\frac{F(\Omega)}{\gamma}}$.

Si $\frac{F(\Omega)}{\gamma} = 0$, $V(F)$ se réduit au cercle-point $\{\Omega\}$.

Si $\frac{F(\Omega)}{\gamma} > 0$, aucun point n'annule F , donc $V(F)$ est vide.

Une fonction circulaire F est **normale** si $\gamma(F) = 1$. Si $\gamma(F) \neq 0$, on se ramène à ce cas en remplaçant F par sa **normalisée** $\frac{1}{\gamma(F)} F$. Pour F normale, on a $F(\mathbf{M}) = \Omega\mathbf{M}^2 + F(\Omega)$. Si $F(\Omega) \leq 0$, $V(F)$ est le cercle de centre Ω , de rayon $R = \sqrt{-F(\Omega)}$ et $F(\mathbf{M})$ est la puissance de $\mathbf{M}$ relativement à ce cercle.

- **Cas où $\gamma(F) = 0$.** Alors $F(\mathbf{M}) = 2\vec{V}_0 \cdot \vec{\mathbf{O}\mathbf{M}} + F(\mathbf{O})$ et F est une fonction affine (9.12, p. 250).

Si $\vec{V}_0 = \vec{0}$, F est constante et $V(F)$ est vide.

Si $\vec{V}_0 \neq \vec{0}$, $V(F)$ est une droite orthogonale à $\vec{V}_0$ partageant le plan en $V(F)$ et les demi-plans ouverts $\{\mathbf{M} \mid F(\mathbf{M}) > 0\}$ et $\{\mathbf{M} \mid F(\mathbf{M}) < 0\}$.

Expression analytique

Pour $\gamma \neq 0$, posons $\Phi(F) = -\gamma F(\Omega)$, du signe de $-\frac{F(\Omega)}{\gamma}$. Ramenant $\mathcal{P}$ à un repère orthonormé $(\mathbf{O}, \vec{i}, \vec{j})$, la fonction circulaire F est de type « polynomial circulaire »

$$\mathbf{M} \begin{pmatrix} x \\ y \end{pmatrix} \mapsto F(\mathbf{M}) = \gamma(x^2 + y^2) + 2\alpha x + 2\beta y + \delta$$

Un calcul immédiat donne

$$\Phi(F) = \|\vec{V}_0\|^2 - \gamma F(\mathbf{O}) = \alpha^2 + \beta^2 - \gamma\delta$$

Cette quantité, indépendante de O , ne dépend que de F . Elle existe aussi si $\gamma = 0$ et vaut alors $\|\vec{V}_0\|^2$. La discussion précédente montre que $V(F)$

- a une infinité de points si $\Phi(F) > 0$: c'est une droite ou un cercle selon que γ est nul ou non,
- est réduit à un point si $\Phi(F) = 0$ et $\gamma \neq 0$,
- est vide si $\Phi(F) < 0$ ou si $\Phi(F) = \gamma = 0$.

Proposition 11.13. *Les fonctions circulaires (en leur adjoignant la fonction nulle) constituent un $\mathbb{R}$ -espace vectoriel de dimension 4. On le note $\mathcal{C}$.*

Démonstration. Soit F et G deux fonctions circulaires définies par

$$F(M) = \gamma(F)OM^2 + \vec{V}_0(F) \cdot \vec{OM} + F(O), \quad G(M) = \gamma(G)OM^2 + \vec{V}_0(G) \cdot \vec{OM} + G(O)$$

O étant une origine quelconque. Soit (λ, μ) un couple de réels, l'application

$$\lambda F + \mu G : M \mapsto (\lambda\gamma(F) + \mu\gamma(G))OM^2 + (\lambda\vec{V}_0(F) + \mu\vec{V}_0(G)) \cdot \vec{OM} + \lambda F(O) + \mu G(O)$$

est une fonction circulaire. L'application $F \mapsto (\gamma(F), \vec{V}_0(F), F(O))$ est un isomorphisme de l'espace vectoriel $\mathcal{C}$ sur $\mathbb{R} \times P \times \mathbb{R}$ de dimension 4.

On peut aussi dire que la donnée d'un repère orthonormé détermine un isomorphisme entre $\mathcal{C}$ et l'espace des polynômes circulaires. $\square$

Exemple : La fonction de Leibnitz F (p. 255) relative à un système de points pondérés (α_i, A_i) est une fonction circulaire telle que $\gamma(F) = \sum \alpha_i$. La discussion relative à cette situation est un cas particulier de la discussion menée plus haut pour les fonctions circulaires.

11.2.4. Axe radical de deux cercles

Proposition 11.14. *Soit deux cercles C_1, C_2 de centres Ω_1, Ω_2 distincts. L'ensemble des points M ayant même puissance par rapport à ces cercles est une droite orthogonale à la droite des centres $(\Omega_1\Omega_2)$. On l'appelle **axe radical** des deux cercles.*

Démonstration. On suppose $C_1 = V(F_1)$ et $C_2 = V(F_2)$ où $F_1(M) = \Omega_1M^2 - R_1^2$ et $F_2(M) = \Omega_2M^2 - R_2^2$. On déduit le résultat de ce que $F = F_1 - F_2$ est la fonction affine $M \mapsto F(M)$ où, I étant milieu de (Ω_1, Ω_2) ,

$$\begin{aligned} F(M) &= \|\vec{\Omega_1M}\|^2 - \|\vec{\Omega_2M}\|^2 - R_1^2 + R_2^2 \\ &= (\vec{\Omega_1M} + \vec{\Omega_2M}) \cdot (\vec{\Omega_1M} - \vec{\Omega_2M}) - R_1^2 + R_2^2 \\ &= 2\vec{IM} \cdot \vec{\Omega_1\Omega_2} - R_1^2 + R_2^2 \end{aligned}$$

$\square$

Corollaire 11.16. Quatre points distincts A, B, C, D sont cocycliques ou alignés si et seulement si $(\widehat{CA, CB}) \equiv (\widehat{DA, DB}) \pmod{\pi}$.

Proposition 11.17. Complément au théorème de l'angle inscrit. Soit A, B deux points distincts, $\theta \in]0, \pi[$ et C_θ le cercle ensemble des M tels que $(\widehat{MA, MB}) \equiv \theta \pmod{\pi}$. La fonction affine $M \mapsto [A, B, M]$ détermine un partage du plan en (9.13, p. 251)

- la droite (AB) ,
- le demi-plan ouvert $\mathcal{P}^+$ des M tels que $[A, B, M] > 0$,
- le demi-plan ouvert $\mathcal{P}^-$ des M tels que $[A, B, M] < 0$,

Le cercle C_θ privé des points A, B se compose des deux arcs de cercles

$$C_\theta^+ = \{M \mid (\widehat{MA, MB}) \equiv \theta \pmod{2\pi}\} = C_\theta \cap \mathcal{P}^+,$$

$$C_\theta^- = \{pM \mid (\widehat{MA, MB}) \equiv \theta + \pi \pmod{2\pi}\} = C_\theta \cap \mathcal{P}^-.$$

Démonstration. On a $[A, B, M] = MA \times MB \times \sin(\widehat{MA, MB})$, du signe de $\sin(\widehat{MA, MB})$ (9.17, p. 254). Si $\theta \in]0, \pi[$, $\sin \theta > 0$, $\sin(\theta + \pi) = -\sin \theta < 0$, d'où

$C_\theta^+ = \{M \mid (\widehat{MA, MB}) \equiv \theta \pmod{2\pi}\} = C_\theta \cap \mathcal{P}^+$ est l'ensemble des points M du cercle C_θ tel que le triangle ABM soit direct,

$C_\theta^- = \{M \mid (\widehat{MA, MB}) \equiv \theta + \pi \pmod{2\pi}\} = C_\theta \cap \mathcal{P}^-$ est l'ensemble des points M du cercle C_θ tel que le triangle ABM soit inverse. $\square$

11.3.2. Applications

Proposition 11.18. Théorème de Simson. Soit un triangle ABC , C le cercle circonscrit, M de projections orthogonales U, V, W sur $(BC), (CA), (AB)$. Pour que U, V, W soient alignés, il faut et il suffit que $M \in C$.

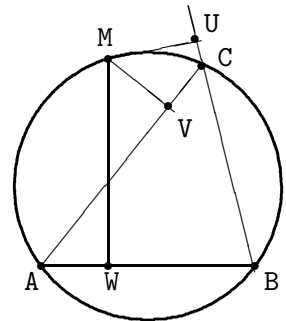
Démonstration. Par cocyclicité des groupes de points (M, A, V, W) , (M, B, W, U) , (M, C, U, V) , on a, modulo π

$$\begin{aligned} (\widehat{UV, UW}) &\equiv (\widehat{UV, UM}) + (\widehat{UM, UW}) \equiv (\widehat{CV, CM}) + (\widehat{BM, BW}) \\ &\equiv (\widehat{CA, CM}) + (\widehat{BM, BA}) \equiv (\widehat{CA, BA}) - (\widehat{CM, BM}) \end{aligned}$$

Donc U, V, W sont alignés si et seulement si on a les propriétés équivalentes

$$\begin{aligned} ((\widehat{UV, UW}) \equiv 0 \pmod{\pi}) &\Leftrightarrow ((\widehat{CA, BA}) \equiv (\widehat{CM, BM}) \pmod{\pi}) \\ &\Leftrightarrow (M \in C) \end{aligned}$$

La droite portant U, V, W s'appelle la **droite de Simson** de M . $\square$



11.4 FAISCEAUX DE CERCLES

11.4.1. Faisceaux à points de base

Définition 11.19. Soit A, B deux points distincts. On appelle **faisceau de cercles à points de base** A, B l'ensemble des cercles passant par A, B . On convient de considérer la droite (AB) comme **cercle dégénéré** du faisceau.

Ces cercles sont centrés sur la médiatrice de (A, B) . Comme A et B ont une puissance nulle par rapport à tout cercle du faisceau, l'axe radical de deux cercles distincts du faisceau est la droite (AB) . Par tout M distinct de A, B passe un unique cercle du faisceau. On a une bijection de $\mathbb{R}/\pi\mathbb{Z}$ sur le faisceau : à θ , on associe $C_\theta = \{M \mid (\widehat{MA, MB}) \equiv \theta \pmod{\pi}\}$. Le « cercle » associé à l'angle nul (modulo π) est l'axe radical (AB) .

11.4.2. Faisceaux de cercles à points limites

Théorème 11.20. Théorème d'Apollonius. Soit A, B deux points distincts du plan et k un réel positif. L'ensemble C_k des M tels que $\frac{MA}{MB} = k$ est

- pour $k = 1$, la médiatrice de (A, B) ,
- pour $k \neq 1$, un cercle centré sur (AB) , coupant le droite (AB) en deux points I, J tels que $[A, B, I, J] = -1$.

L'axe radical de deux cercles C_k et $C_{k'}$ distincts est la médiatrice de (A, B) .

Démonstration. Pour que $M \in C_k$, il faut et il suffit que $MA^2 - k^2 MB^2 = 0$. L'application $M \mapsto MA^2 - k^2 MB^2$ est une fonction de Leibnitz (9.19, p. 255).

Pour $k = 1$, on a la médiatrice de (A, B) .

Pour $k \neq 1$, on a un cercle centré en le barycentre $\Omega = \frac{A - k^2 B}{1 - k^2}$.

Retrouvons ceci autrement.

Supposons $k \neq 1$. On a

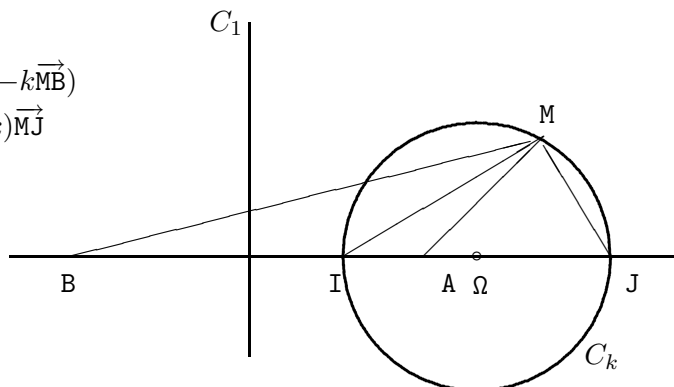
$$\begin{aligned} MA^2 - k^2 MB^2 &= (\vec{MA} + k\vec{MB}) \cdot (\vec{MA} - k\vec{MB}) \\ &= (1+k)\vec{MI} \cdot (1-k)\vec{MJ} \end{aligned}$$

I, J étant les barycentres

$$I = \frac{A+kB}{1+k}, J = \frac{A-kB}{1-k}.$$

On a $M \in C_k$ si et seulement si $\vec{MI} \cdot \vec{MJ} = 0$, c'est-à-dire si M est sur le cercle de diamètre (I, J) (11.7, p. 286).

La division A, B, I, J est harmonique car $\frac{\overline{JA}}{\overline{JB}} = k = -\frac{\overline{IA}}{\overline{IB}}$.



La puissance d'un point M relativement à C_k est

$$C_k(M) = \vec{MI} \cdot \vec{MJ} = \frac{\vec{MA} + k\vec{MB}}{1+k} \cdot \frac{\vec{MA} - k\vec{MB}}{1-k} = \frac{MA^2 - k^2 MB^2}{1-k^2}$$

Pour tout M appartenant à la médiatrice de (A, B) , on a $C_k(M) = MA^2 = MB^2$, quantité indépendante de k . $\square$

Définition 11.21. Par extension, l'axe radical C_1 et les cercles points $\{A\} = C_0$ et $\{B\} = C_\infty$ sont considérés comme faisant partie de cette famille de cercles, appelée **faisceau de cercles à points limites** A, B . On voit alors que par tout point M passe un unique cercle du faisceau à points limites A, B .

Questions :

Deux cercles distincts $C_k, C_{k'}$ d'un faisceau à points limites A, B n'ont pas de point commun car $\frac{MA}{MB}$ n'est pas égal à la fois à k et k' . Inversement, deux cercles sans point commun appartiennent-ils à un même faisceau à points limites ? Pourquoi cette même terminologie de **faisceau** pour des situations apparemment dissemblables : faisceau à points limites, faisceau à points de base ? C'est à la section 11.5.2. qu'il est répondu à ces questions.

Exercice 11.1. On donne les cercles C_1, C_2 par des équations relativement à un repère orthonormé

$$\text{pour } C_1 = V(P_1) \quad , \quad P_1(x, y) = \gamma_1(x^2 + y^2) + 2\alpha_1 x + 2\beta_1 y + \delta_1 = 0$$

$$\text{pour } C_2 = V(P_2) \quad , \quad P_2(x, y) = \gamma_2(x^2 + y^2) + 2\alpha_2 x + 2\beta_2 y + \delta_2 = 0$$

Quelle est l'équation de leur axe radical Δ ?

Exercice 11.2. Sur un cercle C , on donne quatre points A, B, C, M distincts. Étant donné une direction X , les droites $A + X, B + X, C + X$ recoupent C en A', B', C' . Les droites $(MA'), (MB'), (MC')$ coupent $(BC), (CA), (AB)$ en P, Q, R .

1) Montrer que M, P, Q, C sont cocycliques.

2) En déduire que P, Q, R sont alignés.

Exercice 11.3. Soit un triangle ABC , $\mathcal{X}, \mathcal{Y}$ les bissectrices de $(\widehat{CA, CB})$ coupant (AB) en I, J . Montrer que

$$\frac{IA}{IB} = \frac{JA}{JB} = \frac{CA}{CB}$$

Exercice 11.4. 1) Soit Δ une droite, $O \notin \Delta$, A la projection orthogonale de O sur Δ , $A' \in (OA)$ distinct de O , C le cercle de diamètre $[O, A']$. Pour tout $M \in \Delta$, (OM) recoupe C en M' . Montrer que A, A', M, M' sont cocycliques.

2) Soit O un point, $k \neq 0$. On appelle **inversion** de **pôle** O et de **puissance** k l'application $\text{Inv}(O, k)$ de $\mathcal{P}$ privé de O sur lui-même associant à M le point $M' \in (OM)$ tel que $\overline{OM} \times \overline{OM'} = k$. Que deviennent par inversion les droites du plan ?

Exercice 11.5. 1) Soit C un cercle, $O \notin C$, h une homothétie de centre O , $C' = h(C)$. Étant donné deux droites Δ_1 et Δ_2 passant par O coupant C respectivement en M_1, N_1 et M_2, N_2 , soit M'_1, N'_1, M'_2, N'_2 les transformés de M_1, N_1, M_2, N_2 par h . Montrer que M_1, M_2, N'_1, N'_2 sont cocycliques.

2) Que deviennent par une inversion de pôle O les cercles ne passant pas par O ?

11.5 L'ESPACE DES CERCLES DU PLAN

11.5.1. Orthogonalité entre cercles

Considérons l'espace $\mathcal{C}$ de dimension 4 des fonctions circulaires (11.12, p. 289).

Définition 11.22. Soit F définie par :

$$\begin{aligned} F: \mathcal{P} &\longrightarrow \mathbb{R} \\ M &\longmapsto \gamma(F) \|\overrightarrow{OM}\|^2 + 2\overrightarrow{V_0}(F) \cdot \overrightarrow{OM} + F(O) \end{aligned}$$

où O est un point origine. À une fonction circulaire F , on associe l'ensemble $V(F)$ des points M tels que $F(M) = 0$. La quantité $\Phi(F)$ ne dépendant que de F définit une forme quadratique sur $\mathcal{C}$

$$\begin{aligned} \Phi: \mathcal{C} &\longrightarrow \mathbb{R} \\ F &\longmapsto \Phi(F) = \|\overrightarrow{V_0}(F)\|^2 - \gamma(F)F(O) \end{aligned}$$

(1) Si $\Phi(F) > 0$, l'ensemble $V(F)$ est infini : cercle si $\gamma(F) \neq 0$, droite si $\gamma(F) = 0$. Par extension, les droites sont considérées comme cercles particuliers. On l'a déjà fait en considérant l'axe radical d'un faisceau à points de base ou à points limites comme cercle particulier du faisceau.

(2) Si $\Phi(F) < 0$, $V(F)$ est vide.

(3) Si $\Phi(F) = 0$ et $\gamma(F) \neq 0$, $V(F)$ est un cercle-point. On adjoint au plan $\mathcal{P}$ un élément supplémentaire : le **point à l'infini** ∞ . Pour $\Phi(F) = 0$ et $\gamma(F) = 0$, alors Φ est constante sur tout $\mathcal{P}$. Par extension, on dit alors que $V(F) = \{\infty\}$ est le **cercle-point à l'infini**. Les éléments isotropes de $\mathcal{C}$ sont les fonctions circulaires associées aux cercles-points. On y ajoute les fonctions constantes associées au cercle-point $\{\infty\}$.

Définition 11.23. On appellera **cercle généralisé** un ensemble $V(F)$ du type précédent. Un cercle au sens ordinaire est donc un cercle généralisé particulier appelé **cercle à centre**. L'ensemble $\tilde{\mathcal{P}} = \mathcal{P} \cup \{\infty\}$ s'appelle le plan **complété annalagmatique** de $\mathcal{P}$.

Définition 11.24. Soit une origine O . L'application $\Phi: \mathcal{C} \rightarrow \mathbb{R}$ est une forme quadratique sur $\mathcal{C}$ de forme polaire $\mathcal{C} \times \mathcal{C} \rightarrow \mathbb{R}$ (encore notée Φ)

$$(F_1, F_2) \mapsto \Phi(F_1, F_2) = \vec{V}_0(F_1) \cdot \vec{V}_0(F_2) - \frac{1}{2} \left(F_2(O)\gamma(F_1) + F_1(O)\gamma(F_2) \right)$$

Deux cercles généralisés $C_1 = V(F_1)$ et $C_2 = V(F_2)$ sont dits **orthogonaux** si F_1 et F_2 sont deux vecteurs de $\mathcal{C}$ orthogonaux pour la forme quadratique Φ .

Ceci ne dépend pas des fonctions circulaires F_1 et F_2 définissant C_1 et C_2 qui sont les $\lambda_1 F_1$ et $\lambda_2 F_2$ (λ_1, λ_2 non nuls). La nullité de $\Phi(F_1, F_2)$ équivaut à la nullité de $\Phi(\lambda_1 F_1, \lambda_2 F_2) = \lambda_1 \lambda_2 \Phi(F_1, F_2)$. L'orthogonalité entre cercles généralisés est donc bien définie. Décrivons géométriquement cette situation pour les différentes sortes de cercles généralisés.

(1) **Cas de deux droites.** On a

$$F_1(M) = 2\vec{V}_0(F_1) \cdot \vec{OM} + F_1(O), \quad F_2(M) = 2\vec{V}_0(F_2) \cdot \vec{OM} + F_2(O)$$

$$\Phi(F_1, F_2) = \vec{V}_0(F_1) \cdot \vec{V}_0(F_2)$$

$V(F_1)$ et $V(F_2)$ sont des droites orthogonales aux vecteurs $\vec{V}_0(F_1)$ et $\vec{V}_0(F_2)$ et deux droites sont orthogonales au sens précédent si et seulement elles sont orthogonales au sens ordinaire.

(2) **Cas d'un cercle ou droite et d'un cercle-point.** On place l'origine en le centre O du cercle-point : $V(F_1) = \{O\}$. Alors

$$F_1(M) = \gamma(F_1)\|\vec{OM}\|^2, \quad F_2(M) = \gamma(F_2)\|\vec{OM}\|^2 + 2\vec{V}_0(F_2) \cdot \vec{OM} + F_2(O)$$

$$\Phi(F_1, F_2) = -\frac{1}{2}F_2(O)\gamma(F_1)$$

$\Phi(F_1, F_2)$ est nul si et seulement si $F_2(O) = 0$. Un cercle-point $V(F_1) = \{O\}$ et un cercle ou droite $V(F_2)$ sont orthogonaux si et seulement si O appartient à l'ensemble droite ou cercle $V(F_2)$.

(3) **Cas d'un cercle ou droite et du cercle-point $\{\infty\}$.** L'origine est O quelconque. Alors

$$F_1(M) = F_1(O), \quad F_2(M) = \gamma(F_2)\|\vec{OM}\|^2 + 2\vec{V}_0(F_2) \cdot \vec{OM} + F_2(O)$$

$$\Phi(F_1, F_2) = -\frac{1}{2}F_1(O)\gamma(F_2)$$

$\Phi(F_1, F_2) = 0$ si et seulement si $\gamma(F_2) = 0$. Le cercle-point $\{\infty\}$ et le cercle ou droite $V(F_2)$ sont orthogonaux si et seulement si $V(F_2)$ est une droite. On peut donc énoncer : les droites sont les cercles généralisés passant par le point à l'infini.

- (4) **Cas d'un cercle à centre et d'une droite.** On place l'origine 0 en le centre du cercle $V(F_1)$. On a donc

$$F_1(M) = \gamma(F_1) \|\vec{OM}\|^2 + F_1(0), \quad F_2(M) = 2\vec{V_0}(F_2) \cdot \vec{OM} + F_2(0)$$

$$\Phi(F_1, F_2) = -\frac{1}{2}F_2(0)\gamma(F_1)$$

$\Phi(F_1, F_2) = 0$ si et seulement si $F_2(0) = 0$. Un cercle à centre et une droite sont orthogonaux si et seulement si la droite passe par le centre du cercle.

- (5) **Cas de deux cercles à centres.** Cela fait l'objet de la proposition suivante :

Proposition 11.25. Soit C_1, C_2 deux cercles de centres Ω_1, Ω_2 , de rayons R_1, R_2 , $d = \Omega_1\Omega_2$. On a $\Phi(F_1, F_2) = \frac{1}{2}(R_1^2 + R_2^2 - d^2)$ et il y a équivalence entre les énoncés suivants :

- a) C_1 et C_2 sont orthogonaux,
- b) $C_1(\Omega_2) = R_2^2$, (respectivement $C_2(\Omega_1) = R_1^2$),
- c) C_1 et C_2 se coupent et, pour $T \in C_1 \cap C_2$, $(\Omega_1 T)$ et $(\Omega_2 T)$ sont orthogonales,
- d) C_1 et C_2 se coupent et, pour $T \in C_1 \cap C_2$, $(\Omega_1 T)$ est tangente à C_2 (respectivement $(\Omega_2 T)$ est tangente à C_1),
- e) il existe une sécante passant par le centre de l'un coupant ces cercles en quatre points en division harmonique,
- f) toute sécante passant par le centre de l'un coupe ces cercles en quatre points en division harmonique.

Démonstration. Soit F_1 et F_2 les fonctions circulaires normales définissant C_1 et C_2 . Pour tout M et $i = 1, 2$, on a $C_i(M) = F_i(M) = \Omega_i M^2 - R_i^2$. Si 0 est milieu de (Ω_1, Ω_2) , on a $F_i(M) = OM^2 - 2\vec{O\Omega_i} \cdot \vec{OM} + \Omega_i O^2 - R_i^2$, d'où

$$\Phi(F_1, F_2) = \frac{1}{2}(R_1^2 + R_2^2 - d) = \frac{1}{2}(R_1^2 - F_2(\Omega_1)) = \frac{1}{2}(R_2^2 - F_1(\Omega_2))$$

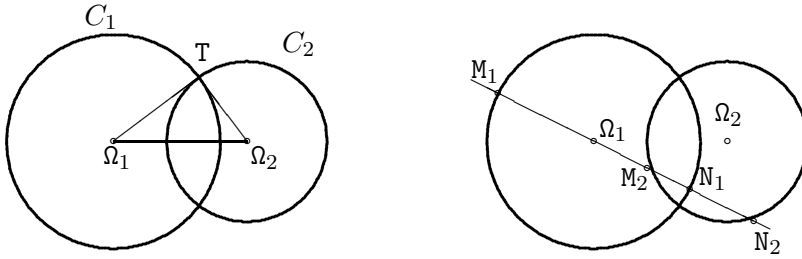
Ceci donne l'équivalence de a), b1) et b2).

Ces cercles se coupent si et seulement si $|R_1 - R_2| \leq d \leq R_1 + R_2$, donc si $|\Phi(F_1, F_2)| \leq 2R_1 R_2$ (11.5, p. 285). Si $\Phi(F_1, F_2) = 0$, l'inégalité est vérifiée et les cercles se coupent. Soit T un point commun. On a

$$F_1(M) = TM^2 - 2\vec{T\Omega_1} \cdot \vec{TM}, \quad F_2(M) = TM^2 - 2\vec{T\Omega_2} \cdot \vec{TM}$$

et $\Phi(F_1, F_2) = \vec{T\Omega_1} \cdot \vec{T\Omega_2}$. On en déduit l'équivalence de a) et c).

L'équivalence de c), d1) et d2) est claire.



Soit $\mathcal{D}$ une sécante passant par Ω_1 coupant C_1 en M_1, N_1 et C_2 en M_2, N_2 ,

$$F_2(\Omega_1) = \overline{\Omega_1 M_2} \cdot \overline{\Omega_1 N_2}, \quad R_1^2 = \Omega_1 M_1^2 = \Omega_1 N_1^2$$

Alors b2) équivaut à $\overline{\Omega_1 M_2} \cdot \overline{\Omega_1 N_2} = \Omega_1 M_1^2 = \Omega_1 N_1^2$, ce qui équivaut à

$$[M_1, N_1, M_2, N_2] = -1$$

par la relation de Newton caractérisant les divisions harmoniques (8.26, p. 217), d'où l'équivalence de b), e) et f). $\square$

Proposition 11.26. *La signature de Φ est $(3, 1)$.*

Démonstration. Soit $C_1 = V(F_1)$, $C_2 = V(F_2)$ deux cercles à centre orthogonaux, P_1, P_2 les points communs. Alors $\{P_1\}$ et $\{P_2\}$ sont orthogonaux à C_1 et C_2 . Posons $\varphi_i(M) = P_i M^2$ pour tout point M . On a une base $(F_1, F_2, \varphi_1, \varphi_2)$ de $\mathcal{C}$ et la décomposition de $\mathcal{C} = \text{Vect}(F_1, F_2) \oplus^\perp \text{Vect}(\varphi_1, \varphi_2)$ en somme directe orthogonale. La matrice de Φ dans cette base est

$$\begin{pmatrix} \Phi(F_1) & 0 & 0 & 0 \\ 0 & \Phi(F_2) & 0 & 0 \\ 0 & 0 & 0 & \Phi(\varphi_1, \varphi_2) \\ 0 & 0 & \Phi(\varphi_1, \varphi_2) & 0 \end{pmatrix}$$

Les restrictions de Φ à $\text{Vect}(F_1, F_2)$ et $\text{Vect}(\varphi_1, \varphi_2)$ sont respectivement définie positive de signature $(2, 0)$ et hyperbolique de signature $(1, 1)$, d'où le résultat. $\square$

11.5.2. Faisceaux de cercles : nouvel éclairage.

La donnée d'un plan vectoriel $\Pi \subset \mathcal{C}$ définit le **faisceau de cercles** $\mathcal{F}(\Pi)$ formé des cercles $V(F)$ où F décrit Π , avec $\Phi(F) \geq 0$ pour que $V(F)$ soit un cercle généralisé non vide (11.22, p. 296). Si donc on choisit deux cercles généralisés $C_1 = V(F_1)$ et $C_2 = V(F_2)$ on aura déterminé un faisceau $\mathcal{F}(C_1, C_2)$ correspondant au plan vectoriel $\text{Vect}(F_1, F_2)$.

Proposition 11.27. *Soit $\mathcal{F}$ un faisceau de cercles et M un point. Alors*

- ou bien M appartient à tout cercle généralisé du faisceau,
- ou bien M appartient à un unique cercle généralisé du faisceau.

Démonstration. Supposons $\mathcal{F}$ défini par $C_1 = V(F_1)$ et $C_2 = V(F_2)$. On cherche les couples (λ_1, λ_2) tels que $\lambda_1 F_1(M) + \lambda_2 F_2(M) = 0$.

Si $M \in C_1 \cap C_2$, $F_1(M) = F_2(M) = 0$, donc tout couple (λ_1, λ_2) convient.

Si $F_1(M)$ et $F_2(M)$ ne sont pas tous deux nuls, les couples (λ_1, λ_2) cherchés sont proportionnels à $(F_2(M), -F_1(M))$. L'unique cercle passant par M est donc $V(F_2(M)F_1 - F_1(M)F_2)$. $\square$

Nous allons maintenant classer les faisceaux de cercles. Tout repose encore sur la forme quadratique Φ , dont la signature est $(3, 1)$. Si $\mathcal{F}(\Pi)$ est un faisceau, il y a trois cas pour la restriction $\Phi|_{\Pi}$ de Φ à Π :

- (1) $\Phi|_{\Pi}$ est de signature $(1, 1)$, les faisceaux définis par ces plans s'appellent **faisceaux à points limites**.
- (2) $\Phi|_{\Pi}$ est de signature $(2, 0)$, les faisceaux définis par ces plans s'appellent **faisceaux à points de base**.
- (3) $\Phi|_{\Pi}$ est dégénérée, c'est-à-dire que le plan vectoriel Π est isotrope : on parle alors de **faisceaux singuliers**.

Avant d'entrer dans l'étude géométrique de ces différents cas, parlons d'orthogonalité. Comme Φ est non dégénérée, l'application involutive $\Pi \mapsto \Pi^\perp$ établit une bijection entre :

- l'ensemble des plans vectoriels P tels que $\Phi|_P$ est de signature $(1, 1)$,
- l'ensemble des plans vectoriels Q tels que $\Phi|_Q$ est de signature $(2, 0)$.

En effet, dans ces deux cas les plans sont non isotropes, la signature est imposée par le théorème de Sylvester (cf. 6.18).

Définition 11.28. Si $\mathcal{F}(\Pi)$ est un faisceau de cercles, $\mathcal{F}(\Pi^\perp)$ s'appelle le **faisceau conjugué**. Pour Π non isotrope, on a la décomposition en somme directe orthogonale $\mathcal{C} = \Pi \oplus^\perp \Pi^\perp$. Donc $\mathcal{F}(\Pi^\perp)$ est formé des cercles orthogonaux à tout cercle de $\mathcal{F}(\Pi)$. On notera $\mathcal{F}^\perp$ le faisceau conjugué au faisceau $\mathcal{F}$.

Un cercle généralisé appartient à $\mathcal{F}^\perp$ si et seulement s'il est orthogonal à deux cercles généralisés distincts de $\mathcal{F}$.

a) Étude des faisceaux à points limites.

On suppose $\Phi|_{\Pi}$ de signature $(1, 1)$. Alors Π a une base formée de deux vecteurs isotropes, fonctions circulaires f et g telles que $\Phi(f) = \Phi(g) = 0$.

- **Cas général** $V(f) = \{A\}$ et $V(g) = \{B\}$ sont deux cercles-points distincts. Supposant f et g normalisées, $f(M) = AM^2$ et $g(M) = BM^2$ pour tout M . Les fonctions circulaires de $\text{Vect}(f, g)$ sont du type

$$F = \lambda f + \mu g : M \longmapsto \lambda AM^2 + \mu BM^2$$

On a clairement $V(F) = \emptyset$ si λ et μ sont non nuls de même signe. Pour λ, μ non nuls de signes contraires,

$$V(F) = \left\{ M \mid \lambda MA^2 + \mu MB^2 = 0 \right\} = \left\{ M \mid \frac{MA}{MB} = k \right\} \text{ où } k = \sqrt{\frac{|\mu|}{|\lambda|}}$$

On trouve la notion de faisceau de cercles à points-limites A, B au sens de la définition 11.21, p. 295. Ce faisceau est noté $\mathcal{F}(A, B)$.

- **Cas de cercles concentriques** c'est le cas où g est constante, c'est-à-dire où $V(g) = \{\infty\}$. On peut supposer que $f(M) = MA^2$ et $g(M) = 1$. Les fonctions circulaires de $\text{Vect}(f, g)$ sont du type

$$F = \lambda f + \mu : M \mapsto \lambda MA^2 + \mu$$

On a $V(F) = \emptyset$ si λ et μ non nuls ont même signe. Pour λ, μ non nuls de signes contraires,

$$V(F) = \left\{ M \mid \lambda MA^2 + \mu = 0 \right\} = \left\{ M \mid MA = R \right\} \text{ où } R = \sqrt{\frac{|\mu|}{|\lambda|}}$$

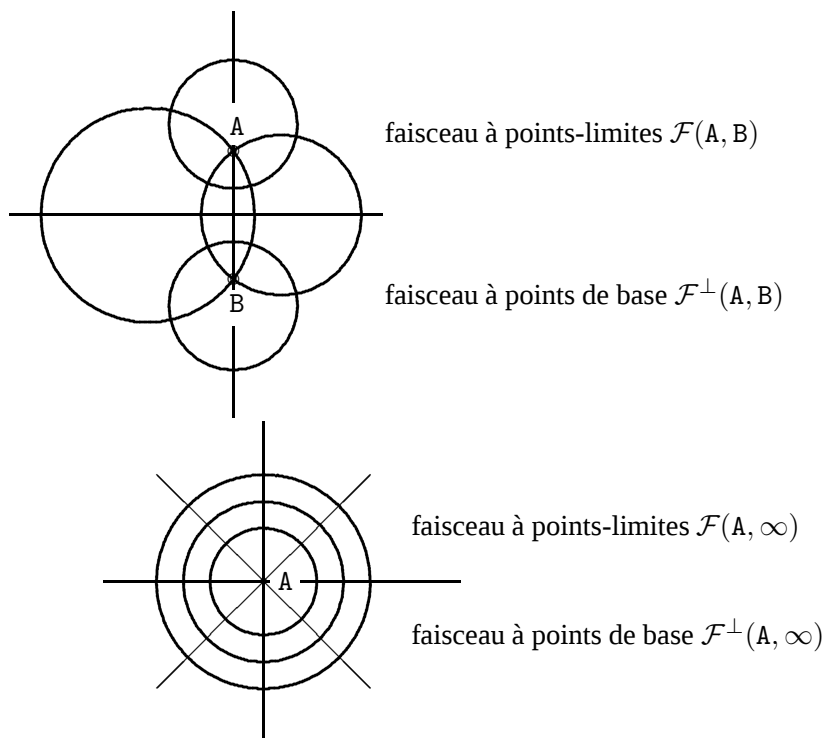
On trouve l'ensemble des cercles de centres A . C'est un cas particulier de faisceau à points-limites, noté $\mathcal{F}(A, \infty)$.

b) Étude des faisceaux à points de base.

Un faisceau $\mathcal{F} = \mathcal{F}(\Pi)$ à points de base est donc un faisceau tel que $\Phi|_{\Pi}$ est de signature $(2, 0)$, donc définie positive. Pour tout $F \in \Pi$, $\Phi(F) > 0$ donc le cercle généralisé $V(F)$ est défini pour tout $F \in \Pi$. Le faisceau conjugué est alors un faisceau à points limites A, B (avec éventuellement $B = \infty$).

- **Cas général** : on a $\mathcal{F}^{\perp} = \mathcal{F}(A, B)$, donc $\mathcal{F}$ est formé des cercles généralisés orthogonaux aux cercles-points $\{A\}$ et $\{B\}$, il s'agit donc des cercles passant par A et B . On retrouve la notion de faisceau à points de base de 11.19, p. 294.
- **Cas de droites concourantes** : c'est le cas où $\mathcal{F}^{\perp} = \mathcal{F}(A, \infty)$ est formé des cercles de centre A . Alors $\mathcal{F}$ est formé des cercles généralisés orthogonaux à $\{A\}$ et $\{\infty\}$, c'est-à-dire des droites passant par A .

La figure suivante décrit ces deux situations.



c) Étude des faisceaux singuliers.

Ce sont donc les faisceaux où $\mathcal{F} = \mathcal{F}(\Pi)$, avec Π isotrope. Alors $\Phi|_\Pi$ est de rang 1. Il existe dans Π une unique droite isotrope Φ -orthogonale à tout élément de Π . Soit f un générateur de cette droite isotrope.

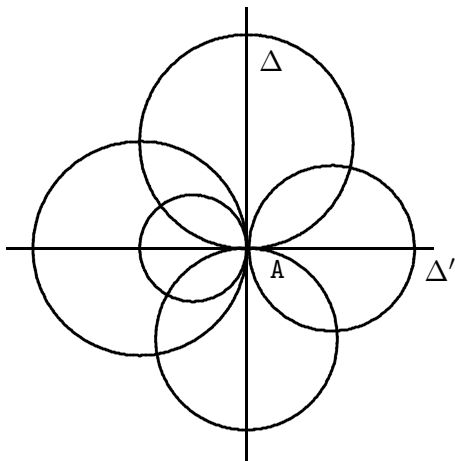
- **Cas général d'un faisceau de cercles tangents** On suppose f non constante, du type $M \mapsto AM^2$, associée au cercle-point $\{A\}$. Tout cercle de $\mathcal{F}$ est orthogonal au cercle-point $\{A\}$, donc passe par A. Deux cercles C, C' distincts de $\mathcal{F}$ passent par A sans avoir un deuxième point B commun, sinon $\mathcal{F}$ serait à points de base A, B. Tous les cercles de $\mathcal{F}$ ont même tangente Δ en A.

Inversement, soit C un cercle tangent en A à Δ (ou $C = \Delta$) et $M \in C$ distinct de A. L'unique cercle de $\mathcal{F}$ passant par M ne peut être que C , d'où $C \in \mathcal{F}$ (11.27, p. 299). Le faisceau conjugué est aussi singulier, formé des cercles tangents en A à la droite Δ' orthogonale à Δ .

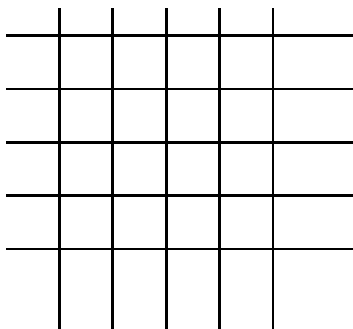
- **Cas d'un faisceau de droites parallèles** On suppose $f = 1$, associée au cercle-point $\{\infty\}$. Les cercles (généralisés) de $\mathcal{F}$ sont orthogonaux à $\{\infty\}$, donc sont des droites. Deux droites de $\mathcal{F}$ n'ont pas de point commun $B \neq \infty$ sinon $\mathcal{F}$ serait le faisceau à points de base $\mathcal{F}(B, \infty)$ des droites passant par B. Donc $\mathcal{F}$ est formé de droites de même direction D et de $\{\infty\}$.

Inversement, soit C une droite de direction D et $M \in C$. L'unique cercle généralisé de $\mathcal{F}$ passant par M ne peut être que C , d'où $C \in \mathcal{F}$.

Le faisceau conjugué est formé des droites de direction $D^\perp$, et les figures suivantes décrivent cette situation.



faisceaux conjugués
de cercles tangents
en A à Δ et Δ'



faisceaux conjugués
de droites parallèles
de directions D et $D^\perp$

Proposition 11.29. Axe radical. Soit $\mathcal{F}$ un faisceau de cercles qui n'est formé, ni de cercles concentriques, ni de droites concourantes ou parallèles. L'axe radical Δ de deux cercles de $\mathcal{F}$ est le même pour toute paire de cercles de $\mathcal{F}$ et est l'unique droite de $\mathcal{F}$.

Démonstration. S'il existait deux droites Δ et Δ' dans $\mathcal{F}$, alors on aurait $\mathcal{F} = \mathcal{F}(\Delta, \Delta')$. Si Δ et Δ' se coupent en A (resp. sont parallèles), $\mathcal{F}$ serait formé des droites passant par A (resp. parallèles à Δ et Δ'). Ces éventualités sont exclues.

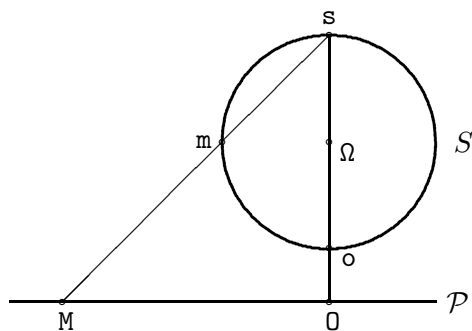
Soit $C_1 = V(F_1)$ et $C_2 = V(F_2)$ deux cercles à centre distincts de $\mathcal{F}$ où les fonctions circulaires F_1, F_2 sont normalisées et 0 une origine

$$\forall M, F_1(M) = OM^2 + 2\vec{V}_1 \cdot \vec{OM} + F_1(0), F_2(M) = OM^2 + 2\vec{V}_2 \cdot \vec{OM} + F_2(0)$$

L'axe radical de C_1 et C_2 est $V(F_1 - F_2)$, cercle généralisé du faisceau. $\square$

11.6 PROJECTION STÉRÉOGRAPHIQUE

Soit $\mathcal{E}$ un espace affine euclidien de dimension 3, E son espace vectoriel associé, $\mathcal{P}$ un plan de $\mathcal{E}$, S une sphère de centre Ω . La perpendiculaire à $\mathcal{P}$ passant par Ω coupe S en deux points s, o avec $s \notin \mathcal{P}$. Les points Ω, o, s ont même projection orthogonale O sur $\mathcal{P}$.



On a une bijection π de la sphère S sur le plan annalagmatique $\check{\mathcal{P}}$ ainsi définie :

- 1) à $m \in S$ distinct de s , on associe le point $\pi(m) = M$ où la droite (sm) coupe $\mathcal{P}$,
- 2) à s on associe le point à l'infini ∞ de $\check{\mathcal{P}}$.

Cette bijection s'appelle la **projection stéréographique** de sommet s de la sphère S sur le plan $\mathcal{P}$. Nous verrons que π établit une bijection entre cercles tracés sur S et cercles généralisés de $\check{\mathcal{P}}$.

11.6.1. Plans et sphères

Comme en géométrie plane pour les droites (9.13, p. 251) et cercles (11.12, p. 289), on décrit un plan (resp. une sphère) comme ensemble des points annulant une fonction affine (resp. circulaire). Étant donné une origine s ,

- une **fonction affine** est une application $f: \mathcal{E} \rightarrow \mathbb{R}$

$$P \mapsto f(P) = \vec{w} \cdot \vec{sP} + f(s)$$

Les fonctions affines forment un espace vectoriel de dimension 4 car isomorphe à $E \times \mathbb{R}$ par l'isomorphisme $f \mapsto (\vec{w}, f(s))$.

- une **fonction circulaire** est une application $G: \mathcal{E} \rightarrow \mathbb{R}$

$$P \mapsto G(P) = \gamma sP^2 + 2\vec{W} \cdot \vec{sP} + G(s)$$

Les fonctions circulaires forment un espace vectoriel de dimension 5 car isomorphe à $\mathbb{R} \times E \times \mathbb{R}$ par l'isomorphisme $G \mapsto (\gamma, \vec{W}, G(s))$.

Une fonction affine f donnée par un couple $(\vec{w}, f(s))$ avec $\vec{w} \neq \vec{0}$ détermine le plan

$$V(f) = \{P \mid f(P) = 0\}$$

Inversement, tout plan est du type $V(f)$ où f est déterminée à facteur réel près.

Comme pour les cercles, l'ensemble des points annulant une fonction circulaire n'est pas toujours une sphère. Ce peut être un plan (si $\gamma = 0$), un point unique, ou l'ensemble vide. Inversement, toute sphère est du type $V(G)$ où G est déterminée

à facteur réel près. Comme pour les cercles, la condition pour que $V(G)$ soit non vide et non réduit à un point (i.e. un plan ou une sphère de rayon non nul) est que $\Phi(G) = \|\vec{W}\|^2 - \gamma G(s) > 0$.

11.6.2. Inversion en dimension 3

Nous avons rencontré la notion d'inversion en géométrie plane (11.4, p. 296). Cette notion n'est pas différente en dimension 3. Voici une autre approche.

Définition 11.30. On appelle **inversion de pôle s et de puissance k** , scalaire non nul, la bijection involutive $\text{Inv}(s, k): P \mapsto Q$ de $\mathcal{E} \setminus \{s\}$ sur lui-même où P, Q, s sont alignés avec

$$\overrightarrow{sP} \times \overrightarrow{sQ} = k, \text{ autrement dit, } \overrightarrow{sP} = \frac{k\overrightarrow{sQ}}{sQ^2}$$

Proposition 11.31. L'inversion $\text{Inv}(s, k)$ établit une bijection entre l'ensemble des plans ne passant pas par s et l'ensemble des sphères passant par s .

Il existe une bijection linéaire $f \mapsto G$ de l'espace vectoriel des fonctions affines sur l'espace vectoriel des fonctions circulaires nulles en s telle que $V(f)$ et $V(G)$ soient transformés l'un de l'autre par $\text{Inv}(s, k)$.

Démonstration. Soit le plan $V(f)$ où $f(P) = \vec{w} \cdot \overrightarrow{sP} + f(s)$ avec $\vec{w} \neq \vec{0}$. Pour que $P \in V(f)$, il faut et il suffit que son inverse Q vérifie

$$\frac{f(s)sQ^2 + k\vec{w} \cdot \overrightarrow{sQ}}{sQ^2} = \vec{w} \cdot \frac{k\overrightarrow{sQ}}{sQ^2} + f(s) = \vec{w} \cdot \overrightarrow{sP} + f(s) = f(P) = 0$$

Soit G la fonction

$$\begin{aligned} Q \mapsto G(Q) &= f(s)sQ^2 + k\vec{w} \cdot \overrightarrow{sQ} \\ &= f(s) \left(\left\| \overrightarrow{sQ} + \frac{k\vec{w}}{2f(s)} \right\|^2 - \frac{k^2\|\vec{w}\|^2}{4f^2(s)} \right) \text{ si } f(s) \neq 0 \end{aligned}$$

C'est une fonction circulaire nulle en s où $V(G)$ est la sphère de centre I défini par $\overrightarrow{sI} = -\frac{k\vec{w}}{2f(s)}$. On a l'équivalence

$$(P \in V(f)) \iff (Q \in V(G))$$

Enfin, il est clair que cette application $f \mapsto G$ est une bijection linéaire entre ces espaces vectoriels. $\square$

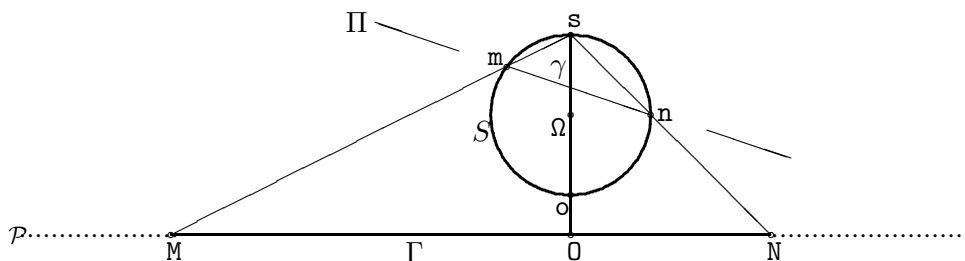
11.6.3. La projection stéréographique

Reprenons les conditions du début de 11.6. Il est clair que la projection stéréographique $\pi: S \setminus \{s\} \rightarrow \mathcal{P}$ est la restriction à S de l'inversion $\text{Inv}(s, k)$ ($k = \overline{so} \times \overline{s0}$) transformant S en $\tilde{\mathcal{P}}$.

Proposition 11.32. *La projection stéréographique π établit une bijection entre l'ensemble des cercles γ tracés sur S et l'ensemble des cercles généralisés Γ de $\mathcal{P}$.*

Pour que Γ soit une droite, il faut et il suffit que γ passe par s .

Il existe une bijection linéaire $\Psi: f \mapsto F$ entre l'espace des fonctions affines f définies sur $\mathcal{E}$ et l'espace des fonctions circulaires F définies sur $\mathcal{P}$ telle que $V(F) = \pi(V(f) \cap S)$.



Démonstration. Soit γ un cercle tracé sur S et Π le plan le contenant. Alors $\Pi = V(f)$ où f est une fonction affine $P \mapsto f(P) = \vec{w} \cdot \vec{sP} + f(s)$. L'inverse de Π est la sphère $\Sigma = V(G)$ où

$$G: Q \mapsto G(Q) = f(s)sQ^2 + k\vec{w} \cdot \vec{sQ}$$

Écrivons $\vec{w} = \vec{u} + \vec{v}$ où $\vec{u} \in P^\perp$, $\vec{v} \in P$. Soit $m \in S$, $M = \pi(m) \in \mathcal{P}$, on a

$$\begin{aligned} G(M) &= f(s)sM^2 + k\vec{w} \cdot \vec{sM} = f(s)(sO^2 + OM^2) + k(\vec{u} + \vec{v}) \cdot (\vec{sO} + \vec{OM}) \\ &= f(s)OM^2 + k\vec{v} \cdot \vec{OM} + f(s)sO^2 + k\vec{u} \cdot \vec{sO} \end{aligned}$$

Autrement dit, la restriction de G au plan $\mathcal{P}$ est la fonction circulaire

$$F: \mathcal{P} \rightarrow \mathbb{R}, M \mapsto F(M) = f(s)OM^2 + k\vec{v} \cdot \vec{OM} + f(s)sO^2 + k\vec{u} \cdot \vec{sO}$$

Pour que $m \in \gamma = V(f) \cap S$, il faut et il suffit que $\pi(m) = M \in V(G) \cap \mathcal{P}$, donc que $M \in \mathcal{P}$ et $F(M) = 0$.

On a les équivalences

$$\begin{aligned} (V(F) \text{ est une droite}) &\Leftrightarrow (f(s) = 0 \text{ et } \vec{v} \neq \vec{0}) \Leftrightarrow (f(s) = 0 \text{ et } \vec{w} \neq \vec{u}) \\ &\Leftrightarrow (\text{le plan } V(f) \text{ passe par } s \text{ et n'est pas tangent à } S) \end{aligned}$$

L'application $\Psi: f \mapsto F$ est bien linéaire, de l'espace des fonctions affines $f: \mathcal{E} \rightarrow \mathbb{R}$ vers l'espace des fonctions circulaires $F: \mathcal{P} \rightarrow \mathbb{R}$. Si $f \in \text{Ker } \Psi$, F est nulle, donc

$f(s) = 0$, $\vec{v} = \vec{0}$ et $\vec{u} \cdot \vec{s0} = 0$ ce qui implique $\vec{u} = \vec{0}$ car $\vec{u}$ et $\vec{s0}$ sont colinéaires, donc $\vec{w} = \vec{0}$ et f est nulle. Comme ces espaces vectoriels sont tous deux de dimension 4, l'injection Ψ est bien bijective.

Remarquons que si le plan $V(f)$ est extérieur à S , γ est vide et $\Gamma = V(F)$ est aussi vide. Le lecteur vérifiera que la condition

$$\Phi(F) = \frac{1}{4}k^2\|v\|^2 - f(s)\left(f(s)s0^2 + k\vec{u} \cdot \vec{s0}\right) > 0$$

exprimant que $V(F)$ est non vide signifie que le plan $V(f)$ coupe S , c'est-à-dire que sa distance au centre Ω est inférieure au rayon de S . $\square$

11.6.4. Faisceaux de cercles dans $\mathcal{P}$ et S

Soit a, b deux points distincts de S , $\Delta_1 = (ab)$ la droite joignant ces points. On notera $\mathcal{X}_1$ l'ensemble des plans Π_1 passant par Δ_1 .

Soit A et B les plans tangents en a et b à S . Deux cas sont à distinguer :

- $a)$ Si a et b ne sont pas diamétralement opposés sur S , les plans A et B ne sont pas parallèles et se coupent suivant une droite Δ_2 . On notera $\mathcal{X}_2$ l'ensemble des plans Π_2 passant par Δ_2 .
- $b)$ Si a et b sont diamétralement opposés sur S , les plans A et B sont parallèles. On notera $\mathcal{X}_2$ l'ensemble des plans parallèles à A et B .

Soit dans $\tilde{\mathcal{P}}$ les points distincts $A = \pi(a)$ et $B = \pi(b)$.

Proposition 11.33. *(i) Si Π_1 décrit $\mathcal{X}_1$, le cercle généralisé $\pi(\Pi_1 \cap S)$ décrit le faisceau à points de base A, B .*

(ii) Si Π_2 décrit l'ensemble des plans de $\mathcal{X}_2$ coupant la sphère S , le cercle généralisé $\pi(\Pi_2 \cap S)$ décrit le faisceau à points-limites A, B .

Lemme 11.34. *Soit deux plans distincts $V(f)$ et $V(g)$ où f et g sont des fonctions affines sur $\mathcal{E}$ non proportionnelles. Si (λ, μ) décrit $\mathbb{R}^2$ (privé du couple nul), $V(\lambda f + \mu g)$ décrit l'ensemble $\mathcal{X}$ de plans*

- *passant par $\Delta = V(f) \cap V(g)$ si ces plans ne sont pas parallèles,*
- *parallèles aux plans $V(f)$ et $V(g)$ si ces plans sont parallèles.*

Démonstration. [du lemme] Il est facile de vérifier que tout plan du type $V(\lambda f + \mu g)$ est dans l'ensemble $\mathcal{X}$.

Inversement, soit Π un plan de $\mathcal{X}$ et P un point de Π non sur $V(f) \cap V(g)$ (ce qui n'est pas une restriction si ces plans sont parallèles). Supposons par exemple $P \notin V(f)$, soit $f(P) \neq 0$. Alors Π est l'unique plan de $\mathcal{X}$ passant par P . Mais le plan $V(g(P)f - f(P)g)$ est aussi un plan de $\mathcal{X}$ passant par P . C'est donc Π qui est bien du type indiqué. $\square$

Démonstration. [de la proposition] Le point (i) est évident.

Soit f et g des fonctions affines sur $\mathcal{E}$ telles que $A = V(f)$ et $B = V(g)$, $F = \Psi(f)$ et $G = \Psi(g)$ les fonctions circulaires sur $\mathcal{P}$. Alors $A \cap S$ et $B \cap S$ sont les cercles-points $\{a\}$ et $\{b\}$ dont les images par π sont les cercles-points $\{A\} = V(F)$ et $\{B\} = V(G)$.

Si $\Pi_2 = V(\lambda f + \mu g) \in \mathcal{X}_2$ coupe S suivant γ_2 , alors $\Gamma_2 = \pi(\gamma_2) = V(\lambda F + \mu G)$ appartient au faisceau $\mathcal{F}(A, B)$ à points-limites A, B .

En revanche, si $\Pi_2 = V(\lambda f + \mu g) \in \mathcal{X}_2$ ne coupe pas S , $V(\lambda F + \mu G)$ est l'ensemble vide.

Tout ceci est encore valable si l'un des points a, b est s . Alors A ou B est le point à l'infini ∞ . On a un faisceau de droites concourantes et un faisceau de cercles concentriques. $\square$

EXERCICES

Exercice 11.6. Soit deux cercles généralisés C_1, C_2 donnés par leurs équations dans un repère orthonormé

$$\text{pour } C_1 = V(P_1) \quad , \quad P_1(x, y) = \gamma_1(x^2 + y^2) + 2\alpha_1x + 2\beta_1y + \delta_1 = 0$$

$$\text{pour } C_2 = V(P_2) \quad , \quad P_2(x, y) = \gamma_2(x^2 + y^2) + 2\alpha_2x + 2\beta_2y + \delta_2 = 0$$

À quelle condition sur les coefficients sont-ils orthogonaux ?

Exercice 11.7. On donne les cercles C_1, C_2 par des équations relativement à un repère orthonormé

$$\text{pour } C_1 = V(P_1) \quad , \quad P_1(x, y) = \gamma_1(x^2 + y^2) + 2\alpha_1x + 2\beta_1y + \delta_1 = 0$$

$$\text{pour } C_2 = V(P_2) \quad , \quad P_2(x, y) = \gamma_2(x^2 + y^2) + 2\alpha_2x + 2\beta_2y + \delta_2 = 0$$

À quelle condition sur les coefficients ces cercles sont-ils tangents (resp. sécants, resp. sans point commun) ?

Exercice 11.8. Réseaux de cercles. À un sous-espace vectoriel $\mathcal{V} \subset \mathcal{C}$ de dimension 3, on associe le **réseau** $\mathcal{R}$ des cercles généralisés $V(F)$ où F décrit $\mathcal{V}$ avec $\Phi(F) \geq 0$.

1) En considérant le Φ -orthogonal de $\mathcal{V}$, montrer qu'il existe en général un point 0 ayant même puissance relativement à tous les cercles de $\mathcal{R}$. Que se passe-t-il s'il n'en est pas ainsi ?

2) Discuter la nature de $\mathcal{R}$ suivant la nature de la restriction de $\Phi|_{\mathcal{V}}$.

3) Que dire de l'ensemble des cercles stables par une inversion $\text{Inv}(0, k)$ (11.4, p. 296) ?

PROBLÈMES

Les corrigés de ces problèmes sont disponibles sur le site de Dunod : www.dunod.com.

11.1. PROBLÈME

Soit un triangle ABC , P, Q, R sur $(BC), (CA), (AB)$ distincts des sommets.

- 1) Montrer que les cercles $(AQR), (BRP), (CPQ)$ concourent en un point M .
- 2) Inversement, étant donné M distinct des sommets, montrer qu'il existe P, Q, R sur $(BC), (CA), (AB)$ tels que M soit sur les cercles $(AQR), (BRP), (CPQ)$. Ces points P, Q, R sont-ils uniques ?
- 3) Montrer que M est sur (ABC) si et seulement si P, Q, R sont alignés.
- 4) Montrer que A, P, M sont alignés si et seulement si B, C, Q, R sont cocycliques.
- 5) Montrer que $(AP), (BQ), (CR)$ concourent en M si et seulement si M est l'orthocentre de ABC .

11.2. PROBLÈME

Soit un triangle $T_0 = A_0B_0C_0$, P un point non situé sur les côtés. On définit par récurrence $T_n = A_nB_nC_n$ où A_n, B_n, C_n sont projections orthogonales de P sur $(B_{n-1}C_{n-1}), (C_{n-1}A_{n-1}), (A_{n-1}B_{n-1})$.

- 1) Montrer que si P n'est pas sur le cercle $(A_0B_0C_0)$, cette construction peut se répéter indéfiniment.
- 2) Montrer que T_{n+3} est transformé de T_n par une similitude directe de centre P .
- 3) Montrer que parmi les trois modèles de triangles T_0, T_1, T_2 obtenus, deux sont identiques si P occupe l'une des positions suivantes relative à T_0 : centre du cercle circonscrit, orthocentre, centre d'un cercle inscrit ou exinscrit.

11.3. PROBLÈME (Steiner)

Soit ABC un triangle, C le cercle circonscrit, H l'orthocentre, M un point, P, Q, R les symétriques de M par rapport aux côtés.

- 1) Montrer que $M \in C$ si et seulement si trois des points P, Q, R, H sont alignés. S'il en est ainsi, ils le sont tous les quatre.
- 2) Discuter suivant la position de M sur C la position relative de H relativement à P, Q, R .

11.4. PROBLÈME (situation orthocentrique)

1) Montrer qu'une base orthogonale de l'espace $\mathcal{C}$ des fonctions circulaires donne lieu à trois cercles généralisés orthogonaux deux à deux.

2) On suppose que ces cercles A, B, C ont des centres A, B, C . Quel rôle joue l'orthocentre D pour ces trois cercles ?

3) Montrer que le triangle ABC a ces trois angles aigus (9.12, p. 258).

4) Inversement, une configuration orthocentrique donne-t-elle lieu à trois cercles orthogonaux deux à deux ?

11.5. PROBLÈME

Soit $\mathcal{P}$ un plan euclidien orienté de plan vectoriel associé P . Les déterminants sont évalués relativement à des bases orthonormales directes de P .

1) Soit la rotation vectorielle rot_θ d'angle θ distinct de $0, \pi$. Montrer que l'application

$$P \rightarrow \mathbb{R}, \vec{V} \mapsto \det(\vec{V}, \text{rot}_\theta(\vec{V}))$$

est une forme quadratique proportionnelle à la forme quadratique euclidienne de P .

2) Soit A, B deux droites vectorielles, σ_A, σ_B les réflexions vectorielles autour de A et B . À quelles conditions sur A et B l'application

$$\vec{V} \mapsto \det(\sigma_A(\vec{V}), \sigma_B(\vec{V}))$$

est-elle une forme quadratique proportionnelle à la forme quadratique euclidienne ?

3) On donne un triangle ABC direct, O le centre du cercle circonscrit. On considère les côtés du triangle $\mathcal{A} = (BC)$, $\mathcal{B} = (CA)$, $\mathcal{C} = (AB)$ de directions A, B, C . On note α, β, γ les réflexions (affines) autour des droites $\mathcal{A}, \mathcal{B}, \mathcal{C}$.

Montrer que l'application

$$F: \mathcal{P} \rightarrow \mathbb{R}, M \mapsto F(M) = [\alpha(M), \beta(M), \gamma(M)]$$

(voir notations du chapitre 9, 9.9, p. 249) est une fonction circulaire.

4) Quel est l'ensemble $V(F)$ des points où F s'annule ? (Utiliser le théorème de Simson.)

5) En déduire que l'ensemble des points M tels que l'aire algébrique du triangle $\alpha(M)\beta(M)\gamma(M)$ est constante est vide ou est un cercle de centre O .

6) Cette aire algébrique peut-elle prendre n'importe quelle valeur réelle ?

7) Discuter de l'orientation du triangle $\alpha(M)\beta(M)\gamma(M)$.

SOLUTIONS DES EXERCICES

Solution 11.1. Pour tout $M(x, y)$, on a $C_1(M) = \frac{1}{\gamma_1}P_1(x, y)$ et $C_2(M) = \frac{1}{\gamma_2}P_2(x, y)$. Pour que $M(x, y) \in \Delta$, il faut et il suffit que $C_1(M) - C_2(M) = 0$, soit, en multipliant par $\gamma_1\gamma_2$

$$2(\alpha_1\gamma_2 - \alpha_2\gamma_1)x + 2(\beta_1\gamma_2 - \beta_2\gamma_1)y + \delta_1\gamma_2 - \delta_2\gamma_1 = 0$$

Si $\alpha_1\gamma_2 - \alpha_2\gamma_1 = \beta_1\gamma_2 - \beta_2\gamma_1 = 0$, les cercles sont concentriques et l'axe radical n'existe pas.

Solution 11.2. 1) La réflexion σ autour de la droite passant par le centre O de C et orthogonale à X échange A et A' , B et B' , C et C' et laisse C stable. On en déduit les congruences modulo π

$$(\widehat{MP, MQ}) \equiv (\widehat{MA', MB'}) \equiv -(\widehat{MA, MB}) \equiv -(\widehat{CA, CB}) \equiv -(\widehat{CQ, CP}) = (\widehat{CP, CQ})$$

d'où la cocyclicité de M, P, Q, C .

2) On a les congruences modulo π

$$\begin{aligned} (\widehat{PQ, X}) &\equiv (\widehat{PQ, PC}) + (\widehat{PC, X}) \equiv (\widehat{MQ, MC}) + (\widehat{BC, X}) \\ &\equiv (\widehat{MB', MC}) + (\widehat{BC, X}) \equiv (\widehat{BB', BC}) + (\widehat{BC, X}) \equiv (\widehat{BB', X}) \equiv 0 \end{aligned}$$

Les points P, Q, R sont donc alignés sur une droite de direction X .

Solution 11.3. On a $[(CA), (CB), \mathcal{X}, \mathcal{Y}] = -1$ donc $[A, B, I, J] = -1$ (8.40, p. 227). Le cercle de diamètre (I, J) est le lieu des points M tels que $\frac{MA}{MB} = \frac{IA}{IB} = \frac{JA}{JB}$. Comme $\mathcal{X}, \mathcal{Y}$ sont orthogonales, C appartient à ce cercle, d'où le résultat.

Solution 11.4. 1) Comme $[O, A']$ est un diamètre de C , la droite $(OM') = (MM')$ est orthogonale à $(A'M')$. Comme (AA') est orthogonale à $\Delta = (MA)$, les points A, A', M, M' sont sur le cercle de diamètre $[A', M]$.

2) Les droites passant par O sont stables. Soit Δ ne passant pas par O , A la projection orthogonale de O sur Δ , A' le transformé de A , C le cercle de diamètre $[O, A']$. Pour tout $M \in \Delta$, les points A, A', M, M' sont sur le cercle Γ de diamètre $[A', M]$ d'où $k = \overline{OA} \times \overline{OA'} = \overline{OM} \times \overline{OM'}$, donc le transformé de Δ par $\text{Inv}(O, k)$ est le cercle C . Les droites ne passant pas par O deviennent les cercles passant par O et inversement.

Solution 11.5. 1) Les vecteurs $\overrightarrow{N_1 N_2}$ et $\overrightarrow{N'_1 N'_2}$ sont colinéaires, d'où, par cocyclicité, les congruences modulo π

$$(\widehat{N_1 N_2, N_1 M'_1}) \equiv (\widehat{N'_1 N'_2, N'_1 M'_1}) \equiv (\widehat{M'_2 N'_2, M'_2 M'_1}) \equiv (\widehat{M'_2 N_2, M'_2 M'_1})$$

ce qui donne la cocyclicité demandée.

2) Soit C ne passant pas par O , $N_1 \in C$, $\Delta_1 = (ON_1)$ recoupant C en M_1 , M'_1 le transformé de N_1 par $\text{Inv}(O, k)$, h l'homothétie de centre O transformant M_1 en M'_1 , $C' = h(C)$. Soit Δ_2 passant par O coupant C en M_2 et N_2 , C' en $M'_2 = h(M_2)$ et $N'_2 = h(N_2)$. Par ce qui précède, N_1, M'_1, N_2, M'_2 sont cocycliques, donc $k = \overline{ON_1} \times \overline{OM'_1} = \overline{ON_2} \times \overline{OM'_2}$, donc M'_2 est transformé de N_2 par $\text{Inv}(O, k)$. Ainsi, C' est transformé de C par $\text{Inv}(O, k)$.

L'inversion $\text{Inv}(O, k)$ induit donc une bijection involutive de l'ensemble des cercles ne contenant pas O sur lui-même.

Solution 11.6. On a $C_1 = V(F_1)$ et $C_2 = V(F_2)$ où, pour $M(x, y)$,

$$F_1(M) = \gamma_1 OM^2 + 2\overrightarrow{V_{1,0}} \cdot \overrightarrow{OM} + \delta_1 \quad \text{où} \quad \overrightarrow{V_{1,0}}(\alpha_1, \beta_1)$$

$$F_2(M) = \gamma_2 OM^2 + 2\overrightarrow{V_{2,0}} \cdot \overrightarrow{OM} + \delta_2 \quad \text{où} \quad \overrightarrow{V_{2,0}}(\alpha_2, \beta_2)$$

La condition est alors $\Phi(F_1, F_2) = 0$, soit (11.24, p. 297)

$$\alpha_1 \alpha_2 + \beta_1 \beta_2 - \frac{1}{2}(\delta_2 \gamma_1 + \delta_1 \gamma_2) = 0$$

Solution 11.7. Soit les fonctions circulaires $F_i : M(x, y) \mapsto P_i(x, y)$. On doit exprimer que le faisceau $\mathcal{F}(C_1, C_2)$ est singulier (resp. à points de base, resp. à points-limites), i.e. que la restriction de Φ au plan vectoriel $\text{Vect}(F_1, F_2)$ est dégénérée (resp. de signature $(2, 0)$, resp. $(1, 1)$). Supposons que C_2 n'est pas un cercle-point. On a

$$\Phi(F_1 + \lambda F_2) = \lambda^2 \Phi(F_2) + 2\lambda \Phi(F_1, F_2) + \Phi(F_1)$$

La restriction de Φ à $\text{Vect}(F_1, F_2)$ est dégénérée (resp. de signature $(2, 0)$, resp. $(1, 1)$) si on a une seule droite isotrope (resp. pas de droite isotrope, resp. deux droites isotropes), i.e. si ce polynôme en λ a une racine double (resp. pas de racine réelle, resp. deux racines réelles distinctes). La discussion porte donc sur le discriminant

$$\Delta(\lambda) = \Phi(F_1, F_2)^2 - \Phi(F_1)\Phi(F_2)$$

$$\Phi(F_1) = \alpha_1^2 + \beta_1^2 - \gamma_1 \delta_1, \quad \Phi(F_2) = \alpha_2^2 + \beta_2^2 - \gamma_2 \delta_2$$

$$\Phi(F_1, F_2) = \alpha_1 \alpha_2 + \beta_1 \beta_2 - \frac{1}{2}(\gamma_1 \delta_2 + \gamma_2 \delta_1)$$

$$\Delta(\lambda) = \left(\alpha_1 \alpha_2 + \beta_1 \beta_2 - \frac{1}{2}(\gamma_1 \delta_2 + \gamma_2 \delta_1) \right)^2 - \left(\alpha_1^2 + \beta_1^2 - \gamma_1 \delta_1 \right) \left(\alpha_2^2 + \beta_2^2 - \gamma_2 \delta_2 \right)$$

C'est l'expression analytique de l'inégalité (i) de la proposition 11.5 (p. 285)

Solution 11.8. 1) Dans $\mathcal{C}$, $\mathcal{V}^\perp$ est une droite vectorielle engendrée par une fonction circulaire f .

a) Supposons $\gamma(f) \neq 0$. Remplaçant f par $\frac{1}{\gamma(f)}f$, on se ramène à $\gamma(f) = 1$. Soit O le centre pour f (11.12, p. 289), on a $f(M) = \|\vec{OM}\|^2 + f(O)$. Soit F une fonction circulaire.

Pour F normale, $F(M) = \|\vec{OM}\|^2 + 2\overrightarrow{V_0(F)} \cdot \vec{OM} + F(O)$, d'où $F \in \mathcal{V}$ si et seulement si $0 = \Phi(f, F) = -\frac{1}{2}(f(O) + F(O))$, i.e. $F(O) = -f(O)$. La puissance de O relativement aux cercles de $\mathcal{R}$ est la même : $-f(O)$. On dit que O est le **centre radical** du réseau $\mathcal{R}$.

Pour $\gamma(F) = 0$, $F(M) = 2\overrightarrow{V_0(F)} \cdot \vec{OM} + F(O)$, d'où $F \in \mathcal{V}$ si et seulement si $0 = \Phi(f, F) = -\frac{1}{2}F(O)$. Les droites de $\mathcal{R}$ sont donc celles qui passent par O .

b) Si $\gamma(f) = 0$, deux cas sont possibles :

- $V(f) = \Delta$ est une droite. Alors $\mathcal{R}$ est formé des cercles centrés sur Δ (y compris les cercles-points) et des droites orthogonales à Δ .
- $f = 1$, soit $V(f) = \{\infty\}$. Alors $\mathcal{R}$ est l'ensemble des droites du plan.

2) Trois cas se présentent pour $\Phi/\mathcal{V}$.

(1) La signature de $\Phi/\mathcal{V}$ est $(2, 1)$, i.e. $\Phi(f) > 0$.

- Dans le cas a), $V(f)$ est un cercle Γ de centre O et $\mathcal{R}$ est formé des cercles orthogonaux à Γ et des droites passant par O .
- Dans le cas b), $V(f)$ est une droite Δ et $\mathcal{R}$ est formé des cercles centrés sur Δ et des droites orthogonales à Δ .

(2) La signature de $\Phi/\mathcal{V}$ est $(3, 0)$, i.e. $\Phi(f) < 0$. Alors aucun cercle généralisé ne correspond à f et le cas b) ne se produit pas. On a un centre radical O avec $f(O) = -R^2 < 0$. Les cercles du réseau $\mathcal{R}$ sont les cercles C tels que $C(O) = -R^2$, c'est-à-dire les cercles qui coupent le cercle de centre O et de rayon R en deux points diamétralement opposés. Tous les faisceaux contenus dans $\mathcal{R}$ sont à points de base, donc deux cercles quelconques de $\mathcal{R}$ se coupent toujours.

(3) $\Phi(f) = 0$, $\Phi/\mathcal{V}$ est dégénérée.

- $V(f)$ est un cercle-point $\{O\}$ et $\mathcal{R}$ est formé des cercles et droites passant par O .
- $V(f) = \{\infty\}$ et $\mathcal{R}$ est l'ensemble des droites du plan (cas b).

3) Les cercles stables par $\text{Inv}(O, k)$ forment un réseau de centre radical O qui a même puissance k relativement à chacun de ces cercles.

Chapitre 12

Coniques

Les coniques sont d'abord envisagées dans un cadre purement affine dans la section 12.1 (genre, positions relatives d'une droite et d'une conique), puis dans un cadre métrique dans la section 12.2 avec les deux points de vue foyer-directrice 12.2.3 et bifocal 12.2.4. Insistons sur l'équation donnée dans le théorème 12.17 (p. 328) qui montre un point de vue local englobant toutes les coniques propres. La scène se passe dans un plan affine $\mathcal{P}$ de plan vectoriel associé P , muni d'un produit scalaire euclidien dans la section 12.2.

12.1 CONIQUES DANS UN PLAN AFFINE

12.1.1. Les trois genres de coniques propres

Définition 12.1. Une *fonction quadratique affine*, en abrégé f.q.a., est une application $F: \mathcal{P} \rightarrow \mathbb{R}$ non identiquement nulle de la forme

$$M \mapsto F(M) = Q(\overrightarrow{OM}) + 2L(\overrightarrow{OM}) + F(O)$$

où O est une origine, $Q: P \rightarrow \mathbb{R}$ une forme quadratique, $L: P \rightarrow \mathbb{R}$ une forme linéaire et $F(O)$ un réel. On notera $\Phi: P \times P \rightarrow \mathbb{R}$ la forme polaire de Q .

Proposition 12.2. L'expression d'une f.q.a. F reste du même type si on change l'origine O . La forme quadratique Q ne dépend pas de O . La forme linéaire, notée L_O , en dépend. La formule de changement d'origine est

$$L_{O'} = L_O - \Phi(\overrightarrow{O'O}, -) = L_O + \Phi(\overrightarrow{OO'}, -)$$

Démonstration. Soit O' une autre origine. On a

$$\begin{aligned} F(M) &= Q(\overrightarrow{OO'} + \overrightarrow{O'M}) + 2L_0(\overrightarrow{OO'} + \overrightarrow{O'M}) + F(O) \\ &= Q(\overrightarrow{O'M}) + 2\Phi(\overrightarrow{OO'}, \overrightarrow{O'M}) + 2L_0(\overrightarrow{O'M}) + Q(\overrightarrow{OO'}) + 2L_0(\overrightarrow{OO'}) + F(O) \\ &= Q(\overrightarrow{O'M}) + 2L_{O'}(\overrightarrow{O'M}) + F(O') \text{ où } L_{O'} = L_0 + \Phi(\overrightarrow{OO'}, -) \end{aligned}$$

□

Expression dans un repère. Soit un repère de $\mathcal{P}$ d'origine O et de base $\mathcal{B} = (\vec{i}, \vec{j})$. La forme quadratique Q et la forme linéaire L_0 sont données par leurs matrices dans la base $\mathcal{B}$:

$$\text{Mat}_{\mathcal{B}}(Q) = \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \quad \text{Mat}_{\mathcal{B}}(L_0) = \begin{pmatrix} e & f \end{pmatrix}, \quad F(O) = g$$

La f.q.a. F a alors pour expression $M(x, y) \mapsto F(M) = p(x, y)$ où p est le polynôme de degré 2

$$p(X, Y) = aX^2 + 2bXY + cY^2 + 2eX + 2fY + g$$

Si on adjoint à l'ensemble des f.q.a. l'application identiquement nulle, on obtient un espace vectoriel $\mathcal{Q}$. La donnée d'un repère $(O, \vec{i}, \vec{j})$ définit un isomorphisme entre $\mathcal{Q}$ et l'espace vectoriel de dimension 6 des polynômes $p(X, Y)$ de degré inférieur ou égal à 2.

Définition 12.3. La conique associée à une f.q.a. F est l'ensemble

$$V(F) = \{M \in \mathcal{P} \mid F(M) = 0\}$$

Si $\mathcal{P}$ est un plan euclidien et si Q est proportionnelle à la forme quadratique définie positive associée au produit scalaire, on trouve la notion de fonction circulaire (11.12, p. 289). Les cercles sont des coniques particulières.

Deux f.q.a. proportionnelles ont même conique associée. La réciproque est fausse : si Q est définie positive, L_0 nulle et $F(O) > 0$, on a $V(F) = \emptyset$, alors que ces conditions ne déterminent pas F à proportionnalité près.

Voici quelques exemples de cas dégénérés.

- $V(F)$ se réduit à un point si Q est définie positive, L_0 et $F(O)$ nuls.
- si $F = fg$ est produit de deux fonctions affines (9.12, p. 250), $V(F)$ est réunion des deux droites $V(f)$ et $V(g)$ (distinctes ou non),
- si Q est nulle, $V(F)$ se réduit à une droite.

Définition 12.4. Une f.q.a. F est dite à **centre** si la forme quadratique Q associée est non dégénérée.

Proposition 12.5. Soit F une f.q.a. à centre. Il existe un unique point Ω appelé **centre** de F , tel que L_Ω soit nulle. Si $V(F)$ est non vide, Ω est centre de symétrie pour $V(F)$.

Démonstration. À la forme quadratique Q de forme polaire Φ correspond une application linéaire $\varphi: P \rightarrow P^*$, $\vec{v} \mapsto \Phi(\vec{v}, -)$. La non-dégénérescence de Q équivaut à la bijectivité de φ . Étant donné $0 \in P$, il existe $\vec{v}_0 \in P$ unique tel que $\Phi(\vec{v}_0, -) = L_0$ et Ω défini par $\Omega = 0 - \vec{v}_0$ est le centre cherché.

Soit un repère d'origine Ω et de base Q -orthogonale $\mathcal{B} = (\vec{i}, \vec{j})$. Alors $V(F)$ est formé des $M(x, y)$ tels que $p(x, y) = 0$ où

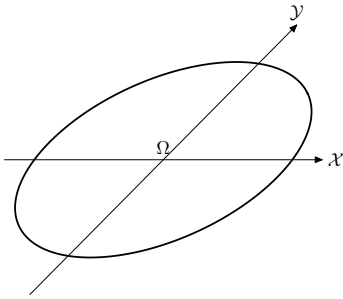
$$p(X, Y) = aX^2 + cY^2 + p(0, 0) \text{ avec } \text{Mat}_{\mathcal{B}}(Q) = \begin{pmatrix} a & 0 \\ 0 & c \end{pmatrix}$$

Comme $p(-X, -Y) = p(X, Y)$, Ω est centre de symétrie pour $V(F)$. $\square$

Définition 12.6. La f.q.a. F est dite **elliptique** si Q est de signature $(0, 2)$ ou $(2, 0)$, c'est-à-dire si Q est définie positive ou définie négative.

Pour la conique associée, les cas de signature $(2, 0)$ ou $(0, 2)$ ne diffèrent pas car $V(-F) = V(F)$. Soit Ω le centre. Supposons F définie positive.

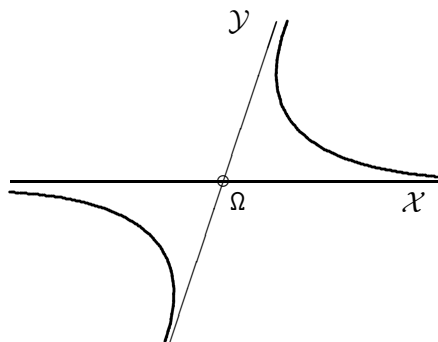
Si $F(\Omega) = p(0, 0) > 0$, $V(F)$ est vide. Si $F(\Omega) = 0$, $V(F) = \{\Omega\}$.



Supposons $F(\Omega) = p(0, 0) < 0$. La forme polaire de Q peut être considérée comme produit scalaire euclidien. Munissant P de ce produit scalaire, $V(F)$ apparaît comme le cercle de centre Ω et de rayon $\sqrt{-F(\Omega)}$. Pour la norme euclidienne associée, $V(F)$ est fermé borné, donc compact. Toutes les normes de P étant équivalentes, $V(F)$ est **compact** pour la topologie déduite de n'importe quelle norme. On dit que $V(F)$ est une **ellipse** de centre Ω .

En remplaçant F par $-\frac{1}{k}F$, on se ramène au cas où $F(\Omega) = -1$. Il existe une base Q -orthonormale $\mathcal{B} = (\vec{i}, \vec{j})$, donc telle que $\text{Mat}_{\mathcal{B}}(Q) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Alors $x^2 + y^2 - 1 = 0$ est équation de $V(F)$ dans le repère d'origine Ω et de base $\mathcal{B}$. C'est l'**équation réduite** de $V(F)$.

Définition 12.7. La f.q.a. F est dite **hyperbolique** si Q est hyperbolique, c'est-à-dire de signature $(1, 1)$.



Soit un repère d'origine le centre Ω de F et de base $\mathcal{B} = (\vec{i}, \vec{j})$ formée de vecteurs Q -isotropes telle que $\text{Mat}_{\mathcal{B}}(Q) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Si $F(\Omega) = 0$, $V(F)$, d'équation $xy = 0$, dégénère en la réunion des droites $\mathcal{X}$ et $\mathcal{Y}$ passant par Ω de vecteurs directeurs $\vec{i}$ et $\vec{j}$. Supposons $F(\Omega) = k \neq 0$. On se ramène à $F(\Omega) = -1$ en remplaçant F par $-\frac{1}{k}F$. On dit que $V(F)$ est une **hyperbole**. L'équation

de $V(F)$ dans le repère $(\Omega, \vec{i}, \vec{j})$ est $xy - 1 = 0$, **équation réduite**¹ de l'hyperbole et $V(F)$ est représentation graphique de la fonction $x \mapsto \frac{1}{x}$. Les droites $\mathcal{X}$ et $\mathcal{Y}$ sont les asymptotes. L'hyperbole est image de l'application injective continue

$$\mathbb{R} \setminus \{0\} \longrightarrow \mathcal{P}, \quad x \longmapsto \mathbb{M} \left(\begin{pmatrix} x \\ \frac{1}{x} \end{pmatrix} \right)$$

Les images des intervalles $] -\infty, 0[$ et $]0, +\infty[$ sont les deux composantes connexes de $V(F)$ et sont les deux **branches** de l'hyperbole. Ce sont des parties fermées, mais non compactes de $\mathcal{P}$.

Définition 12.8. La f.q.a. F est dite **parabolique** si la forme quadratique Q est de rang 1, c'est-à-dire dégénérée non nulle.

Alors Q est du type $Q = \kappa L^2$ où L est une forme linéaire non nulle et $\kappa \neq 0$ un réel. Comme $V(F) = V(-F)$, on peut supposer $\kappa > 0$. Remplaçant L par $\sqrt{\kappa}L$, on se ramène à $Q = L^2$. Étant donné $0 \in \mathcal{P}$, la f.q.a. est $F = L^2 + 2L_0 + F(0)$ où L et L_0 sont deux formes linéaires. La forme polaire de $Q = L^2$ est

$$\begin{aligned} \Phi: P \times P &\longrightarrow \mathbb{R} \\ (\vec{X}, \vec{Y}) &\longmapsto L(\vec{X})L(\vec{Y}) \end{aligned}$$

Lemme 12.9. Si L et L_0 sont proportionnelles, $V(F)$ est, ou bien vide, ou bien réunion de deux droites parallèles, distinctes ou confondues.

Démonstration. Il existe α tel que $L_0 = \alpha L$ et $F = L^2 + 2\alpha L + F(0)$. Posons $p(X) = X^2 + 2\alpha X + F(0)$ de discriminant $\Delta = \alpha^2 - F(0)$. On a $V(F) = \{\mathbb{M} \mid p(L(\vec{OM})) = 0\}$.

Si $\Delta < 0$, $p(X)$ n'a pas de racine réelle et $V(F)$ est vide.

Si $\Delta \geq 0$, $p(X)$ a deux racines x', x'' , distinctes ou non. Alors $V(F)$ est réunion des droites $\mathcal{D}' = \{\mathbb{M} \mid L(\vec{OM}) = x'\}$ et $\mathcal{D}'' = \{\mathbb{M} \mid L(\vec{OM}) = x''\}$ parallèles, de direction la droite vectorielle $D = \text{Ker } L$. Si elles sont confondues en une même droite $\mathcal{D}$, on dit que $V(F)$ est la **droite double** $\mathcal{D}$. $\square$

1. Dans beaucoup d'ouvrages, l'équation réduite d'une hyperbole est $x^2 - y^2 - 1 = 0$, l'origine étant le centre Ω et la base $\mathcal{B}$ Q -orthogonale telle que $\text{Mat}_{\mathcal{B}}(Q) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Supposons les formes linéaires L et L_0 indépendantes, i.e. $\text{Ker } L \neq \text{Ker } L_0$. La formule de changement d'origine (12.2, p. 315) $L_{O'} = L_0 + L(\overrightarrow{OO'})L$ montre que cette indépendance a lieu pour toute origine de $\mathcal{P}$. On dit alors que $V(F)$ est une **parabole**. On a $L_{O'} = L_0$ si et seulement si $\overrightarrow{OO'} \in \text{Ker } L$. La droite vectorielle $\text{Ker } L$ s'appelle la **direction asymptotique**. Voir plus loin la raison de cette terminologie.

Proposition 12.10. *Toute droite de direction la droite vectorielle $\text{Ker } L$ coupe la parabole $V(F)$ en un point unique.*

Démonstration. Soit $\vec{u}$ un vecteur non nul de $\text{Ker } L$. On a

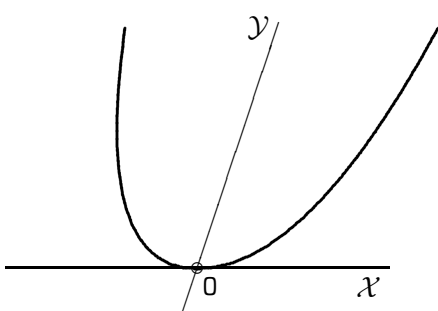
$$F(0 + t\vec{u}) = t^2 L^2(\vec{u}) + 2t L_0(\vec{u}) + F(0) = 2t L_0(\vec{u}) + F(0)$$

Comme $\text{Ker } L \neq \text{Ker } L_0$, on a $L_0(\vec{u}) \neq 0$, donc la droite $0 + \text{Ker } L$ coupe $V(F)$ en l'unique point

$$M_0 = 0 + t_0 \vec{u} \quad \text{où} \quad t_0 = -\frac{F(0)}{2L_0(\vec{u})}$$

Comme 0 est quelconque, toute droite de direction $\text{Ker } L$ coupe $V(F)$ en un unique point. $\square$

Équation réduite d'une parabole



Soit un repère $(0, \vec{i}, \vec{j})$ où $0 \in V(F)$, $\vec{i} \in \text{Ker } L_0$ et $\vec{j} \in \text{Ker } L$ tels que $L(\vec{i}) = -1$ et $L_0(\vec{j}) = \frac{1}{2}$. Pour $M(x, y)$, on a

$$\begin{aligned} F(M) &= L^2(x\vec{i} + y\vec{j}) + 2L_0(x\vec{i} + y\vec{j}) \\ &= -x^2 + y \end{aligned}$$

et $V(F)$ apparaît comme représentation graphique de la fonction $x \mapsto x^2$ et $y - x^2 = 0$ est l'**équation réduite**.

La parabole est image de l'application

$$\mathbb{R} \longrightarrow \mathcal{P}, \quad x \longmapsto M \begin{pmatrix} x \\ x^2 \end{pmatrix}$$

injective continue. Elle est donc connexe, c'est un fermé non compact de $\mathcal{P}$.

Définition 12.11. *Un ensemble $V(F)$ est une **conique propre** s'il est non vide, non réduit à un point et ne contient pas de droites.*

Lemme 12.12. *Si deux f.q.a. F et F' sont telles que $V(F) = V(F')$ est une conique propre, alors F et F' sont proportionnelles.*

Autrement dit, étant donné un repère, si deux polynômes p et q de degré 2 sont tels que $p(x, y) = 0$ et $q(x, y) = 0$ sont équations d'une même conique propre, alors $p(X, Y)$ et $q(X, Y)$ sont des polynômes proportionnels.

Nous admettrons ce lemme, dont la démonstration est « technique ».

Lors des discussions de 12.6, 12.7, 12.8, on a partagé l'ensemble des coniques propres en trois **genres** :

- (1) Le genre **ellipse** d'équation réduite $x^2 + y^2 - 1 = 0$. Une ellipse est connexe, compacte, possède un centre de symétrie.
- (2) Le genre **hyperbole** d'équation réduite $xy - 1 = 0$. Une hyperbole a deux composantes connexes (les branches) non compactes, deux asymptotes, possède un centre de symétrie.
- (3) le genre **parabole** d'équation réduite $y - x^2 = 0$. Une parabole est connexe, non compacte, a une direction asymptotique, n'a pas de centre de symétrie.

Théorème 12.13. *Le groupe affine du plan opère naturellement sur l'ensemble des coniques propres. Cette action détermine trois orbites qui sont les trois genres ellipse, parabole, hyperbole.*

Démonstration. a) Soit $f: \mathcal{P} \rightarrow \mathcal{P}$ bijective affine, $R = (0, \vec{i}, \vec{j})$ un repère, $0_1 = f(0)$, $\vec{f}(\vec{i}) = \vec{i}_1$, $\vec{f}(\vec{j}) = \vec{j}_1$, $M(x, y)$ un point et $M_1 = f(M)$.

$$\vec{0_1 M_1} = \vec{f}(\vec{OM}) = \vec{f}(x\vec{i} + y\vec{j}) = x\vec{f}(\vec{i}) + y\vec{f}(\vec{j}) = x\vec{i}_1 + y\vec{j}_1$$

Donc M_1 a mêmes coordonnées x, y dans le repère $R_1 = (0_1, \vec{i}_1, \vec{j}_1)$ que M dans le repère $R = (0, \vec{i}, \vec{j})$.

b) Soit C une conique propre d'équation réduite $p(x, y) = 0$ dans le repère $R = (0, \vec{i}, \vec{j})$, f une bijection affine, $R_1 = f(R) = (0_1, \vec{i}_1, \vec{j}_1)$ le repère image. Pour tout $M(x, y) \in C$, $p(x, y) = 0$. Par a), $f(M) = M_1$ est sur la conique C_1 d'équation $p(x, y) = 0$ dans le repère R_1 . Donc $f(C) \subset C_1$. Par le même raisonnement appliqué à C_1 et f^{-1} , on a $f^{-1}(C_1) \subset C$, d'où $f(C) = C_1$. Ces coniques sont de même genre car de même équation réduite.

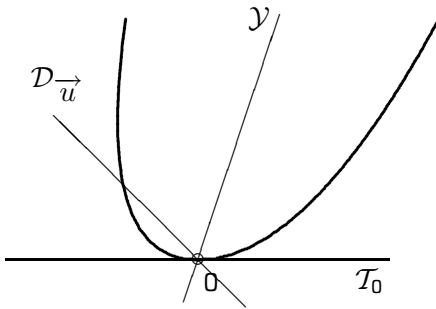
c) Inversement, soit C et C_1 deux coniques de même genre, $R = (0, \vec{i}, \vec{j})$ et $R_1 = (0_1, \vec{i}_1, \vec{j}_1)$ deux repères tels que C ait même équation réduite $p(x, y) = 0$ dans R que C_1 dans R_1 . Il existe une unique bijection affine f telle que $f(0) = 0_1$, $\vec{f}(\vec{i}) = \vec{i}_1$, $\vec{f}(\vec{j}) = \vec{j}_1$. Alors, par b), $f(C) = C_1$.

Le groupe affine opère donc transitivement sur chacun des genres. $\square$

12.1.2. Positions relatives d'une conique et d'une droite

On désigne par F une f.q.a. telle que $V(F) = C$ soit une conique propre. Soit 0 un point de C . Comme $F(0) = 0$, F est de la forme $M \mapsto F(M) = Q(\vec{OM}) + 2L_0(\vec{OM})$. Faisons pivoter autour de 0 la droite $\mathcal{D}_{\vec{u}} = 0 + \mathbb{R}\vec{u}$. Les points de $C \cap \mathcal{D}_{\vec{u}}$ sont donnés par l'équation en λ

$$0 = F(0 + \lambda\vec{u}) = \lambda^2 Q(\vec{u}) + 2\lambda L_0(\vec{u}) = \lambda \left(\lambda Q(\vec{u}) + 2L_0(\vec{u}) \right)$$



La solution $\lambda = 0$ correspond à $0 \in C \cap \mathcal{D}_{\vec{u}}$. On dit que $\mathcal{D}_{\vec{u}}$ est **tangente** si 0 est **racine double**, **sécante** si 0 est **racine simple**. La tangente en 0 est donc la droite $\mathcal{T}_0 = 0 + \text{Ker } L_0$, de direction $\text{Ker } L_0$.

Remarquons que pour $\vec{u} \in \text{Ker } L_0$ non nul, $Q(\vec{u}) \neq 0$ sinon C contiendrait la droite $0 + \text{Ker } L_0$ et serait donc impropre. Dans ces conditions, 0 est l'unique point de $C \cap \mathcal{T}_0$.

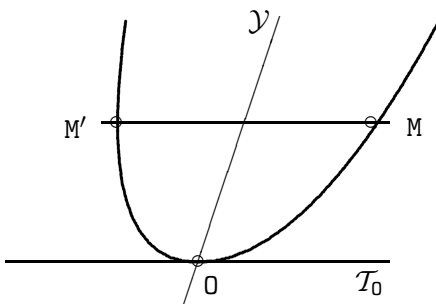
En général, une sécante $\mathcal{D}_{\vec{u}}$ coupe C en deux points distincts, sauf si $\mathcal{D}_{\vec{u}}$ est la droite $\mathcal{Y}$ passant par 0

- parallèle à une asymptote si C est une hyperbole,
- de direction asymptotique si C est une parabole (12.10, p. 319),

car alors $Q(\vec{u}) = 0$, le polynôme en λ est de degré 1 et 0 est alors unique point commun de C et $\mathcal{D}_{\vec{u}}$.

Proposition 12.14. *Dans ces conditions, la symétrie oblique parallèlement à la tangente $\mathcal{T}_0$ autour de la droite passant par 0 et Q -orthogonale à $\mathcal{T}_0$ laisse C stable. Cet axe $\mathcal{Y}$ de symétrie oblique s'appelle le **diamètre passant par 0**.*

- Si C est une ellipse ou hyperbole de centre Ω , c'est la droite (0Ω) ,
- Si C est une parabole de direction asymptotique Y , c'est la droite $0 + Y$.



Démonstration. Soit un repère $(0, \vec{\tau}, \vec{\nu})$ où $\vec{\tau} \in \text{Ker } L_0$ et $\vec{\nu}$ lui est Q -orthogonal. Comme C est propre, $Q(\vec{\tau}) \neq 0$ et $\mathcal{D}_{\vec{\tau}} = \mathcal{T}_0$ est la tangente en 0 à C . L'équation de C est $ax^2 + cy^2 + 2ey = 0$. La symétrie oblique est l'application $M(x, y) \mapsto M'(-x, y)$ qui laisse C stable.

On a $a \neq 0$ sinon C est union de deux droites de même direction $\mathbb{R}\vec{\tau}$.

Si $c \neq 0$, l'équation s'écrit $ax^2 + c\left(y + \frac{e}{c}\right)^2 - \frac{e^2}{c} = 0$ et C est à centre $\Omega(0, -\frac{e}{c})$. Le diamètre passant par 0 est bien (0Ω) .

Si $c = 0$, l'équation est $ax^2 + 2ey = 0$ et C est une parabole de direction asymptotique de vecteur directeur $\vec{\nu}$. $\square$

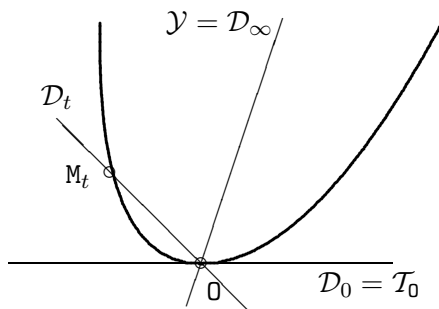
Représentation paramétrique. Dans ces conditions, on a aussi $e \neq 0$. En effet, si $e = 0$, l'équation est $ax^2 + cy^2 = 0$ et C est impropre car C se réduit à $\{0\}$ si a et c ont même signe, se réduit à une droite double si $c = 0$, se compose de deux droites si a et c ont des signes contraires.

La droite $\mathcal{D}_t$ d'équation $y = tx$ coupe C en 0 et en le point

$$M_t \left(x(t) = -\frac{2et}{a + ct^2}, y(t) = -\frac{2et^2}{a + ct^2} \right)$$

C'est une représentation paramétrique de C . Si $t = 0$, $\mathcal{D}_0$ est tangente en 0 au sens des courbes paramétrées car $x'(0) = -\frac{2e}{a}$ et $y'(0) = 0$. La notion de tangente à une conique entre dans le cadre de la notion générale de tangente.

Si t est tel que $a + ct^2 = 0$, le point M_t n'existe plus. C'est le cas si C est une parabole ou une hyperbole et $\mathcal{D}_t$ parallèle à une direction asymptotique.



La droite $\mathcal{D}_\infty$ d'équation $x = 0$ est le diamètre.

Proposition 12.15. *Le complémentaire de C dans $\mathcal{P}$ se partage en :*

- (1) *l'extérieur, ensemble des P par où passent deux tangentes distinctes,*
- (2) *l'intérieur, ensemble des P tels que toute droite passant par P soit sécante.*

Démonstration. Soit $\vec{u}$ non nul et $\mathcal{D}_{\vec{u}}$ la droite passant par P de vecteur directeur $\vec{u}$. Un point $P + \lambda \vec{u}$ de $\mathcal{D}_{\vec{u}}$ est sur C si et seulement si

$$0 = F(P + \lambda \vec{u}) = \lambda^2 Q(\vec{u}) + 2\lambda L_P(\vec{u}) + F(P)$$

La discussion porte sur le signe du discriminant $\Delta_P(\vec{u}) = L_P^2(\vec{u}) - Q(\vec{u})F(P)$. C'est une forme quadratique en $\vec{u}$. Discutons en la signature.

- **Cas d'une parabole.** On a $Q = L^2$ et $\Delta_P(\vec{u}) = L_P^2(\vec{u}) - L^2(\vec{u})F(P)$. C'est une décomposition en somme de carrés dont le premier terme est positif et le deuxième du signe de $F(P)$.
 - Si $F(P) < 0$, alors Δ_P est définie positive, toute droite passant par P est sécante. Alors P est **intérieur** à C .
 - Si $F(P) > 0$, la signature est $(1, 1)$, les deux directions isotropes pour Δ_P sont les directions des tangentes issues de P , point **extérieur** à C .
- **Cas d'une conique à centre Ω .** On a $F(M) = Q(\vec{\Omega M}) + F(\Omega)$. La formule de changement d'origine (12.2, p. 315) donne $L_P = \Phi(\vec{\Omega P}, -)$. La forme polaire de la forme quadratique Δ_P est

$$\delta_P: (\vec{u}, \vec{v}) \mapsto \Phi(\vec{\Omega P}, \vec{u})\Phi(\vec{\Omega P}, \vec{v}) - \Phi(\vec{u}, \vec{v})F(P)$$

Prenons $\vec{u} = \vec{u}_P$ dans $\text{Ker } L_P$, i.e. Q -orthogonal à $\vec{\Omega P}$ et $\vec{v} = \vec{\Omega P}$. On a

$$\delta_P(\vec{u}_P, \vec{\Omega P}) = \Phi(\vec{\Omega P}, \vec{u}_P)Q(\vec{\Omega P}) - \Phi(\vec{u}_P, \vec{\Omega P})F(P) = 0$$

car les deux termes sont nuls. La base $(\vec{u_P}, \vec{\Omega P})$ est Δ_P -orthogonale. Calculons la signature. Comme $F(P) = Q(\vec{\Omega P}) + F(\Omega)$, on a

$$\Delta_P(\vec{u_P}) = -Q(\vec{u_P})F(P), \quad \Delta_P(\vec{\Omega P}) = Q^2(\vec{\Omega P}) - Q(\vec{\Omega P})F(P) = -Q(\vec{\Omega P})F(\Omega)$$

- **Cas d'une ellipse.** Alors Q est définie positive et $F(\Omega) < 0$. La signature est $(1, 1)$ si et seulement si $F(P)$ et $F(\Omega)$ sont de signes contraires. Le centre Ω est intérieur et P est extérieur si et seulement si $F(P) > 0$.
- **Cas d'une hyperbole.** La signature de Q étant $(1, 1)$, $Q(\vec{u_P})$ et $Q(\vec{\Omega P})$ sont de signes contraires. La signature de Δ_P est $(1, 1)$ si et seulement si $F(P)$ a le signe de $F(\Omega)$. Le centre Ω est extérieur à C et P est extérieur si et seulement si $F(P)$ a le signe de $F(\Omega)$.

Pour une hyperbole, l'énoncé n'est pas tout à fait exact car Ω est à l'extérieur, mais il ne passe aucune tangente par le centre. Les droites Δ_Ω -isotropes passant par Ω sont les **asymptotes** ce qui amène à les considérer comme des tangentes particulières, les points de contact étant « à l'infini ». La géométrie projective est le cadre propre pour une justification rigoureuse de ceci.

□

Exercice 12.1. Soit F une f.q.a. telle que $V(F) \neq \emptyset$ et $0 \in V(F)$. Montrer que $V(F)$ est une conique propre si et seulement si $\text{Ker } L_0$ est une droite vectorielle non isotrope pour Q

Exercice 12.2. Soit C une conique propre, P un point extérieur, $\mathcal{D}$ le diamètre passant par P , Δ la parallèle passant par P aux tangentes en les points où $\mathcal{D}$ coupe C , $\mathcal{T}_1, \mathcal{T}_2$ les tangentes à C issues de P en les points M_1 et M_2 .

Montrer que les droites $\Delta, \mathcal{D}, \mathcal{T}_1, \mathcal{T}_2$ forment un faisceau harmonique et que $(M_1 M_2)$ est parallèle à Δ .

Exercice 12.3. polarité

Soit $C = V(F)$ une conique propre.

1) Soit $P \notin C$. Une sécante $\mathcal{D}_{\vec{u}}$ passant par P de vecteur directeur $\vec{u}$ coupe C en M', M'' . Soit Q conjugué harmonique de P relativement à M', M'' .

Montrer que $L_P(\vec{PQ}) + F(P) = 0$.

Montrer que le lieu géométrique de Q si $\mathcal{D}_{\vec{u}}$ pivote autour de P est porté par une droite Δ_P appelée **polaire** de P relativement à C . Montrer que Δ_P est Q -orthogonale au diamètre passant par P .

Le lieu de Q est-il la polaire Δ_P dans son entier ?

2) On dit que P et Q sont **conjugués** relativement à C si $Q \in \Delta_P$. Montrer que c'est une relation symétrique.

3) Peut-on étendre la définition de la polaire aux points de C ?

Rappels sur la théorie des enveloppes Le plan $\mathcal{P}$ étant rapporté à un repère $(O, \vec{i}, \vec{j})$ on donne une application d'un intervalle $I \subset \mathbb{R}$ dans l'ensemble des droites $t \mapsto \mathcal{D}_t$ où $\mathcal{D}_t$ est d'équation $u(t)x + v(t)y + w(t) = 0$. On suppose les fonctions $u: t \mapsto u(t)$, $v: t \mapsto v(t)$, $w: t \mapsto w(t)$ de classe C^1 , les dérivées u' et v' ne s'annulant pas simultanément. Supposons que $\mathcal{D}_t$ soit tangente à une courbe $\mathcal{C}$. Le **point caractéristique** $M(t)$, point de contact de $\mathcal{D}_t$ et $\mathcal{C}$, est l'intersection des droites $\mathcal{D}_t$ et $\mathcal{D}'_t$ (supposée distinctes) d'équation $u'(t)x + v'(t)y + w'(t) = 0$.

Exercice 12.4. Dans ces conditions, on suppose que $u(t), v(t), w(t)$ sont des polynômes de degré 2 en t :

$$u(t) = u_2 t^2 + 2u_1 t + u_0, v(t) = v_2 t^2 + 2v_1 t + v_0, w(t) = w_2 t^2 + 2w_1 t + w_0$$

1) Combien de droites du type $\mathcal{D}_t$ passant par un point donné $M(x, y)$?

2) En déduire qu'en général l'enveloppe des droites $\mathcal{D}_t$ est la conique d'équation

$$(u_1 x + v_1 y + w_1)^2 - (u_2 x + v_2 y + w_2)(u_0 x + v_0 y + w_0) = 0$$

Exercice 12.5. Soit deux droites $\mathcal{X}$ et $\mathcal{Y}$ sécantes en O parcourues par deux points mobiles P(t) et Q(t) animés de mouvements uniformes.

1) On suppose que P et Q ne sont pas en O au même instant. Montrer que la droite $\mathcal{D}_t = (P(t)Q(t))$ enveloppe une parabole tangente à $\mathcal{X}$ et $\mathcal{Y}$ en des points A et B à préciser.

2) Que se passe-t-il si P et Q sont tous deux en O au même instant t_0 ?

Exercice 12.6. Soit $\mathcal{A}$ et $\mathcal{B}$ deux droites distinctes parallèles de direction D , $A \in \mathcal{A}$, $B \in \mathcal{B}$, $\vec{v} \neq \vec{0}$ sur D . Des points P et Q décrivent les droites $\mathcal{A}$ et $\mathcal{B}$ de sorte que le produit $\overline{AP} \times \overline{BQ} = p$ reste constant. Que dire de l'enveloppe de la droite (PQ) ?

12.2 CONIQUES DANS UN PLAN AFFINE EUCLIDIEN

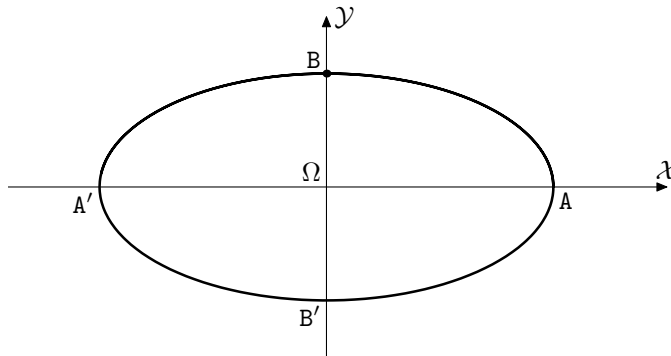
On suppose le plan vectoriel P muni d'un produit scalaire.

12.2.1. Axes d'une conique à centre

Soit $C = V(F)$ une conique à centre Ω et Q la forme quadratique associée.

Si Q est proportionnelle à la forme quadratique euclidienne $\vec{v} \mapsto \|\vec{v}\|^2$, alors C est un cercle.

Dans le cas contraire, il existe deux directions uniques orthogonales à la fois pour la forme quadratique euclidienne et pour Q . Munissons ces directions de vecteurs unitaires $\vec{i}, \vec{j}$, on a une base $(\vec{i}, \vec{j})$ à la fois orthonormale et Q -orthogonale. Dans le repère orthonormé $(\Omega, \vec{i}, \vec{j})$, l'équation de C est du type $Ux^2 + Vy^2 + W = 0$ où U, V, W sont non nuls car C est propre.



- (1) On suppose Q définie positive, donc C est une **ellipse**. Alors U et V sont positifs et W négatif. En échangeant éventuellement $\vec{i}$ et $\vec{j}$, on peut supposer $0 < U < V$. La droite $\mathcal{X} = \Omega + \mathbb{R}\vec{i}$ coupe C en les **sommets** A, A' . L'équation devient

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$$

où $a = \sqrt{-\frac{W}{U}} = \overline{\Omega A} = -\overline{\Omega A'}$, $b = \sqrt{-\frac{W}{V}}$. Comme C n'est pas un cercle, puisque $b < a$, les seuls axes de symétrie sont :

- l'**axe focal** $\mathcal{X} = \Omega + \mathbb{R}\vec{i}$ coupant C en $A(a, 0)$ et $A'(-a, 0)$,
- l'**axe non focal** $\mathcal{Y} = \Omega + \mathbb{R}\vec{j}$ coupant C en $B(b, 0)$ et $B'(-b, 0)$.

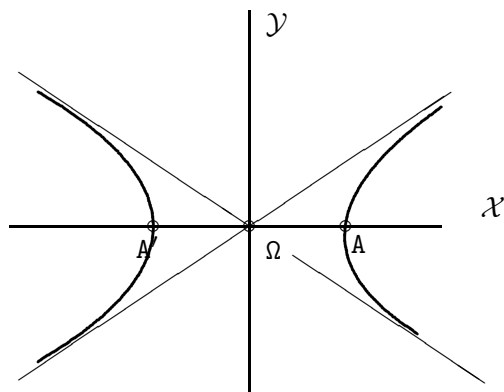
- (2) On suppose Q de signature $(1, 1)$, donc U, V de signes contraires, C est une **hyperbole**. En échangeant éventuellement $\vec{i}$ et $\vec{j}$, on peut supposer U et W de signes contraires, et donc que $\mathcal{X} = \Omega + \mathbb{R}\vec{i}$ coupe C en deux points : les **sommets** A, A' . L'équation devient

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0 \quad \text{où} \quad a = \sqrt{-\frac{W}{U}} = \overline{\Omega A} = -\overline{\Omega A'}, \quad b = \sqrt{\frac{W}{V}}$$

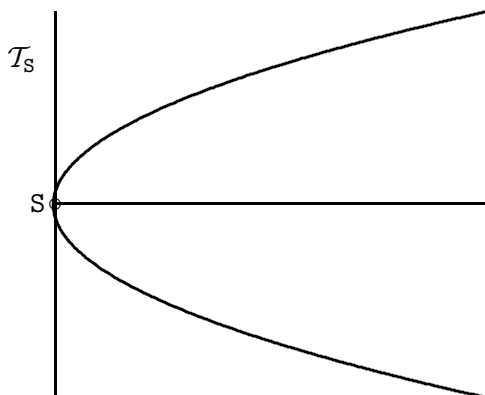
Les seuls axes de symétrie sont :

- l'**axe focal** $\mathcal{X} = \Omega + \mathbb{R} \vec{i}$ coupant C en $A(a, 0)$ et $A'(-a, 0)$
- l'**axe non focal** $\mathcal{Y} = \Omega + \mathbb{R} \vec{j}$ ne coupant pas C .

Ces processus ne sont pas ambigus. Les longueurs a, b sont donc **intrinsèquement** associées à C . Les équations obtenues sont les **équations métriques réduites** de C . À la différence de ce qui se passe pour le groupe affine (12.13, p. 320), l'opération du groupe des isométries détermine une **infinité** d'orbites, chacune d'elles correspondant à une équation métrique réduite.



12.2.2. Axe d'une parabole



Soit C une parabole d'équation réduite (affine) $y - x^2 = 0$. La dérivée $x \mapsto 2x$ de $x \mapsto x^2$ est une bijection $\mathbb{R} \rightarrow \mathbb{R}$, donc pour toute direction T distincte de la direction asymptotique, il existe un unique point $M \in C$ tel que la tangente en M ait la direction T . On appelle **sommet** de C l'unique point $S \in C$ où la tangente est orthogonale à la direction asymptotique. L'unique axe de symétrie est la droite passant par le sommet parallèle à la direction asymptotique appelée **axe** de la parabole C .

Soit un repère orthonormé $(S, \vec{i}, \vec{j})$ où $\vec{i}, \vec{j}$ sont unitaires et vecteurs directeurs de la direction asymptotique et de la tangente au sommet T_s . L'équation de C est de la forme $y^2 - 2px = 0$. En remplaçant éventuellement $\vec{i}$ par $-\vec{i}$, on peut supposer $p > 0$. On a alors l'**équation métrique réduite** de la parabole. Ici encore, l'action du groupe des isométries détermine sur l'ensemble des paraboles une infinité d'orbites, chacune d'elles correspondant à une valeur du **paramètre** p .

12.2.3. Définition par foyers et directrices

Quels exemples connaît-on de f.q.a. ayant une description géométrique ?

Étant donné un point O , la fonction $M \mapsto OM^2$ est une f.q.a. Une combinaison linéaire de telles f.q.a. est une fonction de Leibnitz (9.19, p. 255), donc une fonction circulaire (11.12, p. 289). Les coniques ainsi obtenues sont des droites ou des cercles. Il faut donc considérer d'autres f.q.a.

Étant donné une droite $\mathcal{D}$, la fonction associant à M le carré de sa distance à $\mathcal{D}$ est une f.q.a. (9.15, p. 252).

En combinant des f.q.a. de ces deux types, on peut espérer obtenir n'importe quelle f.q.a., donc n'importe quelle conique. Essayons la plus simple de ces combinaisons.

Définition 12.16. Soit e un réel strictement positif, F un point, $\mathcal{D}$ une droite ne passant pas par F . La fonction $F: M \mapsto MF^2 - e^2MH^2$, où H est la projection orthogonale de M sur $\mathcal{D}$, est une f.q.a. On dit que la conique associée

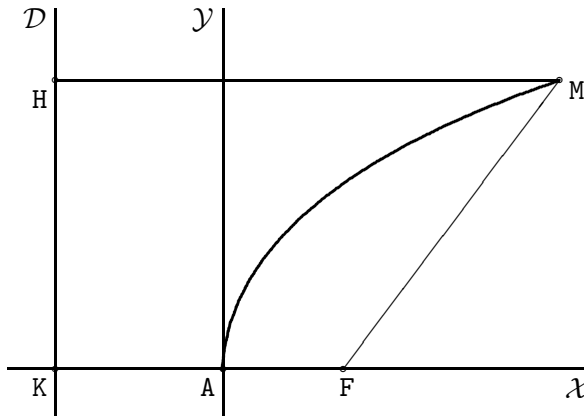
$$V(F) = \{M \mid MF^2 - e^2MH^2 = 0\} = \{M \mid \frac{MF}{MH} = e\}$$

est de **foyer** F , de **directrice** $\mathcal{D}$ et d'**excentricité** e .

Plusieurs questions se posent :

- la conique $V(F)$ est-elle propre ?
- si oui, quel est son genre ?
- toute conique propre est-elle de ce type ?

Équation d'une conique donnée par foyer et directrice. Soit K la projection orthogonale de F sur $\mathcal{D}$ et $A \in [K, F]$ tel que $\frac{AF}{AK} = -e$.



Soit un repère orthonormé $(A, \vec{i}, \vec{j})$ où $\vec{i}$ est unitaire, colinéaire de même sens que $\overrightarrow{AF}$ et $\mathcal{X} = A + \mathbb{R} \vec{i}$ est la droite passant par le foyer F orthogonale à la directrice $\mathcal{D}$. Posant $\overline{KA} = k > 0$, on a $\overline{AF} = ek$. Pour $M(x, y)$, de projection H sur $\mathcal{D}$, on a

$$\begin{aligned} MF^2 - e^2MH^2 &= (x - ek)^2 + y^2 - e^2(x + k)^2 \\ &= (1 - e^2)x^2 + y^2 - 2ek(1 + e)x \end{aligned}$$

Posons $p = ek(1 + e)$. L'équation devient $(1 - e^2)x^2 + y^2 - 2px = 0$ où $p > 0$ est une longueur et $e > 0$ un nombre pur.

Inversement, étant donné un repère orthonormé $(A, \vec{i}, \vec{j})$, une longueur $p > 0$, un nombre pur $e \geq 0$, soit $C_{e,p}$ la conique d'équation

$$(1 - e^2)x^2 + y^2 - 2px = 0$$

Deux cas se présentent :

1) Si $e > 0$, $C_{e,p}$ est la conique d'excentricité e , de foyer $F(\frac{p}{e+1}, 0)$, de directrice $\mathcal{D}$ d'équation $x = -\frac{p}{e(e+1)}$. En effet, posant $k = \frac{p}{e(e+1)}$, soit la droite $\mathcal{D}$ d'équation $x = -k$ et le point $F(ek, 0)$. Le calcul précédent montre que $C_{e,p}$ est l'ensemble des M dont le rapport des distances à F et $\mathcal{D}$ est e .

2) Si $e = 0$, $C_{0,p}$ est le cercle de centre $F(\frac{p}{e+1}, 0)$ et de rayon p . En effet, l'équation devient $x^2 + y^2 - 2px = 0$. La longueur k n'existe pas et $\mathcal{D}$ non plus, mais $F(p, 0)$ est le centre et p est le rayon.

Théorème 12.17. *Étant donné un repère orthonormé $(A, \vec{i}, \vec{j})$, une longueur $p > 0$ et un nombre pur $e \geq 0$, soit $C_{e,p}$ la conique d'équation*

$$(1 - e^2)x^2 + y^2 - 2px = 0$$

La conique $C_{e,p}$ est propre, hyperbole si $e > 1$, parabole si $e = 1$, ellipse si $0 \leq e < 1$.

Pour $e > 0$, $C_{e,p}$ admet A pour sommet et $\mathcal{X} = A + \mathbb{R}\vec{i}$ pour axe focal.

Pour $e = 0$, $C_{0,p}$ est un cercle de rayon p tangent en A à $\mathcal{Y} = A + \mathbb{R}\vec{j}$.

Pour toute conique propre C , il existe un unique couple (e, p) tel que C soit isométrique à $C_{e,p}$.

Démonstration. Pour $e = 1$, $C_{1,p}$ est la parabole d'équation $y^2 - 2px = 0$, de sommet A , d'axe $\mathcal{X}$, de paramètre p . Toute parabole est bien isométrique à une unique parabole $C_{1,p}$ (12.2.2., p. 326).

Pour $e \neq 1$, $C_{e,p}$ est une conique à centre. L'équation s'écrit

$$\begin{aligned} \left(x - \frac{p}{1-e^2}\right)^2 + \frac{y^2}{1-e^2} - \frac{p^2}{(1-e^2)^2} &= 0 \\ x_1^2 + \frac{y_1^2}{1-e^2} - \frac{p^2}{(1-e^2)^2} &\quad \text{où} \quad x_1 = x - \frac{p}{1-e^2}, \quad y_1 = y \\ \frac{x_1^2(1-e^2)^2}{p^2} - \frac{y_1^2(1-e^2)}{p^2} - 1 &= 0 \end{aligned}$$

Le centre est donc Ω de coordonnées $(\frac{p}{1-e^2}, 0)$ dans le repère initial $(A, \vec{i}, \vec{j})$.

- Pour $0 \leq e < 1$, c'est l'équation réduite métrique d'une **ellipse**. Les longueurs a, b relatives à l'équation métrique réduite sont

$$a = \frac{p}{1-e^2}, \quad b = \frac{p}{\sqrt{1-e^2}}, \quad b \leq a$$

Pour $0 < e < 1$, on a $b < a$ et l'axe focal est $\mathcal{X}$.

Inversement, la donnée de (a, b) permet de déterminer (e, p) , donc toute ellipse est isométrique à une unique conique $C_{e,p}$ où $0 \leq e < 1$:

$$e = \sqrt{1 - \frac{b^2}{a^2}}, \quad p = \frac{b^2}{a}$$

- Pour $e > 1$, c'est l'équation réduite métrique d'une **hyperbole** d'axe focal $\mathcal{X}$. Les longueurs a, b relatives à l'équation métrique réduite sont

$$a = \frac{p}{e^2 - 1}, \quad b = \frac{p}{\sqrt{e^2 - 1}}$$

Inversement, la donnée de (a, b) permet de déterminer (e, p) , donc toute hyperbole est isométrique à une unique conique $C_{e,p}$ où $e > 1$:

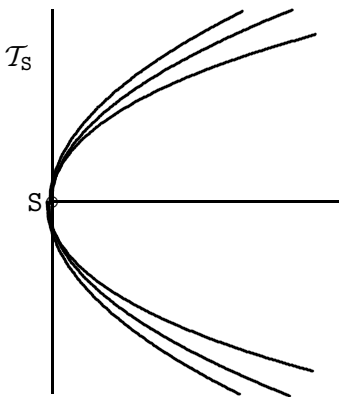
$$e = \sqrt{1 + \frac{b^2}{a^2}}, \quad p = \frac{b^2}{a}$$

Dans les deux cas $e < 1$ et $e > 1$, le centre est donné par $\overline{A\Omega} = \frac{p}{1-e^2}$, à droite ou à gauche de A selon que C est ellipse ou hyperbole. $\square$

Il résulte de ce qui précède que l'orbite d'une conique propre C pour l'action du groupe des isométries est déterminée par le couple (e, p) .

Le **paramètre** $p > 0$ est une longueur, l'**excentricité** $e \geq 0$ est un nombre pur. La terminologie **axe focal** pour l'axe de symétrie $\mathcal{X} = A + \mathbb{R} \vec{i}$ vient de ce que le ou les foyers sont sur $\mathcal{X}$.

Remarque : Dans l'étude des coniques propres, nous avons rencontré deux cas exceptionnels : les paraboles qui n'ont pas de centre, les cercles qui n'ont pas de directrice. En revanche, l'équation du théorème précédent est valable dans **tous les cas** et le couple (p, e) détermine une conique propre à isométrie près. C'est l'intérêt de cette équation.



Supposons p donné et faisons varier e continûment de 0 à l'infini, $C_{e,p}$ étant représentée dans le repère fixe $(A, \vec{i}, \vec{j})$, par exemple sur un écran d'ordinateur. Que l'observateur placé en l'origine ne s'attende pas à constater dans la forme de $C_{e,p}$ un changement brutal quand e franchit la valeur 1.

Pour $e \neq 1$, le centre Ω est donné par $\overline{A\Omega} = \frac{p}{1-e^2}$. Si e croît de 0 à 1, le centre Ω s'éloigne à droite vers l'infini, n'existe plus pour $e = 1$, reparait à gauche et se rapproche de A si e tend vers $+\infty$. Pour e très proche de 1, l'observateur ne pourra pas constater le

genre de $C_{e,p}$ sur son écran, le centre étant hors de l'écran tellement éloigné à gauche ou à droite qu'on ne verra pas de différence entre ellipse, parabole et hyperbole.

12.2.4. Définition bifocale des coniques à centre

Les deux couples foyer-directrice. Soit C une conique de paramètre p et d'excentricité e distincte de 0 et 1 : C est une conique à centre qui n'est pas un cercle. On rapporte le plan à un repère orthonormé $(A, \vec{i}, \vec{j})$ tel que l'équation de C soit

$$(1 - e^2)x^2 + y^2 - 2px = 0$$

Par les calculs de la preuve du théorème 12.17 (p. 328), le foyer F , la directrice $\mathcal{D}$ et le centre Ω sont donnés par (axe focal $\mathcal{X}$ orienté de A vers F)

$$\overline{A\Omega} = \frac{p}{1 - e^2}, \quad \overline{AF} = \frac{p}{1 + e}, \quad \text{équation de } \mathcal{D} : x = -\frac{p}{e(e + 1)}$$

La symétrie autour du centre Ω transforme $A, F, \mathcal{D}$ en $A', F', \mathcal{D}'$, d'où un deuxième couple foyer-directrice $F', \mathcal{D}'$. Il n'en existe pas d'autre car C n'a que deux sommets A, A' et les valeurs de l'excentricité e et du paramètre p déterminent le foyer et la directrice à partir du sommet A ou A' choisi.

Théorème 12.18. Soit C une conique de centre Ω , de foyers F, F' , de sommets A, A' , $a = \Omega A = \Omega A'$. Alors C est l'ensemble des M tels que

- 1) $MF + MF' = 2a$ si C est une ellipse,
- 2) $MF - MF' = \pm 2a$ si C est une hyperbole.

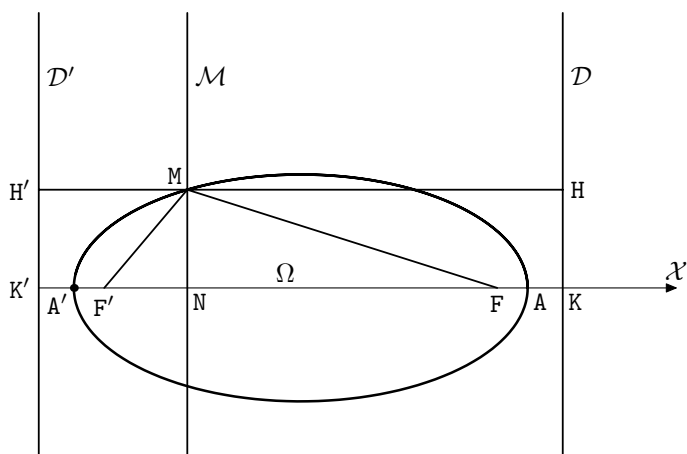
Démonstration. a) Soit $\mathcal{D}$ et $\mathcal{D}'$ les directrices relatives à F et F' coupant l'axe focal $\mathcal{X}$, orienté de A vers F , en K et K' . Par les formules précédentes

$$\begin{aligned} \overline{MK} &= \overline{AK} - \overline{A\Omega} = -\frac{p}{e(e + 1)} - \frac{p}{1 - e^2} = \frac{p}{e(e^2 - 1)} = \frac{\overline{\Omega A}}{e} \\ \overline{MK'} &= \overline{AK'} - \overline{A\Omega} = -\frac{p}{e + 1} - \frac{p}{1 - e^2} = \frac{pe}{e^2 - 1} = e\overline{\Omega A} \end{aligned}$$

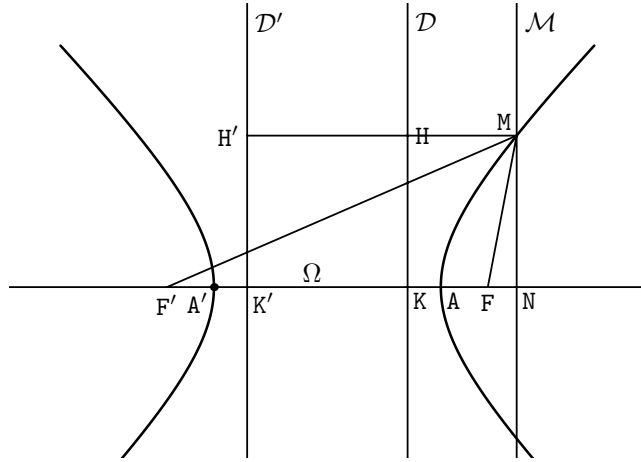
Orientons maintenant l'axe focal $\mathcal{X}$ de Ω vers A . Alors

$$\overline{AK} = a = -\overline{AK'} \quad , \quad \overline{MK} = \frac{a}{e} = -\overline{MK'} \quad , \quad \overline{MF} = ea = -\overline{MF'}$$

Cas de l'ellipse :



Cas de l'hyperbole :



b) Soit M un point, H et H' ses projections orthogonales sur $\mathcal{D}$ et $\mathcal{D}'$. On a

$$\frac{2a}{e} = \overline{K'K} = \overline{H'H} = \overline{MH} - \overline{MH'}$$

Deux cas se présentent :

- (1) M est entre H' et H, c'est-à-dire intérieur à la bande du plan $\mathcal{P}$ limitée par $\mathcal{D}'$ et $\mathcal{D}$. Alors $\overline{MH'} + \overline{MH} = \overline{H'H} = \frac{2a}{e}$.
- (2) M est extérieur au segment $[H', H]$, c'est-à-dire extérieur à la bande du plan $\mathcal{P}$ limitée par $\mathcal{D}'$ et $\mathcal{D}$. Alors $\overline{MH'} - \overline{MH} = \pm \overline{H'H} = \pm \frac{2a}{e}$.

c) Soit $M \in C$. On a $\overline{MF} = e\overline{MH}$ et $\overline{MF'} = e\overline{MH'}$. Soit le repère orthonormé $(\Omega, \vec{i}, \vec{j})$ où $\mathcal{X} = \Omega + \mathbb{R}\vec{i}$ est l'axe focal.

- (1) $0 < e < 1$, C est une ellipse d'équation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$$

Pour $M(x, y)$ sur C , $x^2 \leq a^2 < \left(\frac{a}{e}\right)^2$, donc $|x| \leq a < \frac{a}{e}$, et M est intérieur à la bande limitée par $\mathcal{D}$ et $\mathcal{D}'$. On a donc

$$\overline{MF} + \overline{MF'} = e(\overline{MH} + \overline{MH'}) = 2a$$

- (2) $1 < e$, C est une hyperbole d'équation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0$$

Pour $M(x, y)$ sur C , $x^2 \geq a^2 > \left(\frac{a}{e}\right)^2$, donc $|x| \geq a > \frac{a}{e}$ et M est extérieur à la bande limitée par $\mathcal{D}$ et $\mathcal{D}'$. On a donc

$$\overline{MF} - \overline{MF'} = e(\overline{MH} - \overline{MH'}) = \pm 2a$$

d) Reste la réciproque : si C est une ellipse (resp. une hyperbole) et si $MF + MF' = 2a$ (resp. $MF - MF' = \pm 2a$), a-t-on $M \in C$? Soit M un point quelconque, N sa projection orthogonale sur $\mathcal{X}$. On a

$$\begin{aligned} MF^2 - MF'^2 &= (MF - MF')(MF + MF') \\ &= (\overrightarrow{MF} - \overrightarrow{MF'}) \cdot (\overrightarrow{MF} + \overrightarrow{MF'}) = 2\overrightarrow{F'F} \cdot \overrightarrow{NN} = 4ea\overrightarrow{NN} \end{aligned}$$

- (1) C est une ellipse. Supposons $MF + MF' = 2a$. Alors $2a(MF - MF') = 4ea\overrightarrow{NN}$, donc $2e\Omega N = |MF - MF'| \leq FF' = 2ea$, d'où $\Omega N \leq a$, donc $N \in [A', A]$ est intérieur à C . La perpendiculaire $\mathcal{M}$ en N à $\mathcal{X}$ coupe C en deux points symétriques autour de $\mathcal{X}$. Soit M_1 celui de ces points situé sur la demi-droite limitée par N sur $\mathcal{M}$ contenant M et montrons que $M = M_1$.

On a :

$$\begin{aligned} MF &= \sqrt{NM^2 + NF^2}, \quad MF' = \sqrt{NM^2 + NF'^2}, \quad MF + MF' = 2a \\ M_1F &= \sqrt{NM_1^2 + NF^2}, \quad M_1F' = \sqrt{NM_1^2 + NF'^2}, \quad M_1F + M_1F' = 2a \end{aligned}$$

L'application $y \mapsto \sqrt{y^2 + NF^2} + \sqrt{y^2 + NF'^2}$ est strictement croissante, donc injective. On en déduit $NM = NM_1$, d'où $M = M_1$ appartient à C .

- (2) C est une hyperbole. Supposons $MF - MF' = \pm 2a$. Alors $\pm 2a(MF + MF') = 4ea\overrightarrow{NN}$, donc $2e\Omega N = MF + MF' \geq FF' = 2ea$, d'où $\Omega N \geq a$, donc $N \notin [A', A]$ est extérieur à C . La perpendiculaire $\mathcal{M}$ en N à $\mathcal{X}$ coupe C en deux points symétriques autour de $\mathcal{X}$. Soit M_1 celui de ces points situé sur la demi-droite limitée par N sur $\mathcal{M}$ contenant M et montrons que $M = M_1$.

On a :

$$\begin{aligned} MF &= \sqrt{NM^2 + NF^2}, \quad MF' = \sqrt{NM^2 + NF'^2}, \quad MF - MF' = 2\epsilon a \\ M_1F &= \sqrt{NM_1^2 + NF^2}, \quad M_1F' = \sqrt{NM_1^2 + NF'^2}, \quad M_1F - M_1F' = 2\epsilon a \end{aligned}$$

$\epsilon = \pm 1$ ne dépendant que de la place de la projection orthogonale N de M et M_1 sur $\mathcal{X}$. L'application

$$y \mapsto \sqrt{y^2 + NF^2} - \sqrt{y^2 + NF'^2} = \frac{NF^2 - NF'^2}{\sqrt{y^2 + NF^2} + \sqrt{y^2 + NF'^2}}$$

est strictement monotone (dénominateur croissant), donc injective. On en déduit $NM = NM_1$, d'où $M = M_1$ appartient à C .

□

Théorème 12.19. Soit C une conique de centre Ω et de foyers F, F' .

(i) La tangente $\mathcal{T}_M$ en un point M de C est

- bissectrice extérieure de $(\overrightarrow{MF}, \overrightarrow{MF'})$ si C est une ellipse,
- bissectrice intérieure de $(\overrightarrow{MF}, \overrightarrow{MF'})$ si C est une hyperbole.

(ii) Le symétrique de F relativement à $\mathcal{T}_M$ décrit le cercle F' de centre F' et de rayon $2a$ appelé **cercle directeur** de C relatif à F' .

(iii) Les projections orthogonales H et H' de F et F' sur $\mathcal{T}_M$ décrivent le cercle Ω de centre Ω et de rayon a appelé **cercle principal** de C .

Démonstration. (i) Commençons par le lemme suivant :

Lemme 12.20. Soit une fonction vectorielle $t \mapsto \vec{V}(t)$ non nulle, dérivable, de dérivée $t \mapsto \frac{d\vec{V}(t)}{dt}$. Alors la fonction numérique $t \mapsto \|\vec{V}(t)\|$ est dérivable de dérivée égale à $\frac{d\vec{V}(t)}{dt} \cdot \vec{U}(t)$ où $\vec{U}(t)$ est le vecteur unitaire colinéaire à $\vec{V}(t)$ et de même sens.

Soit $f(t) = \|\vec{V}(t)\|$ et $g(t) = f^2(t) = \|\vec{V}(t)\|^2 = \vec{V}(t) \cdot \vec{V}(t)$. On a $\frac{dg(t)}{dt} = 2\vec{V}(t) \cdot \frac{d\vec{V}(t)}{dt}$, d'où l'on déduit le lemme :

$$\frac{df(t)}{dt} = \frac{1}{2\sqrt{g(t)}} \frac{dg(t)}{dt} = \frac{\vec{V}(t)}{\|\vec{V}(t)\|} \cdot \frac{d\vec{V}(t)}{dt} = \vec{U}(t) \cdot \frac{d\vec{V}(t)}{dt}$$

Soit une représentation paramétrique (p. 321) $t \mapsto M(t)$. Alors $\vec{\tau}(t) = \frac{d\overrightarrow{FM}(t)}{dt}$ est un vecteur tangent. On a $\overrightarrow{F'M}(t) = \overrightarrow{F'F} + \overrightarrow{FM}(t)$, d'où $\frac{d\overrightarrow{F'M}(t)}{dt} = \frac{d\overrightarrow{FM}(t)}{dt} = \vec{\tau}(t)$. Soit $\vec{U}(t)$ et $\vec{U}'(t)$ les vecteurs unitaires colinéaires de même sens à $\overrightarrow{FM}(t)$ et $\overrightarrow{F'M}(t)$ (attention, $\vec{U}'(t)$ n'est pas la dérivée de $\vec{U}(t)$!). Le lemme énonce que les fonctions numériques $t \mapsto FM(t)$ et $t \mapsto F'M(t)$ sont dérivables de dérivées $\vec{\tau}(t) \cdot \vec{U}(t)$ et $\vec{\tau}(t) \cdot \vec{U}'(t)$.

(1) Si C est une ellipse, on a $FM(t) + F'M(t) = 2a$. En dérivant, on obtient $\vec{\tau}(t) \cdot (\vec{U}(t) + \vec{U}'(t)) = 0$. D'où $\vec{\tau}(t)$ est orthogonal à $\vec{U}(t) + \vec{U}'(t)$,

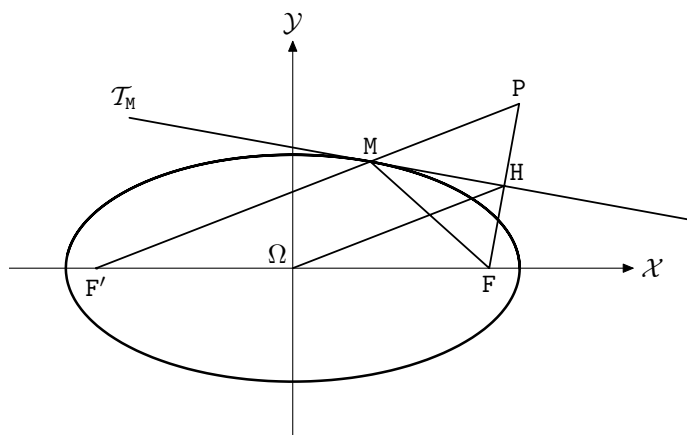
vecteur directeur de la bissectrice intérieure de $(\overrightarrow{FM}(t), \overrightarrow{F'M}(t))$. Donc $\vec{\tau}(t)$ est vecteur directeur de la bissectrice extérieure de cet angle.

(2) Si C est une hyperbole, on a $FM(t) - F'M(t) = \pm 2a$. En dérivant² on obtient $\vec{\tau}(t) \cdot (\vec{U}(t) - \vec{U}'(t)) = 0$. D'où $\vec{\tau}(t)$ est orthogonal à $\vec{U}(t) - \vec{U}'(t)$, vec-

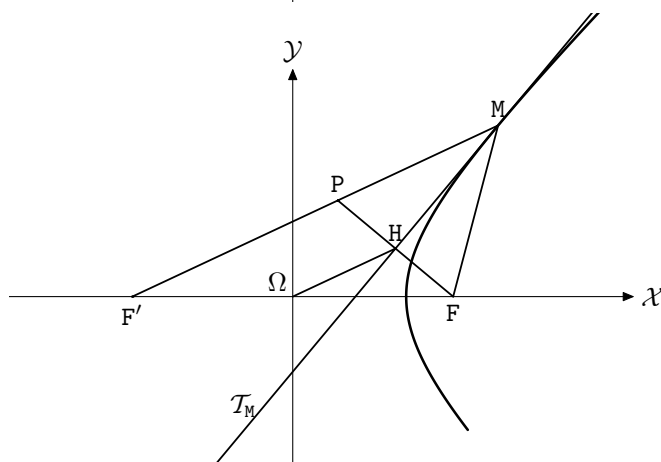
teur directeur de la bissectrice extérieure $(\overrightarrow{FM}(t), \overrightarrow{F'M}(t))$. Donc $\vec{\tau}(t)$ est vecteur directeur de la bissectrice intérieure de cet angle.

2. L'ensemble de définition de cette fonction est union d'intervalles ouverts sur lesquels elle est constante : soit $2a$, soit $-2a$, d'où la dérivabilité sur l'ensemble de définition.

Cas de l'ellipse :



Cas de l'hyperbole :



(ii) La tangente $\mathcal{T}_M$ est donc bissectrice extérieure ou intérieure de $\widehat{(\overrightarrow{FM(t)}, \overrightarrow{F'M(t)})}$ selon que C est ellipse ou hyperbole. La réflexion autour de $\mathcal{T}_M$ transforme F en $P \in (F'M)$.

- (1) Si C est une ellipse, P est sur la demi-droite d'origine M ne contenant pas F' , donc M est entre F' et P . On a $2a = FM + F'M = PM + F'M = PF'$.
- (2) Si C est une hyperbole, P est sur la demi-droite d'origine M contenant F' , donc M n'est pas entre F' et P . On a $2a = |FM - F'M| = |PM - F'M| = PF'$.

Dans les deux cas, P est sur le cercle directeur F' de centre F' et de rayon $2a$. On a ainsi une application $M \mapsto P$ de C vers Γ .

Inversement, soit $P \in F'$. En général, la médiatrice de (F, P) coupe $(F'P)$ en M , unique point, ce qui prouve l'injectivité de l'application $C \rightarrow F'$. Ces droites sont parallèles si (FP) et $(F'P)$ sont orthogonales, i.e. si (FP) est tangente en P à F' .

- (1) Si C est une ellipse, $FF' < FM + F'M = 2a$, donc F est intérieur à F' et ce cas ne se produit pas. L'application $M \mapsto P$ est bijective de C sur F' .

(2) Si C est une hyperbole, F est extérieur à F' . Il existe deux positions P_1 et P_2 ne correspondant à aucun point de C . On démontre que les médiatrices $\mathcal{A}_1$ et $\mathcal{A}_2$ sont les asymptotes de C .

(iii) Ce point se déduit de (ii) au moyen de l'homothétie de centre F et de rapport $\frac{1}{2}$ qui transforme P en H et le cercle directeur F' en le cercle principal Ω . L'assertion relative à la projection orthogonale H' de F' sur $\mathcal{T}_M$ vient de ce que les deux foyers jouent le même rôle. $\square$

Remarque : Dans cette section, l'excentricité a été supposée distincte de 0 et 1. Cependant, les cercles (excentricité 0) entrent dans ce cadre : les foyers sont confondus avec le centre et C coïncide avec son cercle principal. En revanche, le cas des paraboles échappe à ce point de vue.

12.2.5. Propriétés spéciales à la parabole

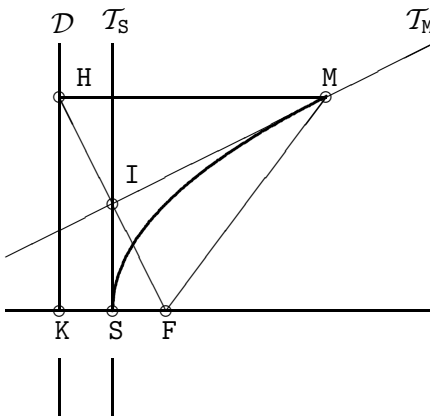
Proposition 12.21. Soit une parabole C de foyer F , de directrice $\mathcal{D}$ et de sommet S . Pour $M \in C$, soit H la projection orthogonale de M sur $\mathcal{D}$ et I le milieu de (F, H) .

(i) La tangente $\mathcal{T}_M$ est la médiatrice de (H, F) .

(ii) Le symétrique de F relativement à $\mathcal{T}_M$ décrit $\mathcal{D}$ si M décrit C .

(iii) La projection orthogonale I de F sur $\mathcal{T}_M$ décrit la tangente au sommet $\mathcal{T}_S$ si M décrit C .

Démonstration. Cette proposition est l'homologue du théorème 12.19 (p. 333). La droite passant par M et orthogonale à $\mathcal{D}$ joue le rôle de (MF') . La directrice $\mathcal{D}$ joue le rôle du cercle directeur de centre F' . La tangente au sommet joue le rôle du cercle principal.



(i) Soit une représentation paramétrique $t \mapsto M(t)$ de C . Dans cette figure, les points M, H, I dépendent de t . Le vecteur $\frac{d\vec{FM}}{dt} = \vec{\tau}$ est tangent. Soit $\vec{j}$ un vecteur unitaire de $\mathcal{D}$ et K la projection orthogonale de F sur $\mathcal{D}$. On a $\vec{HM} = \vec{FM} + \vec{KF} - \vec{KH}$. Comme $\frac{d\vec{KH}}{dt}$ est colinéaire à $\vec{j}$, il existe $\lambda(t)$ tel que $\frac{d\vec{HM}}{dt} = \frac{d\vec{FM}}{dt} - \frac{d\vec{KH}}{dt} = \vec{\tau} - \lambda(t)\vec{j}$. Dérivons $FM^2 - HM^2 = 0$. On a $2\vec{FM} \cdot \frac{d\vec{FM}}{dt} - 2\vec{HM} \cdot \frac{d\vec{HM}}{dt} = 0$, soit $(\vec{FM} - \vec{HM}) \cdot \vec{\tau} = 0$, i.e. $\vec{FH} \cdot \vec{\tau} = 0$. La tangente $\mathcal{T}_M$ est donc hauteur, donc axe de symétrie, du triangle **isocèle** FHM , c'est la médiatrice de (H, F) .

(ii) Par (i), la réflexion d'axe $\mathcal{T}_M$ transforme F en $H \in \mathcal{D}$. Inversement, l'application $M \mapsto H$ étant bijective de C sur $\mathcal{D}$ (12.10, p. 319), H décrit $\mathcal{D}$ si M décrit C .

(iii) Par (i), $\mathcal{T}_S$ est médiatrice de (F, K) donc l'homothétie $H_{F, \frac{1}{2}}$ transforme $\mathcal{D}$ et H en $\mathcal{T}_S$ et I , d'où I décrit $\mathcal{T}_S$ si M décrit C . $\square$

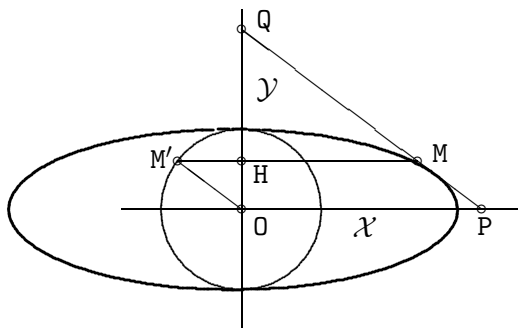
Démonstration. Le cercle de centre O et de rayon a est d'équation $x^2 + y^2 - a^2 = 0$. Un point $M(x, y)$ de A devient par l'affinité d'axe $\mathcal{X}$ et de rapport $\frac{b}{a}$ le point $M'(x', y')$ où $x' = x$ et $y' = \frac{b}{a}y$. On a donc

$$a^2 = x^2 + y^2 = x'^2 + \frac{a^2}{b^2}y'^2 \quad \text{soit} \quad \frac{x'^2}{a^2} + \frac{y'^2}{b^2} - 1 = 0$$

On a donc $f(A) \subset C$. On a de même $f^{-1}(C) \subset A$, d'où $f(A) = C$. On montre de même que $f(B) = C$. $\square$

Proposition 12.25. Bande de papier Soit $\mathcal{X}$ et $\mathcal{Y}$ deux droites se coupant orthogonalement. Les points O , P et Q varient sur $\mathcal{X}$ et $\mathcal{Y}$ de sorte que la longueur $PQ = d$ soit constante. Soit k un réel distinct de 0 et 1. Le point $M \in (PQ)$ tel que $\frac{PM}{PQ} = k$ décrit une ellipse d'axes $\mathcal{X}$ et $\mathcal{Y}$.

Ainsi, P , Q , M sont fixés sur une bande de papier mobile, P et Q étant assujettis à se déplacer sur les droites fixes $\mathcal{X}$ et $\mathcal{Y}$.



Démonstration. Soit H la projection orthogonale de M sur $\mathcal{Y}$ et M' le point où la parallèle à (PQ) passant par O coupe (HM) . La considération du parallélogramme $OPMM'$ donne $\overrightarrow{OM'} = \overrightarrow{PM}$. La longueur PM étant constante, égale à $|k|d$, M' est sur le cercle de centre O et de rayon $|k|d$.

L'homothétie de centre H transformant M en M' transforme Q en O , d'où

$$\frac{\overline{HM}}{\overline{HM'}} = \frac{\overline{QM}}{\overline{OM'}} = \frac{\overline{QM}}{\overline{PM}} = \frac{k-1}{k}$$

L'affinité orthogonale d'axe $\mathcal{Y}$ et de rapport $\frac{k-1}{k}$ transforme le cercle de centre O et de rayon $r = |k|d$ en l'ellipse de centre O , d'axes portés par $\mathcal{Y}$ de longueur $2|k|d$, porté par $\mathcal{X}$ de longueur $|1 - k|d$.

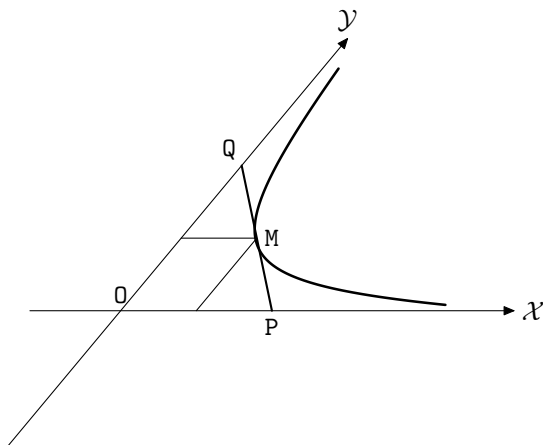
Pour $k = \frac{1}{2}$, M , milieu de $[P, Q]$, décrit le cercle de centre O et de rayon $\frac{1}{2}d$. $\square$

12.2.7. Propriétés spéciales à l'hyperbole

Parmi les coniques propres, la spécificité des hyperboles est d'avoir des asymptotes. La proposition suivante en est une illustration.

Proposition 12.26. Soit $\mathcal{X}$ et $\mathcal{Y}$ deux droites sécantes en O , P et Q deux points décrivant $\mathcal{X}$ et $\mathcal{Y}$ de sorte que l'aire algébrique $\frac{1}{2}[O, P, Q] = A$ soit constante. Alors le milieu M de (P, Q) décrit une hyperbole C et (PQ) est la tangente en M à C .

Démonstration. Soit $\vec{i}$ et $\vec{j}$ des vecteurs directeurs de $\mathcal{X}$ et $\mathcal{Y}$ tels que le déterminant de $\vec{i}, \vec{j}$ relativement à une base orthonormale directe soit 1.



Soit $P \in \mathcal{X}$, $Q \in \mathcal{Y}$. Les coordonnées du milieu M de (P, Q) sont $x = \frac{1}{2}\overline{OP}$ et $y = \frac{1}{2}\overline{OQ}$. On a (9.9, p. 249) $[O, P, Q] = \overline{OP} \times \overline{OQ} = 4xy$ et l'aire algébrique du triangle OPQ est $2xy$. Alors M décrit l'hyperbole C d'équation $2xy = A$ dans le repère $(O, \vec{i}, \vec{j})$. En calculant la dérivée de la fonction $x \mapsto \frac{A}{2x}$, on voit que (PQ) est la tangente en M à C . $\square$

EXERCICES

Exercice 12.7. Soit C et C' deux coniques propres. Montrer qu'il existe une similitude directe transformant C en C' si et seulement si C et C' ont même excentricité. (utiliser l'exercice 8.14 p. 234)

Exercice 12.8. Soit C une parabole, F son foyer, $\mathcal{D}$ sa directrice.

Montrer que par tout $J \in \mathcal{D}$ passent deux tangentes $\mathcal{T}, \mathcal{T}'$ à C et qu'elles sont orthogonales (considérer les réflexions σ et σ' d'axes $\mathcal{T}$ et $\mathcal{T}'$).

Montrer que les points de contact M et M' sont alignés avec F .

Exercice 12.9. Soit C une conique de foyer F , directrice $\mathcal{D}$, excentricité $e > 0$. Montrer que F est intérieur à C et que pour toute sécante passant par F coupant C en M et M' et $\mathcal{D}$ en I , on a une division harmonique $[F, I, M, M'] = -1$, autrement dit la directrice $\mathcal{D}$ est polaire du foyer F (12.3, p. 323).

Exercice 12.10. Soit C une conique de foyer F , de directrice $\mathcal{D}$, d'excentricité e . Soit $M \in C$. On suppose que la tangente $\mathcal{T}_M$ coupe $\mathcal{D}$ en J . Montrer (par dérivation) que $\overrightarrow{FJ}$ et $\overrightarrow{FM}$ sont orthogonaux.

Exercice 12.11. Soit deux droites $\mathcal{D}$ et Δ sécantes en I , A, B deux points distincts de Δ non sur $\mathcal{D}$.

1) Montrer que le lieu géométrique des foyers F des coniques C passant par A, B et de directrice $\mathcal{D}$ est un cercle Γ .

2) Soit A et B les cercles de centre A et B tangents à $\mathcal{D}$. Montrer que A, B, Γ appartiennent à un même faisceau $\mathcal{F}$ de cercles (considérer des fonctions circulaires α, β, γ définissant les cercles A, B, Γ , voir Chap.11).

3) Discuter le genre de C selon la position de F sur Γ

Exercice 12.12. Soit C une conique à centre de foyer F, F' , de cercle principal Ω , $M \in C$, H, H' les projections orthogonales de F, F' sur la tangente $\mathcal{T}_M$. Montrer que pour tout $M \in C$, $\overline{FH} \times \overline{F'H'} = a^2 - c^2$ où a est rayon de Ω et $c = \Omega F$.

Exercice 12.13. Soit C une conique à centre de foyers F, F' , F' le cercle directeur de centre F' , A un point extérieur à C .

1) Construire les tangentes $\mathcal{T}_1, \mathcal{T}_2$ à C issues de A (considérer les intersections P_1, P_2 du cercle A de centre A passant par F et de F' , voir 12.19, p. 333).

2) Montrer que les angles de droites $(\widehat{\mathcal{T}_1, \mathcal{T}_2})$ et $(\widehat{AF, AF'})$ ont mêmes bissectrices (considérer les réflexions $\sigma, \sigma', \sigma_1, \sigma_2$ d'axes $(AF), (AF'), \mathcal{T}_1, \mathcal{T}_2$).

3) Quel est l'ensemble Γ des points A tels que les tangentes $\mathcal{T}_1, \mathcal{T}_2$ issues de A soient orthogonales ?

4) A-t-on des résultats semblables pour les paraboles ?

Exercice 12.14. Soit un triangle ABC . Montrer que les foyers des paraboles tangentes aux trois côtés sont sur le cercle circonscrit au triangle privé des sommets (utiliser le théorème de Simson 11.20, p. 294) et que les directrices passent par l'orthocentre (voir droite de Steiner – problème 11.3, p. 309)

PROBLÈME

Le corrigé de ce problème est disponible sur le site de Dunod : www.dunod.com.

12.1. PROBLÈME

On donne deux points distincts A, A' d'un plan affine $\mathcal{P}$.

1) Soit $\varphi: P \rightarrow P$ une application linéaire, $\mathcal{B}$ une base de P . Montrer que l'application suivante est une f.q.a.

$$F: \mathcal{P} \longrightarrow \mathcal{P}, \quad M \longmapsto F(M) = \det_{\mathcal{B}} \left(\varphi(\overrightarrow{AM}), \overrightarrow{A'M} \right)$$

2) Montrer que $V(F)$ est non vide. En déduire que c'est une conique propre si et seulement si φ est bijective et $\overrightarrow{AA'}$ non vecteur propre de φ .

Montrer que $V(F)$ passe par A et A' et préciser les tangentes en ces points.

3) Soit f une bijection affine $\mathcal{P} \rightarrow \mathcal{P}$ transformant A en A' et ne laissant pas la droite (AA') stable. À une droite $\mathcal{D}$ passant par A , on associe l'intersection M de $\mathcal{D}$ et $\mathcal{D}' = f(\mathcal{D})$. Quel est le lieu de M si $\mathcal{D}$ pivote autour de A ?

4) Soit $\chi_{\varphi}(X)$ le polynôme caractéristique de φ . Montrer que

- $V(F)$ est une ellipse si χ_{φ} n'a pas de racine réelle.
- $V(F)$ est une hyperbole si χ_{φ} a deux racines réelles distinctes et que les directions asymptotiques sont les droites vectorielles propres.
- $V(F)$ est une parabole si χ_{φ} a une racine réelle double.

5) Montrer que l'application $\varphi \mapsto F$ est une bijection linéaire de l'espace $\mathcal{L}(P)$ des bijection linéaires $P \rightarrow P$ sur l'espace des f.q.a. nulles en A, A' .

En déduire que si C est une conique propre passant par A et A' , alors C peut être générée de cette façon.

6) Soit C une conique propre, A, A' deux points distincts, M_1, M_2, M_3, M_4 quatre points distincts de C tels que $(AM_1), (AM_2), (AM_3), (AM_4)$ forment un faisceau harmonique. Montrer que les quatre droites $(A'M_1), (A'M_2), (A'M_3), (A'M_4)$ forment aussi un faisceau harmonique (Si A et M_i sont confondus, on convient que (AM_i) est la tangente en A).

Montrer qu'on peut donner sens à la notion : quatre points d'une conique propre sont en division harmonique. On parle alors de **quadrangle harmonique** de C .

Montrer que A, A', M_1, M_2 sont en division harmonique sur C si et seulement si (M_1M_2) et les tangentes en A et A' concourent ou sont parallèles (utiliser 12.3, p. 323).

7) On suppose $V(F)$ propre. Montrer que le milieu O de (A, A') est centre de $V(F)$ si et seulement si $\varphi^2 = \varphi \circ \varphi$ est une homothétie vectorielle (on montrera que $\overrightarrow{AA'}$ doit être propre pour φ^2 , puis que, n'étant pas propre pour φ , φ^2 doit être une homothétie vectorielle).

8) On suppose $\mathcal{P}$ euclidien. Retrouver le théorème de l'angle inscrit (11.15, p. 292) en prenant pour φ une rotation vectorielle.

9) On suppose $\mathcal{P}$ euclidien. Montrer que $V(F)$ est une hyperbole équilatère de centre le milieu de (A, A') si φ est une réflexion vectorielle. Préciser les asymptotes.

SOLUTIONS DES EXERCICES

Solution 12.1. Pour tout M , $F(M) = Q(\overrightarrow{OM}) + 2L_0(\overrightarrow{OM})$ car $F(O) = 0$.

- 1) Cas où Q est nulle. Alors $V(F) = 0 + \text{Ker } L_0$ est une droite, donc est impropre.
- 2) Cas où L_0 est nulle, alors $F(M) = Q(\overrightarrow{OM})$.
- Si Q est hyperbolique, i.e. de signature $(1, 1)$, alors $Q = L_1 L_2$ (L_1, L_2 formes linéaires) et $V(F) = (0 + \text{Ker } L_1) \cup (0 + \text{Ker } L_2)$ est union de deux droites distinctes.
- Si Q est dégénérée de rang 1, alors $Q = L^2$ (L forme linéaire) et $V(F) = 0 + \text{Ker } L$ est une droite (double).
- Si Q est de signature $(2, 0)$ ou $(0, 2)$, $V(F) = \{0\}$, singleton.

Dans tous ces cas, $V(F)$ est impropre.

3) Cas où $\text{Ker } L_0$ est une droite vectorielle Q -isotrope. Alors $V(F)$ contient la droite $0 + \text{Ker } L_0$, donc est impropre.

4) Reste à montrer que si $\text{Ker } L_0$ est une droite vectorielle non Q -isotrope, alors $V(F)$ est propre. Si $V(F)$ était impropre, elle se réduirait à $\{0\}$ ou se composerait d'une ou deux droites dont l'une passerait par 0 (voir 12.1.1 où tous les cas ont été envisagés). Faisons pivoter autour de 0 une droite $\mathcal{D}_{\vec{u}}$ de vecteur directeur $\vec{u} \neq \vec{0}$. Les points de $\mathcal{D}_{\vec{u}} \cap V(F)$ sont du type $0 + \lambda \vec{u}$, λ étant tel que $P(\lambda) = \lambda^2 Q(\vec{u}) + 2\lambda L_0(\vec{u}) = 0$.

Or $(L_0(\vec{u}) = 0) \Rightarrow (Q(\vec{u}) \neq 0)$, pour aucun $\vec{u}$, $P(\lambda)$ n'est identiquement nul. Donc $V(F)$ ne contient pas de droite passant par 0. Cette équation en λ a une solution non nulle si $\vec{u} \notin \text{Ker } L_0$ donc $V(F)$ ne se réduit pas à $\{0\}$. D'où $V(F)$ est propre.

Solution 12.2. La symétrie oblique d'axe $\mathcal{D}$, de direction $\overline{\Delta}$, laisse P fixe et C stable (12.14, p. 321). Elle échange donc $\mathcal{T}_1, M_1$ et $\mathcal{T}_2, M_2$, d'où $(M_1 M_2)$ est parallèle à Δ .

Le milieu I de (M_1, M_2) est sur $\mathcal{D}$, donc $[\mathcal{T}_1, \mathcal{T}_2, \mathcal{D}, \Delta] = -1$ (8.27, p. 218).

Solution 12.3. 1) Choisissons P pour origine. On écrit pour tout M la formule $F(M) = Q(\overrightarrow{PM}) + 2L_P(\overrightarrow{PM}) + F(P)$. Les points $M' = P + \lambda' \overrightarrow{u}$ et $M'' = P + \lambda'' \overrightarrow{u}$ sont obtenus en prenant pour λ' et λ'' les racines du trinôme

$$\lambda^2 Q(\overrightarrow{u}) + 2\lambda L_P(\overrightarrow{u}) + F(P)$$

Le point $Q = P + \mu \overrightarrow{u}$ est obtenu en écrivant

$$\frac{\lambda'}{\lambda''} = \frac{\overrightarrow{PM'}}{\overrightarrow{PM''}} = -\frac{\overrightarrow{QM'}}{\overrightarrow{QM''}} = -\frac{\lambda' - \mu}{\lambda'' - \mu}$$

$$\text{d'où } \mu = \frac{2\lambda'\lambda''}{\lambda' + \lambda''} = -\frac{F(P)}{L_P(\overrightarrow{u})} \text{ et } \overrightarrow{PQ} = -\frac{F(P)}{L_P(\overrightarrow{u})} \overrightarrow{u}$$

Cette expression ne dépend pas du vecteur $\overrightarrow{u}$ directeur de la sécante. Prenant $\overrightarrow{u} = \overrightarrow{PQ}$, on obtient

$$-\frac{F(P)}{L_P(\overrightarrow{PQ})} = 1 \text{ soit } L_P(\overrightarrow{PQ}) + F(P) = 0$$

Pour deux positions Q_1 et Q_2 de Q, on trouve

$$L_P(\overrightarrow{PQ_1}) = -F(P) = L_P(\overrightarrow{PQ_2}) \text{ soit } L_P(\overrightarrow{Q_1Q_2}) = 0$$

Le lieu est donc porté par une droite Δ_P de direction $\text{Ker } L_P$.

Soit une sécante $\mathcal{D}_1$ coupant C en M'_1, M''_1 , $Q_1 \in \mathcal{D}_1$ tel que $[P, Q_1, M'_1, M''_1] = -1$. La symétrie oblique d'axe le diamètre passant par P et de direction Q -orthogonale laisse C stable et P fixe (12.14, p. 321). Elle transforme $\mathcal{D}_1, Q_1, M'_1, M''_1$ en $\mathcal{D}_2, Q_2, M'_2, M''_2$ avec $[P, Q_2, M'_2, M''_2] = -1$. Donc $\overrightarrow{Q_1Q_2} \in \text{Ker } L_P$ est Q -orthogonal au diamètre passant par P.

Si P est intérieur à C , toute droite passant par P est sécante et le lieu est toute la polaire Δ_P .

Si P est extérieur, on peut mener de P deux tangentes en T_1 et T_2 . Le lieu est le segment $[T_1, T_2]$ si C est une ellipse, une parabole, une hyperbole avec T_1, T_2 sur la même branche, deux demi-droites d'origines T_1, T_2 si C est une hyperbole avec T_1, T_2 sur deux branches différentes.

2) Si $Q \in \Delta_P$ et si (PQ) coupe C en M', M'' , on a $[P, Q, M', M''] = -1$, donc $[Q, P, M', M''] = -1$ et $P \in \Delta_Q$.

Ce raisonnement ne convient pas si (PQ) n'est pas sécante. On a $L_P(\overrightarrow{PQ}) + F(P) = 0$. Montrons que $L_Q(\overrightarrow{QP}) + F(Q) = 0$. On a (12.2, p. 315) $L_Q = L_P + \Phi(\overrightarrow{PQ}, -)$, d'où

$$\begin{aligned} L_Q(\overrightarrow{QP}) + F(Q) &= L_P(\overrightarrow{QP}) + \Phi(\overrightarrow{PQ}, \overrightarrow{QP}) + Q(\overrightarrow{PQ}) + 2L_P(\overrightarrow{PQ}) + F(P) \\ &= -L_P(\overrightarrow{PQ}) - Q(\overrightarrow{PQ}) + Q(\overrightarrow{PQ}) + 2L_P(\overrightarrow{PQ}) + F(P) \\ &= L_P(\overrightarrow{PQ}) + F(P) = 0 \end{aligned}$$

3) Supposons $P \in C$ et disons que Q est conjugué de P si $L_P(\overrightarrow{PQ}) + F(P) = 0$. Comme $F(P) = 0$, Q est conjugué de P si et seulement si $L_P(\overrightarrow{PQ}) = 0$, i.e. $Q \in \Delta_P = P + \text{Ker } L_P$. La polaire de P est alors la tangente T_P à C en P .

Solution 12.4. 1) Les droites $\mathcal{D}_t$ passant par le point $M(x, y)$ sont telles que t soit racine du polynôme

$$\begin{aligned} D_{x,y}(T) &= u(T)x + v(T)y + w(T) \\ &= T^2(u_2 + v_2 + w_2) + 2T(u_1 + v_1 + w_1) + u_0 + v_0 + w_0 \end{aligned}$$

Le discriminant est

$$\Delta(x, y) = (u_1x + v_1y + w_1)^2 - (u_2x + v_2y + w_2)(u_0x + v_0y + w_0)$$

On a donc deux, une ou zéro droites $\mathcal{D}_t$ passant par $M(x, y)$ selon que $\Delta(x, y)$ est positif, nul ou négatif, soit selon la place M relativement à la conique C d'équation $\Delta(x, y) = 0$.

2) Si $M(x, y)$ est point caractéristique de $\mathcal{D}_t$, il vérifie

$$D_{x,y}(t) = u(t)x + v(t)y + w(t) = 0 \text{ et } D'_{x,y}(t) = u'(t)x + v'(t)y + w'(t) = 0$$

donc t est racine commune du polynôme $D_{x,y}(T) = u(T)x + v(T)y + w(T)$ et de sa dérivée $D'_{x,y}(T) = u'(T)x + v'(T)y + w'(T)$, c'est-à-dire que t est racine multiple de $D_{x,y}(T)$. Ceci équivaut à $\Delta(x, y) = 0$.

Il existe bien sûr de nombreux cas singuliers : le polynôme $D_{x,y}(T)$ peut être de degré 1 pour certains points (x, y) ; les droites $\mathcal{D}_t$ et $\mathcal{D}'_t$ peuvent être confondues pour certains t . Nous n'entrerons pas dans cette discussion.

Solution 12.5. On se place dans un repère d'origine O et de base $\mathcal{B} = (\vec{i}, \vec{j})$ où $\vec{i}$ et $\vec{j}$ sont vecteurs directeurs de $\mathcal{X}$ et $\mathcal{Y}$. Les points mobiles sont donnés par leurs coordonnées $P(t)(at + b, 0)$ et $Q(t)(0, ct + d)$. Ils se trouvent en O aux instants $-\frac{b}{a}$ pour P , $-\frac{d}{c}$ pour Q .

La droite $\mathcal{D}_t$ est d'équation $\frac{x}{at+b} + \frac{y}{ct+d} - 1 = 0$, soit

$$act^2 + t(ad + bc - cx - ay) + bd - dx - by = 0$$

Le discriminant $\Delta(x, y)$ de ce polynôme en t est

$$\begin{aligned} &(ad + bc - cx - ay)^2 - 4ac(bd - dx - by) \text{ soit} \\ &(cx + ay)^2 - 2((ad + bc)(cx + ay) + ac(dx + by)) + (ad + bc)^2 - 4abcd \end{aligned}$$

C'est l'expression analytique d'une f.q.a. F

- de forme quadratique $Q = L^2$ où L est la forme linéaire définie par

$$\text{Mat}_{\mathcal{B}}(L) = \begin{pmatrix} c & a \end{pmatrix}$$

- de forme linéaire L_0 avec $\text{Mat}_{\mathcal{B}}(L_0) = - \begin{pmatrix} bc^2 + 2acd & da^2 + 2acb \end{pmatrix}$,
- $F(0) = (ad + bc)^2 - 4abcd$

Le déterminant de L et L_0 est

$$\det \begin{pmatrix} c & bc^2 + 2acd \\ a & da^2 + 2acb \end{pmatrix} = ac(bc - ad)$$

a et c sont non nuls car ce sont les vitesses de $P(t)$ et $Q(t)$ qui ne sont pas fixes.

1) On a $-\frac{b}{a} \neq -\frac{d}{c}$, donc $ad - bc \neq 0$. Les formes linéaires L et L_0 sont indépendantes. L'enveloppe est la parabole C d'équation $\Delta(x, y) = 0$.

Pour $t = -\frac{b}{a}$, $P(t)$ est en O , $\mathcal{D}_t$ est en $\mathcal{Y}$ qui est donc tangente à C . On vérifie que c'est en le point B où se trouve Q à cet instant. De même C est tangente à $\mathcal{X}$ en la position A qu'occupe P à l'instant $-\frac{d}{c}$.

2) Si P, Q sont en O au même instant $t_0 = -\frac{b}{a} = -\frac{d}{c}$, l'équation de $\mathcal{D}_t$ est

$$\frac{x}{a(t-t_0)} + \frac{y}{b(t-t_0)} - 1 = 0 \text{ soit } bx + ay + t_0 - t = 0$$

de vecteur directeur $(-a, b)$, donc de direction fixe.

Solution 12.6. Soit le repère $(O, \vec{u}, \vec{v})$ où O est milieu de (A, B) et $\vec{u}$ colinéaire à (AB) . Soit les coordonnées $(a, 0)$ pour A , $(-a, 0)$ pour B , (a, t) pour P , $(-a, \frac{p}{t})$ pour Q . La droite $\mathcal{D}_t = (PQ)$ est d'équation

$$(x-a)(t - \frac{p}{t}) - (y-t)2a = 0 \text{ soit } t^2(x+a) - 2ayt - p(x-a) = 0$$

Le discriminant de ce polynôme en t est

$$\Delta(x, y) = a^2y^2 + (x^2 - a^2)p = x^2p + y^2a^2 - pa^2 = 0$$

On a une conique C de centre O , tangente en A à $\mathcal{A}$ et en B à $\mathcal{B}$. C'est une ellipse si $p > 0$, une hyperbole si $p < 0$.

Dans le cas $p = -b^2 < 0$ où C est hyperbole, les droites $\mathcal{D}_b$ et $\mathcal{D}_{-b}$ passent par O . On montre que ce sont les asymptotes de C .

Solution 12.7. a) Les similitudes conservant l'orthogonalité et les rapports des distances, on voit facilement que la transformée par une similitude s d'une conique propre C de foyer F , directrice $\mathcal{D}$, excentricité e est la conique propre de foyer $s(F)$, directrice $s(\mathcal{D})$ et de même excentricité e .

Inversement, soit C' une conique d'excentricité e , de foyer F' et de directrice $\mathcal{D}'$. Soit K et K' les projections orthogonales de F et F' sur $\mathcal{D}$ et $\mathcal{D}'$. Il existe une unique similitude directe s transformant (F, K) en (F', K') . Alors $s(\mathcal{D}) = \mathcal{D}'$, donc $s(C) = C'$.

Le rapport de similitude est $\frac{KF'}{KF} = \frac{p'}{p}$ où p' et p sont les paramètres.

b) Soit C un cercle de centre Ω et de rayon p , s une similitude directe de rapport λ . Alors $s(C)$ est le cercle de centre $s(\Omega)$ et de rayon λp .

L'excentricité e détermine donc la **forme** de la conique C . Pour e donné, le paramètre détermine la **taille** de C .

Solution 12.8. Supposons qu'il existe une tangente $\mathcal{T}$ passant par J. Soit M son point de contact et H la projection orthogonale de M sur $\mathcal{D}$. Alors $(JM) = \mathcal{T}$ est la médiatrice de (F, H) , donc $JF = JH$.

Pour mener les tangentes à $\mathcal{C}$ issues de J, on construit le cercle de centre J et de rayon JF. Il coupe $\mathcal{D}$ en H et H', diamétralement opposés. Les médiatrices de (F, H) et (F, H') sont les deux tangentes $\mathcal{T}$ et $\mathcal{T}'$ issues de J.

La réflexion σ d'axe $\mathcal{T}$ transforme H en F. La réflexion σ' d'axe $\mathcal{T}'$ transforme F en H'. Ainsi, $\sigma' \circ \sigma$ est la rotation de centre J transformant H en H'. Comme J est milieu de (H, H') , $\sigma' \circ \sigma$ est la symétrie de centre J, donc $\mathcal{T}$ et $\mathcal{T}'$ sont orthogonales.

La réflexion σ transforme (MH) en (MF) . La réflexion σ' transforme $(M'F)$ en $(M'H')$. La symétrie $S_J = \sigma' \circ \sigma$ transforme la droite (MH) en la droite $(M'H')$, donc les droites $(M'F)$ et (MF) ont même transformée $(M'H')$ par σ' . Elles sont donc confondues, d'où l'alignement de M, F, M'.

Solution 12.9. a) Soit Δ une droite passant par F coupant $\mathcal{D}$ en I, θ l'angle non orienté aigu de Δ et $\mathcal{D}$. Pour tout $M \in \Delta$ de projection orthogonale H sur $\mathcal{D}$, on a $MH = MI \sin \theta$. Donc $M \in \mathcal{C}$ si et seulement si $\frac{MF}{MI} = e \sin \theta$. Si $e \sin \theta \neq 1$ on a deux points M, M' sur Δ de projections orthogonales H, H' sur $\mathcal{D}$ tels que

$$\frac{\overrightarrow{MF}}{\overrightarrow{MI}} = -\frac{\overrightarrow{MF}}{\overrightarrow{MI}} = e \sin \theta \quad \text{d'où} \quad \frac{M'F}{M'H} = \frac{MF}{MH} = e$$

Ainsi, Δ coupe $\mathcal{C}$ en M, M' conjugués harmoniques relativement à F et I.

Si Δ_0 est parallèle à $\mathcal{D}$, soit h la distance de Δ et $\mathcal{D}$. Alors Δ_0 coupe $\mathcal{C}$ en M_0, M'_0 tels que $M_0F = M'_0F = eh$, donc F est le milieu de (M_0, M'_0) .

Solution 12.10. Soit H la projection orthogonale de M sur $\mathcal{D}$. Soit une représentation paramétrique $t \mapsto M(t)$. Le vecteur $\vec{\tau} = \frac{d\vec{FM}}{dt}$ est tangent. Dérivons la relation $FM^2 - e^2HM^2 = 0$. On a

$$0 = \vec{FM} \cdot \frac{d\vec{FM}}{dt} - e^2 \vec{HM} \cdot \left(\frac{d\vec{FM}}{dt} - \frac{d\vec{FH}}{dt} \right) = (\vec{FM} - e^2 \vec{HM}) \cdot \vec{\tau}$$

car $\vec{HM} \cdot \frac{d\vec{FH}}{dt} = 0$. Par définition de J, $\vec{\tau}$ et $\vec{JM}$ sont colinéaires. On a donc

$$0 = (\vec{FM} - e^2 \vec{HM}) \cdot \vec{JM} = \vec{FM} \cdot (\vec{JF} + \vec{FM}) - e^2 \vec{HM} \cdot (\vec{JH} + \vec{HM})$$

Comme $FM^2 - e^2HM^2 = 0$, il reste $0 = \vec{FM} \cdot \vec{JF} - e^2 \vec{HM} \cdot \vec{JH} = \vec{FM} \cdot \vec{JF}$.

Solution 12.11. 1) Soit H et K les projections orthogonales de A et B sur $\mathcal{D}$.

Si F est foyer d'une conique C d'excentricité e , on a $\frac{AF}{AH} = e = \frac{BF}{BK}$, donc (en utilisant l'homothétie de centre I transformant A et H en B et K)

$$\frac{BF}{AF} = \frac{BK}{AH} = \frac{IK}{IH} = \frac{IB}{IA} = \frac{JB}{JA}$$

où J est conjugué harmonique de I relativement à A, B . Donc F est sur le cercle Γ de diamètre $[I, J]$ (11.20, p. 294).

Inversement, si $F \in \Gamma$ et $F \notin \mathcal{D}$, on a ces égalités, donc la conique de foyer F , directrice $\mathcal{D}$, excentricité $\frac{AF}{AH} = e = \frac{BF}{BK}$ passe par A et B .

2) Les cercles A, B sont tangents en H, K à $\mathcal{D}$. Soit les fonctions circulaires α, β, γ définies par

$$\alpha(M) = AM^2 - AH^2, \beta(M) = BM^2 - BK^2, \gamma(M) = AH^2BM^2 - BK^2AM^2$$

Alors $A = V(\alpha), B = V(\beta), \Gamma = V(\gamma)$ et $\gamma = AH^2\beta - BK^2\alpha$. Les fonctions circulaires α, β, γ appartiennent à un même plan vectoriel, donc A, B, Γ appartiennent à un même faisceau de cercles $\mathcal{F}$.

3) Soit $F \in \Gamma$. Comme $\frac{AF}{AH} = e = \frac{BF}{BK}$, F a même position par rapport aux cercles A et B .

(1) Si F extérieur à A et B , $e > 1$, C est une hyperbole,

(2) Si F intérieur à A et B , $e < 1$, C est une ellipse,

(3) Si $F \in A \cap B$, $e = 1$, C est une parabole.

Comme $I \in \mathcal{D}$, tangente commune à A et B , I est extérieur à A, B et l'arc Γ_h des points de Γ extérieurs à A et B est non vide. Si $F \in \Gamma_h \setminus \mathcal{D}$, C est une hyperbole, donc le cas 1 est toujours possible.

Si les cercles A et B n'ont pas de point commun, i.e. si $\mathcal{F}$ est à points limites, C est toujours une hyperbole car $\Gamma = \Gamma_h$. C'est en particulier le cas si A et B ne sont pas dans le même demi-plan limité par $\mathcal{D}$.

Si les cercles A et B se coupent en F_1, F_2 , $\mathcal{F}$ est à points de base F_1, F_2 . Les arcs de Γ limités par F_1, F_2 sont Γ_e et Γ_h , intérieur et extérieur à A et B .

(1) Si $F \in \Gamma_h$, C est une hyperbole,

(2) Si $F \in \Gamma_e$, C est une ellipse,

(3) Si $F \in \{F_1, F_2\}$, C est une parabole.

Solution 12.12. On sait (12.19, p. 333) que H, H' sont sur le cercle principal Ω . Soit K le point où (FH) recoupe Ω , symétrique de H' par rapport au centre Ω . On a donc $\overline{FH} \times \overline{F'H'} = -\overline{FH} \times \overline{FK} = -\Omega(F) = a^2 - c^2$.

Solution 12.13. 1) Les symétriques P_1, P_2 de F par rapport à $\mathcal{T}_1, \mathcal{T}_2$ sont sur F' (12.19, p. 333). Ce sont donc les intersections des cercles F' et A .

2) On a $\sigma_1(F) = P_1, \sigma'(P_1) = P_2, \sigma_2(P_2) = F$. La réflexion $\sigma_2 \circ \sigma' \circ \sigma_1$ laisse F fixe, donc $\sigma_2 \circ \sigma' \circ \sigma_1 = \sigma$ et $\sigma_2 \circ \sigma' = \sigma \circ \sigma_1$ est la rotation de centre A et d'angle $2(\widehat{AF', \mathcal{T}_2}) \equiv 2(\widehat{\mathcal{T}_1, AF})$ modulo 2π . Soit τ une réflexion échangeant (AF) et (AF') . Elle change les angles en leurs opposés, donc échange $\mathcal{T}_1$ et $\mathcal{T}_2$, d'où le résultat.

3) La rotation $\sigma_2 \circ \sigma_1$, de centre A , d'angle $2(\widehat{\mathcal{T}_1, \mathcal{T}_2})$ (modulo 2π), envoie P_1 en P_2 . Les tangentes $\mathcal{T}_1, \mathcal{T}_2$ sont orthogonales si et seulement si (P_1, P_2) sont diamétralement opposés sur A , i.e. si A est sur l'axe radical de A et F' , i.e. si

$$-AF^2 = A(A) = \overline{AP_1} \times \overline{AP_2} = F'(A) = AF'^2 - 4a^2$$

i.e. si $AF^2 + AF'^2 = 4a^2$. Posons $2c = FF'$. Soit Ω le centre de C . On a

$$\begin{aligned} AF^2 + AF'^2 &= \|\vec{A\Omega} + \vec{\Omega F}\|^2 + \|\vec{A\Omega} + \vec{\Omega F'}\|^2 \\ &= 2A\Omega^2 + 2\vec{A\Omega} \cdot (\vec{\Omega F} + \vec{\Omega F'}) + \Omega F^2 + \Omega F'^2 = 2A\Omega^2 + 2c^2 \end{aligned}$$

Par un calcul simple (voir 12.17, 328), on montre que $e = \frac{c}{a}$. D'où $A \in \Gamma$ si et seulement si $A\Omega^2 = 2a^2 - c^2 = (2 - e^2)a^2$.

- si $e < \sqrt{2}$ ce qui est le cas des ellipses et hyperboles de faible excentricité, Γ est le **cercle orthoptique** de centre Ω et de rayon $a\sqrt{2 - e^2}$.
- Si $e > \sqrt{2}$, Γ est vide.
- Si $e = \sqrt{2}$, C est une **hyperbole équilatère** i.e. d'asymptotes $\mathcal{X}, \mathcal{Y}$ orthogonales. En les considérant comme tangentes avec points de contact à l'infini, Γ se réduit à $\{\Omega\}$.

4) Pour étendre ces résultats à la parabole, il faut remplacer F' par la directrice $\mathcal{D}$, la droite (AF') étant remplacée par la droite orthogonale à $\mathcal{D}$ passant par A . L'ensemble Γ est alors la directrice (exercice 12.8, p. 338).

Solution 12.14. Soit C une telle parabole, F son foyer, $\mathcal{D}$ sa directrice. Les projections orthogonales A', B', C' de F sur les tangentes $(BC), (CA), (AB)$ sont sur la tangente au sommet Δ de C (12.21, p. 335). Par le théorème de Simson (11.18, p. 293), F est sur le cercle (ABC) et Δ est la droite de Simson de F .

On sait que l'homothétie $H_{F, \frac{1}{2}}$ transforme la droite de Simson Δ de F en la droite de Steiner qui passe par l'orthocentre (voir le problème 11.6.3, p. 309).

Inversement, soit F sur le cercle (ABC) distinct de A, B, C et Δ sa droite de Simson. Alors la parabole de foyer F et de tangente au sommet Δ est bien tangente aux côtés du triangle (12.21, p. 335). Même raisonnement pour la directrice.

Chapitre 13

Nombres complexes et géométrie

Le corps $\mathbb{C}$ des nombres complexes a une structure de plan vectoriel euclidien orienté. Ceci permet de transporter dans $\mathbb{C}$ les problèmes de géométrie plane pour les résoudre par calcul. Pour ce faire, il faut disposer d'un dictionnaire mettant en regard les phénomènes géométriques et leurs traductions algébriques dans $\mathbb{C}$. Cette problématique est développée dans les sections 13.1 et 13.2.

L'adjonction d'un élément ∞ à $\mathbb{C}$ permet une application à la géométrie des cercles développée au chapitre 11. Cette partie est longuement développée dans les sections 13.3 et 13.4. C'est l'occasion de découvrir de nouvelles transformations qui ne font pas partie de la géométrie affine : les homographies et antihomographies qui constituent le groupe circulaire.

13.1 LE CORPS $\mathbb{C}$ COMME PLAN GÉOMÉTRIQUE

13.1.1. Le plan vectoriel euclidien orienté $\mathbb{C}$

Rappelons comment on définit sur $\mathbb{C}$ la structure canonique de **plan vectoriel euclidien orienté** :

- l'inclusion $\mathbb{R} \subset \mathbb{C}$ détermine sur $\mathbb{C}$ une structure naturelle de $\mathbb{R}$ -espace vectoriel de dimension 2,

- l'application **module** $Z \mapsto |Z|$ est la norme euclidienne ; le produit scalaire de $Z = X + iY$ et $Z' = X' + iY'$ est

$$XX' + YY' = \Re(\overline{Z}Z') = \frac{1}{2}(\overline{Z}Z' + \overline{Z'}Z)$$

- la base orthonormale $(1, i)$ est déclarée **directe** ; le déterminant de $Z = X + iY$ et $Z' = X' + iY'$ relativement à $(1, i)$ (et donc relativement à toute base orthonormale directe) est

$$\det \begin{pmatrix} X & X' \\ Y & Y' \end{pmatrix} = XY' - YX' = \Im(\overline{Z}Z') = \frac{1}{2i}(\overline{Z}Z' - \overline{Z'}Z)$$

L'**argument** d'un complexe $Z \neq 0$ est l'angle $\arg(Z) \equiv \widehat{(1, Z)}$ modulo 2π . Le groupe $\mathbf{SO}(\mathbb{C})$ est formé des applications $Z \mapsto aZ$ où $|a| = 1$. Il est canoniquement isomorphe au groupe multiplicatif $\mathbb{U}$ des nombres complexes de module 1. Autrement dit, a étant de la forme $e^{i\theta}$, l'application $\mathbb{R}$ -linéaire $Z \mapsto e^{i\theta}Z$ est la rotation vectorielle d'angle θ .

Le groupe, isomorphe à $\mathbb{C}^\times$, des applications $Z \mapsto aZ$ ($a \neq 0$) est le groupe des similitudes vectorielles directes. Pour $a = \rho e^{i\theta}$, l'application $Z \mapsto aZ$ est la similitude vectorielle directe de rapport ρ et d'angle θ .

13.1.2. Le plan affine euclidien orienté $\mathbb{C}$

On peut aussi considérer $\mathbb{C}$ comme plan **affine** euclidien orienté. Reportons nous à la définition d'espace affine (8.1, p. 204) : le plan vectoriel (euclidien orienté) associé est $\mathbb{C}$ lui-même et l'opération du plan vectoriel $\mathbb{C}$ sur le plan affine $\mathbb{C}$ est l'addition interne $(b, z) \mapsto z + b$ de $\mathbb{C}$. Autrement dit, les translations sont les applications $z \mapsto z + b$ où z est considéré comme un point et b comme un vecteur.

Proposition 13.1. *Une application de $\mathbb{C}$ dans $\mathbb{C}$ est une similitude directe pour la structure affine euclidienne de $\mathbb{C}$ si et seulement si elle est du type $z \mapsto az + b$ où $a \neq 0$.*

Autrement dit, le groupe $\text{Sim}^+(\mathbb{C})$ des similitudes directes s'identifie au groupe affine $\mathbf{GA}(\mathbb{C})$ de $\mathbb{C}$ considéré comme droite affine sur le corps $\mathbb{C}$.

Démonstration. Montrons que ces applications sont des similitudes directes.

- Pour $a = 1$, on a toutes les translations.
- Pour $a = \rho e^{i\theta} \neq 1$, l'application a un unique point fixe obtenu en résolvant l'équation $z_0 = az_0 + b$, c'est le point d'axe $z_0 = \frac{b}{1-a}$. Soit $z \neq z_0$ et $z' = az + b$ son image. Considérons les "vecteurs" $Z = z - z_0$ et $Z' = z' - z_0$. On a $Z' = (az + b) - (az_0 + b) = a(z - z_0) = aZ$, donc $|Z'| = \rho|Z|$ et $\widehat{(Z, Z')} \equiv \arg(a) \equiv \theta$ modulo 2π . L'application est bien la similitude directe de centre z_0 , de rapport ρ et d'angle θ .

Comme ρ et θ peuvent être pris quelconques, de même que z_0 (en prenant $b = (1 - a)z_0$), toute similitude directe est de ce type. $\square$

13.1.3. Le plan affine $\mathbb{C}$

On va voir maintenant $\mathbb{C}$ comme simple plan affine, autrement dit on oublie la structure euclidienne. On se propose de décrire les applications affines par un énoncé analogue à la proposition 13.1

Proposition 13.2. *Une application $\mathbb{C} \rightarrow \mathbb{C}$ est une application affine si et seulement si elle est du type $z \mapsto az + b\bar{z} + c$.*

Démonstration. Ces applications sont clairement affines. La question est : toute application affine est-elle de ce type ? Cela revient à montrer le lemme suivant :

Lemme 13.3. *Une application de $\mathbb{C}$ dans $\mathbb{C}$ est $\mathbb{R}$ -linéaire si et seulement si elle est du type $f_{a,b} : Z \mapsto aZ + b\bar{Z}$ où $(a, b) \in \mathbb{C}^2$.*

Ces applications sont clairement $\mathbb{R}$ -linéaires et forment un sous- $\mathbb{R}$ -espace vectoriel de l'espace $\mathcal{L}_{\mathbb{R}}(\mathbb{C})$. Comme $\dim_{\mathbb{R}}(\mathbb{C}) = 2$, on a $\dim_{\mathbb{R}} \mathcal{L}_{\mathbb{R}}(\mathbb{C}) = 2 \times 2 = 4$. L'application $(a, b) \mapsto f_{a,b}$ est injective, $\mathbb{R}$ -linéaire de $\mathbb{C}^2$ dans $\mathcal{L}_{\mathbb{R}}(\mathbb{C})$. Mais $\dim_{\mathbb{R}}(\mathbb{C}^2) = 4$, cette application est donc bijective par égalité des dimensions des espaces de départ et d'arrivée, le lemme est démontré.

Soit $f : \mathbb{C} \rightarrow \mathbb{C}$ une application affine : f est déterminée par son application linéaire associée, qui est de la forme $f_{a,b}$ par le lemme, et par l'image $c = f(0)$ (8.13, p. 209). Elle est donc bien du type annoncé. $\square$

13.2 UTILISATION DE $\mathbb{C}$ EN GÉOMÉTRIE AFFINE PLANE

Soit $\mathcal{P}$ un plan affine euclidien orienté de plan vectoriel associé P . Nous allons voir comment la donnée d'un repère orthonormé direct $R = (0, \vec{u}, \vec{v})$ permet d'« algébriser » une situation géométrique.

13.2.1. Affixes et images

Si on considère les bijections

$$\begin{aligned} \mathbb{C} &\longrightarrow P \\ Z = X + iY &\longmapsto \vec{V} = X\vec{u} + Y\vec{v} \\ \mathbb{C} &\longrightarrow \mathcal{P} \\ z = x + iy &\longmapsto M = 0 + x\vec{u} + y\vec{v} \end{aligned}$$

On dit que $\vec{V}$ et M sont les **images** de Z et z respectivement et que Z et z sont les **affixes** de $\vec{V}$ et M . Ces notions sont relatives au repère R . Il n'y a pas d'ambiguïté sur les notions d'affixes. En revanche la notion d'**image** est plus ambiguë et on devrait préciser **image vectorielle** pour $\vec{V}$ et **image ponctuelle** pour M , ce que nous ferons s'il y a risque de confusion.

Une situation géométrique de $\mathcal{P}$ ou de P peut être ainsi traduite dans $\mathbb{C}$ et étudiée par calcul. On a déjà rencontré cette démarche. Le choix d'une base en algèbre linéaire ou d'un repère en géométrie affine permet de transporter un problème géométrique dans $\mathbb{R}^n$ pour l'étudier numériquement.

Ici, l'outil numérique est sophistiqué car la structure algébrique de $\mathbb{C}$ est plus riche que celle de $\mathbb{R}^2$: on y dispose de la **multiplication interne** du corps $\mathbb{C}$. Les deux coordonnées réelles d'un point sont remplacées par *un seul* nombre complexe, l'affixe, ce qui laisse prévoir une simplification des calculs. Un objectif de ce chapitre est d'élaborer un « dictionnaire » mettant en correspondance phénomènes géométriques de $\mathcal{P}$ et leurs **expressions complexes** dans $\mathbb{C}$. Ainsi, les propositions 13.1 et 13.2 donnent les expressions complexes des similitudes directes et des applications affines quelconques.

13.2.2. Équations complexes de droites

Lemme 13.4. (i) Toute forme linéaire $L: P \rightarrow \mathbb{R}$ est d'expression complexe $Z \mapsto \bar{a}Z + a\bar{Z}$ où $a \in \mathbb{C}$ est unique, affixe d'un vecteur orthogonal à $\text{Ker } L$.

(ii) Toute fonction affine $f: \mathcal{P} \rightarrow \mathbb{R}$ est d'expression complexe $z \mapsto \bar{a}z + a\bar{z} + b$ où $(a, b) \in \mathbb{C} \times \mathbb{R}$ est unique.

Démonstration. (i) On sait que L est du type $\vec{V} \mapsto \vec{A} \cdot \vec{V}$ où $\vec{A}$ est unique orthogonal à $\text{Ker } L$. Pour $\vec{V} = X\vec{u} + Y\vec{v}$ et $\vec{A} = \alpha\vec{u} + \beta\vec{v}$, on a (13.1.1., p. 349)

$$\vec{A} \cdot \vec{V} = X\alpha + Y\beta = \bar{a}Z + a\bar{Z}$$

où $Z = X + iY$ est affixe de $\vec{V}$ et $a = \frac{1}{2}(\alpha + i\beta)$ est affixe de $\frac{1}{2}\vec{A}$.

(ii) Une fonction affine (9.12, p. 250) f est définie par la donnée de la forme linéaire associée L et l'image de l'origine $f(0) = b \in \mathbb{R}$:

$$f(M) = L(\vec{OM}) + f(0) = \bar{a}z + a\bar{z} + b$$

où z est affixe de M . □

Proposition 13.5. Toute droite affine $\mathcal{D}$ a une équation complexe du type $\bar{a}z + a\bar{z} + b = 0$ où $(a, b) \in \mathbb{C}^* \times \mathbb{R}$ est unique à facteur réel près ; a est l'affixe d'un vecteur orthogonal. Inversement, une telle équation est celle d'une droite.

Démonstration. On sait (9.13, p. 251) qu'une droite est l'ensemble $\mathcal{D} = V(f)$ des points annulant une fonction affine f , unique à facteur (réel) près, et réciproquement. □

13.2.3. Équations complexes de cercles

Proposition 13.6. L'équation $az\bar{z} + \bar{b}z + b\bar{z} + c = 0$, où $(a, c) \in \mathbb{R}^* \times \mathbb{R}$ et $b \in \mathbb{C}$ est celle d'un cercle si $|b|^2 - ac \geq 0$. Tout cercle a une telle équation, le triplet (a, b, c) est unique à facteur réel près.

Démonstration. Soit L la forme linéaire d'expression complexe $Z \mapsto \bar{b}Z + b\bar{Z}$. Considérons l'application

$$F: M \mapsto F(M) = a\|\vec{OM}\|^2 + L(\vec{OM}) + c = az\bar{z} + \bar{b}z + b\bar{z} + c$$

où z est affixe de M . C'est une fonction circulaire (11.12, p. 289). Voyons à quelle condition $V(F)$ est non vide. On écrit

$$\begin{aligned} F(M) &= a\left(z\bar{z} + \frac{\bar{b}}{a}z + \frac{b}{a}\bar{z} + \frac{c}{a}\right) = a\left(z + \frac{b}{a}\right)\left(\bar{z} + \frac{\bar{b}}{a}\right) + \frac{ac - |b|^2}{a^2} \\ &= a\left(\left|z + \frac{b}{a}\right|^2 - \frac{|b|^2 - ac}{a^2}\right) \end{aligned}$$

Si $|b|^2 - ac < 0$, $V(F)$ est vide.

Si $|b|^2 - ac \geq 0$, $V(F)$ est le cercle de centre Ω d'affixe $-\frac{b}{a}$ et de rayon $\frac{1}{|a|}\sqrt{|b|^2 - ac}$, réduit à son centre si $|b|^2 - ac = 0$.

Inversement, tout cercle donné par son centre et son rayon admet une telle équation. $\square$

Exercice 13.1. Pantographe. Soit un parallélogramme articulé $OACB$ ($\vec{OA} + \vec{OB} = \vec{OC}$) où O est fixe, les longueurs $OA = BC$ et $OB = AC$ sont constantes. Deux plaques triangulaires rigides MAC et CBM' , directement semblables, sont fixées sur les segments $[A, C]$ et $[C, B]$. Montrer que l'on passe de M à M' par une similitude fixe directe de centre O .

Exercice 13.2. 1) Soit A, B, C d'affixes a, b, c . On note j le nombre $e^{\frac{2i\pi}{3}}$. Montrer les équivalences :

$$\begin{aligned} (\text{ABC équilatéral direct}) &\Leftrightarrow (a + bj + cj^2 = 0) \\ (\text{ABC équilatéral}) &\Leftrightarrow ((a - b)^2 + (b - c)^2 + (c - a)^2 = 0) \end{aligned}$$

2) Soit un polynôme $P(X) = X^3 + \lambda X^2 + \mu X + \nu$ à coefficients complexes. À quelle condition sur λ, μ, ν les racines de $P(X)$ sont-elles sommets d'un triangle équilatéral de $\mathbb{C}$?

3) Soit $T = PQR$ un triangle quelconque. On construit à l'extérieur de T les triangles équilatéraux de côtés $[Q, R]$, $[R, P]$, $[P, Q]$. Soit A, B, C les centres de ces triangles. Montrer que le triangle ABC est équilatéral (théorème de Napoléon).

Exercice 13.3. Soit M_1, M_2 d'affixes z_1, z_2 distincts. Donner une équation complexe de la droite (M_1M_2) . On la mettra sous la forme de la nullité d'un déterminant d'ordre 3 (l'équation étant $\bar{a}z + a\bar{z} + b = 0$, exprimer que z_1, z_2 la vérifient ; considérer un autre point d'affixe z sur la droite et utiliser le fait qu'un système linéaire homogène ayant une solution non triviale est de déterminant nul).

Exercice 13.4. Soit un triangle direct ABC avec les affixes a, b, c . On construit les triangles équilatéraux PBC, QCA, RAB, extérieurement au triangle ABC. Soit p, q, r les affixes de P, Q, R.

1) Montrer que les vecteurs $\overrightarrow{AP}, \overrightarrow{BQ}, \overrightarrow{CR}$ ont même norme et font entre eux les angles $\pm \frac{2\pi}{3}$.

2) Montrer que les droites (AP), (BQ), (CR) sont concourantes.

Exercice 13.5. On donne trois points M_1, M_2, M_3 . Par la même méthode que dans l'exercice 13.3, donner l'équation complexe du cercle passant par ces trois points (s'ils ne sont pas alignés).

Exercice 13.6. Soit ABC un triangle équilatéral.

1) Montrer que pour tout M, on a $MA + MB \geq MC$.

2) Montrer que $MA + MB = MC$ si et seulement si M est sur l'arc du cercle (ABC) limité par A et B, ne contenant pas C (on réfléchira à la question suivante : à quelle condition sur les arguments de z_1 et z_2 l'inégalité $|z_1 + z_2| \leq |z_1| + |z_2|$ est-elle une égalité ?).

Exercice 13.7. Soit $\mathcal{E}$ un espace euclidien orienté de dimension 3, $(0, \vec{e}_1, \vec{e}_2, \vec{e}_3)$ un repère orthonormé direct, A, B, C trois points tels que $\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC}$ soient deux à deux orthogonaux et de même norme. Autrement dit, ce sont trois arêtes d'un cube de sommet O. Soit $\mathcal{P}$ le plan passant par O de direction $P = \text{Vect}(\vec{e}_1, \vec{e}_2)$.

1) Soit a, b, c les affixes des projections orthogonales A', B', C' de A, B, C sur $\mathcal{P}$. Montrer que $a^2 + b^2 + c^2 = 0$ (considérer la matrice de $\overrightarrow{OA}, \overrightarrow{OB}, \overrightarrow{OC}$ dans la base $\mathcal{B} = (\vec{e}_1, \vec{e}_2, \vec{e}_3)$ et utiliser les propriétés des matrices orthogonales).

2) Montrer que si O est isobarycentre de A', B', C' , alors $A'B'C'$ est équilatéral (utiliser l'exercice 13.2).

3) Comment envisager la réciproque de 1) ?

Exercice 13.8. Soit M', M'' deux points mobiles animés de mouvements circulaires uniformes de même vitesse angulaire ω sur des cercles C', C'' de centres A', A'' , de rayons R', R'' . Prenons la droite des centres pour axe réel. Si a', a'' sont les affixes (réels) des centres, les affixes de M', M'' sont donc

$$z' = a' + R' e^{i(\omega t + \varphi')}, \quad z'' = a'' + R'' e^{i(\omega t + \varphi'')}$$

1) Quel est le mouvement du milieu M de (M', M'') ?

2) À quelle condition M est-il immobile ?

Exercice 13.9. Soit $\varphi: P \rightarrow P$ une application linéaire de représentation complexe $Z \mapsto aZ + b\bar{Z}$.

1) Montrer que φ est inversible si et seulement si $|a| \neq |b|$.

2) Dans le cas général, calculer $\det(\varphi)$.

3) Quel est le polynôme caractéristique de φ ? Montrer que φ a des valeurs propres réelles si et seulement si l'image A de a se trouve dans la bande comprise entre les deux droites horizontales tangentes au cercle de centre l'origine O et de rayon $|b|$.

Exercice 13.10. Soit M', M'' deux points mobiles animés de mouvements circulaires uniformes de vitesses angulaires opposées ω et $-\omega$ sur des cercles C', C'' de centres A', A'' , de rayons R', R'' . Prenons la droite des centres pour axe réel. Si a', a'' sont les affixes (réels) des centres, les affixes de M', M'' sont

$$z' = a' + R'e^{i(\omega t + \varphi')}, \quad z'' = a'' + R''e^{i(-\omega t + \varphi'')}$$

Montrer que la trajectoire du milieu M de (M', M'') est en général une ellipse dont on déterminera les éléments de réduction. La trajectoire de M peut-elle être un segment? (On pourra poser $u(t) = e^{i\omega t}$ et étudier l'application faisant passer de $u(t)$ à l'affixe $z(t)$ de M).

13.3 LA GÉOMÉTRIE DES CERCLES ET $\mathbb{C}$

Cette section est liée au chapitre 11. Le cadre est un plan affine euclidien orienté $\mathcal{P}$, de plan vectoriel associé P . On le complétera en le **plan annalagmatique** $\tilde{\mathcal{P}} = \mathcal{P} \cup \{\infty\}$ en ajoutant le **point à l'infini** ∞ (11.22, p. 296).

13.3.1. Birapport de quatre points

Définition 13.7. Soit z_1, z_2, z_3, z_4 quatre nombres complexes distincts. On appelle **birapport** de ces nombres rangés dans cet ordre le nombre

$$[z_1, z_2, z_3, z_4] = \frac{z_3 - z_1}{z_4 - z_1} : \frac{z_3 - z_2}{z_4 - z_2} \quad (\text{où le signe : signifie « divisé par »})$$

Ces nombres étant distincts, le birapport n'est jamais égal à 0 ou 1. On vérifie immédiatement que

$$[z_1, z_2, z_3, z_4] = [z_2, z_1, z_4, z_3] = [z_3, z_4, z_1, z_2] = [z_4, z_3, z_2, z_1]$$

Le birapport ne change pas en intervertissant les couples (z_1, z_2) et (z_3, z_4) ou simultanément les termes de chaque couples. Comme on a $4! = 24$ ordonnancements possibles des z_i , les 24 valeurs sont égales 4 à 4 et on a $\frac{24}{4} = 6$ valeurs possibles du birapport en changeant l'ordre. Par exemple, en intervertissant z_1 et z_2 , ou z_3 et z_4 , le birapport est changé en son inverse.

On notera $\overline{\mathbb{C}}$ l'ensemble $\mathbb{C} \cup \{\infty\}$. L'élément infini ∞ est l'abscisse du point à l'infini (noté aussi ∞) du plan annalagmatique $\tilde{\mathcal{P}} = \mathcal{P} \cup \{\infty\}$. Prolongeons la définition du birapport à $\overline{\mathbb{C}}$ en posant

$$[\infty, z_2, z_3, z_4] = [z_2, \infty, z_4, z_3] = [z_3, z_4, \infty, z_2] = [z_4, z_3, z_2, \infty] = \frac{z_4 - z_2}{z_3 - z_2}$$

Ces valeurs sont les limites du birapport si z_1 tend vers l'infini alors que les trois autres abscisses z_i sont fixes.

Proposition 13.8. Soit quatre points M_1, M_2, M_3, M_4 . Le birapport de leurs abscisses z_1, z_2, z_3, z_4 relativement à un repère orthonormé direct ne dépend pas de ce repère. On pose alors $[M_1, M_2, M_3, M_4] = [z_1, z_2, z_3, z_4]$, c'est le **birapport** des quatre points M_i .

Démonstration. Soit $R = (O, \vec{u}, \vec{v})$ et $R' = (O', \vec{u}', \vec{v}')$ deux repères, b l'abscisse de O' dans R , $\theta \equiv (\vec{u}, \vec{u}') \equiv (\vec{v}, \vec{v}')$ modulo 2π . Si z et z' sont les abscisses du même point M dans R et R' , la relation¹ liant z et z' est $z = e^{i\theta} z' + b$. Pour trois indices i, j, k , on voit que $\frac{z'_i - z'_j}{z'_i - z'_k} = \frac{z_i - z_j}{z_i - z_k}$, donc on a égalité des birapports $[z_1, z_2, z_3, z_4] = [z'_1, z'_2, z'_3, z'_4]$, y compris si l'un des z_i est ∞ . $\square$

Proposition 13.9. Le birapport de quatre points alignés est leur birapport au sens du chapitre 8.

Démonstration. Si les M_i sont sur une même droite $\mathcal{D}$, prenons $O \in \mathcal{D}$ et $\vec{u}$ un vecteur unitaire directeur de $\mathcal{D}$. L'abscisse de M_i est $z_i = \overline{OM_i}$ et on a

$$[z_1, z_2, z_3, z_4] = \frac{z_3 - z_1}{z_4 - z_1} : \frac{z_3 - z_2}{z_4 - z_2} = \frac{\overline{M_1 M_3}}{\overline{M_1 M_4}} : \frac{\overline{M_2 M_3}}{\overline{M_2 M_4}}$$

ce qui est le birapport au sens du chapitre 8 (8.1.8., p. 217). $\square$

1. La même démonstration marche encore si on change l'unité de longueur, c'est-à-dire si on remplace le produit scalaire de P par un autre proportionnel (8.32, p. 223). La formule de changement de repère est alors $z = az' + b$ où $a \neq 0$ n'est pas toujours de module 1.

13.3.2. Faisceaux de cercles à points de base

Théorème 13.10. *Pour que quatre points distincts M_1, M_2, M_3, M_4 soient cocycliques au sens généralisé, il faut et il suffit que leur birapport $[M_1, M_2, M_3, M_4]$ soit réel.*

Voir le corollaire du théorème de l'angle inscrit (p. 293).

Démonstration. Soit z_1, z_2, z_3, z_4 les affixes relativement à un repère orthonormé direct et $\rho = [M_1, M_2, M_3, M_4] = [z_1, z_2, z_3, z_4]$.

1) Supposons qu'aucun de ces points n'est ∞ . On a

$$\rho = \frac{z_3 - z_1}{z_4 - z_1} : \frac{z_3 - z_2}{z_4 - z_2}$$

d'où les congruences modulo 2π

$$\arg(\rho) \equiv \arg\left(\frac{z_3 - z_1}{z_4 - z_1}\right) - \arg\left(\frac{z_3 - z_2}{z_4 - z_2}\right) = (\widehat{M_1M_4, M_1M_3}) - (\widehat{M_2M_4, M_2M_3})$$

Pour que ρ soit réel, il faut et il suffit que son argument soit nul modulo π , donc qu'on ait $(\widehat{M_1M_4, M_1M_3}) \equiv (\widehat{M_2M_4, M_2M_3})$ modulo π . Comme cette congruence est modulo π , on peut s'exprimer en termes d'angles de droites :

$$(\widehat{M_1M_4, M_1M_3}) \equiv (\widehat{M_2M_4, M_2M_3}) \text{ modulo } \pi$$

ce qui équivaut à ce que les quatre points soient cocycliques ou alignés (p. 293), donc appartiennent à un même cercle généralisé.

2) Supposons qu'un de ces points soit ∞ , par exemple $M_4 = \infty$. On a

$$\rho = [z_1, z_2, z_3, \infty] = \frac{z_3 - z_1}{z_3 - z_2}$$

d'où les congruences modulo 2π

$$\arg(\rho) \equiv \arg\left(\frac{z_3 - z_1}{z_3 - z_2}\right) \equiv (\widehat{M_2M_3, M_1M_3})$$

Pour que ρ soit réel, il faut et il suffit que son argument soit nul modulo π , donc que $(\widehat{M_2M_3, M_1M_3}) \equiv 0$, c'est-à-dire que les trois points M_1, M_2, M_3 soient alignés. Ceci équivaut à ce que M_1, M_2, M_3, ∞ soient cocycliques au sens généralisé car les droites sont les cercles généralisés passant par ∞ . $\square$

Corollaire Soit A, B deux points distincts du plan annalagmatique et $\mathcal{F}$ le faisceau de cercles à points de base A, B . Pour tout M_0 distinct de A, B , le cercle de $\mathcal{F}$ passant par M_0 est :

$$C_0 = \left\{ M \in \check{\mathcal{P}} \mid [A, B, M_0, M] \in \mathbb{R} \right\}$$

13.3.3. Faisceaux de cercles à points-limite

Théorème 13.11. Soit A, B deux points distincts du plan annalagmatique et $\mathcal{F}$ le faisceau de cercles (généralisés) de points limites A, B . Pour que deux points distincts P, Q appartiennent au même cercle de $\mathcal{F}$, il faut et il suffit que le birapport $[A, B, P, Q]$ soit de module 1.

Démonstration. Soit P, Q deux points distincts entre eux et de A, B , a, b, p, q les affixes de A, B, P, Q . Posons $\rho = [A, B, P, Q] = [a, b, p, q]$.

1) Supposons qu'aucun des points A, B n'est ∞ . On a

$$|\rho| = \frac{|p-a|}{|p-b|} : \frac{|q-a|}{|q-b|} = \frac{PA}{PB} : \frac{QA}{QB}$$

Par le théorème 11.20 (p. 298), pour que P et Q soient sur le même cercle de $\mathcal{F}$ il faut et il suffit que $\frac{PA}{PB} = \frac{QA}{QB}$, donc que $|\rho| = 1$.

2) Supposons que $B = \infty$. On a

$$|\rho| = \frac{|p-a|}{|q-a|} = \frac{PA}{QA}$$

On sait que les cercles de $\mathcal{F}$ sont les cercles de centre A (p. 301). Pour que P et Q soient sur le même cercle de $\mathcal{F}$, il faut et il suffit que $PA = QA$, donc que $|\rho| = 1$. $\square$

13.3.4. Quadrangle harmonique

Définition 13.12. On dit que quatre points distincts A, B, P, Q forment un **quadrangle harmonique** si $[A, B, P, Q] = -1$

Comme -1 est réel, un quadrangle harmonique est formé de points alignés ou cocycliques. S'ils sont alignés, ils forment une division harmonique au sens du chapitre 8 (13.9 et 8.25, p. 216).

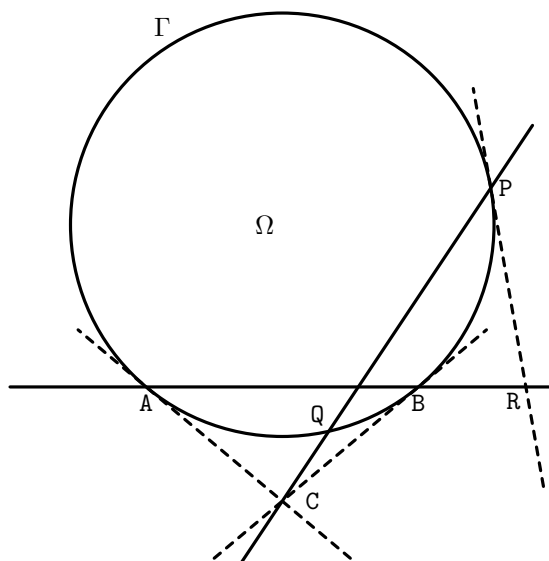
En particulier, pour $Q = \infty$, on a $[A, B, P, \infty] = -1$ si et seulement si P est **milieu** de (A, B) . En effet, $[a, b, p, \infty] = \frac{p-a}{p-b}$ a la valeur -1 si et seulement si $2p = a + b$.

Il nous reste à étudier le cas où les points ne sont pas alignés.

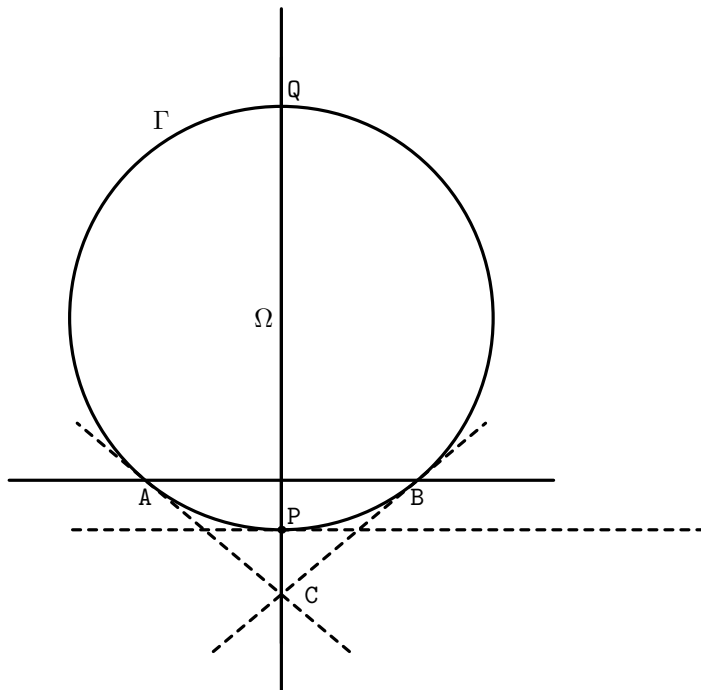
Proposition 13.13. Soit A, B, P, Q quatre points distincts d'un cercle Γ de centre O . On a équivalence entre

- (i) $[A, B, P, Q] = -1$,
- (ii) la droite (AB) et les tangentes en P, Q à Γ sont concourantes ou parallèles.
- (iii) la droite (PQ) et les tangentes en A, B à Γ sont concourantes ou parallèles,

Cas général :



Le cas particulier où P et Q sont diamétralement opposés :



Le point R est envoyé à l'infini.

Démonstration. Le nombre -1 est le seul nombre complexe distinct de 1 qui soit à la fois réel et de module 1 .

Par le théorème 13.11, on a $[A, B, P, Q] = -1$ si et seulement si P et Q sont sur un même cercle γ du faisceau à points limites A, B , donc si ce sont les intersections de Γ et d'un cercle γ centré en un point $R \in (AB)$ orthogonal à Γ (éventuellement la médiatrice de (A, B)). Il revient au même de dire que les tangentes en P, Q à Γ se rencontrent en R sur (AB) , ou sont parallèles à (AB) . Ceci donne l'équivalence de (i) et (ii). Comme dans un quadrangle harmonique, les paires A, B et P, Q jouent le même rôle l'une par rapport à l'autre, on a aussi l'équivalence de (i) et (iii). $\square$

Exercice 13.11. On donne quatre points distincts du plan annalagmatique M_1, M_2, M_3, M_4 et on pose $\rho = [M_1, M_2, M_3, M_4]$. Montrer que la modification de l'ordre des M_i donne les valeurs suivantes pour le birapport :

$$\begin{aligned} \rho &= [M_1, M_2, M_3, M_4], & 1 - \frac{1}{\rho} &= [M_2, M_3, M_1, M_4], & \frac{1}{1 - \rho} &= [M_3, M_1, M_2, M_4] \\ \frac{1}{\rho} &= [M_2, M_1, M_3, M_4], & \frac{\rho}{\rho - 1} &= [M_3, M_2, M_1, M_4], & 1 - \rho &= [M_1, M_3, M_2, M_4] \end{aligned}$$

Exercice 13.12. Soit A, B, C, D quatre points distincts d'un même cercle. Les six valeurs deux à deux inverses du birapport sont réelles (13.10, p. 357). Montrer que parmi elles, deux sont négatives et quatre positives (faire le produit des trois valeurs de la première ligne de l'exercice précédent). Montrer que si $[A, B, C, D] < 0$, alors A et B sont de part et d'autre de (CD) .

Exercice 13.13. Théorème de Ptolémée. On donne quatre points distincts A, B, C, D de $\mathcal{P}$. Montrer l'inégalité

$$AB \times CD \leq BC \times AD + AC \times BD$$

avec égalité si et seulement si A, B, C, D sont sur un même cercle ou alignés, les segments $[A, B]$ et $[C, D]$ ayant une intersection non vide. On remarquera que $1 = \rho + (1 - \rho) = \frac{1}{\rho} + (1 - \frac{1}{\rho}) = \frac{1}{1 - \rho} + \frac{\rho}{\rho - 1}$.

Exercice 13.14. Relation de Newton. Soit A, B, P, Q quatre points distincts d'affixes a, b, p, q .

1) Montrer que $[A, B, P, Q] = -1$ si et seulement si

$$2(ab + pq) = (a + b)(p + q)$$

2) Soit I, J les milieux de (A, B) et de (P, Q) . Montrer qu'on a équivalence entre les conditions

(i) $[A, B, P, Q] = -1$,

(ii) (AB) est bissectrice de $(\overrightarrow{IP}, \overrightarrow{IQ})$ et $IP \times IQ = IA^2 = IB^2$,

(iii) (PQ) est bissectrice de $(\overrightarrow{JA}, \overrightarrow{JB})$ et $JA \times JB = JP^2 = JQ^2$.

Exercice 13.15. Quadrangle équiانharmonique. 1) Soit ρ et ρ' de module 1. Montrer que $\rho + \rho' = 1$ si et seulement si l'un des deux nombres ρ, ρ' est $-j$ et l'autre $-j^2$ (remarquer que les parties imaginaires de ρ et ρ' doivent être opposées).

2) Soit quatre points A, B, C, D d'affixes a, b, c, d . Montrer que les six valeurs du birapport (13.11) sont toutes de module 1 si et seulement si

$$\rho = 1 - \frac{1}{\rho} = \frac{1}{1 - \rho} = -j \text{ ou } -j^2$$

S'il en est ainsi, on dit que (A, B, C, D) est un quadrangle **équiانharmonique**.

3) Montrer qu'il en est ainsi si et seulement si

$$AB \times CD = BC \times AD = CA \times BD$$

4) Que dire de trois points formant avec ∞ un quadrangle équiانharmonique ?

5) Soit A, B, C trois points distincts. Construire deux points D' et D'' tels que (A, B, C, D') et (A, B, C, D'') soient équiانharmoniques.

Exercice 13.16. Soit un triangle ABC, Γ le cercle circonscrit, a, b, c les affixes des sommets. On pose

$$\alpha = \frac{c - b}{\bar{c} - \bar{b}}, \quad \beta = \frac{a - c}{\bar{a} - \bar{c}}, \quad \gamma = \frac{b - a}{\bar{b} - \bar{a}}$$

Pour tout M d'affixe z , on note A', B', C' les symétriques de M relativement aux côtés (BC), (CA), (AB) et a', b', c' leurs affixes.

1) Montrer les formules

$$\begin{aligned} a' &= \alpha(\bar{z} - \bar{b}) + b = \alpha(\bar{z} - \bar{c}) + c \\ b' &= \beta(\bar{z} - \bar{c}) + c = \beta(\bar{z} - \bar{a}) + a \\ c' &= \gamma(\bar{z} - \bar{a}) + a = \gamma(\bar{z} - \bar{b}) + b \end{aligned}$$

2) En déduire $[A', B', C', \infty] = \frac{a' - c'}{b' - c'} = [\bar{a}, \bar{b}, \bar{c}, \bar{z}]$

3) En déduire que $M \in \Gamma$ si et seulement si A', B', C' sont alignés (voir problème 11.3, p. 309). Montrer que C' est milieu de (A', B') si et seulement si $[A, B, C, M] = -1$.

4) Montrer que A'B'C' est équilatéral si et seulement si A, B, C, M est équiانharmonique.

13.4 GROUPE CIRCULAIRE

13.4.1. Les homographies

Définition 13.14. On appelle **homographie numérique** de $\mathbb{C}$ toute fonction du type $h: z \mapsto \frac{az+b}{cz+d}$ où a, b, c, d sont des coefficients complexes tels que $ad - bc \neq 0$.

Pour $c = 0$, on retrouve les similitudes directes pour la structure de plan affine euclidien de $\mathbb{C}$ (13.1, p. 350).

Pourquoi imposer la condition $ad - bc \neq 0$? Si $ad - bc = 0$, il existe λ tel que $a = \lambda c$ et $b = \lambda d$ et $h(z) = \lambda$ pour tout $z \neq -\frac{d}{c}$, la fonction h est constante.

Si $c \neq 0$, h n'est pas définie en $-\frac{d}{c}$. On prolonge h en une bijection de $\overline{\mathbb{C}}$ sur lui-même en posant

- si $c \neq 0$, $h(-\frac{d}{c}) = \infty$ et $h(\infty) = \frac{a}{c}$,
- si $c = 0$, $h(\infty) = \infty$.

Les similitudes directes sont donc les homographies laissant fixe ∞ .

On vérifie que la composée de deux homographies est une homographie et que l'inverse de $h: z \mapsto \frac{az+b}{cz+d}$ est $h^{-1}: z \mapsto \frac{-dz+b}{cz-a}$. Les homographies numériques forment le **groupe circulaire direct numérique** noté $H^+(\mathbb{C})$.

Théorème 13.15. (i) Les homographies sont les applications de $\overline{\mathbb{C}}$ dans $\overline{\mathbb{C}}$ conservant le birapport.

(ii) Étant donné deux triplets (z_1, z_2, z_3) et (z'_1, z'_2, z'_3) où les z_i (resp. les z'_i) sont distincts, il existe une homographie unique h telle que $h(z_i) = z'_i$ pour les trois indices.

Démonstration. 1) Montrons qu'une homographie $h: z \mapsto \frac{az+b}{cz+d}$ conserve le birapport. C'est immédiat si $c = 0$. Supposons $c \neq 0$. Alors

$$h(z) = \frac{\frac{a}{c}(cz+d) - \frac{ad}{c} + b}{cz+d} = \frac{a}{c} - \frac{ad-bc}{c(cz+d)}$$

On voit que h est composée des applications

$$z \mapsto cz + d, \quad z \mapsto \frac{ad-bc}{cz}, \quad z \mapsto \frac{a}{c} - z$$

Le lecteur vérifiera que chacune d'elles conserve le birapport. Ceci établit la moitié de (i) car il restera à montrer qu'une application conservant le birapport est une homographie.

2) Étant donné les deux triplets z_1, z_2, z_3 et z'_1, z'_2, z'_3 , on va montrer qu'il existe une unique application h conservant le birapport transformant z_1, z_2, z_3 en z'_1, z'_2, z'_3 et que c'est une homographie. Ceci établira (ii) et la moitié manquante de (i).

Si h conserve le birapport et transforme les z_i en les z'_i , on aura pour tout z distinct de z_1, z_2, z_3

$$[z_1, z_2, z_3, z] = [z'_1, z'_2, z'_3, h(z)]$$

En remplaçant les birapports par leurs expressions, on constate que cette relation permet le calcul de $h(z)$ (ce qui prouve l'unicité de h qui est bien déterminée), que la formule obtenue pour $h(z)$ est bien celle d'une homographie et enfin qu'elle est encore valable si z est l'un des z_i .

En particulier, si une homographie laisse fixe trois points, c'est l'identité. $\square$

Définition 13.16. *Étant donné un plan affine euclidien $\mathcal{P}$ et son complété annalagmatique, on appelle **homographie** de $\mathcal{P}$ toute application de $\tilde{\mathcal{P}}$ dans lui-même conservant le birapport.*

Proposition 13.17. (i) Une application h de $\tilde{\mathcal{P}}$ dans lui-même est une homographie si et seulement si son expression complexe dans un repère (orthonormé direct) est une homographie numérique. S'il en est ainsi, l'expression complexe de h dans n'importe quel repère est une homographie numérique.

(ii) Les homographies de $\tilde{\mathcal{P}}$ forment un groupe pour la composition appelé **groupe circulaire direct** de $\mathcal{P}$ noté $H^+(\tilde{\mathcal{P}})$.

Démonstration. Soit R un repère, h une application de $\tilde{\mathcal{P}}$ dans lui-même, h_R l'expression complexe de h . Soit M_1, M_2, M_3, M_4 des points, M'_1, M'_2, M'_3, M'_4 les images par h , z_1, z_2, z_3, z_4 et z'_1, z'_2, z'_3, z'_4 les affixes de M_1, M_2, M_3, M_4 et M'_1, M'_2, M'_3, M'_4 . Alors $z'_i = h_R(z_i)$ et on a (13.8, p. 356)

$$[M_1, M_2, M_3, M_4] = [z_1, z_2, z_3, z_4] \quad \text{et} \quad [M'_1, M'_2, M'_3, M'_4] = [z'_1, z'_2, z'_3, z'_4]$$

Il est clair que h conserve le birapport si et seulement si h_R conserve le birapport c'est-à-dire si h_R est une homographie numérique. Le repère R étant quelconque, la preuve de (i) est finie.

Le fait que les homographies forment un groupe vient de ce qu'il en est ainsi pour l'ensemble des homographies numériques. $\square$

Théorème 13.18. *Les homographies transforment cercles généralisés en cercles généralisés et conservent l'orthogonalité.*

Cette propriété explique la terminologie de **groupe circulaire**.

Démonstration. Soit h une homographie et C un cercle. Soit A, B, C trois points de C , A', B', C' les images de A, B, C par h , C' le cercle passant par A', B', C' . Pour tout $M \in C$ d'image M' , on a $[A, B, C, M] = [A', B', C', M']$. Ce nombre est réel, donc $M' \in C'$ (13.10, p. 357) et $h(C) \subset C'$. On montre de même que $h^{-1}(C') \subset C$, d'où $h(C) = C'$.

Soit C, Γ deux cercles orthogonaux de transformés C', Γ' . Soit $A \in C, A \notin \Gamma$. Alors C est orthogonal aux cercles du faisceau $\mathcal{F}$ contenant Γ et le cercle-point $\{A\}$. Comme $A \notin \Gamma, \mathcal{F}$ est à points limites A, B et $B \in C$. Soit A', B' les transformés de A, B . Ils sont sur $C' = h(C)$.

Soit M, N deux points de Γ , M', N' leurs transformés qui sont sur Γ' . Alors $[A, B, M, N] = [A', B', M', N']$ est de module 1 (13.11, p. 358) donc Γ' appartient au faisceau $\mathcal{F}'$ à points limites A', B' . Le cercle C' passant par ces points appartient au faisceau conjugué $\mathcal{F}'^\perp$ à points de base A', B' , donc est orthogonal à Γ' . $\square$

13.4.2. Points fixes d'une homographie

Cherchons les points fixes de l'homographie numérique $h: z \mapsto \frac{az+b}{cz+d}$.

1) Cas d'une similitude directe : $c = 0$ et $d \neq 0$ car $ad - bc \neq 0$. On a $h(z) = \frac{a}{d}z + \frac{b}{d}$ et ∞ est point fixe.

- Pour $a = d$, $h(z) = z + \frac{b}{d}$. Si $b = 0$, h est l'identité, sinon la similitude est une translation et ∞ est l'unique point fixe.
- Pour $a \neq d$, c'est une similitude de centre le point d'affixe $\frac{b}{d-a}$; il y a donc deux points fixes en comptant ∞ .

2) Cas où $c \neq 0$, $h(\infty) = \frac{a}{c} \neq \infty$. On trouve les points fixes en étudiant l'équation $h(z) = z$, c'est-à-dire les racines dans $\mathbb{C}$ du trinôme $cX^2 - (a-d)X - b$, de discriminant

$$\Delta(h) = (a-d)^2 + 4bc = (a+d)^2 - 4(ad-bc)$$

- Pour $\Delta(h) \neq 0$ il y a deux points fixes distincts.
- Pour $\Delta(h) = 0$, il y a un point fixe unique $\frac{a-d}{2c}$. On dit que h est une **élation**.

Remarquons que pour $c = 0$, h est une similitude directe avec ∞ pour point fixe ; ce sera une élation si et seulement si ∞ est seul point fixe, c'est-à-dire si h est une translation : $a = d$, $(a-d)^2 = \Delta(h) = 0$.

Pour une homographie h sur un plan $\mathcal{P}$, on applique les résultats de cette discussion à la représentation complexe de h dans un repère.

Proposition 13.19. Soit h une homographie avec deux points fixes U, V (distincts). Pour tout M , le birapport $[U, V, M, h(M)]$ est indépendant de M .

Démonstration. Raisonnons numériquement. Soit u et v les affixes de U, V . Soit z_0 distinct de u, v . Alors $z_1 = h(z_0) \neq z_0$. Posons $\rho = [u, v, z_0, z_1]$. Pour tout z , définissons $h'(z)$ par la relation $[u, v, z, h'(z)] = \rho$. Ceci permet le calcul de $h'(z)$ en fonction de u, v, z . On constate que l'expression obtenue est celle d'une homographie. On a $h' = h$ car h' et h coïncident en u, v, z_0 (13.15, p. 362). Comme h' a la propriété, le résultat est établi. $\square$

Exemple : Soit la similitude $z \mapsto az + b$ où $a \neq 0$. On a les points doubles $v = \frac{b}{1-a}$ et ∞ . Alors

$$[\infty, v, z, az + b] = \frac{az + b - v}{z - v} = \frac{(az + b) - (av + b)}{z - v} = a$$

13.4.3. Involutions

Définition 13.20. Une homographie h , distincte de l'identité, est une **involution** si elle est elle-même sa propre inverse.

Proposition 13.21. Une homographie h est une involution si et seulement si elle a deux points fixes distincts U, V et si $[U, V, M, h(M)] = -1$ pour tout M .

Démonstration. Montrons qu'une homographie ayant cette propriété est une involution. On prend U, V tels que $[U, V, M, h(M)] = -1$ pour tout M . Comme -1 est son propre inverse, on a

$$[U, V, h(M), M] = \frac{1}{[U, V, M, h(M)]} = -1 = [U, V, h(M), h(h(M))]$$

donc, $M = h(h(M))$, d'où $h^2 = \text{Id}$.

Inversement, soit h une involution et U un point fixe (il en existe au moins un). Soit M_0 un point distinct de son image M_1 , V tel que $[U, V, M_0, M_1] = -1$ et h' l'involution de points fixes U et V . Alors h et h' coïncident en U, M_0, M_1 , donc $h = h'$ (13.15, p. 362) a bien la propriété. $\square$

Description géométrique des involutions Montrons qu'on peut construire l'image d'un point quelconque par une involution. Notons U, V les points fixes de l'involution h . Pour tout M , $h(M) = M'$ est tel que U, V, M, M' est un quadrangle harmonique (13.13, p. 358).

Si M est sur la droite (UV) , c'est le conjugué harmonique de M relativement à U, V au sens du chapitre 8 (8.25, p. 216).

Sinon, on construit le cercle $\Gamma = (UVM)$. On joint M à l'intersection R des tangentes en U, V à Γ . Cette droite recoupe Γ en $M' = h(M)$.

Si un point fixe est $U = \infty$, alors h est une similitude involutive, c'est-à-dire la **symétrie centrale** de centre V .

13.4.4. Les antihomographies

Nous allons étudier un autre type de transformations qui jouent pour les homographies le rôle que jouent les antidéplacements pour les déplacements.

Définition 13.22. On appelle **antihomographie** toute application f du plan annalagmatique $\tilde{\mathcal{P}}$ dans lui-même transformant les birapports en leurs conjugués : pour M_1, M_2, M_3, M_4 distincts on a

$$[f(M_1), f(M_2), f(M_3), f(M_4)] = \overline{[M_1, M_2, M_3, M_4]}$$

Théorème 13.23. (i) Les réflexions sont des antihomographies.

(ii) Étant donné une réflexion σ , l'application $f \mapsto f \circ \sigma$ (resp. $f \mapsto \sigma \circ f$) est involutive et échange l'ensemble $H^+(\check{\mathcal{P}})$ des homographies et l'ensemble $H^-(\check{\mathcal{P}})$ des antihomographies.

(iii) La réunion des ensembles $H^+(\check{\mathcal{P}})$ et $H^-(\check{\mathcal{P}})$ est un groupe noté $H(\check{\mathcal{P}})$ appelé **groupe circulaire** de $\mathcal{P}$, dont $H^+(\check{\mathcal{P}})$ est un sous-groupe d'indice 2.

(iv) Les homographies et antihomographies transforment cercles généralisés en cercles généralisés et conservent l'orthogonalité.

Démonstration. (i) Soit σ la réflexion d'axe une droite Δ de $\mathcal{P}$. Pour que l'énoncé ait un sens, il faut préciser comment on prolonge σ à $\check{\mathcal{P}}$. On posera $\sigma(\infty) = \infty$. Soit un repère orthonormé direct $(O, \vec{u}, \vec{v})$ où $O \in \Delta$ et $\vec{u}$ est vecteur directeur de Δ . Autrement dit, Δ est l'axe des réels. La représentation complexe de σ est alors $z \mapsto \bar{z}$. C'est bien une antihomographie.

(ii) Ce point est immédiat à vérifier. Il en résulte que la représentation complexe d'une antihomographie est $z \mapsto \frac{az+b}{c\bar{z}+d}$ avec $ad - bc \neq 0$.

(iii) Soit f, g dans $H(\check{\mathcal{P}})$. Étudions $g \circ f$:

- si f conserve le birapport et g conserve le birapport, alors $g \circ f$ conserve le birapport, donc est une homographie,
- si f conserve le birapport et g conjugue le birapport, alors $g \circ f$ conjugue le birapport, donc est une antihomographie,
- si f conjugue le birapport et g conserve le birapport, alors $g \circ f$ conjugue le birapport, donc est une antihomographie
- si f conjugue le birapport et g conjugue le birapport, alors $g \circ f$ conserve le birapport, donc est une homographie.

On en déduit facilement que $H(\check{\mathcal{P}})$ est un groupe et que $H^+(\check{\mathcal{P}})$ est sous-groupe d'indice 2.

(iv) On connaît le résultat pour les homographies (13.18, p. 363). Il est très facile à vérifier pour les réflexions. Il est donc vrai pour les antihomographies. $\square$

Définition 13.24. Étant donné un point $O \in \mathcal{P}$ et un nombre $k \neq 0$, on appelle **inversion** de pôle O et de puissance k l'application notée $\text{Inv}(O, k)$ involutive de $\check{\mathcal{P}}$ sur lui-même ainsi définie :

- $\text{Inv}(O, k)$ échange O et ∞ ,
- à M distinct de O et de ∞ , $\text{Inv}(O, k)$ associe le point M' de la droite (OM) tel que $\overline{OM} \times \overline{OM'} = k$.

Nous avons déjà rencontré cette transformation en exercice au chapitre 11 (11.4, p. 296).

Proposition 13.25. (i) Les inversions sont des antihomographies.

(ii) Les antihomographies qui sont involutives sont les réflexions et les inversions.

Démonstration. (i) Soit l'inversion $\text{Inv}(0, k)$. Prenons un repère orthonormé direct d'origine 0. On voit facilement que l'application $z \mapsto \frac{k}{\bar{z}}$ pour $z \neq 0$ et échangeant 0 et ∞ est représentation complexe de cette inversion, d'où le résultat.

(ii) Il est clair que les réflexions et les inversions sont involutives. Inversement, soit f une antihomographie telle que $f = f^{-1}$.

- Supposons $f(\infty) = \infty$. Prenons un repère et la représentation complexe de f encore notée f . Elle est du type $z \mapsto \frac{a\bar{z}+b}{c\bar{z}+d}$. La condition $f(\infty) = \infty$ donne $c = 0$. Alors $d \neq 0$ car $ad - bc \neq 0$. En remplaçant a, b par $\frac{a}{d}, \frac{b}{d}$, on se ramène au cas où $f(z) = a\bar{z} + b$. C'est donc une application affine (13.2, p. 351). Comme f est involutive et conserve les milieux, elle laisse fixe le milieu 0 de deux points homologues. Plaçons l'origine en 0. Alors $b = 0$ et $f(z) = a\bar{z}$. Comme f est sa propre inverse, on a $z = f(f(z)) = a\bar{a}z$ pour tout z . Ceci impose $|a| = 1$ et il existe θ tel que $a = e^{i\theta}$. Alors f est composée des applications de représentations complexes $z \mapsto \bar{z}$ et $z \mapsto e^{i\theta}z$, c'est-à-dire d'une réflexion d'axe passant par 0 et d'une rotation de centre 0. C'est donc une réflexion.
- Supposons que f échange 0 et ∞ . La représentation complexe de f dans un repère d'origine 0, encore notée f , est $z \mapsto \frac{a\bar{z}+b}{c\bar{z}+d}$ avec $a = 0$ pour $f(\infty) = 0$ et $d = 0$ pour $f(0) = \infty$. La représentation complexe est donc $z \mapsto \frac{b}{c\bar{z}}$. L'inverse f^{-1} a pour représentation complexe $z \mapsto \frac{\bar{b}}{c\bar{z}}$. Comme $f = f^{-1}$, on a $\frac{b}{c} = \frac{\bar{b}}{c} = k$ où k est réel car $\bar{\bar{k}} = k$. D'où $f = \text{Inv}(0, k)$. $\square$

EXERCICES

Exercice 13.17. Montrer qu'une homographie h transforme un faisceau de cercles en un faisceau de cercles de même nature.

Exercice 13.18. Soit A, B deux points distincts et h une homographie échangeant A et B. Montrer que h est une involution (considérer les points fixes de h et h^2 et utiliser 13.15, p. 362)

Exercice 13.19. Point de Frégier d'une involution sur un cercle. Soit h une involution de points fixes U, V.

1) Montrer qu'un cercle généralisé est stable par h si et seulement si il appartient au faisceau $\mathcal{F}(U, V)$ de points-limites U, V ou au faisceau $\mathcal{F}^\perp(U, V)$ à points de base U, V.

2) On suppose que C est un cercle à centre 0 stable par h . Montrer que pour tout $M \in C$, la droite Δ_M joignant M et $h(M)$ passe par un point fixe (ou est de direction fixe) F appelé **point de Frégier** situé sur la droite des centres du faisceau $\mathcal{F}(U, V)$ ou $\mathcal{F}^\perp(U, V)$ auquel C appartient (distinguer les deux cas).

Exercice 13.20. Groupe de Klein. Soit h_1, h_2 deux involutions distinctes de points fixes U_1, V_1 pour h_1 , U_2, V_2 pour h_2 . On pose $h_3 = h_2 \circ h_1$

1) Montrer qu'on a équivalence entre les conditions

- (i) h_1 et h_2 commutent,
- (ii) h_3 est une involution,
- (iii) h_2 échange U_1 et V_1 ,
- (iv) h_1 échange U_2 et V_2 .

2) Dans ces conditions, montrer que $\{\text{Id}, h_1, h_2, h_3\}$ est un groupe commutatif.

3) Étant donné un ensemble Q de quatre points distincts, montrer qu'il existe un tel groupe d'involutions laissant stable Q .

Exercice 13.21. Soit A un point de $\check{\mathcal{P}}$ et G_A le groupe des homographies admettant A pour point fixe.

1) Montrer que G_A est un sous-groupe de $H^+(\check{\mathcal{P}})$ conjugué du groupe des similitudes directes de $\mathcal{P}$.

2) Quelles sont les conjuguées dans G_A des translations de $\mathcal{P}$?

3) Pour tout $h \neq \text{Id}$ de G_A , on a les éventualités :

- Si $h \in G_A$ a un deuxième point fixe $B \in \check{\mathcal{P}}$, pour tout M , la valeur du birapport $[A, B, M, h(M)]$ est une constante $\rho(h)$ (13.19, p. 364).
- Si h est une élation et n'a donc pas de deuxième point fixe, on pose $\rho(h) = 1$.

Montrer que l'application $h \mapsto \rho(h)$ est un morphisme de groupe $G_A \rightarrow \mathbb{C}^\times$.

Exercice 13.22. Soit un ensemble $T = \{A_1, A_2, A_3\}$ de trois points distincts.

1) Montrer que le groupe $G(T)$ des homographies laissant T stable est isomorphe à S_3 (utiliser 13.15), p. 362).

2) Montrer que ces groupes $G(T)$ forment une classe de conjugaison de sous-groupes de $H^+(\check{\mathcal{P}})$.

3) On suppose que T est un triangle de cercle circonscrit Γ . Construire les points fixes des trois involutions de $G(T)$ (utiliser l'exercice 13.19).

4) En déduire un autre triangle T' tel que $G(T) = G(T')$ (considérer les points fixes des involutions de $G(T)$).

5) Construire T' si $A_1 = \infty$.

Exercice 13.23. 1) Soit Q un quadrangle harmonique. Montrer que le groupe des homographies laissant Q stable est diédral de cardinal 8.

2) Montrer que les groupes $G(Q)$ forment une classe de conjugaison de sous-groupes de $H^+(\check{\mathcal{P}})$.

Exercice 13.24. 1) Soit Q un quadrangle équiharmonique. Montrer que le groupe des homographies laissant stable Q est isomorphe au groupe alterné $\mathcal{A}_4$.

2) Montrer que les groupes $G(Q)$ forment une classe de conjugaison de sous-groupes de $H^+(\check{\mathcal{P}})$.

PROBLÈMES

Les corrigés de ces problèmes sont disponibles sur le site de Dunod : www.dunod.com.

13.1. PROBLÈME

Dans un plan affine euclidien orienté $\mathcal{P}$ rapporté à un repère, on donne un polygone à n sommets $A = (A_0, \dots, A_{n-1})$ d'affixes $a_0, \dots, a_{n-1}$. On note $\text{Mat}(A)$ la matrice à une colonne dont les coefficients sont $a_0, \dots, a_{n-1}$.

Soit la racine primitive n -ième de l'unité $\omega = \exp(\frac{2i\pi}{n})$ et $P = (P_0, \dots, P_{n-1})$ le polygone régulier direct convexe dont les sommets ont pour affixes $1, \omega, \dots, \omega^{n-1}$.

Il est commode pour les notations de supposer que l'indice varie dans $\mathbb{Z}/n\mathbb{Z}$ au lieu de $\{0, 1, \dots, n-1\}$ de sorte que $A_{(n-1)+1} = A_0$.

1) Dans cette question, $n = 4$. On considère les conditions

$$(1) \quad a_0 + ia_1 - a_2 - ia_3 = 0$$

$$(2) \quad a_0 - a_1 + a_2 - a_3 = 0$$

Quelles traductions géométriques ont ces relations pour $A = (A_0, A_1, A_2, A_3)$ (la condition (1) seule, la condition (2) seule, les conditions (1) et (2) ensemble) ?

2) Soit la matrice à $n-2$ lignes et n colonnes

$$\Omega = (\omega^{hk}) = \begin{pmatrix} 1 & \omega & \omega^2 & \dots & \omega^{n-1} \\ 1 & \omega^2 & \omega^4 & \dots & \omega^{2(n-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \omega^{n-2} & \omega^{(n-2)2} & \dots & \omega^{(n-2)(n-1)} \end{pmatrix}$$

interprétée comme application linéaire de $\mathbb{C}^n$ dans $\mathbb{C}^{n-2}$.

Montrer que $\text{Ker } \Omega$ est engendré par les matrices à une colonne U et V de coefficients $1, \dots, 1$ et $1, \omega, \dots, \omega^{n-1}$.

En déduire que A est un polygone régulier direct convexe si et seulement si on a la relation matricielle $\Omega \text{ Mat}(A) = O$ (exprimer que A est régulier direct convexe si et seulement si il se déduit de P par une similitude directe).

3) On dira que A est **affinement** régulier convexe s'il se déduit de P par une application affine.

Interprétation géométrique pour $n = 4$.

Montrer qu'il en est ainsi si et seulement si on a la relation matricielle $\Omega' \text{ Mat}(A) = O$ où Ω' est la matrice à n colonnes et $n - 3$ lignes se déduisant de Ω en enlevant la première ligne.

4) On suppose $n = 4$. Soit $A = (A_0, A_1, A_2, A_3)$ quelconque.

- On pose $a'_h = \frac{a_{h+1} + a_h}{2}$. Quelle propriété possède $A' = (A'_0, A'_1, A'_2, A'_3)$?
- On pose $a''_h = \frac{ia_{h+1} - a_h}{i-1}$. Quelle propriété possède $A'' = (A''_0, A''_1, A''_2, A''_3)$?
- On pose $b'_h = \frac{ia'_{h+1} - a'_h}{i-1}$ d'image B'_h . Quelle propriété possède $B' = (B'_0, B'_1, B'_2, B'_3)$?
- On pose $b''_h = \frac{a''_{h+1} + a''_h}{2}$ d'image B''_h . Quelle propriété possède $B'' = (B''_0, B''_1, B''_2, B''_3)$?

Faire des figures.

5) Soit $u \neq 1$ une racine n -ième de l'unité. On forme $A' = (A'_0, \dots, A'_{n-1})$ d'affixes $a'_0, \dots, a'_{n-1}$ où $a'_h = \frac{ua_{h+1} - a_h}{u-1}$.

Comment les polygones A et A' sont-ils liés géométriquement ?

Montrer que $a'_0 + ua'_1 + \dots + u^{n-1}a'_{n-1} = 0$.

Montrer que si $v \neq 1$ est une autre racine n -ième de l'unité telle que

$$a_0 + va_1 + \dots + v^{n-1}a_{n-1} = 0,$$

alors on a encore $a'_0 + va'_1 + \dots + v^{n-1}a'_{n-1} = 0$.

13.2. PROBLÈME

Soit $\mathcal{E}$ un espace affine euclidien orienté de dimension 3 d'espace vectoriel associé E , $\mathcal{P}$ un plan de $\mathcal{E}$, S une sphère de centre Ω , de rayon R , tangente en un point O à $\mathcal{P}$, s le point de S diamétralement opposé à O , π la projection stéréographique de sommet s de S sur $\tilde{\mathcal{P}} = \mathcal{P} \cup \{\infty\}$. Soit $G^+(S)$ (resp. $G(S)$) le groupe des rotations (resp. isométries) de $\mathcal{E}$ laissant S stable. On se propose d'étudier les bijections de $\tilde{\mathcal{P}}$ sur lui-même du type $\pi \circ f \circ \pi^{-1}$ où $f \in G^+(S)$ (resp. $f \in G(S)$).

1) Montrer que $G^+(S)$ et $G(S)$ sont isomorphes à $\mathbf{SO}(E)$ et $\mathbf{O}(E)$.

2) On oriente la droite (Os) de O vers s . Soit la rotation $\text{rot}(Os, \theta)$. Montrer que $\pi \circ \text{rot}(Os, \theta) \circ \pi^{-1}$ est une homographie dont on donnera les points fixes et le birapport (13.19, p. 364).

3) Soit a, b diamétralement opposés sur S , $A = \pi(a)$, $B = \pi(b)$ et $\rho_{(ab)}$ le retournement d'axe (ab) . Montrer que $\pi \circ \rho_{(ab)} \circ \pi^{-1}$ est l'involution de points fixes A, B (utiliser 11.33, p. 307, 13.13, p. 358 et 13.21, p. 365).

4) En déduire que les $\pi \circ f \circ \pi^{-1}$ sont des homographies si f décrit $G^+(S)$ (utiliser 8.43, p. 230).

5) Soit f une rotation d'axe (ab) orienté de a vers b et d'angle θ . Montrer que $h = \pi \circ f \circ \pi^{-1}$ est une homographie dont on donnera les points fixes et le birapport (on pourra montrer que f est conjuguée d'une rotation d'axe (0s) par un retournement).

6) Montrer que pour qu'une homographie h soit du type $\pi \circ f \circ \pi^{-1}$ où $f \in G^+(S)$, il faut et il suffit que h ait deux points fixes A, B homologues par l'inversion $\text{Inv}(0, -4R^2)$ et que le birapport de h soit de module 1.

7) Donner un exemple de sous-groupe fini de $H^+(\tilde{P})$ isomorphe à S_4 (utiliser 10.13, p. 275).

SOLUTIONS DES EXERCICES

Solution 13.1. L'origine étant 0, soit a, b, c, z, z' les affixes de A, B, C, M, M'. La similitude (variable) des plaques MAC et CBM' s'exprime par l'existence de θ et ρ tels que

$$\widehat{(\overrightarrow{AM}, \overrightarrow{AC})} \equiv \theta \equiv \widehat{(\overrightarrow{BC}, \overrightarrow{BM'})} \pmod{2\pi} \quad \text{et} \quad \frac{AC}{AM} = \rho = \frac{BM'}{BC}$$

On en déduit

$$\frac{c-a}{z-a} = k = \frac{z'-b}{c-b} \quad \text{où} \quad k = \rho e^{i\theta}$$

Comme $c-a=b$ et $c-b=a$, on en déduit $z' = kz$.

Si on fixe des crayons en M et M', lorsque le crayon M décrit une figure $\mathcal{F}$, la crayon M' décrit la figure se déduisant de $\mathcal{F}$ par la similitude de centre 0, d'angle θ , de rapport ρ .

Solution 13.2. 1) Le triangle ABC est équilatéral direct si et seulement si $\text{rot}(B, \frac{\pi}{3})$ transforme C en A, ce qui est équivalent à

$$\frac{a-b}{c-b} = e^{\frac{i\pi}{3}} = -j^2$$

autrement dit à $a+bj+cj^2=0$. Le triangle est équilatéral indirect si $a+bj^2+cj=0$, et il sera équilatéral si et seulement si

$$(a+bj+cj^2)(a+bj^2+cj) = a^2+b^2+c^2-bc-ca-ab=0$$

2) Pour que les images des racines a, b, c forment un triangle équilatéral, il faut et il suffit que

$$0 = a^2 + b^2 + c^2 - bc - ca - ab = (a+b+c)^2 - 3(bc+ca+ab)$$

Les relations entre coefficients et racines donnent $bc+ca+ab = \mu$, et $a+b+c = -\lambda$. La condition est donc $\mu = \frac{1}{3}\lambda^2$.

3) Soit p, q, r, a, b, c les affixes de P, Q, R, A, B, C. Supposons T direct. On a

$$AQ = AR \text{ et } (\widehat{AQ, AR}) \equiv -\frac{2\pi}{3}, \text{ soit } \frac{r-a}{q-a} = e^{-\frac{2i\pi}{3}} = j^2$$

et deux autres formules analogues. On a donc

$$a(j^2 - 1) = qj^2 - r, \quad b(j^2 - 1) = rj^2 - p, \quad c(j^2 - 1) = pj^2 - q$$

$$(j^2 - 1)(a + bj + cj^2) = qj^2 - r + r - pj + pj - qj^2 = 0$$

d'où $a + bj + cj^2 = 0$.

Solution 13.3. Cette équation est du type $\bar{a}z + a\bar{z} + b = 0$ où $a \neq 0$ et $b \in \mathbb{R}$. Soit M d'affixe z sur cette droite. On aura

$$\begin{cases} \bar{a}z_1 + a\bar{z}_1 + b = 0 \\ \bar{a}z_2 + a\bar{z}_2 + b = 0 \\ \bar{a}z + a\bar{z} + b = 0 \end{cases}$$

Le système linéaire homogène en x, y, t

$$\begin{cases} z_1x + \bar{z}_1y + t = 0 \\ z_2x + \bar{z}_2y + t = 0 \\ zx + \bar{z}y + t = 0 \end{cases}$$

a une solution $(\bar{a}, a, b)$, non triviale car $a \neq 0$. On a donc

$$\det \begin{pmatrix} z_1 & \bar{z}_1 & 1 \\ z_2 & \bar{z}_2 & 1 \\ z & \bar{z} & 1 \end{pmatrix} = 0$$

soit $z(\bar{z}_1 - \bar{z}_2) - \bar{z}(z_1 - z_2) + z_1\bar{z}_2 - z_2\bar{z}_1 = 0$. En multipliant par i , on obtient l'équation classique d'une droite contenant M_1, M_2 de façon évidente. On a donc bien l'équation de (M_1M_2) .

Solution 13.4. 1) Comme P et A ne sont pas dans le même demi-plan limité par (BC), PBC est indirect et on passe de B à C par la rotation de centre P et d'angle $-\frac{\pi}{3}$. On a donc $c - p = -j(b - p)$, donc $p = -j^2b - jc$ et de même $q = -j^2c - ja$ et $r = -j^2a - jb$. Les trois vecteurs ont même norme et font entre eux des angles $\pm \frac{2\pi}{3}$ car on a

$$a - p = a + j^2b + jc, \quad b - q = b + j^2c + ja = j(a - p), \quad c - r = c + j^2a + jb = j^2(a - p)$$

2) Les équations de (AP), (BQ), (CR) sont (ex.13.3)

$$z(\bar{a} - \bar{p}) - \bar{z}(a - p) + a\bar{p} - \bar{a}p = 0$$

$$z(\bar{b} - \bar{q}) - \bar{z}(b - q) + b\bar{q} - \bar{b}q = 0$$

$$z(\bar{c} - \bar{r}) - \bar{z}(c - r) + c\bar{r} - \bar{c}r = 0$$

où on a

$$\begin{aligned} a\bar{p} - \bar{a}p &= j a \bar{b} + j^2 a \bar{c} - j^2 \bar{a} b + j \bar{a} c \\ b\bar{q} - \bar{b}q &= j b \bar{c} + j^2 b \bar{a} - j^2 \bar{b} c - j \bar{b} a \\ c\bar{r} - \bar{c}r &= j c \bar{a} + j^2 c \bar{b} - j^2 \bar{c} a - j \bar{c} b \end{aligned}$$

On constate alors que la somme des premiers membres des équations est identiquement nulle. Tout point commun aux deux premières droites est aussi sur la troisième. Ces trois droites sont concourantes.

Solution 13.5. La méthode est exactement la même que pour l'exercice 13.3. On trouve pour équation

$$\det \begin{pmatrix} z_1 \bar{z}_1 & z_1 & \bar{z}_1 & 1 \\ z_2 \bar{z}_2 & z_2 & \bar{z}_2 & 1 \\ z_3 \bar{z}_3 & z_3 & \bar{z}_3 & 1 \\ z \bar{z} & z & \bar{z} & 1 \end{pmatrix} = 0$$

Solution 13.6. 1) Soit a, b, c, z les affixes de A, B, C, M . Supposons ABC direct, on a $a + bj + cj^2 = 0$ (ex.13.2). D'où $(z - a) + j(z - b) + j^2(z - c) = 0$ et $(z - a) + j(z - b) = -j^2(z - c)$. L'inégalité triangulaire donne

$$\begin{aligned} MC &= |-j^2(z - c)| = |(z - a) + j(z - b)| \\ &\leq |(z - a)| + |j(z - b)| = MA + MB \end{aligned}$$

2) L'inégalité $|z_1 + z_2| \leq |z_1| + |z_2|$ est une égalité si et seulement si les images vectorielles de z_1 et z_2 sont colinéaires de même sens. L'inégalité précédente est une égalité si et seulement si les conditions équivalentes ci-dessous sont satisfaites

$$\left(\arg(z - a) = \arg(j(z - b)) \right) \Leftrightarrow \left(\arg\left(\frac{z - a}{z - b}\right) = \arg(j) = \frac{2\pi}{3} \right)$$

Ceci équivaut à l'appartenance de M à l'arc indiqué.

Solution 13.7. 1) Soit la matrice

$$M = \text{Mat}_B(\vec{OA}, \vec{OB}, \vec{OC}) = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

Posons $\|\vec{OA}\| = \|\vec{OB}\| = \|\vec{OC}\| = l$, longueur de l'arête du cube. Alors $\frac{1}{l}M$ est une matrice orthogonale. On a

$$\begin{aligned} a^2 + b^2 + c^2 &= (a_1 + ia_2)^2 + (b_1 + ib_2)^2 + (c_1 + ic_2)^2 \\ &= \left((a_1^2 + b_1^2 + c_1^2) - (a_2^2 + b_2^2 + c_2^2) \right) + 2i(a_1a_2 + b_1b_2 + c_1c_2) \end{aligned}$$

La transposée d'une matrice orthogonale est orthogonale, donc les vecteurs colonnes $\vec{V}_1, \vec{V}_2, \vec{V}_3$ de la transposée tM sont de même norme et deux à deux orthogonaux. On a donc

$$a^2 + b^2 + c^2 = (\|\vec{V}_1\|^2 - \|\vec{V}_2\|^2) + 2i\vec{V}_1 \cdot \vec{V}_2 = 0$$

2) Si 0 est isobarycentre du triangle projeté, on a

$$0 = (a + b + c)^2 = a^2 + b^2 + c^2 + 2ab + 2bc + 2ca = 2(ab + bc + ca)$$

On a alors $(a - b)^2 + (b - c)^2 + (c - a)^2 = 0$, donc a, b, c sont bien affixes des sommets d'un triangle équilatéral.

3) Considérons les images A', B', C' de a, b, c dans $\mathcal{P}$. On suppose $a^2 + b^2 + c^2 = 0$. Il est évident que si A, B, C sont trois points de $\mathcal{E}$ ayant A', B', C' pour projetés, sans autre hypothèse, nous ne parviendrons pas à prouver que ce sont les arêtes d'un cube (prendre par exemple A proche de A' et B très loin sur la verticale de B' , nous aurons $OA < OB$).

Il reste à interpréter géométriquement la condition algébrique $a^2 + b^2 + c^2 = 0$. Envisageons la réciproque ainsi : existe-t-il A, B, C de projetés A', B', C' tels que $\vec{OA}, \vec{OB}, \vec{OC}$ soient les arêtes d'un cube ?

Posons $a = a_1 + ia_2, b = b_1 + ib_2, c = c_1 + ic_2$. On cherche a_3, b_3, c_3 tels que la matrice

$$M = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

soit proportionnelle à une matrice orthogonale, i.e. que la transposée tM soit proportionnelle à une matrice orthogonale. Cela revient à compléter la matrice

$$\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{pmatrix} \text{ par une troisième colonne } \begin{pmatrix} a_3 \\ b_3 \\ c_3 \end{pmatrix}$$

de façon à obtenir une matrice proportionnelle à une matrice orthogonale. La condition $a^2 + b^2 + c^2 = 0$ exprime que les vecteurs

$$\vec{V}_1 = a_1\vec{e}_1 + b_1\vec{e}_2 + c_1\vec{e}_3 \quad \text{et} \quad \vec{V}_2 = a_2\vec{e}_1 + b_2\vec{e}_2 + c_2\vec{e}_3$$

sont orthogonaux de même norme l . On cherche $\vec{V}_3 = a_3\vec{e}_1 + b_3\vec{e}_2 + c_3\vec{e}_3$ de norme l orthogonaux à $\vec{V}_1$ et $\vec{V}_2$. On a deux solutions opposées

$$\vec{V}_3 = \pm \frac{1}{l} \vec{V}_1 \wedge \vec{V}_2 = \pm (a_3\vec{e}_1 + b_3\vec{e}_2 + c_3\vec{e}_3)$$

$$\text{où} \quad a_3 = \frac{1}{l}(b_1c_2 - c_1b_2), \quad b_3 = \frac{1}{l}(c_1a_2 - a_1c_2), \quad c_3 = \frac{1}{l}(a_1b_2 - b_1a_2)$$

$$\text{et} \quad l = \sqrt{a_1^2 + b_1^2 + c_1^2} = \sqrt{a_2^2 + b_2^2 + c_2^2}$$

Solution 13.8.

1) L'affixe de M est

$$z = \frac{1}{2}(z' + z'') = \frac{1}{2}(a' + a'') + \frac{1}{2}(R'e^{i\varphi'} + R''e^{i\varphi''})e^{i\omega t}$$

Le mouvement de M est donc circulaire uniforme de vitesse angulaire ω sur un cercle de centre A, milieu de (A', A'') , de rayon $R = \frac{1}{2}|R'e^{i\varphi'} + R''e^{i\varphi''}|$.

2) Pour que M soit fixe, il faut et il suffit que $R = 0$, donc que $\frac{R'}{R''} = -\exp(i(\varphi'' - \varphi'))$. Comme $\frac{R'}{R''}$ est réel positif et $-\exp(i(\varphi'' - \varphi'))$ est de module 1, ceci équivaut à $R' = R''$ et $\varphi'' = \varphi' + \pi$.

Solution 13.9. 1) Si φ n'est pas injective, il existe $z_0 \neq 0$ tel que $az_0 = -b\bar{z}_0$, d'où $|a| \times |z_0| = |b| \times |z_0|$, d'où $|a| = |b|$.

Inversement, si $|b| = |a|$, il existe θ tel que $b = e^{i\theta}a$. Le noyau est l'ensemble des vecteurs dont l'affixe z vérifie $z + \bar{z}e^{i\theta} = 0$, soit $\arg(z) \equiv \theta - \arg(z) + \pi$ modulo 2π , c'est-à-dire $\arg(z) \equiv \frac{\theta}{2} + \frac{\pi}{2}$ modulo π . Donc $\text{Ker}(\varphi)$ est une droite vectorielle.

2) La base $(1, i)$ de $\mathbb{C}$ devient $(a + b, ia - ib)$, donc (13.1.1., p. 349)

$$\det(\varphi) = \det(a + b, ia - ib) = \frac{1}{2i}((\bar{a} + \bar{b})(ia - ib) - (-i\bar{a} + i\bar{b})(a + b)) = |a|^2 - |b|^2$$

3) Le réel λ est valeur propre si et seulement si l'application $z \mapsto (a - \lambda)z + b\bar{z}$ est de déterminant nul, soit $|a - \lambda| = |b|$. Le polynôme caractéristique est

$$(a - X)(\bar{a} - X) - |b|^2 = X^2 - (a + \bar{a})X + |a|^2 - |b|^2$$

Pour trouver les $\lambda \in \mathbb{R}$ tels que $|a - \lambda| = |b|$, on trace la droite horizontale passant par A qui coupera (ou pas) le cercle de centre O et de rayon $|b|$ en les images des $a - \lambda$. Si c'est le cas, c'est-à-dire si $|\text{Im}(a)| \leq |b|$, on a deux valeurs propres réelles λ', λ'' .

Solution 13.10. L'affixe de M est

$$z = \frac{1}{2}(z' + z'') = \frac{1}{2}(a' + a'') + \frac{1}{2}(R'e^{i\varphi'}u(t) + R''e^{i\varphi''}\bar{u}(t))$$

On passe de u à z par une application affine dont l'application linéaire associée est $u \mapsto K'u + K''\bar{u}$ où $K' = \frac{1}{2}R'e^{i\varphi'}$ et $K'' = \frac{1}{2}R''e^{i\varphi''}$. Dans $\mathbb{C}$, u décrit le cercle trigonométrique et z décrit l'ellipse transformée de ce cercle par cette application affine. Son centre est le transformé de 0, donc $\frac{1}{2}(a' + a'')$, affixe du milieu A de (A', A'') .

Cette ellipse dégénère en un segment si et seulement si cette application est non injective, donc (ex.13.9) si $|K'| = |K''|$, donc si $R' = R''$.

La distance de M au centre est

$$AM = |z - a| = \frac{1}{2} \left| R' \exp i(\omega t + \varphi') + R'' \exp i(-\omega t + \varphi'') \right|$$

- Elle est maximum si $R' \exp i(\omega t + \varphi')$ et $R'' \exp i(-\omega t + \varphi'')$ ont même argument modulo 2π , donc si $\omega t + \varphi' = -\omega t + \varphi'' = \frac{1}{2}(\varphi' + \varphi'')$ ce qui donne la direction du grand axe. La longueur du grand axe est donc $\frac{1}{2}(R' + R'')$.
- Elle est minimum si les arguments de $R' \exp i(\omega t + \varphi')$ et $R'' \exp i(-\omega t + \varphi'')$ diffèrent de π modulo 2π , donc si $\omega t + \varphi' = -\omega t + \varphi'' + \pi = \frac{1}{2}(\varphi' + \varphi'' + \pi)$ ce qui donne la direction du petit axe. La longueur du petit axe est donc $\frac{1}{2}|R' - R''|$.

Solution 13.11. C'est un calcul. On peut le simplifier en supposant $M_4 = \infty$ et en mettant l'origine en M_1 . Pour justifier cette supposition, on peut considérer une homographie envoyant M_4 à l'infini et en s'appuyant sur la conservation des birapports par les homographies (13.15, p. 362).

Remarquons que l'on a ainsi toutes les valeurs possibles du birapport qui sont égales quatre à quatre. Les modifications provoquées par les permutations paires (resp. impaires) sont sur la première (resp. seconde) ligne.

Solution 13.12. On a les trois valeurs $\rho, 1 - \frac{1}{\rho}, \frac{1}{1-\rho}$ dont le produit est -1 et leurs inverses. Donc on a deux nombres positifs et un négatif parmi ces trois nombres et de même parmi les inverses.

Si $[A, B, C, D] < 0$, l'argument du birapport est π , d'où la congruence modulo 2π

$$(\widehat{BC}, \widehat{BD}) \equiv (\widehat{AC}, \widehat{AD}) + \pi \quad \text{modulo } 2\pi$$

ce qui donne le résultat.

Solution 13.13. Soit $\rho = [A, B, C, D]$. Par l'exercice 13.11, on a $\frac{1}{1-\rho} = [C, A, B, D]$ et $\frac{\rho}{\rho-1} = [B, C, A, D]$, d'où

$$\left| \frac{1}{1-\rho} \right| = |[C, A, B, D]| = \frac{BC \times DA}{DC \times BA}, \quad \left| \frac{\rho}{\rho-1} \right| = |[B, C, A, D]| = \frac{AC \times DB}{AB \times DC}$$

On a $1 = \left| \frac{1}{1-\rho} + \frac{\rho}{\rho-1} \right| \leq \left| \frac{1}{1-\rho} \right| + \left| \frac{\rho}{\rho-1} \right|$, d'où

$$AB \times CD \leq BC \times AD + AC \times BD$$

avec égalité si et seulement si $\frac{1}{1-\rho}$ et $\frac{\rho}{\rho-1}$ ont même argument modulo 2π . Le rapport étant $-\rho$, ceci équivaut à $\rho < 0$ et on peut utiliser l'exercice précédent.

Solution 13.14. 1) $[A, B, P, Q] = \frac{p-a}{p-b} : \frac{q-a}{q-b}$, donc le birapport est -1 si et seulement si $\frac{p-a}{p-b} + \frac{q-a}{q-b} = 0$. On obtient la formule en développant.

2) Plaçons l'origine en I. Alors $a + b = 0$ et $ab = -a^2 = -b^2$. Le quadrangle est harmonique si et seulement si $pq = +a^2 = +b^2$ ce qui équivaut aux deux relations

$$|p| \times |q| = |a|^2 = |b|^2$$

$$\arg(p) + \arg(q) \equiv 2\arg(a) \equiv 2\arg(b) \text{ modulo } 2\pi$$

C'est la traduction complexe de (ii). On a de même l'équivalence de (i) et (iii).

Solution 13.15. 1) Il est clair que $-j, -j^2$ vérifient la propriété $-j^2 - j = 1$, $|-j^2| = |-j| = 1$.

Inversement si ρ et ρ' de module 1 vérifient $\rho + \rho' = 1$, on a $\Im(\rho) + \Im(\rho') = 0$, donc, ou bien $\rho' = -\rho$, ou bien $\rho' = \bar{\rho} = \rho^{-1}$. La première éventualité n'est pas compatible avec $\rho + \rho' = 1$. Donc $\rho' = \rho^{-1} = 1 - \rho$, d'où $\rho^2 - \rho + 1 = 0$ ce qui donne le résultat.

2) Si les six valeurs du birapport sont de module 1, c'est le cas de ρ et $1 - \rho$, donc nécessairement ρ est $-j$ ou $-j^2$. On vérifie que c'est suffisant. Il n'y a plus que deux valeurs du birapport avec par exemple

$$\rho = 1 - \frac{1}{\rho} = \frac{1}{1 - \rho} = -j \quad \text{et} \quad \frac{1}{\rho} = \frac{\rho}{\rho - 1} = 1 - \rho = -j^2$$

3) Cette relation traduit que les birapports sont de module 1.

4) Si $[a, b, c, \infty] = -j$, alors $\frac{c-a}{c-b} = -j$. On passe donc de B à A par la rotation de centre C et d'angle $\arg(-j) \equiv -\frac{\pi}{3}$. Le triangle ABC est donc équilatéral direct et inversement.

5) Considérons le faisceau $\mathcal{F}(A, B)$ de cercles à points limites A, B. Par 13.11 (p. 358), D est sur le cercle C de $\mathcal{F}(A, B)$ passant par C. Il est de même sur le cercle A de $\mathcal{F}(B, C)$ passant par A et sur le cercle B de $\mathcal{F}(C, A)$ passant par D. S'il est sur deux de ces cercles, il est sur le troisième à cause de la question 1). Ces trois cercles appartiennent à un même faisceau à points de base D', D'' qui sont les points cherchés.

Soit $\Gamma = (ABC)$. Pour construire C , on prend la tangente en C à Γ qui coupe (AB) en le centre de C .

Solution 13.16. 1) On a $\frac{a'-b}{a'-c} = \frac{\bar{z}-\bar{b}}{\bar{z}-\bar{c}}$ et deux autres relations analogues. On en déduit

$$a' = \alpha(\bar{z} - \bar{b}) + b = \alpha(\bar{z} - \bar{c}) + c$$

$$b' = \beta(\bar{z} - \bar{c}) + c = \beta(\bar{z} - \bar{a}) + a$$

$$c' = \gamma(\bar{z} - \bar{a}) + a = \gamma(\bar{z} - \bar{b}) + b$$

2) La poursuite de ce calcul aboutit à

$$[a', b', c', \infty] = \frac{a' - c'}{b' - c'} = [\bar{a}, \bar{b}, \bar{c}, \bar{z}]$$

3) On a $M \in \Gamma$ si et seulement si $[A, B, C, M]$ est réel, donc si et seulement si $[A', B', C', \infty] = [A, B, C, M]$. On retrouve l'alignement de A', B', C' sur la droite de Steiner de M avec en plus la localisation exacte. Ainsi, C' est milieu de (A', B') si et seulement si (A, B, C, M) est harmonique.

4) De même, le triangle $A'B'C'$ est équilatéral si et seulement si (A, B, C, M) est équi-harmonique. *Un quadrangle est équi-harmonique si et seulement si les symétriques d'un sommet par rapport aux côtés du triangle formé par les trois autres, forment un triangle équilatéral.*

Solution 13.17. Soit A, B deux points distincts de transformés A', B' .

Soit un cercle C du faisceau $\mathcal{F}^\perp(A, B)$ à points de base A, B . Alors A, B étant sur C , A', B' sont sur $C' = h(C)$, donc $C' \in \mathcal{F}^\perp(A', B')$. On montre la réciproque en faisant le même raisonnement pour h^{-1} .

Soit Γ un cercle du faisceau à points-limites $\mathcal{F}(A, B)$. Alors Γ est orthogonal à tout $C \in \mathcal{F}^\perp(A, B)$, donc (13.18, p. 363) $\Gamma' = h(\Gamma)$ est orthogonal à tout $C' \in \mathcal{F}^\perp(A', B')$, donc appartient à $\mathcal{F}(A', B')$. On montre la réciproque en utilisant h^{-1} .

Soit $\mathcal{F}$ un faisceau singulier de cercles n'ayant en commun qu'un seul point A (un faisceau de droites parallèles si $A = \infty$). Pour toute paire C_1, C_2 de $\mathcal{F}$, $\{A\} = C_1 \cap C_2$, donc $\{h(A)\} = h(C_1) \cap h(C_2)$. Les cercles $h(C)$ sont tous tangents en $h(A)$ (ou sont des droites parallèles si $A = \infty$), donc sont dans un même faisceau singulier $\mathcal{F}'$. En utilisant h^{-1} , on montre que tout $C' \in \mathcal{F}'$ est du type $h(C)$ où $C \in \mathcal{F}$.

Solution 13.18. On sait (13.4.2., p. 364) que h a au moins un point fixe U . Alors A, B, U sont trois points fixes pour $h^2 = h \circ h$, d'où $h^2 = \text{Id}$ (13.15, p. 362) et h est une involution.

Solution 13.19. 1) Soit C un cercle de $\mathcal{F}(U, V)$ (resp. $\mathcal{F}^\perp(U, V)$). Pour tout M de C , on a

$$[U, V, M, h(M)] = -1,$$

donc $h(M) \in C$ car -1 est de module 1 (resp. réel), donc $h(C) = C$.

Inversement, soit C un cercle tel que $h(C) = C$.

Si $U \in C$, pour tout $M \in C$, $[U, V, M, h(M)] = -1$ est réel, donc $V \in C$ et $C \in \mathcal{F}^\perp(U, V)$. Supposons U, V non sur C . Le faisceau $\mathcal{F}(U, C)$ est stable et est à points-limites car il contient le cercle-point $\{U\}$ avec $U \notin C$. Comme $h(U) = U$, le deuxième point-limite est fixe par h , donc est V . Donc $C \in \mathcal{F}(U, V)$.

2) a) Supposons que $C \in \mathcal{F}^\perp(U, V)$. Pour tout M , on a $[U, V, M, h(M)] = -1$, donc Δ_M et les tangentes en U et V à C passent par le même point F ou sont parallèles (13.13, p. 358).

Inversement, soit C un cercle, F un point extérieur (éventuellement ∞). L'application associant à $M \in C$ le point M' où la droite (FM) recoupe C . C'est bien la restriction à C de l'involution de points fixes U, V , points de contact des tangentes à C issues de F .

b) Supposons $C \in \mathcal{F}(U, V)$. Soit $M \in C$, F le point où Δ_M coupe (UV) , Γ_M le cercle de $\mathcal{F}^\perp(U, V)$ passant par M . Il passe par U, V, M et $h(M)$. Soit O le centre de C et R le rayon. On a

$$\begin{aligned} OF^2 - R^2 &= C(F) = \overline{FM} \times \overline{Fh(M)} = \Gamma_M(F) \\ &= \overline{FU} \times \overline{FV} = (\overline{OU} - \overline{OF})(\overline{OV} - \overline{OF}) \\ \text{D'où} \quad \overline{OF}(\overline{OU} + \overline{OV}) &= R^2 + \overline{OU} \times \overline{OV} = R^2 + \Gamma_M(O) = 2R^2 \end{aligned}$$

Ceci montre que F est fixe.

Le point F est intérieur à C , sinon, (réciproque de a)) les points fixes U, V seraient les points de contact des tangentes à C issues de F et on serait dans le cas a). Le lecteur retrouvera cela en montrant que $OF < R$ (utiliser la division harmonique formée par U, V et les points où (UV) coupe C).

Solution 13.20. 1) On a $h_3^{-1} = (h_2 \circ h_1)^{-1} = h_1 \circ h_2$, d'où l'équivalence (i) $\Leftrightarrow$ (ii).

La conjuguée $h_2 \circ h_1 \circ h_2^{-1} = h_2 \circ h_1 \circ h_2$ est l'involution de points fixes $h_2(U_1)$ et $h_2(V_1)$. Pour que h_2 et h_1 commutent, il faut et il suffit que $h_2 \circ h_1 \circ h_2^{-1} = h_1$, donc que h_2 laisse stable $\{U_1, V_1\}$. Comme $h_2 \neq h_1$, U_1, V_1 ne sont pas fixes par h_2 . Donc $\{U_1, V_1\}$ est stable par h_2 si et seulement si h_2 échange U_1 et V_1 , d'où l'équivalence (i) $\Leftrightarrow$ (iii) et de même l'équivalence (i) $\Leftrightarrow$ (iv).

2) Si ces conditions sont remplies $h_1 \circ h_3 = h_1 \circ (h_1 \circ h_2) = h_2$. On voit facilement que l'on a un groupe commutatif avec $h_i \circ h_j = h_k$ pour i, j, k distincts.

3) Soit $Q = \{M_0, M_1, M_2, M_3\}$. Pour i, j, k distincts dans $1, 2, 3$, il existe une homographie h_i telle que (13.15, p. 362)

$$h_i(M_0) = M_i, \quad h_i(M_i) = M_0, \quad h_i(M_j) = M_k$$

Comme h_i échange M_0 et M_i , c'est une involution, (ex.13.18) donc h_i échange M_j et M_k . Par le théorème 13.15 (p. 362), l'action des h_i sur Q les définit entièrement. Par son action sur Q , le groupe $\{\text{Id}, h_1, h_2, h_3\}$ s'identifie au sous-groupe de Klein de S_4 .

Solution 13.21. 1) Le groupe des similitudes directes n'est autre que G_∞ . Pour toute homographie f telle que $f(A) = \infty$, on a $G_\infty = fG_A f^{-1}$.

2) Les translations sont les homographies ayant ∞ pour seul point fixe. Leurs conjuguées dans G_A sont donc les élations de seul point fixe A .

3) Soit h de deuxième point fixe B . Posons $O = f(B)$. Alors $s = f \circ h \circ f^{-1}$ est une similitude directe de centre O . Pour tout M , on a

$$\begin{aligned}\rho(h) &= [A, B, M, h(M)] = [f(A), f(B), f(M), f(h(M))] \\ &= [\infty, O, f(M), (f \circ h \circ f^{-1})(f(M))] = [\infty, O, M', s(M')]\end{aligned}$$

où $M' = f(M)$. Ceci n'est autre que le complexe associé à la similitude directe $s = f \circ h \circ f^{-1}$ dont l'expression complexe dans un repère orthonormé direct est $z \mapsto \rho(h)z + b$ où b est affixe de l'image $s(O)$. On en déduit le résultat.

On dit que $\rho(h)$ est le **birapport** de h . Ceci présente une ambiguïté car cette notion dépend de l'ordre que l'on met sur les points fixes, les birapports $[A, B, M, h(M)]$ et $[B, A, M, h(M)]$ étant inverses l'un de l'autre. Si on s'intéresse au groupe G_A , le point fixe A joue un rôle particulier et, convenant de le mettre en premier, l'ambiguïté disparaît.

Solution 13.22. 1) C'est une conséquence du théorème 13.15 : étant donnés deux triplets ordonnés de trois points distincts, il existe une unique homographie transformant l'un en l'autre.

2) Soit T et T' deux ensembles de trois points. Comme il existe six façons d'ordonner chaque ensemble, il existe six homographies transformant T en T' . Soit h l'une d'elles. On a $hG(T)h^{-1} = G(T')$, d'où le résultat.

3) Il existe une unique involution u_i laissant A_i fixe et échangeant A_j et A_k (i, j, k distincts). Le point de Frégier F_i de u_i est l'intersection de la tangente en A_i à Γ et de la droite $(A_j A_k)$ (ex.13.19). L'autre point fixe A'_i est point de contact de l'autre tangente à Γ issue de F_i .

4) On a donc $[A_j, A_k, A_i, A'_i] = -1$. Le groupe $G(T)$ opère sur $\{1, 2, 3\}$: pour $i \in \{1, 2, 3\}$ et $g \in G(T)$, on définit $g(i)$ par $g(A_i) = A_{g(i)}$. On a alors

$$\begin{aligned}-1 &= [A_j, A_k, A_i, A'_i] = [g(A_j), g(A_k), g(A_i), g(A'_i)] \\ &= [A_{g(j)}, A_{g(k)}, A_{g(i)}, g(A'_i)] \\ \text{et } -1 &= [A_{g(j)}, A_{g(k)}, A_{g(i)}, A'_{g(i)}]\end{aligned}$$

On en déduit $g(A'_i) = A'_{g(i)}$. Le groupe $G(T)$ laisse donc stable $T' = \{A'_1, A'_2, A'_3\}$. Il est clair que les ensembles T et T' jouent le même rôle l'un par rapport à l'autre.

5) Si $A_1 = \infty$, A'_1 est milieu de (A_2, A_3) , et A_2 (resp. A_3) est milieu de (A_3, A'_3) (resp. (A_2, A'_2)).

Solution 13.23. Soit un quadrangle harmonique $Q = \{M_0, M_1, M_2, M_3\}$. C'est une disposition particulière des **paires** de points $\{M_0, M_1\}$ et $\{M_2, M_3\}$ l'une par rapport à l'autre : on peut intervertir deux paires et intervertir les deux points d'une paire sans changer le birapport -1 égal à son inverse :

$$\begin{aligned} -1 &= [M_0, M_1, M_2, M_3] = [M_1, M_0, M_3, M_2] = [M_2, M_3, M_0, M_1] = [M_3, M_2, M_1, M_0] \\ &= [M_1, M_0, M_2, M_3] = [M_0, M_1, M_3, M_2] = [M_3, M_2, M_0, M_1] = [M_2, M_3, M_1, M_0] \end{aligned}$$

Il en résulte huit homographies laissant Q stable formant le groupe $G(Q)$. Si Q et Q' sont deux quadrangles harmoniques, il existe huit homographies transformant Q en Q' . Si h est l'une d'elles, $G(Q') = hG(Q)h^{-1}$.

Soit $Q = \{M_0, M_1, M_2, M_3\}$ un carré, M_0 et M_1 (resp. M_2 et M_3) étant deux sommets opposés. Le groupe cyclique d'ordre 4 des rotations laissant Q stable est un sous-groupe de $G(Q)$. En revanche, les quatre réflexions laissant Q stable sont des anti-homographies et n'appartiennent donc pas à $G(Q)$.

Cependant, l'involution de points fixes M_0, M_1 (resp. M_2, M_3) appartient à $G(Q)$ et opère sur Q comme la réflexion autour de la diagonale (M_0M_1) (resp. (M_2M_3)).

Soit Γ le cercle circonscrit au carré, M'_0, M'_1 les points de Γ où la tangente est parallèle à (M_0M_2) et (M_1M_3) . L'involution de points fixes M'_0, M'_1 appartient à $G(Q)$ et agit sur Q comme la réflexion autour de la médiatrice de $[M_0, M_2]$ et $[M_1, M_3]$. De même pour l'involution échangeant M_0, M_2 avec M_3, M_1 .

Les groupes $G(Q)$ sont donc diédraux.

Solution 13.24. Si on a $[M_0, M_1, M_2, M_3] = -j$, alors toute permutation paire de ces points conserve le birapport et toute permutation impaire le change en $-j^2$. La méthode est la même que pour l'exercice précédent. Si σ est une permutation paire de $\{0, 1, 2, 3\}$, on a $[M_0, M_1, M_2, M_3] = [M_{\sigma(0)}, M_{\sigma(1)}, M_{\sigma(2)}, M_{\sigma(3)}]$ et il existe une homographie unique h_σ telle que $h_\sigma(M_i) = M_{\sigma(i)}$. L'application $\sigma \mapsto h_\sigma$ est l'isomorphisme $\mathcal{A}_4 \rightarrow G(Q)$.

Enfin, il est clair que $hG(Q)h^{-1} = G(h(Q))$, donc ces groupes forment une classe de conjugaison de $H^+(\tilde{\mathcal{P}})$.

Index

A

- action de groupe 130
- action fidèle 131
- action transitive 131
- adjoint 164
- affinité orthogonale 336
- algorithme d'Euclide 28, 45
- angle 189–191
- angle inscrit 292
- annalagmatique 297, 355
- anneau 25
 - commutatif 25
 - de Boole 46
 - euclidien 40
 - factoriel 41, 63
 - intègre 28
 - principal 40
 - produit 35
 - quotient 31
- annulateur 35
- antidéplacement 226
- antihomographie 365
- application 3
- associé 36
- automorphisme intérieur 133
- axe focal 325
- axe radical 291
- axiome du choix 9

B

- barycentre 244
- base orthonormale 163
- birapport 217, 355
- birapport d'une homographie 380
- bissectrice 227
- bon ordre 9
- borne supérieure 9
- branches d'une hyperbole 318

C

- caractéristique d'un anneau 33
- cardinal 5
- centre 126, 290, 316
 - de gravité 247
- cercle
 - généralisé 297
 - orthoptique 347
 - principal 333
- cercle directeur 333
- Céva 219
- coefficient dominant 60
- cône isotrope 155
- congruence 13
- conique propre 319
- conjugaison 212
- contenu 63
- coordonnées 206

corps 28
algébriquement clos 62
des fractions 47
fini 75
critère d'Eisenstein 65, 75
critères d'irréductibilité 65
cube 274

D

décomposition
d'Iwasawa 197
de Dunford 97
LU 196
polaire 196
QR 196
degré total 68
demi-plan 251
dénombrable 5
déplacement 226
dilatation 179, 235
dimension 206
directrice 327
discriminant 74
diviseur de zéro 27
diviseurs élémentaires 108, 129
division euclidienne 31, 40, 60
division harmonique 217
droite
affine 207
d'Euler 225
de Simson 293
dualité 108

E

élation 364
ellipse 317
endomorphisme
autoadjoint 165
cyclique 105
normal 167
ensemble 1
inductif 10
quotient 11
vide 1
équation
d'un hyperplan 208
normale d'un cercle 288
réduite 317, 319
équipotent 5
espace
affine 204
hermitien 163
vectoriel euclidien 159

excentricité 327, 329
exposant 126

F

facteurs invariants 127
faisceau
conjugués 300
de cercles 299
harmonique 218
à points de base 294, 301
à points-limites 294, 300
famille 4
fonction
affine 250, 304
circulaire 304
circulaire normale 290
de Möbius 75
quadratique affine ou f.q.a. 315
forme bilinéaire 147
antisymétrique 150
non dégénérée 148
symétrique 150
forme définie 155
forme quadratique 155
forme sesquilinéaire hermitienne 161
foyer 327

G

groupe
circulaire 362, 366
commutatif 126
cyclique 121
de Klein 130
des angles 190
diédral 122, 126
du rectangle 130
linéaire 181
orthogonal 183
produit 124
spécial linéaire 181
spécial orthogonal 187

H

homographie 362
homothétie 183, 210
hyperbole 318
hyperplan affine 207

I

idéal 30
bilatère 30, 31
maximal 37
premier 37

- principal 37
- trivial 30
- idéaux
 - étrangers 39
- indicateur d'Euler 47
- indice 120
- inégalité de Cauchy-Schwarz 159
- inégalité de Minkowski 160
- intersection 2
- intérieur d'un triangle 253
- inversible 26
- inversion 296
- involution 365
- irréductible 36
- isotrope 153

L

- lemme
 - d'Euclide 14, 36
 - de Gauss 13, 44, 63
 - des noyaux 94

M

- masse 245
- matrice
 - compagnon 88
 - congruentes 149
 - hermitiennes 161
 - orthogonale 163, 188
 - unitaire 163
- maximal 9
- Ménélaüs 219
- milieu 205
- morphisme
 - d'anneaux 29
 - de Frobenius 34
 - de groupes 120
- multi-indice 68

N

- nilpotent 35
- normalisateur 134
- norme hermitienne 162
- norme subordonnée 110
- noyaux itérés 109

O

- orbite 132
- orientation 222
 - du plan 189
- origine 206
- orthocentre 224
- orthogonalité 151, 297

P

- parabole 319
- parallèles 208
- parallélogramme 205
- paramètre 329
- p.g.c.d. 14, 43, 45
- plan affine 207
- point de Frégier 367
- point pondéré 245
- polaire 323
- polynôme
 - antisymétrique 73
 - caractéristique 85
 - circulaire 288
 - cyclotomique 65, 74
 - minimal 85
 - primitif 63
 - symétrique 69
 - symétrique élémentaire 69
- polynômes
 - annulateurs 86
- p.p.c.m. 14, 43, 45
- premier 36
- produit
 - cartésien 3
 - d'une famille 4
 - mixte 267
 - scalaire 159
 - scalaire hermitien 161
 - vectorel 269
- projection stéréographique 304
- puissance 287

Q

- quadrangle harmonique 358
- quadrilatère complet 234
- quotient de sous-espaces 102

R

- racines multiples 61
- réduction de Frobenius 102
- réduction de Gauss 157
- réduction de Jordan 99
- réflexion 185, 187, 226
- régulier 27
- relation
 - binaire 3
 - d'ordre 8
 - d'équivalence 11
 - de Bezout 44
 - de Chasles 205
- repère
 - affine 206

barycentrique 247
représentation paramétrique 207
retournement 185, 195, 230
réunion 4
rotation 187

S

segment 221
semilinéarité 161
signature 157
similitude directe 232
simplexe 247
situation orthocentrique 225
sommes de Newton 73
sommet 325
sous-anneau 28
sous-espaces caractéristiques 96
sous-espaces cycliques 103
sous-groupe distingué 123
sous-groupes 119
spectre 86
stabilisateur 132
stathme 40

T

tangente 321
tétraèdre
 orthocentrique 273
 régulier 274
 équifacial 277
théorème
 chinois 38, 46
 d'Appolonius 294
 d'isomorphisme 33, 124
 de Bézout 13
 de Cantor 7

 de Cantor-Bernstein 6
 de Cayley 132
 de Cayley-Hamilton 86
 de correspondance 32, 124
 de d'Alembert-Gauss 62
 de Hadamard 93
 de Krull 39
 de Lagrange 120
 de Napoléon 353
 de Simson 293
 de Steinitz 63
 de Sylow 134
 de Sylvester 157
 de Wedderburn 66
 de Zermelo 9
 de Zorn 10, 39
spectral 166
transformation antisymétrique 195
transformation orthogonale 183
translation 204
transvection 179, 235
triangle direct 249
triangle orthique 225
trigonalisable 92

U

union 2
unité 26

V

valeur propre 85
valuation 44
vecteur propre 86
vectorialisation 205
vissage 230

Bibliographie

- [1] Michel ALESSANDRI. *Thèmes de Géométrie*. Dunod, 1999.
- [2] Marcel BERGER. *Géométrie*. Cedic-Nathan, 1973.
- [3] Claude CARREGA. *Théorie des corps*. Hermann, 1975.
- [4] H.S.M. COXETER. *Introduction to Geometry*. John Wiley Sons, 1989.
- [5] Jean de BIASI. *Mathématiques pour le CAPES et l'Agrégation Interne*. Ellipses, 1995.
- [6] Jean DELCOURT. *Théorie des Groupes*. Dunod, 2001.
- [7] Jean DELCOURT. *Problèmes de Mathématiques*. Dunod, 2004.
- [8] Michel DEMAZURE. *Cours d'Algèbre*. Cassini, 1997.
- [9] Alain BOUVIER et Denis RICHARD. *Groupes*. Hermann, 1968.
- [10] Marie-Noëlle GRAS et Georges GRAS. *Algèbre fondamentale, Arithmétique*. Ellipses, 2004.
- [11] Serge FRANCINO et Hervé GIANELLA. *Exercices de Mathématiques pour l'Agrégation*. Masson, 1993.
- [12] Jean-Marie ARNAUDIES et José BERTIN. *Groupes et Géométrie*. Ellipses, 1996.
- [13] H.S.M. COXETER et S.L. GREITZER. *Redécouvrons la Géométrie*. Dunod, 1971.
- [14] Jean FRESNEL. *Méthodes Modernes en Géométrie*. Hermann, 1996.
- [15] Rémi GOBLOT. *Algèbre linéaire*. Masson, 1995.
- [16] Rémi GOBLOT. *Thèmes de géométrie*. Masson, 1998.
- [17] Rémi GOBLOT. *Algèbre commutative*. Dunod, 2001.
- [18] Paul HALMOS. *Naive set theory*. Springer, 1986.
- [19] Rached MNEIMNE. *Éléments de Géométrie*. Cassini, 1997.
- [20] Jean-Marie MONIER. *Géométrie, MPSI, MP*. Dunod, 3^e édition, 2003.
- [21] Daniel PERRIN. *Cours d'algèbre*. Ellipses, 1996.
- [22] Marc ROGALSKI. *Carrefours entre Analyse, Algèbre et Géométrie*. Ellipses, 2001.
- [23] Romain VIDONNE. *Groupe circulaire, rotations et quaternions*. Ellipses, 2001.