

Algèbre linéaire 1

Licence 1 MIPI

2025 – 2026

Sommaire

Notations	3
Préambule	4
1 Nombres complexes	6
1.1 Introduction	6
1.1.a Équations cubiques et méthode de Cardan	6
1.1.b Résolution des cas irréductibles par Bombelli	7
1.1.c Unité imaginaire et nombres complexes	8
1.2 Forme algébrique	9
1.2.a Somme et représentation géométrique	10
1.2.b Produit et quotient de nombres complexes	11
1.2.c Conjugué et module d'un nombre complexe	13
1.3 Équations du 2 nd degré	15
1.3.a Racine carrée d'un nombre complexe	16
1.3.b Résolution des équations du 2 nd degré dans \mathbb{C}	17
1.4 Exponentielle et forme trigonométrique	19
1.4.a Exponentielle imaginaire	19
1.4.b Arguments d'un nombre complexe	21
1.4.c Exponentielle d'un nombre complexe	24
1.5 Applications à la trigonométrie	25
1.5.a Formules de linéarisation	25
1.5.b Formules d'angle multiple	26
1.5.c Factorisation de l'angle moitié	27
1.6 Racines n -ièmes	28
1.6.a Racines n -ièmes de l'unité	28
1.6.b Racines n -ièmes d'un nombre complexe	29
1.6.c Équations de degré n	30
2 Résolution de systèmes linéaires	31
2.1 Systèmes linéaires	31
2.2 Opérations élémentaires	32
2.3 Systèmes triangulaires	34
2.4 Systèmes échelonnés	35
2.5 Algorithme d'élimination de Gauss	37
2.6 Exemples	39
3 Espaces vectoriels	41
3.1 Vecteurs en géométrie euclidienne	41
3.1.a Définition géométrique des vecteurs	41
3.1.b Vecteurs, droites et plans de l'espace	43
3.1.c Vecteurs en dimension n	45

3.2	Structure d'espace vectoriel	47
3.2.a	Définition axiomatique	47
3.2.b	Exemples fondamentaux	49
3.3	Sous-espaces vectoriels	50
3.3.a	Caractérisation des sous-espaces vectoriels	50
3.3.b	Intersection de sous-espaces	52
3.3.c	Sous-espace engendré par une partie	52
3.4	Familles de vecteurs	54
3.4.a	Familles génératrices	54
3.4.b	Familles libres et familles liées	54
3.4.c	Base d'un espace vectoriel	57
3.5	Théorie de la dimension	59
3.5.a	Existence de bases	59
3.5.b	Dimension d'un espace vectoriel	60
3.6	Sommes de sous-espaces, sommes directes et supplémentaires	63
3.6.a	Sommes de sous-espaces	63
3.6.b	Somme directe de sous-espaces vectoriels	64
3.6.c	Supplémentaires d'un sous-espace vectoriel	65
4	Applications linéaires	68
4.1	Définition et premiers exemples	68
4.1.a	Définition des applications linéaires	68
4.1.b	Exemples d'applications linéaires	69
4.2	Noyau, image et équations linéaires	70
4.2.a	Noyau et image d'une application linéaire	71
4.2.b	Structure des solutions des équations linéaires	72
4.3	Applications linéaires et dimension	74
4.3.a	Image d'une famille libre, d'une famille génératrice, d'une base	74
4.3.b	Rang d'une application linéaire	74
4.3.c	Caractérisation d'une application linéaire	77
A	Notions générales	79
A.1	Théorie des ensembles	79
A.2	Relations binaires	81
A.3	Applications	83
A.4	Structures algébriques	84
B	Construction de \mathbb{C}	87
B.1	Objectif	87
B.2	Structure de corps	88
B.3	Sous-corps isomorphe à \mathbb{R}	90
B.4	Unité imaginaire et forme algébrique	91
	Index	92

Notations

Logique

\Rightarrow : implication (« si ..., alors ... »).

\Leftrightarrow : équivalence logique (« ... si et seulement si ... »).

\forall : quantificateur universel (« quel que soit ... »).

\exists : quantificateur existentiel (« il existe (au moins) un ... »).

$\exists!$: quantificateur existentiel unique (« il existe un unique ... »).

$a := b$: a est défini par b .

Théorie des ensembles

\emptyset : ensemble vide.

$x \in E$: x appartient à E .

$F \subset E$: F est inclus dans E .

$\mathcal{P}(E)$: ensemble des parties de E .

$E \cap F$: intersection de E et F .

$E \cup F$: réunion de E et F .

$E \setminus F$: ensemble E privé de F .

F^c : complémentaire de F .

$E \times F$: produit cartésien de E et F .

E^n : puissance cartésienne de E .

$\mathcal{F}(E, F)$: ensemble des applications de E dans F .

$\{x \in E \mid P(x)\}$: ensemble défini en compréhension.

$\{f(x) : x \in E\}$: ensemble défini par image directe.

Ensembles de nombres

\mathbb{N} : ensemble des entiers positifs ou nuls.

\mathbb{Z} : ensemble des entiers relatifs.

\mathbb{Q} : ensemble des nombres rationnels.

\mathbb{R} : ensemble des nombres réels.

\mathbb{C} : ensemble des nombres complexes.

A^* : ensemble des nombres non nuls de A , où $A \in \{\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}\}$.

$[a, b]$: intervalle réel fermé, c.-à-d. l'ensemble $\{x \in \mathbb{R} \mid a \leq x \leq b\}$.

$]a, b[$: intervalle réel ouvert, c.-à-d. l'ensemble $\{x \in \mathbb{R} \mid a < x < b\}$.

$\llbracket a, b \rrbracket$: intervalle d'entiers, c.-à-d. l'ensemble $\{n \in \mathbb{Z} \mid a \leq n \leq b\}$.

Préambule

Introduction à l'algèbre linéaire

L'algèbre linéaire trouve son origine dans l'étude des systèmes linéaires, en lien avec la géométrie analytique initiée au XVII^e siècle par René Descartes (1596-1650). Dans cette approche de la géométrie, les points de l'espace sont repérés par des coordonnées, tandis que les droites et les plans sont décrits par des équations linéaires. Déterminer l'intersection de droites et de plans se traduit alors par la résolution de systèmes d'équations linéaires. La géométrie analytique crée un pont entre la géométrie et l'algèbre : d'un côté, elle permet de résoudre des problèmes de géométrie en utilisant des méthodes algébriques; d'un autre côté, elle fournit une représentation géométrique de certains problèmes algébriques dits *linéaires*.

L'algèbre linéaire prend son essor au cours du XIX^e siècle avec l'invention par Carl Friedrich Gauss (1777-1855) d'une méthode de résolution systématique des systèmes linéaires, et avec la mise en lumière progressive de la structure d'*espace vectoriel*. Cette notion permet de généraliser les principes du calcul vectoriel à des espaces de dimensions quelconques, et même à des objets non géométriques comme les polynômes, les suites ou les fonctions. Les espaces vectoriels viennent remplacer les axiomes d'Euclide pour définir la géométrie (qu'elle soit euclidienne, affine ou projective).

L'algèbre linéaire est la branche des mathématiques qui étudie les problèmes linéaires et leur structure sous-jacente. Elle comprend l'étude des systèmes d'équations linéaires, des espaces vectoriels, des applications linéaires et du calcul matriciel. L'algèbre linéaire trouve des applications dans de nombreuses branches des mathématiques : géométrie, théorie des nombres, équations différentielles, analyse fonctionnelle, statistiques, etc. C'est un domaine incontournable pour les futurs mathématicien·nes, informaticien·nes, physicien·nes et ingénieur·es.

Ce cours débute par l'étude de l'ensemble des nombres complexes (chapitre 1), sans supposer de connaissances sur ce sujet à priori. Bien que l'étude des nombres complexes ne relève pas strictement de l'algèbre linéaire, les deux ne sont pas sans liens. De plus, ces nombres sont nécessaires dans de nombreux domaines des mathématiques, il est indispensable de les maîtriser. Le cours d'algèbre linéaire 1 proprement dit couvre la résolution des systèmes d'équations linéaires par la méthode du pivot de Gauss (chapitre 2), la définition des espaces vectoriels (chapitre 3) et l'étude des applications linéaires (chapitre 4). L'étude des matrices et de leur lien avec les applications linéaires et les systèmes linéaires sera l'objet du cours d'algèbre linéaire 2.

Organisation de ce document de cours

Le contenu du cours est divisé en différents items numérotés, répartis en 7 catégories : *définition*, *remarque*, *exemple*, *proposition*, *théorème*, *lemme* et *corollaire*. Rappelons le sens de ces quatre derniers termes :

- **Proposition** : nom générique pour désigner un résultat mathématique démontré¹.
- **Théorème** : proposition importante, centrale, d'une théorie mathématique.

1. Dans les textes mathématiques, le sens du mot « proposition » diffère de celui employé en logique. En logique, une proposition est un énoncé correctement formulé, pouvant être vrai ou faux.

- **Lemme** : proposition, souvent technique, qui vise à préparer la démonstration d'une autre proposition ou d'un théorème.
- **Corollaire** : proposition qui est la conséquence directe d'une proposition ou d'un théorème plus général ou plus important.

Les différents items de ce cours sont numérotés selon le chapitre et la section auxquels ils appartiennent. Par exemple, l'item « Proposition 2.4.8 » est une proposition se trouvant au chapitre 2, section 4, et constitue le 8^e item de cette section.

Expliquons enfin la signification des deux symboles \square et \blacktriangleleft employés dans ce document :

- Le symbole \square marque la fin d'une démonstration. C'est un symbole standard qui remplace le sigle CQFD (« ce qu'il fallait démontrer ») traditionnellement utilisé pour indiquer la fin d'une démonstration.
- Le symbole \blacktriangleleft marque la fin d'un item. Ce n'est pas un symbole standard. Dans ce cours, il sert à mieux délimiter les différents items et faciliter la lecture.

La mise en page de ce cours est inspirée du livre numérique *An Infinite Descent into Pure Mathematics* de Clive Newstead.

Modalités du contrôle des connaissances

L'évaluation de cette UE comporte un examen en janvier et du contrôle continu. Il y aura 3 devoirs sur table au cours du semestre sur les horaires de TD, seules vos 2 meilleures notes sont prises en compte. La note finale est le maximum entre l'examen seul, et la moyenne pondérée entre l'examen (coeff. 2) et le contrôle continu (coeff. 1) :

$$\text{Note finale} = \max\left(\text{CT}; \frac{2}{3} \text{CT} + \frac{1}{3} \text{CC}\right),$$

où CT (contrôle terminal) est la note de l'examen et CC est la moyenne des deux meilleurs devoirs.

Référence

- Grifone, J. (2019). *Algèbre linéaire* (6^e éd.). Cépaduès-Éditions.

Chapitre 1

Nombres complexes

1.1 Introduction

Au XVI^e siècle, l'étude des équations polynômiales a conduit les mathématicien·nes à introduire des nombres dont le carré est négatif. Ces nombres ne peuvent évidemment pas être des nombres réels, il s'agit d'un nouveau type de nombres qu'on appelle des nombres complexes. Aujourd'hui, ces nombres sont devenus un outil essentiel dans de nombreuses branches des mathématiques, mais aussi en physique et en ingénierie.

L'objectif de ce chapitre est d'introduire les nombres complexes et leur manipulation : on verra comment effectuer les opérations de bases sur ces nombres, ainsi que leur interprétation géométrique ; on s'intéressera à l'extraction des racines carrées des nombres complexes, et plus généralement des racines n -ièmes ; on étendra la définition de la fonction exponentielle aux nombres complexes et on étudiera ses applications à la trigonométrie.

1.1.a Équations cubiques et méthode de Cardan

L'idée de racines carrées de nombres négatifs apparaît au XVI^e siècle en Italie pour résoudre les équations cubiques, c'est-à-dire les équations polynomiales de degré 3. Pour comprendre pourquoi il a été nécessaire d'introduire de tels nombres, on va détailler sur deux exemples le fonctionnement de la méthode de résolution, appelée **méthode de Cardan**, du nom de Jérôme Cardan (1501-1576), le premier à l'avoir publiée¹.

Exemple 1.1.1. Considérons l'équation cubique suivante :

$$x^3 - 6x - 9 = 0. \quad (1.1)$$

Le principe de la méthode de Cardan est d'introduire deux variables u et v telles que $u + v = x$ pour remplacer x dans l'équation (1.1), puis d'ajouter une deuxième équation liant u et v afin de simplifier la première équation. On obtient ainsi un système de deux équations sur les variables u et v pouvant être résolu facilement. En substituant x par $u + v$ dans l'équation (1.1), on obtient :

$$\begin{aligned} (u + v)^3 - 6(u + v) - 9 = 0 &\iff u^3 + v^3 + 3u^2v + 3uv^2 - 6(u + v) - 9 = 0 && \text{(on développe le cube)} \\ &\iff u^3 + v^3 + 3uv(u + v) - 6(u + v) - 9 = 0 && \text{(on factorise par } 3uv\text{)} \\ &\iff u^3 + v^3 + (u + v)(3uv - 6) - 9 = 0. && \text{(on factorise par } u + v\text{)} \end{aligned}$$

En imposant la condition supplémentaire $3uv - 6 = 0$, l'équation ci-dessus se réduit à $u^3 + v^3 = 9$. De plus, la condition $3uv - 6 = 0$ est équivalente à $uv = 2$. En prenant le cube de cette équation, on


1. mais pas le premier à l'avoir découverte. En effet, cette méthode a d'abord été découverte par Scipione del Ferro (1465-1526), et plus tard par Tartaglia (1499-1557) de façon indépendante. Mais aucun ne la publie, préférant la garder secrète pour mieux en tirer avantage.

obtient finalement le système suivant pour u et v :

$$\begin{cases} u^3 + v^3 = 9 \\ u^3 v^3 = 8. \end{cases}$$

Ainsi, u^3 et v^3 sont solutions d'un système somme-produit, ce sont donc les racines du polynôme² :

$$X^2 - 9X + 8.$$

On vérifie facilement que 1 et 8 sont les deux racines de ce polynôme, donc $u^3 = 1$ et $v^3 = 8$, c'est-à-dire $u = 1$ et $v = 2$. Une³ solution de l'équation (1.1) est donc $x = 1 + 2 = 3$, comme on peut facilement le vérifier. 

Plus généralement, la méthode de Cardan permet de résoudre les équations cubiques de la forme :

$$x^3 + px + q = 0. \quad (1.2)$$

En écrivant x sous la forme $u + v$ et en imposant la condition $3uv + p = 0$, résoudre (1.2) revient à chercher des nombres u et v tels que leurs cubes u^3 et v^3 soient solutions du système somme-produit :

$$\begin{cases} u^3 + v^3 = -q \\ u^3 v^3 = -\frac{p^3}{27}, \end{cases} \quad (1.3)$$

ce qui revient à dire que u^3 et v^3 sont racines de $X^2 + qX - \frac{p^3}{27}$. Si le discriminant de ce polynôme est positif, alors l'équation (1.2) admet une solution réelle donnée par $u + v$ avec :

$$u := \sqrt[3]{-\frac{q}{2} + \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}, \quad v := \sqrt[3]{-\frac{q}{2} - \sqrt{\left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3}}.$$

En revanche, si le discriminant du polynôme est strictement négatif, alors le système (1.3) n'a pas de solutions réelles et il semble que la méthode de Cardan ne soit pas applicable. Ces équations cubiques qui donnent lieu à des systèmes somme-produit insolubles sont appelées des **cas irréductibles** de la méthode de Cardan.

On pourrait penser que les cas irréductibles sont des équations sans solutions. Pourtant, une équation cubique de la forme (1.2) admet nécessairement au moins une solution réelle. En effet, on a :

$$\lim_{x \rightarrow +\infty} x^3 + px + q = +\infty, \quad \lim_{x \rightarrow -\infty} x^3 + px + q = -\infty,$$

donc d'après le théorème des valeurs intermédiaires, la fonction $x \mapsto x^3 + px + q$ doit s'annuler au moins une fois sur \mathbb{R} . La méthode est donc incomplète : il y a des équations de degré 3 admettant des solutions que la méthode ne parvient pas à résoudre.

1.1.b Résolution des cas irréductibles par Bombelli

C'est à Raphaël Bombelli (1526-1572) qu'on doit la résolution des équations de degré 3 dans les cas irréductibles, grâce à l'utilisation de racines carrées de nombres négatifs. Illustrons sa démarche en reprenant l'exemple historique qu'il a étudié.

2. deux nombres ont pour somme s et pour produit p si et seulement si ils sont racines du polynôme $X^2 - sX + p$.
3. il s'avère que c'est l'unique solution, mais en général une équation cubique peut avoir jusqu'à trois solutions réelles.

Exemple 1.1.2. Considérons l'équation :

$$x^3 = 15x + 4. \quad (1.4)$$

Bombelli remarque que $x = 4$ est une solution de l'équation et va chercher à la retrouver à l'aide des formules de Cardan. Cette équation est de la forme $x^3 + px + q = 0$ avec $p = -15$ et $q = -4$. D'après la méthode de Cardan, on doit chercher la solution sous la forme $x = u + v$, où u et v sont tels que u^3 et v^3 soient racines du polynôme $X^2 - 4X + 125$. Or, le discriminant de ce polynôme est -484 , donc il n'admet pas de racines réelles : c'est un cas irréductible de la méthode de Cardan.

L'idée de Bombelli est d'appliquer les formules de Cardan coûte que coûte, même si des racines carrées de nombres négatifs apparaissent, et d'effectuer les calculs comme s'il s'agissait de nombres ordinaires. Formellement, les racines de $X^2 - 4X + 125$ s'écrivent :

$$\frac{4 + \sqrt{-484}}{2} \quad \text{et} \quad \frac{4 - \sqrt{-484}}{2},$$

où « $\sqrt{-484}$ » désigne un nombre dont le carré serait égal à -484 . Puisque $\sqrt{484} = 22$, on simplifie $\sqrt{-484}$ en $22\sqrt{-1}$, et les racines s'écrivent $2 + 11\sqrt{-1}$ et $2 - 11\sqrt{-1}$. La suite de la méthode consiste à déterminer u et v tels que :

$$u^3 = 2 + 11\sqrt{-1}, \quad v^3 = 2 - 11\sqrt{-1}.$$

Bombelli remarque⁴ que $u := 2 + \sqrt{-1}$ et $v := 2 - \sqrt{-1}$ conviennent. En effet, on a :

$$\begin{aligned} (2 + \sqrt{-1})^3 &= 2^3 + 3 \times 2^2 \times \sqrt{-1} + 3 \times 2 \times (\sqrt{-1})^2 + (\sqrt{-1})^3 \\ &= 8 + 12\sqrt{-1} + 6 \times (-1) + (\sqrt{-1})^2 \times \sqrt{-1} \\ &= 8 + 12\sqrt{-1} - 6 - \sqrt{-1} \\ &= 2 + 11\sqrt{-1}, \end{aligned}$$

et on vérifie de même que $v^3 = 2 - 11\sqrt{-1}$. Ainsi, u et v sont des racines cubiques de respectivement $2 + 11\sqrt{-1}$ et $2 - 11\sqrt{-1}$. Enfin, leur somme vaut :

$$u + v = 2 + \sqrt{-1} + 2 - \sqrt{-1} = 4,$$

ce qui est bien la solution de l'équation (1.4) qu'avait remarquée Bombelli. ◀

Ce que montre Bombelli avec cet exemple, c'est qu'il est possible d'utiliser la méthode de Cardan pour résoudre les équations cubiques dans les cas irréductibles, mais il faut pour cela accepter d'utiliser des racines carrées de nombres négatifs dans les calculs.

1.1.c Unité imaginaire et nombres complexes

À la suite de Bombelli, de plus en plus de mathématicien·nes ont utilisé des nombres comportant une racine carrée d'un nombre négatif, c'est-à-dire des nombres de la forme :

$$a + \sqrt{-b}, \quad a \in \mathbb{R} \text{ et } b \in \mathbb{R}_+.$$

Ces nombres sont baptisées **nombres imaginaires** par René Descartes (1586-1650), ce qui restera leur nom jusqu'au XIX^e siècle, où Carl Friedrich Gauss (1777-1855) les renomme **nombres complexes**. Cependant, l'écriture des racines carrées de nombres négatifs sous la forme « $\sqrt{-b}$ », en plus d'être perturbante, peut être source d'erreur. Par exemple, on peut être tenté d'écrire :

$$-1 = \sqrt{-1} \times \sqrt{-1} = \sqrt{(-1) \times (-1)} = \sqrt{1} = 1, \quad \text{absurde!}$$

4. Bombelli n'explique pas comment il est parvenu à ce résultat. En pratique, il n'est pas si évident de déterminer une racine cubique d'un nombre complexe donné.

L'erreur réside dans l'application de la formule $\sqrt{a} \times \sqrt{b} = \sqrt{ab}$, qui est fautive si a et b sont négatifs. Pour cette raison, il est préférable d'éviter la notation « $\sqrt{-b}$ » où $b \geq 0$, et de réserver le symbole⁵ « $\sqrt{}$ » aux nombres réels positifs, ce qu'on fera dans le reste de ce cours.

C'est Leonhard Euler (1707-1783) qui a proposé de remplacer « $\sqrt{-1}$ » par la lettre i , appelée **unité imaginaire**. Ainsi, les racines carrées d'un nombre négatif $-b$ s'écrivent $i\sqrt{b}$ et $-i\sqrt{b}$, c'est-à-dire comme des multiples réels de i . Le nombre i suffit donc pour écrire tous les nombres complexes, ceux-ci étant la somme d'un nombre réel et d'un multiple réel de i .

1.2 Forme algébrique

Commençons par définir de façon axiomatique les nombres complexes. Le théorème suivant affirme l'existence d'un ensemble contenant les nombres réels et une racine carrée de -1 . Il existe plusieurs façons de construire un tel ensemble à partir de l'ensemble \mathbb{R} , l'une d'entre elles est donnée dans l'annexe B. Mais les détails de ces constructions n'ont aucune importance pour décrire et manipuler les nombres complexes : les propriétés de ces nombres ne dépendent que des axiomes. On note \mathbb{C} n'importe laquelle de ces constructions.

Dans cette section, on décrit les opérations de base sur les nombres complexes, ainsi que leur représentation géométrique.

Théorème 1.2.1. Il existe un ensemble \mathbb{C} , appelé ensemble des **nombres complexes**, qui contient l'ensemble des nombres réels et qui vérifie les propriétés suivantes :

1. \mathbb{C} est muni d'une addition $+$ et d'une multiplication \times qui prolongent celles des nombres réels, qui ont les mêmes propriétés (associativité, commutativité, distributivité), et qui font de \mathbb{C} un corps⁶.
2. \mathbb{C} contient un élément noté i , appelé **unité imaginaire** et vérifiant l'égalité $i^2 = -1$.
3. Tout élément de \mathbb{C} s'écrit sous la forme $a + ib$ avec $a, b \in \mathbb{R}$.

Proposition 1.2.2. L'écriture d'un nombre complexe sous la forme $a + ib$ avec $a, b \in \mathbb{R}$ est unique. ◀

Démonstration. Soit $z \in \mathbb{C}$ et soient $a, b, a', b' \in \mathbb{R}$ tels que $z = a + ib$ et $z = a' + ib'$. Montrons que nécessairement $a = a'$ et $b = b'$.

On a $a + ib = a' + ib'$ donc $a - a' = i(b' - b)$. En élevant au carré, on obtient $(a - a')^2 = -(b' - b)^2$. Or le membre de gauche est positif tandis que le membre de droite est négatif, donc ils sont nuls tous les deux, c'est-à-dire $a = a'$ et $b = b'$. Par conséquent, l'écriture est unique. ◻

Définition 1.2.3. Soit $z \in \mathbb{C}$ et soient $a, b \in \mathbb{R}$ tels que $z = a + ib$.

- L'expression $a + ib$ est appelée la **forme algébrique** du nombre complexe z .
- Le réel a est appelé la **partie réelle** de z et on le note $\Re(z)$.
- Le réel b est appelé la **partie imaginaire** de z et on le note $\Im(z)$.
- Les nombres de la forme ib avec $b \in \mathbb{R}$, c'est-à-dire les nombres complexes dont la partie réelle est nulle, sont appelés des **nombres imaginaires purs**. L'ensemble des nombres imaginaires purs se note $i\mathbb{R}$.

Remarque 1.2.4. La partie imaginaire d'un nombre complexe z est un nombre réel : c'est le coefficient de i dans la forme algébrique de z . Par exemple, dans le nombre complexe $z := 2 - 3i$, on a $\Re(z) = 2$ et $\Im(z) = -3$ (pas de i dans la partie imaginaire!). ◀

5. ce symbole est appelé le **radical**.

6. Pour le dire rapidement, un **corps** est un ensemble dans lequel on peut effectuer les 4 opérations arithmétiques de base (addition, soustraction, multiplication et division) et où ces opérations ont les mêmes propriétés algébriques qu'on leur connaît pour les nombres réels. Pour une définition plus détaillée, consulter l'annexe A.

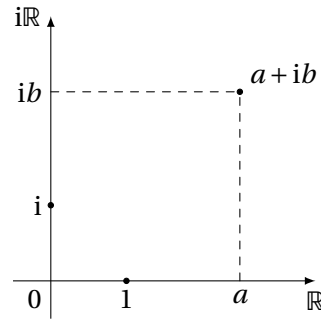


Figure 1.1 – Plan complexe

1.2.a Somme et représentation géométrique

L'addition des nombres complexes s'effectue en ajoutant entre elles les parties réelles et les parties imaginaires. Par exemple, la somme de $2 + 3i$ et $4 - 5i$ est :

$$(2 + 3i) + (4 - 5i) = 2 + 4 + 3i - 5i = 6 - 2i.$$

De même, la multiplication d'un nombre complexe par un nombre réel s'obtient par distributivité, par exemple :

$$-4 \times (2 - 3i) = -8 + 12i.$$

Proposition 1.2.5. Pour tous nombres complexes z et w , et pour tout réel λ , on a :

- (i) $\Re(z + w) = \Re(z) + \Re(w)$ et $\Im(z + w) = \Im(z) + \Im(w)$.
- (ii) $\Re(\lambda z) = \lambda \Re(z)$ et $\Im(\lambda z) = \lambda \Im(z)$. ◀

Démonstration. Soient $z = a + ib$ et $w = a' + ib'$ deux nombres complexes, et soit $\lambda \in \mathbb{R}$. Calculons la somme de z et w :

$$z + w = (a + ib) + (a' + ib') = a + ib + a' + ib' = (a + a') + i(b + b'),$$

d'où les égalités sur les parties réelles et imaginaires de $z + w$. Calculons maintenant le produit de λ et z :

$$\lambda z = \lambda(a + ib) = (\lambda a) + i(\lambda b),$$

d'où les égalités sur les parties réelles et imaginaires de λz . □

Géométriquement, on est habitué à représenter les nombres réels par les points d'une droite graduée, qu'on appelle la droite réelle. Les nombres complexes étant caractérisés par deux nombres réels, à savoir leur partie réelle et leur partie imaginaire, il est naturel de les représenter par les points d'un plan. On représente les nombres imaginaires purs par les points d'une seconde droite graduée, perpendiculaire à la droite réelle, qu'on appelle la droite des imaginaires purs. Ces droites forment un repère orthonormé du plan : la droite réelle est l'axe des abscisses et la droite des imaginaires purs est l'axe des ordonnées. Chaque nombre complexe z est alors représenté par le point du plan dont l'abscisse est la partie réelle de z et l'ordonnée est la partie imaginaire de z (cf. figure 1.1).

Définition 1.2.6. On munit le plan d'un repère orthonormé. À tout nombre complexe $z := a + ib$, on associe le point M de coordonnées (a, b) .

- Le point M est appelé l'**image** du nombre z , et le nombre z est appelé l'**affiche** de du point M .
- Le plan dans lequel les points représentent les nombres complexes est appelé le **plan complexe**.
- L'axe des abscisses est l'axe des réels et l'axe des ordonnées est l'axe des imaginaires purs.

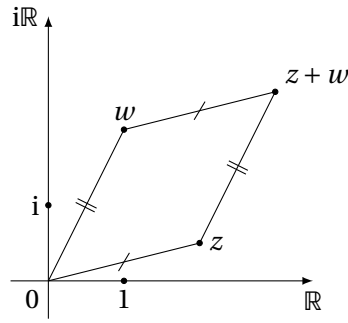


Figure 1.2 – Addition dans le plan complexe

On associe de même à tout complexe $a+ib$ le vecteur du plan de coordonnées $\begin{pmatrix} a \\ b \end{pmatrix}$. Ainsi, un nombre complexe peut être représenté aussi bien par un point que par un vecteur. Dans le cas des vecteurs, cette représentation s'étend aux opérations vectorielles⁷ : l'addition des nombres complexes correspond à l'addition vectorielle, et la multiplication des nombres complexes par des nombres réels correspond à la multiplication des vecteurs par des scalaires.

Proposition 1.2.7. On se place dans le plan complexe.

- (i) Pour tous vecteurs \vec{u} et \vec{v} d'affixes respectifs $z_{\vec{u}}$ et $z_{\vec{v}}$, leur somme $\vec{u} + \vec{v}$ a pour affixe :

$$z_{\vec{u}+\vec{v}} = z_{\vec{u}} + z_{\vec{v}}.$$

- (ii) Pour tout vecteur \vec{u} d'affixe $z_{\vec{u}}$ et pour tout réel λ , le vecteur $\lambda\vec{u}$ a pour affixe :

$$z_{\lambda\vec{u}} = \lambda z_{\vec{u}}.$$

- (iii) Pour tous points A et B d'affixes respectifs z_A et z_B , le vecteur \overrightarrow{AB} a pour affixe $z_B - z_A$. ◀

La démonstration est une simple vérification. Le fait que l'addition des nombres complexes correspondent à l'addition vectorielle fournit une interprétation géométrique de l'addition : dans le plan complexe, la somme de z et w est telle que les nombres 0 , z , $z+w$ et w forment un parallélogramme (cf. figure 1.2).

1.2.b Produit et quotient de nombres complexes

Pour multiplier deux nombres complexes, on procède par double distributivité, en n'oubliant pas que $i^2 = -1$. Par exemple, le produit de $3+2i$ et $4+3i$ est :

$$(3+2i)(4+3i) = 12+9i+8i-6 = 6+17i.$$

Attention, la partie réelle (resp. imaginaire) d'un produit n'est pas le produit des parties réelles (resp. imaginaires) des facteurs.

Proposition 1.2.8. Pour tous $z, w \in \mathbb{C}$, on a :

$$\Re(zw) = \Re(z)\Re(w) - \Im(z)\Im(w), \quad \Im(zw) = \Re(z)\Im(w) + \Im(z)\Re(w). \quad \blacktriangleleft$$

Démonstration. Soient $z := a+ib$ et $w := a'+ib'$ des nombres complexes, calculons leur produit :

$$\begin{aligned} zw &= (a+ib)(a'+ib') \\ &= aa' + iab' + ia'b + i^2bb' && \text{(distributivité)} \\ &= (aa' - bb') + i(ab' + a'b), && (i^2 = -1) \end{aligned}$$

d'où les égalités annoncées sur les parties réelles et imaginaires de zw . ◻

7. en utilisant le vocabulaire des chapitres 3 et 4, on dit que les nombres complexes forment un \mathbb{R} -espace vectoriel de dimension 2, et que l'application $a+ib \mapsto (a, b)$ est un isomorphisme entre \mathbb{C} et \mathbb{R}^2 .

L'addition et la multiplication des nombres complexes ayant les mêmes propriétés que celles des nombres réels, de nombreuses formules sont généralisables quand on passe des réels aux complexes. Par exemple, la multiplication étant distributive par rapport à l'addition, les identités remarquables restent vraies pour les nombres complexes :

$$\begin{aligned}(a+b)^2 &= a^2 + 2ab + b^2, \\ (a-b)^2 &= a^2 - 2ab + b^2, \\ a^2 - b^2 &= (a-b)(a+b).\end{aligned}$$

Plus généralement, la **formule du binôme de Newton** est encore valable dans \mathbb{C} . Ajoutons également la formule de factorisation de la différence de puissances n -ièmes, ainsi que la formule donnant la somme des puissances d'un nombre.

Proposition 1.2.9 (formule du binôme de Newton). Pour tous $a, b \in \mathbb{C}$ et pour tout $n \in \mathbb{N}$, on a :

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

où les nombres $\binom{n}{k} := \frac{n!}{k!(n-k)!}$ sont les coefficients binomiaux. ▶

Proposition 1.2.10 (factorisation de $a^n - b^n$). Pour tous $a, b \in \mathbb{C}$ et pour tout $n \in \mathbb{N}$, on a :

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

Proposition 1.2.11 (somme géométrique). Pour tout $q \in \mathbb{C} \setminus \{1\}$ et pour tout $n \in \mathbb{N}$, on a :

$$\sum_{k=0}^n q^k = \frac{1-q^{n+1}}{1-q}. \quad \text{▶}$$

Les démonstrations de ces formules sont identiques au cas réel.

La division est un peu plus délicate. Par définition, diviser par un nombre complexe non nul z est équivalent à multiplier par son **inverse**, c'est-à-dire le nombre complexe noté z^{-1} vérifiant l'égalité $z \times z^{-1} = 1$. Avant de savoir diviser, il faut savoir déterminer la forme algébrique de l'inverse d'un nombre complexe.

Proposition 1.2.12. L'inverse de tout nombre $z = a + ib \in \mathbb{C}^*$ est donné par :

$$z^{-1} = \frac{a - ib}{a^2 + b^2}.$$

Par conséquent, les parties réelles et imaginaires de l'inverse de z sont données par :

$$\Re(z^{-1}) = \frac{\Re(z)}{\Re(z)^2 + \Im(z)^2}, \quad \Im(z^{-1}) = \frac{-\Im(z)}{\Re(z)^2 + \Im(z)^2}. \quad \text{▶}$$

Démonstration. Soit $z := a + ib \in \mathbb{C}^*$, c'est-à-dire tel que $(a, b) \neq (0, 0)$. L'idée est de multiplier z par $a - ib$ et de reconnaître une identité remarquable :

$$(a + ib)(a - ib) = a^2 - (ib)^2 = a^2 + b^2.$$

En divisant par $a^2 + b^2$, qui est non nul car $(a, b) \neq (0, 0)$, on obtient :

$$z \times \frac{a - ib}{a^2 + b^2} = 1,$$

ce qui prouve que l'inverse de z est $\frac{a-ib}{a^2+b^2}$. □

Si z et w sont deux nombres complexes et si w est non nul, alors leur quotient est par définition :

$$\frac{z}{w} := z \times w^{-1}.$$

Une fois que l'on sait calculer la forme algébrique de l'inverse d'un nombre complexe, effectuer une division revient à calculer un produit, comme on l'a vu précédemment.

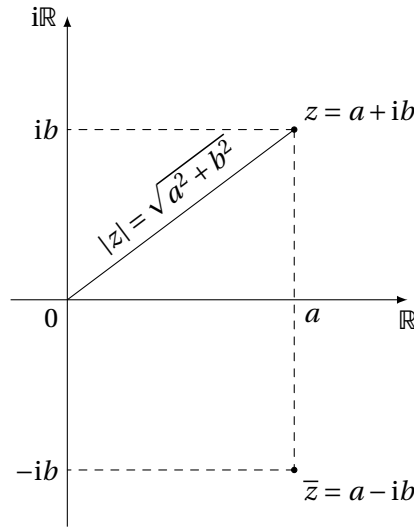


Figure 1.3 – Conjugaison et module dans le plan complexe

1.2.c Conjugué et module d'un nombre complexe

Le calcul de l'inverse d'un nombre complexe $a+ib$ fait apparaître deux quantités importantes : $a-ib$ et $a^2 + b^2$. Ces deux nombres associés à $a+ib$ jouent un rôle important dans l'étude des nombres complexes par la suite.

Définition 1.2.13. Soit $z := a+ib \in \mathbb{C}$.

- On appelle **conjugué** de z le nombre complexe défini par :

$$\bar{z} := \Re(z) - i\Im(z) = a - ib.$$

- On appelle **module** de z le nombre réel positif défini par :

$$|z| := \sqrt{\Re(z)^2 + \Im(z)^2} = \sqrt{a^2 + b^2}.$$

Le lien fondamental entre le conjugué et le module est la relation $z \times \bar{z} = |z|^2$, ce qui nous a permis de montrer que l'inverse de z est donné par $z^{-1} = \frac{\bar{z}}{|z|^2}$ (cf. proposition 1.2.12). Plus généralement, pour mettre sous forme algébrique le quotient de deux nombres complexes, on multiplie le numérateur et le dénominateur par le conjugué du dénominateur :

$$\forall z \in \mathbb{C}, \forall w \in \mathbb{C}^*, \quad \frac{z}{w} = \frac{z \times \bar{w}}{w \times \bar{w}} = \frac{z \times \bar{w}}{|w|^2}.$$

Par exemple, la forme algébrique du quotient de $2+3i$ par $1-4i$ est :

$$\frac{2+3i}{1-4i} = \frac{(2+3i)(1+4i)}{1^2+4^2} = \frac{2+8i+3i-12}{17} = -\frac{10}{17} + \frac{11}{17}i.$$

Remarque 1.2.14. Dans le plan complexe, le conjugué d'un nombre complexe z est son symétrique par rapport à l'axe des réels, et son module est égal à la distance entre 0 et z (cf. figure 1.3). Le module $|z|$ est aussi égal à la norme du vecteur du plan d'affixe z . Par conséquent, si A et B sont deux points du plan d'affixes respectifs z_A et z_B , alors $|z_B - z_A|$ est égal la distance entre A et B . En effet, le complexe $z_B - z_A$ est l'affixe du vecteur \overrightarrow{AB} , d'où :

$$|z_B - z_A| = \|\overrightarrow{AB}\| = AB. \quad \blacktriangleleft$$

Remarque 1.2.15. Le module d'un nombre réel coïncide avec sa valeur absolue⁸, c'est pourquoi on les note de la même façon. ◀

En considérant la somme et la différence d'un nombre complexe avec son conjugué, on trouve les relations suivantes avec la partie réelle et la partie imaginaire.

Proposition 1.2.16. Pour tout $z \in \mathbb{C}$, sa partie réelle et sa partie imaginaire sont données par :

$$\Re(z) = \frac{z + \bar{z}}{2}, \quad \Im(z) = \frac{z - \bar{z}}{2i}. \quad \blacktriangleleft$$

Démonstration. Soit $z := a + ib \in \mathbb{C}$, son conjugué s'écrit $\bar{z} = a - ib$. Calculons leur somme et leur différence :

$$\begin{aligned} z + \bar{z} &= (a + ib) + (a - ib) = 2a, \\ z - \bar{z} &= (a + ib) - (a - ib) = 2ib. \end{aligned}$$

On en déduit que $a = \frac{z + \bar{z}}{2}$ et que $b = \frac{z - \bar{z}}{2i}$. ◻

Ces relations permettent de caractériser les nombres réels et les nombres imaginaires purs à partir de leur conjugué : les nombres réels sont les nombres complexes égaux à leur conjugué, et les nombres imaginaires purs sont les nombres complexes opposés à leur conjugué.

Corollaire 1.2.17. Pour tout $z \in \mathbb{C}$, on a :

$$(i) \quad z \in \mathbb{R} \iff \Im(z) = 0 \iff z = \bar{z}. \quad (ii) \quad z \in i\mathbb{R} \iff \Re(z) = 0 \iff z = -\bar{z}. \quad \blacktriangleleft$$

Étudions maintenant les propriétés de la conjugaison et du module vis-à-vis des opérations (addition, multiplication, division) dans \mathbb{C} . Commençons par la conjugaison. Celle-ci a la propriété remarquable de commuter avec les opérations de \mathbb{C} : le conjugué d'une somme est la somme des conjugués, le conjugué d'un produit est le produit des conjugués, etc. On dit que la conjugaison est un automorphisme de corps de \mathbb{C} .

Proposition 1.2.18. Pour tous $z, w \in \mathbb{C}$ et pour tout $\lambda \in \mathbb{R}$, on a :

$$\begin{aligned} (i) \quad \bar{\bar{z}} &= z. & (iv) \quad \overline{z \times w} &= \bar{z} \times \bar{w}. \\ (ii) \quad \overline{z + w} &= \bar{z} + \bar{w}. \\ (iii) \quad \overline{\lambda z} &= \lambda \bar{z}. & (v) \quad \overline{z / w} &= \bar{z} / \bar{w} \text{ si } w \neq 0. \end{aligned} \quad \blacktriangleleft$$

La démonstration est une simple vérification. Poursuivons avec les propriétés du module. Celui-ci hérite du bon comportement de la conjugaison par rapport à la multiplication et à la division : le module d'un produit (resp. d'un quotient) est le produit (resp. le quotient) des modules.

Proposition 1.2.19. Pour tous $z, w \in \mathbb{C}$, on a :

$$\begin{aligned} (i) \quad |\bar{z}| &= |z|. & (iii) \quad |z / w| &= |z| / |w| \text{ si } w \neq 0. \\ (ii) \quad |z \times w| &= |z| \times |w|. & (iv) \quad |z| = 0 &\iff z = 0. \end{aligned} \quad \blacktriangleleft$$

Démonstration. Soit $z := a + ib \in \mathbb{C}$ avec $a, b \in \mathbb{R}$, et soit $w \in \mathbb{C}$.

(i) On a $|\bar{z}| = \sqrt{a^2 + (-b)^2} = \sqrt{a^2 + b^2} = |z|$.

(ii) On calcule le carré du module à l'aide du conjugué :

$$|zw|^2 = zw \times \overline{zw} = z\bar{z} \times w\bar{w} = |z|^2 \times |w|^2.$$

En prenant la racine carrée (les modules sont positifs), on obtient $|zw| = |z| \times |w|$.

(iii) La propriété du quotient se démontre de la même façon.

(iv) On a $|z| = 0 \iff a^2 + b^2 = 0 \iff a = b = 0 \iff z = 0$. ◻

8. on rappelle que pour tout $a \in \mathbb{R}$, $\sqrt{a^2} = |a|$.

En revanche, le module d'une somme n'est pas la somme des modules, mais on a tout de même la majoration du premier par la deuxième : c'est ce qu'on appelle l'**inégalité triangulaire**.

Proposition 1.2.20 (inégalités triangulaires). Pour tous $z, w \in \mathbb{C}$ on a :

- (i) $|\Re(z)| \leq |z|$ et $|\Im(z)| \leq |z|$.
- (ii) $|z + w| \leq |z| + |w|$.
- (iii) $||z| - |w|| \leq |z \pm w| \leq |z| + |w|$. ▶

Démonstration. Soient z et w des nombres complexes.

(i) Pour tous $a, b \in \mathbb{R}$, on a l'inégalité $|a| \leq \sqrt{a^2 + b^2}$. On en déduit les inégalités $|\Re(z)| \leq |z|$ et $|\Im(z)| \leq |z|$.

(ii) Calculons le carré du module de $z + w$ à l'aide du conjugué :

$$\begin{aligned}
 |z + w|^2 &= (z + w)(\overline{z + w}) \\
 &= z\overline{z} + z\overline{w} + w\overline{z} + w\overline{w} \\
 &= |z|^2 + 2\Re(z\overline{w}) + |w|^2 && (Z + \overline{Z} = 2\Re(Z)) \\
 &\leq |z|^2 + 2|z\overline{w}| + |w|^2 && (|\Re(Z)| \leq |Z|) \\
 &\leq |z|^2 + 2|z||w| + |w|^2 && (|\overline{w}| = |w|) \\
 &= (|z| + |w|)^2.
 \end{aligned}$$

En prenant la racine carrée de chaque côté (les modules sont des réels positifs), on obtient l'inégalité triangulaire $|z + w| \leq |z| + |w|$.

(iii) Pour démontrer l'inégalité de droite, on applique l'inégalité triangulaire aux nombres z et $-w$:

$$|z - w| \leq |z| + |-w| = |z| + |w|.$$

Pour montrer l'inégalité de gauche, on commence par appliquer l'inégalité triangulaire aux nombres $z - w$ et w :

$$|z| = |z - w + w| \leq |z - w| + |w| \implies |z - w| \geq |z| - |w|.$$

On démontre de la même façon que $|z - w| \geq |w| - |z|$. Par conséquent, on a :

$$|z - w| \geq \max(|z| - |w|, |w| - |z|) = ||z| - |w||.$$

Ainsi, on a montré que $||z| - |w|| \leq |z - w| \leq |z| + |w|$. En remplaçant w par $-w$, on obtient :

$$||z| - |w|| \leq |z + w| \leq |z| + |w|,$$

puisque $|-w| = |w|$. □

1.3 Équations du 2nd degré

Maintenant qu'on connaît mieux les opérations de base dans \mathbb{C} , on va s'intéresser à des opérations plus élaborées : l'extraction de racines n -ièmes et l'exponentielle. Commençons par la recherche de racines carrées et la résolutions d'équations du 2nd degré, puisque ce sont ces questions qui ont mené à l'invention des nombres complexes initialement.

1.3.a Racine carrée d'un nombre complexe

Définition 1.3.1. On appelle **racine carrée** d'un nombre complexe z tout nombre complexe w tel que $w^2 = z$.

La raison d'être des nombres complexes est de donner des racines carrées aux nombres réels négatifs. Il s'avère que les nombres complexes eux-mêmes admettent des racines carrées dans \mathbb{C} .

Proposition 1.3.2. Tout nombre complexe non nul admet exactement deux racines carrées opposées dans \mathbb{C} . ▶

Démonstration. Soit $z := a + ib \in \mathbb{C}$. On donne ici la méthode permettant de calculer la forme algébrique des racines carrées de z . On verra à la section 1.6 une autre méthode pour obtenir les racines carrées (et plus généralement les racines n -ièmes) de z .

On cherche les racines carrées de z sous forme algébrique $w := x + iy$. Puisque $w^2 = x^2 - y^2 + 2xyi$, résoudre l'équation $w^2 = z$ revient à résoudre le système d'inconnues $(x, y) \in \mathbb{R}^2$:

$$\begin{cases} x^2 - y^2 = a \\ 2xy = b. \end{cases}$$

Une façon de résoudre ce système est de lui ajouter l'équation redondante $|w|^2 = |z|$, c'est-à-dire $x^2 + y^2 = r$ avec $r := \sqrt{a^2 + b^2}$ le module de z . On fait ainsi apparaître un système linéaire en (x^2, y^2) , qu'on résout par combinaisons :

$$\begin{cases} x^2 - y^2 = a \\ x^2 + y^2 = r \\ 2xy = b \end{cases} \iff \begin{cases} x^2 = \frac{1}{2}(r + a) & L_1 + \frac{1}{2}(L_1 + L_2) \\ y^2 = \frac{1}{2}(r - a) & L_2 - \frac{1}{2}(L_1 + L_2) \\ 2xy = b. \end{cases}$$

En prenant les racines carrées⁹ dans les deux premières équations, on trouve deux valeurs possibles opposées pour x , de même pour y , soit quatre valeurs possibles pour le couple (x, y) . C'est l'équation $2xy = b$ qui permet de trancher :

- Si $b \geq 0$, alors x et y sont de même signe, donc les deux solutions sont :

$$\begin{cases} x = \sqrt{\frac{1}{2}(r + a)} \\ y = \sqrt{\frac{1}{2}(r - a)} \end{cases} \quad \text{et} \quad \begin{cases} x = -\sqrt{\frac{1}{2}(r + a)} \\ y = -\sqrt{\frac{1}{2}(r - a)}. \end{cases}$$

- Si $b \leq 0$, alors x et y sont de signe opposé, donc les deux solutions sont :

$$\begin{cases} x = \sqrt{\frac{1}{2}(r + a)} \\ y = -\sqrt{\frac{1}{2}(r - a)} \end{cases} \quad \text{et} \quad \begin{cases} x = -\sqrt{\frac{1}{2}(r + a)} \\ y = \sqrt{\frac{1}{2}(r - a)}. \end{cases}$$

Dans tous les cas, on trouve deux solutions opposées. □

Remarque 1.3.3. Le radical « $\sqrt{}$ » ne s'utilise que pour désigner la racine carrée d'un nombre réel positif, on ne l'emploie jamais pour désigner une racine carrée d'un nombre complexe quelconque. En effet, si x est un réel positif alors il admet deux racines carrées : l'une est positive et l'autre est négative. On désigne par \sqrt{x} la racine carrée positive de x , l'autre étant $-\sqrt{x}$.

9. puisque $|a| \leq \sqrt{a^2 + b^2}$, les nombres $\sqrt{a^2 + b^2} \pm a$ sont positifs et on peut extraire leurs racines carrées dans \mathbb{R} .

En revanche, si $z \in \mathbb{C} \setminus \mathbb{R}_+$, alors ses racines carrées ne peuvent pas être distinguées aussi facilement (il n'y a pas de relation d'ordre naturelle¹⁰ dans \mathbb{C}). On ne saurait donc pas laquelle des racines doit être notée « \sqrt{z} » et laquelle doit être notée « $-\sqrt{z}$ ».

Exemple 1.3.4. Déterminons les racines carrées de $z := -8 - 6i$. Soit $w := x + iy$ tel que $w^2 = z$, alors (x, y) est solution de :

$$(x + iy)^2 = -8 - 6i \iff \begin{cases} x^2 - y^2 = -8 \\ 2xy = -6. \end{cases}$$

On ajoute l'équation $|w|^2 = |z|$, c'est-à-dire $x^2 + y^2 = 10$:

$$\begin{cases} x^2 - y^2 = -8 \\ x^2 + y^2 = 10 \\ 2xy = -6. \end{cases} \iff \begin{cases} x^2 = 1 \\ y^2 = 9 \\ 2xy = -6. \end{cases}$$

Puisque $xy < 0$, les nombres x et y sont de signe opposé, donc les solutions sont $(1, -3)$ et $(-1, 3)$. Les racines carrées de z sont donc $1 - 3i$ et $-1 + 3i$.

1.3.b Résolution des équations du 2nd degré dans \mathbb{C}

Nous pouvons à présent donner les formules de résolution des équations du 2nd degré à coefficients complexes. Ces formules généralisent naturellement celles vues au lycée dans le cas réel, à la différence qu'il existe toujours des solutions dans \mathbb{C} .

Théorème 1.3.5. Soient a, b et c des nombres complexes avec $a \neq 0$. Considérons l'équation du 2nd degré d'inconnue z :

$$az^2 + bz + c = 0, \tag{1.5}$$

et notons $\Delta := b^2 - 4ac$ son discriminant.

- Si $\Delta \neq 0$, alors (1.5) admet deux solutions distinctes dans \mathbb{C} :

$$z_1 = \frac{-b - \delta}{2a}, \quad z_2 = \frac{-b + \delta}{2a},$$

où δ une racine carrée de Δ .

- Si $\Delta = 0$, alors (1.5) admet une seule solution dans \mathbb{C} :

$$z_0 = -\frac{b}{2a},$$

appelée **racine double** du polynôme $az^2 + bz + c$.

Démonstration. Écrivons le polynôme $az^2 + bz + c$ sous forme canonique :

$$\begin{aligned} az^2 + bz + c &= a \left(z^2 + \frac{b}{a}z + \frac{c}{a} \right) \\ &= a \left[\left(z + \frac{b}{2a} \right)^2 - \frac{b^2 - 4ac}{4a^2} \right] \\ &= a \left[\left(z + \frac{b}{2a} \right)^2 - \frac{\Delta}{4a^2} \right]. \end{aligned}$$

10. dans \mathbb{C} , il n'existe pas de relation d'ordre totale qui prolonge celle sur \mathbb{R} et qui soit compatible avec la multiplication. En effet, supposons par l'absurde qu'il en existe une : si $i \geq 0$ pour cette relation, alors en multipliant par i chaque membre, on aurait $i^2 \geq 0$, c'est-à-dire $-1 \geq 0$, absurde ! Si au contraire $i \leq 0$ pour cette relation, alors $i^2 \geq 0$ (le sens de l'inégalité change car on multiplie par $i \leq 0$), et obtient une nouvelle fois $-1 \geq 0$, absurde !

Soit δ une racine carrée de Δ , alors $\frac{\Delta}{4a^2} = \left(\frac{\delta}{2a}\right)^2$. On a donc :

$$\begin{aligned} az^2 + bz + c &= a \left[\left(z + \frac{b}{2a} \right)^2 - \left(\frac{\delta}{2a} \right)^2 \right] \\ &= a \left[\left(z + \frac{b}{2a} - \frac{\delta}{2a} \right) \left(z + \frac{b}{2a} + \frac{\delta}{2a} \right) \right] \\ &= a \left(z - \frac{-b+\delta}{2a} \right) \left(z - \frac{-b-\delta}{2a} \right). \end{aligned}$$

Maintenant que nous avons factorisé le polynôme $az^2 + bz + c$, nous pouvons résoudre l'équation (1.5) :

$$\begin{aligned} az^2 + bz + c = 0 &\iff a \left(z - \frac{-b+\delta}{2a} \right) \left(z - \frac{-b-\delta}{2a} \right) = 0 && \text{(équation produit nul)} \\ &\iff z = \frac{-b+\delta}{2a} \quad \text{ou} \quad z = \frac{-b-\delta}{2a}. \end{aligned}$$

Les solutions de l'équation (1.5) sont donc $\frac{-b+\delta}{2a}$ et $\frac{-b-\delta}{2a}$. Elles sont distinctes si $\delta \neq 0$ (donc si $\Delta \neq 0$) et elles sont égales si $\delta = 0$ (donc si $\Delta = 0$). \square

Exemple 1.3.6. Considérons l'équation :

$$z^2 + (1-i)z + (2+i) = 0.$$

Son discriminant vaut :

$$\Delta = (1-i)^2 - 4(2+i) = -2i - 8 - 4i = -8 - 6i.$$

On a calculé à l'exemple 1.3.4 que les racines carrées de Δ sont $1-3i$ et $-1+3i$. Les solutions de l'équation sont donc :

$$z_1 = \frac{-1+i-1+3i}{2} = -1+2i, \quad z_2 = \frac{-1+i+1-3i}{2} = -i. \quad \blacktriangleleft$$

Le théorème 1.3.5 nous permet d'approfondir la résolution des équations du 2nd degré à coefficients réels : dans le cas d'un discriminant négatif, il existe maintenant des solutions dans \mathbb{C} .

Corollaire 1.3.7. Soit a , b et c des nombres réels avec $a \neq 0$. Considérons l'équation du 2nd degré d'inconnue x :

$$ax^2 + bx + c = 0, \tag{1.6}$$

et notons $\Delta := b^2 - 4ac$ son discriminant.

- Si $\Delta > 0$, alors (1.6) admet deux solutions distinctes dans \mathbb{R} :

$$x_1 = \frac{-b - \sqrt{\Delta}}{2a}, \quad x_2 = \frac{-b + \sqrt{\Delta}}{2a}.$$

- Si $\Delta = 0$, alors (1.6) admet une seule solution dans \mathbb{R} :

$$x_0 = -\frac{b}{2a},$$

appelée **racine double** du polynôme $ax^2 + bx + c$.

- Si $\Delta < 0$, alors (1.6) n'admet pas de solutions dans \mathbb{R} , mais elle admet deux solutions conjuguées dans \mathbb{C} :

$$z_1 = \frac{-b - i\sqrt{|\Delta|}}{2a}, \quad z_2 = \frac{-b + i\sqrt{|\Delta|}}{2a}. \quad \blacktriangleleft$$

Exemple 1.3.8. Reprenons l'équation $x^2 - 4x + 125 = 0$ qui apparaissait lors de la résolution de l'équation cubique étudiée par Bombelli (exemple 1.1.2). Le discriminant de ce polynôme est :

$$\Delta = (-4)^2 - 4 \times 1 \times 125 = -484.$$

Il est négatif, donc l'équation admet deux solutions complexes conjuguées :

$$z_1 = \frac{4 + i\sqrt{484}}{2} = 2 + 11i, \quad z_2 = \frac{4 - i\sqrt{484}}{2} = 2 - 11i.$$

1.4 Exponentielle et forme trigonométrique

Si la forme algébrique est bien adaptée au calcul d'additions et de multiplications par des réels, les calculs de multiplications et de divisions entre nombres complexes sont pénibles, et leur sens géométrique nous échappe. Le calcul de puissances est particulièrement inefficace. En effet, d'après la formule du binôme de Newton, la puissance n -ième d'un nombre complexe $z := a + ib$ est :

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

Les parties réelles et imaginaires de z^n sont alors des expressions polynomiales en a et b difficiles à analyser. Calculer une racine n -ième d'un nombre complexe à partir de sa forme algébrique est alors très difficile, voire impossible, au-delà du cas $n = 2$ (racines carrées). Si on souhaite parvenir à calculer des racines n -ièmes de nombres complexes, il va nous falloir une meilleure façon de calculer et de comprendre la multiplication dans \mathbb{C} .

1.4.a Exponentielle imaginaire

Définition 1.4.1. On note \mathbb{U} l'ensemble des nombres complexes de module 1 :

$$\mathbb{U} = \{z \in \mathbb{C} \mid |z| = 1\}.$$

Remarque 1.4.2. Dans le plan complexe, cet ensemble est le cercle de centre 0 et de rayon 1, c'est-à-dire le **cercle trigonométrique**.

Cette remarque est fondamentale pour caractériser les nombres complexes de module 1. Puisque ceux-ci sont situés sur le cercle trigonométrique, leur partie réelle et leur partie imaginaire sont par définition le cosinus et le sinus de l'angle orienté (mesuré en radians) qu'ils forment avec la demi droite des réels positifs (cf. figure 1.4). Un élément de \mathbb{U} est donc repéré par un réel θ . Enfin, ce réel n'est pas unique puisque deux angles dont les mesures diffèrent d'un multiple de 2π repère le même point sur le cercle trigonométrique.

Proposition 1.4.3. Pour tout $z \in \mathbb{C}$, on a :

$$z \in \mathbb{U} \iff \exists \theta \in \mathbb{R}, z = \cos(\theta) + i\sin(\theta).$$

De plus, le réel θ est unique à 2π près, c'est-à-dire que pour tous réels θ et θ' , on a :

$$\cos(\theta) + i\sin(\theta) = \cos(\theta') + i\sin(\theta') \iff \exists k \in \mathbb{Z}, \theta = \theta' + 2k\pi.$$

Définition 1.4.4. Lorsque deux nombres θ et θ' sont égaux à 2π près, on dit que θ est **congru à θ' modulo 2π** . On note cette relation :

$$\theta \equiv \theta' \pmod{2\pi} \stackrel{\text{def.}}{\iff} \exists k \in \mathbb{Z}, \theta = \theta' + 2k\pi.$$

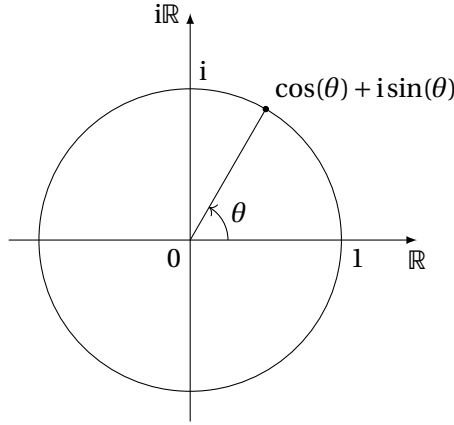


Figure 1.4 – Cercle trigonométrique \mathbb{U} dans le plan complexe.

L'écriture des nombres complexes de module 1 en termes de cosinus et de sinus rend le calcul des produits simple, du fait des formules d'addition de cosinus et sinus :

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

En effet, considérons la fonction $f(\theta) := \cos(\theta) + i \sin(\theta)$, alors pour tous $\theta, \theta' \in \mathbb{R}$, le produit de $f(\theta)$ et $f(\theta')$ est :

$$\begin{aligned}f(\theta) \times f(\theta') &= (\cos(\theta) + i \sin(\theta))(\cos(\theta') + i \sin(\theta')) \\ &= [\cos(\theta) \cos(\theta') - \sin(\theta) \sin(\theta')] + i[\cos(\theta) \sin(\theta') + \sin(\theta) \cos(\theta')] \\ &= \cos(\theta + \theta') + i \sin(\theta + \theta') \\ &= f(\theta + \theta'),\end{aligned}\tag{1.7}$$

où (1.7) est le résultat des formules d'addition des fonctions trigonométriques. La fonction f a donc la propriété de transformer les sommes en produits. Or dans \mathbb{R} , on sait que cette propriété est caractéristique des fonctions exponentielles. On peut donc vouloir écrire la fonction f sous la forme d'une exponentielle, la question étant : de quelle nombre ? Pour le savoir, calculons la dérivée de f :

$$\begin{aligned}f'(\theta) &= \cos'(\theta) + i \sin'(\theta) \\ &= -\sin(\theta) + i \cos(\theta) \\ &= i^2 \sin(\theta) + i \cos(\theta) \\ &= i(\cos(\theta) + i \sin(\theta)) \\ &= i f(\theta).\end{aligned}\tag{1.9}$$

On voit que f satisfait l'équation différentielle $f' = i f$. De plus, $f(0) = \cos(0) + i \sin(0) = 1$, donc f est la solution du problème de Cauchy :

$$\begin{cases} y' = i y \\ y(0) = 1. \end{cases}$$

Or, si $k \in \mathbb{R}$, on sait que la solution du problème de Cauchy :

$$\begin{cases} y' = k y, \\ y(0) = 1, \end{cases}$$

est la fonction $x \mapsto e^{kx}$. Par analogie, on décide d'écrire la fonction f sous la forme d'une exponentielle imaginaire « $e^{i\theta}$ ».

θ	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\cos(\theta)$	1	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{1}{2}$	0
$\sin(\theta)$	0	$\frac{1}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	1

Table 1.1 – Cosinus et sinus remarquables.

Définition 1.4.5. Pour tout $\theta \in \mathbb{R}$, on définit l'**exponentielle** du nombre imaginaire pur $i\theta$ par :

$$e^{i\theta} := \cos(\theta) + i\sin(\theta).$$

Comme on l'a vu, l'exponentielle imaginaire a la propriété de transformer les sommes en produits, comme l'exponentielle réelle. Plus généralement, elles ont les mêmes propriétés algébriques.

Proposition 1.4.6. Pour tous $\theta, \theta' \in \mathbb{R}$ et pour tout $n \in \mathbb{Z}$, on a :

$$\begin{aligned} \text{(i)} \quad e^{i(\theta+\theta')} &= e^{i\theta} e^{i\theta'} & \text{(iii)} \quad e^{i(\theta-\theta')} &= \frac{e^{i\theta}}{e^{i\theta'}} & \text{(v)} \quad |e^{i\theta}| &= 1. \\ \text{(ii)} \quad \overline{e^{i\theta}} &= e^{-i\theta} = \frac{1}{e^{i\theta}} & \text{(iv)} \quad (e^{i\theta})^n &= e^{in\theta} & \text{(vi)} \quad \frac{de^{i\theta}}{d\theta} &= ie^{i\theta}. \end{aligned}$$

Démonstration. Soient θ et θ' des nombres réels.

(i) C'est une réécriture de l'égalité (1.8).

(ii) On utilise la parité des fonctions cos et sin :

$$\overline{e^{i\theta}} = \cos(\theta) - i\sin(\theta) = \cos(-\theta) + i\sin(-\theta) = e^{-i\theta}.$$

Par ailleurs, d'après (i), on a $e^{i\theta} e^{-i\theta} = e^{i(\theta-\theta)} = e^{i \cdot 0} = 1$, donc $e^{-i\theta} = \frac{1}{e^{i\theta}}$.

(iii) On utilise les propriétés (i) et (ii) :

$$e^{i(\theta-\theta')} = e^{i\theta} e^{-i\theta'} = e^{i\theta} \frac{1}{e^{i\theta'}} = \frac{e^{i\theta}}{e^{i\theta'}}.$$

(iv) Par récurrence immédiate.

(v) $|e^{i\theta}| = \cos^2(\theta) + \sin^2(\theta) = 1$.

(vi) Déjà démontré, cf. égalité (1.9). □

Il est essentiel de bien connaître les propriétés et les valeurs remarquables des fonctions cosinus et sinus. Dans la table 1.1, on rappelle les valeurs des cosinus et des sinus de quelques angles remarquables dans l'intervalle $[0, \frac{\pi}{2}]$ (premier quadrant du cercle trigonométrique). On déduit d'autres valeurs remarquables à l'aide de symétries dans le cercle trigonométrique (cf. figure 1.5). En particulier, il est toujours utile de connaître l'écriture sous forme d'une exponentielle imaginaire des nombres ± 1 et $\pm i$:

$$e^{2i\pi} = e^{0i} = 1, \quad e^{i\pi} = -1, \quad e^{i\frac{\pi}{2}} = i, \quad e^{-i\frac{\pi}{2}} = -i.$$

1.4.b Arguments d'un nombre complexe

L'exponentielle imaginaire nous permet d'écrire les nombres complexes d'une nouvelle façon. En effet, si z est un nombre complexe non nul, alors $\frac{z}{|z|}$ est un nombre complexe de module 1, car :

$$\left| \frac{z}{|z|} \right| = \frac{|z|}{|z|} = 1.$$

Puisque $\frac{z}{|z|} \in \mathbb{U}$, il existe $\theta \in \mathbb{R}$ unique à 2π près tel que $\frac{z}{|z|} = e^{i\theta}$, et par conséquent, $z = |z| e^{i\theta}$.

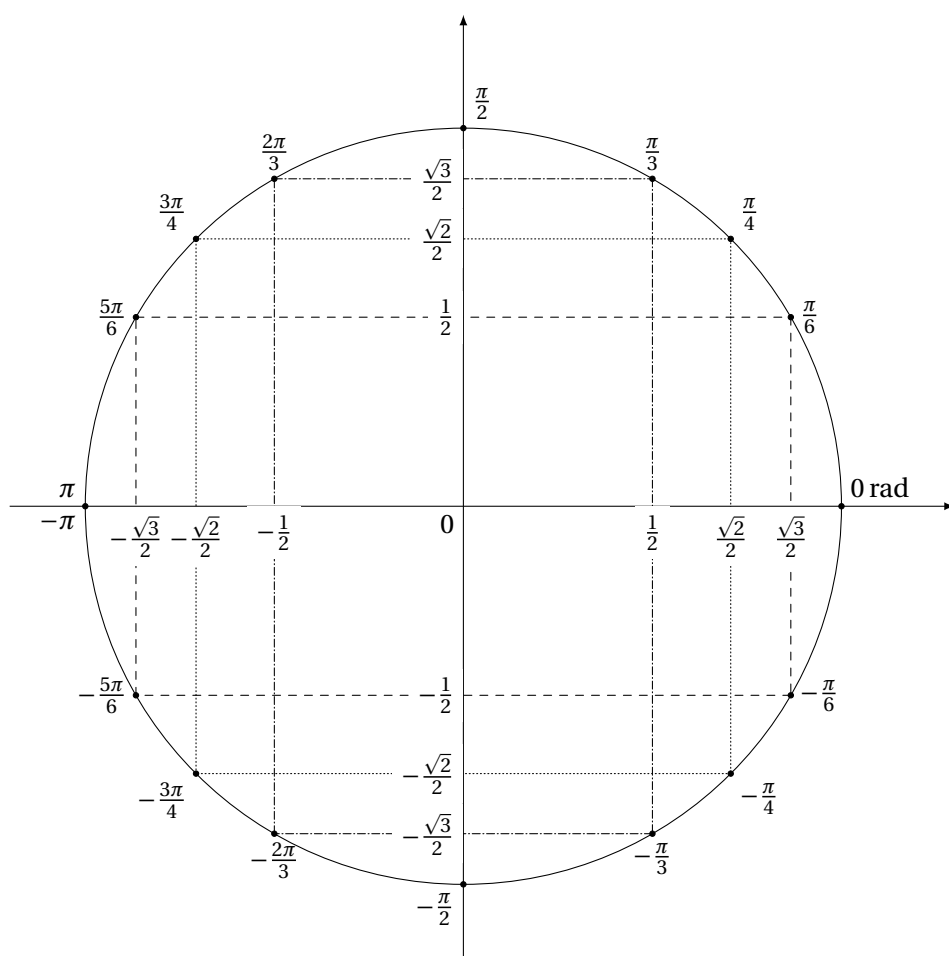


Figure 1.5 – Angles remarquables sur le cercle trigonométrique.

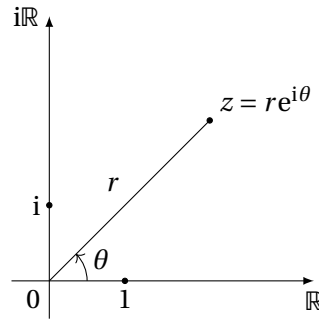


Figure 1.6 – Forme trigonométrique d'un nombre complexe.

Définition 1.4.7. Soit $z \in \mathbb{C}^*$, notons $r := |z| \in \mathbb{R}_+^*$ son module et soit $\theta \in \mathbb{R}$ tel que $z = re^{i\theta}$.

- L'expression $re^{i\theta}$ est appelée une **forme trigonométrique** de z .
- Le nombre réel θ est appelé un **argument** de z , on le note $\arg(z)$.
- L'unique argument de z appartenant à $]-\pi, \pi]$ est appelé l'**argument principal** de z .
- Le couple (r, θ) est appelé un couple de **coordonnées polaires** du point d'affixe z .

Remarque 1.4.8.

1. Un nombre complexe non nul admet une infinité d'arguments dans \mathbb{R} , qui sont tous égaux à 2π près.
2. Dans l'écriture $z = re^{i\theta}$, le nombre r doit être un réel positif (ce nombre est le module de z), sinon ce n'est pas une forme trigonométrique. Par exemple, le nombre $z := -2e^{i\frac{\pi}{3}}$ n'est pas sous forme trigonométrique. Pour obtenir celle-ci, on utilise l'égalité $-1 = e^{i\pi}$:

$$-2e^{i\frac{\pi}{3}} = e^{i\pi} \times 2e^{i\frac{\pi}{3}} = 2e^{i(\frac{\pi}{3} + \pi)} = 2e^{i\frac{4\pi}{3}}.$$

Le module de z est donc 2 et un argument de z est $\frac{4\pi}{3}$.

3. Le nombre 0 n'admet pas de forme trigonométrique. En effet, on peut écrire $0 = |0|e^{i\theta}$ pour n'importe quel $\theta \in \mathbb{R}$, on devrait donc accepter que tous les réels sont des arguments de 0. On décide plutôt que 0 n'a pas d'arguments, et donc pas de forme trigonométrique. ◀

Remarque 1.4.9. Géométriquement, l'argument de z est une mesure en radians de l'angle orienté entre le vecteur d'affixe 1 et le vecteur d'affixe z (cf. figure 1.6). ◀

Exemple 1.4.10. Cherchons la forme trigonométrique de $z := -2 + 2i$. On commence par calculer son module :

$$|z| = \sqrt{(-2)^2 + 2^2} = \sqrt{8} = 2\sqrt{2}.$$

On divise z par son module :

$$\frac{z}{|z|} = -\frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}} = -\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}.$$

On cherche un argument de z , c'est-à-dire un nombre réel θ tel que :

$$\begin{cases} \cos(\theta) = -\frac{\sqrt{2}}{2} \\ \sin(\theta) = \frac{\sqrt{2}}{2} \end{cases}$$

Le réel $\theta = \frac{3\pi}{4}$ convient (c'est l'argument principal de z). Par conséquent on a :

$$\frac{z}{|z|} = \cos\left(\frac{3\pi}{4}\right) + i\sin\left(\frac{3\pi}{4}\right) = e^{i\frac{3\pi}{4}}.$$

Une forme trigonométrique de z est $2\sqrt{2}e^{i\frac{3\pi}{4}}$. ◀

De la même façon que la partie réelle et la partie imaginaire caractérisent un nombre complexe, le module et l'argument caractérisent un nombre complexe non nul : deux nombres complexes non nuls sont égaux si et seulement si ils ont le même module et le même argument modulo 2π .

Proposition 1.4.11. Pour tous $z, w \in \mathbb{C}^*$, on a :

$$z = w \iff |z| = |w| \quad \text{et} \quad \arg(z) \equiv \arg(w) \pmod{2\pi}. \quad \blacktriangleleft$$

Démonstration. Soient $z, w \in \mathbb{C}^*$, alors on a :

$$\begin{aligned} z = w &\iff |z| = |w| \quad \text{et} \quad z = w \\ &\iff |z| = |w| \quad \text{et} \quad |z|e^{i\arg(z)} = |w|e^{i\arg(w)} \\ &\iff |z| = |w| \quad \text{et} \quad e^{i\arg(z)} = e^{i\arg(w)} \\ &\iff |z| = |w| \quad \text{et} \quad \arg(z) \equiv \arg(w) \pmod{2\pi}. \quad \square \end{aligned}$$

Au vu des propriétés du module et de l'argument (proposition 1.2.19 et corollaire 1.4.13), la forme trigonométrique est particulièrement adaptée au calcul de produits et de quotients. En effet, le module d'un produit (resp. d'un quotient) est le produit (resp. le quotient) des modules, et l'argument d'un produit (resp. d'un quotient) est la somme (resp. la différence) des arguments modulo 2π .

Proposition 1.4.12. Pour tous nombres complexes non nuls $z := re^{i\theta}$ et $w := r'e^{i\theta'}$ sous forme trigonométrique, leur produit et leur quotient sont donnés par :

$$zw = rr'e^{i(\theta+\theta')}, \quad \frac{z}{w} = \frac{r}{r'}e^{i(\theta-\theta')}.$$

De plus, si n est un entier relatif, alors :

$$z^n = r^n e^{in\theta}. \quad \blacktriangleleft$$

Corollaire 1.4.13. Pour tous $z, w \in \mathbb{C}^*$ et pour tout $n \in \mathbb{Z}$, on a :

- | | |
|---|--|
| (i) $\arg(zw) \equiv \arg(z) + \arg(w) \pmod{2\pi}.$ | (iii) $\arg\left(\frac{z}{w}\right) \equiv \arg(z) - \arg(w) \pmod{2\pi}.$ |
| (ii) $\arg(\bar{z}) \equiv \arg(z^{-1}) \equiv -\arg(z) \pmod{2\pi}.$ | (iv) $\arg(z^n) \equiv n \times \arg(z) \pmod{2\pi}.$ \blacktriangleleft |

Remarque 1.4.14. Nous pouvons à présent donner un sens géométrique à la multiplication dans \mathbb{C} . Multiplier un nombre complexe z par une exponentielle imaginaire $e^{i\theta}$ revient à ajouter θ à l'argument de z , sans changer son module. Géométriquement, z subit donc une rotation autour de 0 et d'angle θ (le signe de θ donnant le sens de la rotation). Plus généralement, si on multiplie un nombre complexe par $re^{i\theta}$, celui-ci subit une rotation de centre 0 et d'angle θ , puis une homothétie de centre 0 et de rapport r (cf. figure 1.7). L'égalité $i^2 = -1$ trouve ainsi un sens géométrique : puisque multiplier par i correspond à une rotation d'un quart de tour, la multiplication par i^2 correspond à une rotation d'un demi-tour, tout comme la multiplication par -1 . \blacktriangleleft

1.4.c Exponentielle d'un nombre complexe

On connaît déjà l'exponentielle des nombres réels et on a défini précédemment l'exponentielle des nombres imaginaires purs. Puisque tout nombre complexe est la somme d'un nombre réel a et d'un nombre imaginaire pur ib , et puisqu'on attend de l'exponentielle qu'elle transforme les sommes en produits, on définit naturellement l'exponentielle du nombre complexe $a + ib$ comme le produit de e^a et e^{ib} .

Définition 1.4.15. Pour tout $z := a + ib \in \mathbb{C}$, on définit l'**exponentielle** de z par :

$$e^z := e^{\Re(z)} e^{i\Im(z)} = e^a e^{ib} = e^a (\cos b + i \sin b).$$

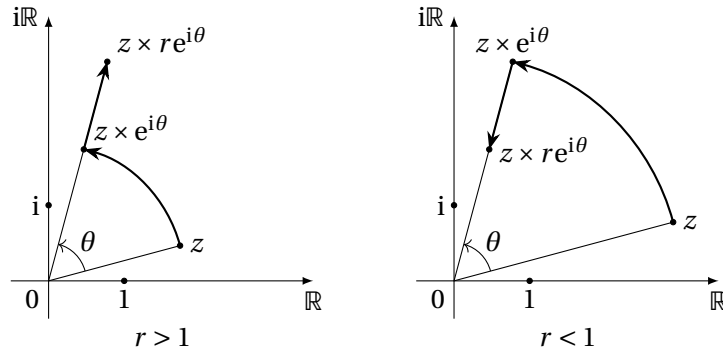


Figure 1.7 – Multiplication dans le plan complexe.

Remarquons que l'exponentielle de z est donnée directement sous forme trigonométrique. Les propriétés algébriques de l'exponentielle complexe sont les mêmes que celles des exponentielles réelles et imaginaires.

Proposition 1.4.16. Pour tous $z, w \in \mathbb{C}$ et pour tout $n \in \mathbb{Z}$, on a :

- (i) $|e^z| = e^{\Re(z)} > 0$. En particulier, l'exponentielle complexe ne s'annule jamais.
- (ii) $\arg(e^z) \equiv \Im(z) \bmod{2\pi}$.
- (iii) L'exponentielle complexe est $2i\pi$ -périodique : $e^z = e^w \iff \exists k \in \mathbb{Z}, z = w + 2i\pi k$.
- (iv) $e^{z+w} = e^z \times e^w$.
- (v) $e^{-z} = \frac{1}{e^z}$.
- (vi) $e^{z-w} = \frac{e^z}{e^w}$.
- (vii) $(e^z)^n = e^{nz}$.

Les démonstrations ne posent aucune difficulté.

1.5 Applications à la trigonométrie

Par définition de l'exponentielle imaginaire, on a $e^{i\theta} = \cos(\theta) + i\sin(\theta)$. Par conséquent, les fonctions cosinus et sinus sont respectivement la partie réelle et la partie imaginaire de l'exponentielle imaginaire. Or, on a vu dans la proposition 1.2.16 qu'on pouvait exprimer les parties réelles et imaginaires d'un nombre complexe en fonction de lui-même de son conjugué. Compte tenu du fait que le conjugué de $e^{i\theta}$ est $e^{-i\theta}$, on obtient les formules suivantes, appelées **formules d'Euler**, qui permettent d'exprimer $\cos(\theta)$ et $\sin(\theta)$ en fonction des exponentielles imaginaires $e^{i\theta}$ et $e^{-i\theta}$.

Proposition 1.5.1 (formules d'Euler). Pour tout $\theta \in \mathbb{R}$, on a :

$$\cos(\theta) = \Re(e^{i\theta}) = \frac{e^{i\theta} + e^{-i\theta}}{2}, \quad \sin(\theta) = \Im(e^{i\theta}) = \frac{e^{i\theta} - e^{-i\theta}}{2i}.$$

La combinaison des formules d'Euler avec les propriétés algébriques de l'exponentielle imaginaire permettent de démontrer simplement de nombreuses propriétés des fonctions trigonométriques.

1.5.a Formules de linéarisation

Une première application des formules d'Euler est la **linéarisation** des produits de $\cos(\theta)$ et $\sin(\theta)$. La linéarisation consiste à transformer une expression de la forme $\cos^n(\theta) \sin^p(\theta)$ avec $n, p \in \mathbb{N}$, en une combinaison linéaire de $\cos(k\theta)$ et $\sin(k\theta)$ pour $k \in \llbracket 0, n+p \rrbracket$, c'est-à-dire une expression de la forme :

$$\sum_{k=0}^{n+p} a_k \cos(k\theta) + b_k \sin(k\theta),$$

où $(a_k)_{0 \leq k \leq n+p}$ et $(b_k)_{0 \leq k \leq n+p}$ sont des coefficients réels. L'expression linéarisée ne contient alors plus aucun produit ni puissances de cosinus et de sinus. Pour linéariser, on applique les trois étapes suivantes :

1. On commence par utiliser les formules d'Euler :

$$\cos^n(\theta) \sin^p(\theta) = \left(\frac{e^{i\theta} + e^{-i\theta}}{2} \right)^n \left(\frac{e^{i\theta} - e^{-i\theta}}{2i} \right)^p.$$

2. On développe le membre de droite à l'aide de la formule du binôme de Newton et en utilisant les propriétés de l'exponentielle imaginaire : $(e^{i\theta})^k = e^{ik\theta}$ et $e^{ik\theta} e^{i\ell\theta} = e^{i(k+\ell)\theta}$.
3. On regroupe les exponentielles conjuguées pour faire apparaître $\cos(k\theta)$ et $\sin(k\theta)$ à l'aide des formules d'Euler : $e^{ik\theta} + e^{-ik\theta} = 2\cos(k\theta)$ et $e^{ik\theta} - e^{-ik\theta} = 2i\sin(k\theta)$.

Exemple 1.5.2 (linéarisation de $\sin^2(\theta)$). Pour tout $\theta \in \mathbb{R}$, on a :

$$\begin{aligned} \sin^2(\theta) &= \left(\frac{e^{i\theta} - e^{-i\theta}}{2i} \right)^2 && \text{(formule d'Euler)} \\ &= \frac{(e^{i\theta})^2 - 2e^{i\theta}e^{-i\theta} + (e^{-i\theta})^2}{-4} && \text{(on développe)} \\ &= \frac{e^{2i\theta} + e^{-2i\theta} - 2}{-4} && \text{(on regroupe les termes conjugués)} \\ &= \frac{2\cos(2\theta) - 2}{-4} && \text{(formule d'Euler)} \\ &= \frac{1}{2} - \frac{1}{2}\cos(2\theta). \end{aligned}$$

La linéarisation est utilisée notamment pour déterminer les primitives de produits de fonctions trigonométriques. Par exemple, les primitives de la fonction $\sin^2(x)$ sont de la forme $\frac{1}{2}x - \frac{1}{4}\sin(2x) + C$.

1.5.b Formules d'angle multiple

Il est parfois utile de délinéariser des expressions trigonométriques. On utilise pour cela les formules d'angle multiple, qui expriment $\cos(n\theta)$ et $\sin(n\theta)$ en fonction de produits et de puissances de $\cos(\theta)$ et $\sin(\theta)$. C'est encore une fois l'exponentielle imaginaire qui permet de résoudre facilement ce problème. Pour tout $n \in \mathbb{Z}$, on a vu que $(e^{i\theta})^n = e^{in\theta}$ (proposition 1.4.6). En réécrivant ces exponentielles imaginaires en fonction de cosinus et sinus, on obtient l'égalité suivante, appelée **formule de Moivre**.

Proposition 1.5.3 (formule de Moivre). Pour tout $\theta \in \mathbb{R}$ et pour tout $n \in \mathbb{Z}$, on a :

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

La formule de Moivre permet d'obtenir de façon systématique les formules d'angle multiple :

1. On écrit la formule de Moivre : $\cos(n\theta) + i\sin(n\theta) = (\cos(\theta) + i\sin(\theta))^n$.
2. On développe le membre de droite à l'aide de la formule du binôme de Newton.
3. Par identification, $\cos(n\theta)$ est égal à la partie réelle du résultat, et $\sin(n\theta)$ est égal à sa partie imaginaire.

Exemple 1.5.4 (formules de duplication). Pour tout $\theta \in \mathbb{R}$, on a :

$$\begin{aligned} \cos(2\theta) + i\sin(2\theta) &= (\cos(\theta) + i\sin(\theta))^2 && \text{(formule de Moivre)} \\ &= \cos^2(\theta) - \sin^2(\theta) + 2i\cos(\theta)\sin(\theta). && \text{(on développe)} \end{aligned}$$

Par identification des parties réelles et imaginaires, on obtient les **formules de duplication** :

$$\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta), \quad \sin(2\theta) = 2\cos(\theta)\sin(\theta).$$

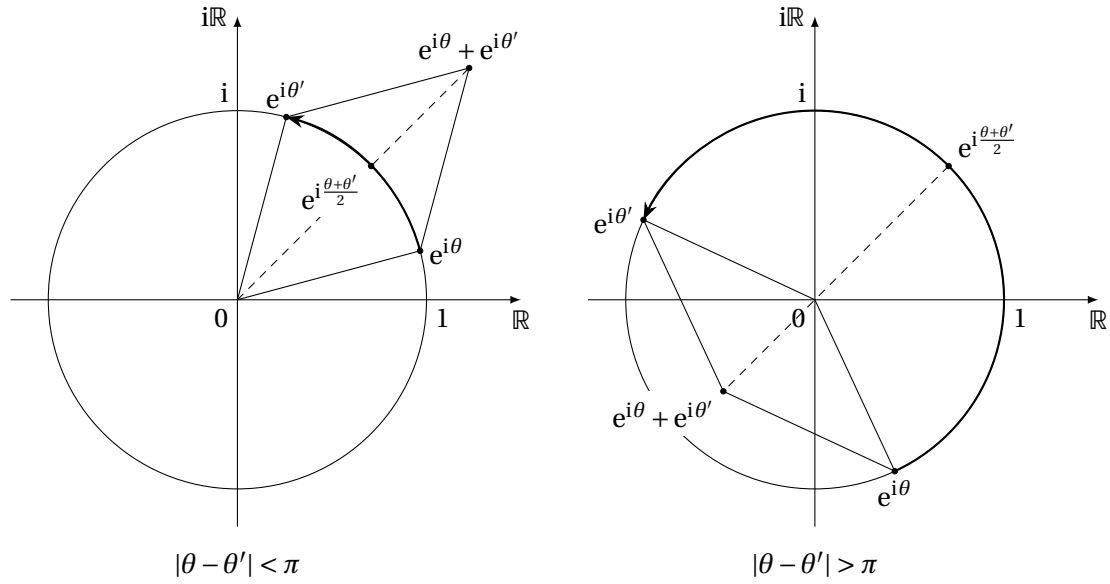


Figure 1.8 – Factorisation de l'angle moitié.

1.5.c Factorisation de l'angle moitié

Une autre utilisation des formules d'Euler est la technique appelée **factorisation de l'angle moitié**. Cette technique permet de mettre (presque) sous forme trigonométrique la somme ou la différence de deux exponentielle imaginaire $e^{i\theta}$ et $e^{i\theta'}$.

En écrivant $\theta = \frac{\theta+\theta'}{2} + \frac{\theta-\theta'}{2}$ et $\theta' = \frac{\theta+\theta'}{2} - \frac{\theta-\theta'}{2}$, on peut mettre en facteur $e^{i\frac{\theta+\theta'}{2}}$ dans la somme $e^{i\theta} + e^{i\theta'}$:

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

Selon le signe du cosinus, on obtient une forme trigonométrique de $e^{i\theta} + e^{i\theta'}$, ou son opposée. Supposons par exemple que $\theta, \theta' \in]-\pi, \pi]$, c'est-à-dire que θ et θ' sont les arguments principaux.

- Si $|\theta - \theta'| < \pi$, alors $\cos\left(\frac{\theta-\theta'}{2}\right)$ est positif, donc $2 \cos\left(\frac{\theta-\theta'}{2}\right) e^{i\frac{\theta+\theta'}{2}}$ est une forme trigonométrique de $e^{i\theta} + e^{i\theta'}$. Par conséquent, un argument de $e^{i\theta} + e^{i\theta'}$ est $\frac{\theta+\theta'}{2}$.
- Si $|\theta - \theta'| > \pi$, alors $\cos\left(\frac{\theta-\theta'}{2}\right)$ est négatif, donc $2 \cos\left(\frac{\theta-\theta'}{2}\right) e^{i\frac{\theta+\theta'}{2}}$ est l'opposé d'une forme trigonométrique de $e^{i\theta} + e^{i\theta'}$. Par conséquent, un argument de $e^{i\theta} + e^{i\theta'}$ est $\frac{\theta+\theta'}{2} + \pi$.

La figure 1.8 illustre la factorisation de l'angle moitié dans le plan complexe. On peut également utiliser la factorisation de l'angle moitié pour une différence de deux exponentielle imaginaires, on obtient alors :

$$e^{i\theta} - e^{i\theta'} = 2i \sin\left(\frac{\theta-\theta'}{2}\right) e^{i\frac{\theta+\theta'}{2}}.$$

Proposition 1.5.5 (factorisation de l'angle moitié). Pour tous $\theta, \theta' \in \mathbb{R}$, on a :

$$(i) \quad e^{i\theta} + e^{i\theta'} = 2 \cos\left(\frac{\theta-\theta'}{2}\right) e^{i\frac{\theta+\theta'}{2}}. \quad (ii) \quad e^{i\theta} - e^{i\theta'} = 2i \sin\left(\frac{\theta-\theta'}{2}\right) e^{i\frac{\theta+\theta'}{2}}. \quad \blacktriangleleft$$

Un cas particulier de factorisation de l'angle moitié souvent rencontré est le cas où $\theta' = 0$. On a dans ce cas les égalités :

$$e^{i\theta} + 1 = 2 \cos\left(\frac{\theta}{2}\right) e^{i\frac{\theta}{2}}, \quad e^{i\theta} - 1 = 2i \sin\left(\frac{\theta}{2}\right) e^{i\frac{\theta}{2}}.$$

En prenant la partie réelle et la partie imaginaire dans les formules de factorisation de l'angle moitié, on obtient les **formules de factorisation** de cosinus et sinus.

Corollaire 1.5.6 (formules de factorisation). Pour tous $\theta, \theta' \in \mathbb{R}$, on a :

$$\begin{aligned} \text{(i)} \quad \cos(\theta) + \cos(\theta') &= 2 \cos\left(\frac{\theta+\theta'}{2}\right) \cos\left(\frac{\theta-\theta'}{2}\right). & \text{(iii)} \quad \sin(\theta) + \sin(\theta') &= 2 \sin\left(\frac{\theta+\theta'}{2}\right) \cos\left(\frac{\theta-\theta'}{2}\right). \\ \text{(ii)} \quad \cos(\theta) - \cos(\theta') &= -2 \sin\left(\frac{\theta+\theta'}{2}\right) \sin\left(\frac{\theta-\theta'}{2}\right). & \text{(iv)} \quad \sin(\theta) - \sin(\theta') &= 2 \cos\left(\frac{\theta+\theta'}{2}\right) \sin\left(\frac{\theta-\theta'}{2}\right). \end{aligned} \quad \blacktriangleleft$$

1.6 Racines n -ièmes

On a à présent tous les ingrédients pour étudier l'extraction de racines n -ièmes des nombres complexes. Mais avant de voir comment déterminer les racines n -ièmes d'un nombre complexe quelconque, on commence par étudier les racines n -ièmes de l'unité, c'est-à-dire les racines n -ièmes de 1. Par la suite, on verra que si on connaît les racines n -ièmes de l'unité et une racine n -ième particulière d'un nombre complexe, on peut déduire toutes ses autres racines n -ièmes.

1.6.a Racines n -ièmes de l'unité

Définition 1.6.1. Soit $n \in \mathbb{N}^*$, on appelle **racine n -ième de l'unité** tout nombre $\zeta \in \mathbb{C}^*$ tel que :

$$\zeta^n = 1.$$

L'ensemble des racines n -ièmes de l'unité est noté \mathbb{U}_n .

Dans \mathbb{R} , l'étude des racines n -ièmes de l'unité n'est pas très palpitante, puisqu'il n'y a que deux cas de figure : si n est pair, les racines n -ièmes de l'unité sont 1 et -1 , et si n est impair, la seule racine n -ième de l'unité est 1. Dans \mathbb{C} , la question possède une réponse plus riche : il y a n racines n -ièmes de l'unité.

Théorème 1.6.2. Pour tout $n \in \mathbb{N}^*$, il existe exactement n racines n -ièmes de l'unité et celles-ci sont données par :

$$\mathbb{U}_n = \left\{ e^{\frac{2i\pi k}{n}} : k \in \llbracket 0, n-1 \rrbracket \right\}.$$

Autrement dit, en notant $\zeta_n := e^{\frac{2i\pi}{n}}$, les racines n -ièmes de l'unité sont $1, \zeta_n, \zeta_n^2, \dots, \zeta_n^{n-1}$.

Démonstration. On cherche les solutions de $\zeta^n = 1$ sous forme trigonométrique $\zeta = re^{i\theta}$ avec $r > 0$ et $\theta \in [0, 2\pi[$:

$$\begin{aligned} \zeta^n = 1 &\iff r^n e^{in\theta} = 1 \times e^{0i} \\ &\iff r^n = 1 \text{ et } \exists k \in \mathbb{Z}, n\theta = 0 + 2k\pi & \text{(proposition 1.4.11)} \\ &\iff r = 1 \text{ et } \exists k \in \mathbb{Z}, \theta = \frac{2k\pi}{n}. \end{aligned}$$

Les racines n -ième de l'unité sont donc les $e^{\frac{2ik\pi}{n}}$ où $k \in \mathbb{Z}$. Enfin, puisque $\theta \in [0, 2\pi[$, l'entier k ne peut prendre que les valeurs de 0 à $n-1$. \square

Exemple 1.6.3. Regardons quelles sont les racines n -ièmes de l'unité pour des petites valeurs de n :

- $n = 2$: les deux racines carrées de l'unité sont 1 et -1 .
- $n = 3$: les trois racines cubiques de l'unité sont 1, j et j^2 , où j est défini par :

$$j := e^{\frac{2i\pi}{3}} = -\frac{1}{2} + i\frac{\sqrt{3}}{2}, \quad j^2 = \bar{j} = -\frac{1}{2} - i\frac{\sqrt{3}}{2}.$$

- $n = 4$: les quatre racines 4^{es} de l'unité sont 1, i , -1 et $-i$. \blacktriangleleft

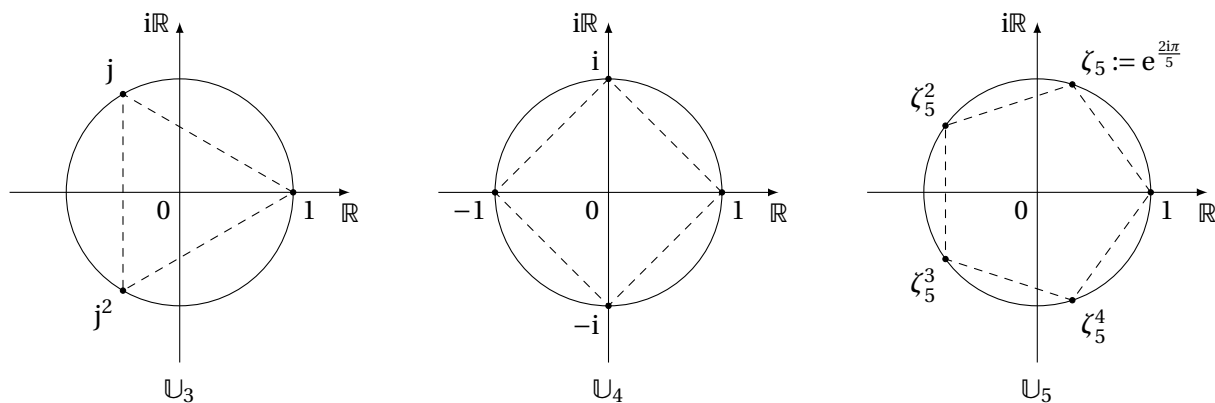


Figure 1.9 – Représentations de \mathbb{U}_3 , \mathbb{U}_4 et \mathbb{U}_5 dans le plan complexe.

Remarque 1.6.4. Dans le plan complexe, les racines n -ièmes de l'unité forment un polygone régulier inscrit dans le cercle unité, et dont l'un des sommets est 1 (cf. figure 1.9). ◀

Proposition 1.6.5. Pour tout entier $n \geq 2$ et pour tout $\zeta \in \mathbb{U}_n \setminus \{1\}$, on a :

$$\sum_{k=0}^{n-1} \zeta^k = 0. \quad \blacktriangleleft$$

Démonstration. Il suffit d'utiliser la formule donnant la somme des puissances du nombre complexe $\zeta \neq 1$ (proposition 1.2.11) :

$$\sum_{k=0}^{n-1} \zeta^k = \frac{1 - \zeta^n}{1 - \zeta} = 0,$$

car $\zeta^n = 1$ puisque ζ est une racine n -ième de l'unité. □

1.6.b Racines n -ièmes d'un nombre complexe

Définition 1.6.6. Soit $n \in \mathbb{N}^*$. On appelle **racine n -ième** d'un nombre complexe z tout nombre complexe w tel que $w^n = z$.

Tout comme pour les racines n -ièmes de l'unité, un nombre complexe non nul admet exactement n racines n -ièmes dans \mathbb{C} .

Théorème 1.6.7. Soit $n \in \mathbb{N}^*$ et soit $z = re^{i\theta} \in \mathbb{C}^*$ un nombre complexe non nul sous forme trigonométrique. Le nombre z admet exactement n racines n -ièmes et celles-ci sont données par :

$$\forall k \in \llbracket 0, n-1 \rrbracket, \quad w_k := r^{\frac{1}{n}} e^{i\left(\frac{\theta+2k\pi}{n}\right)} = r^{\frac{1}{n}} e^{i\frac{\theta}{n}} \times \zeta_n^k,$$

où $\zeta_n := e^{\frac{2i\pi}{n}}$.

Remarque 1.6.8. Les racines n -ième d'un nombre complexe z s'obtiennent en multipliant une racine n -ième particulière de z par les racines n -ièmes de l'unité. ◀

Démonstration. Soit $z = re^{i\theta}$ un nombre complexe non nul sous forme trigonométrique et posons $w_0 = r^{\frac{1}{n}} e^{i\frac{\theta}{n}}$. Vérifions que w_0 est une racine n -ième de z :

$$w_0^n = \left(r^{\frac{1}{n}} e^{i\frac{\theta}{n}}\right)^n = r^{\frac{1}{n} \times n} e^{i\frac{\theta}{n} \times n} = re^{i\theta} = z.$$

Cherchons les autres racines n -ièmes de z en résolvant l'équation $w^n = z$:

$$\begin{aligned}
 w^n = z &\iff w^n = w_0^n \\
 &\iff \frac{w^n}{w_0^n} = 1 && (w_0^n = z \neq 0) \\
 &\iff \left(\frac{w}{w_0}\right)^n = 1 \\
 &\iff \frac{w}{w_0} \in \mathbb{U}_n \\
 &\iff \exists k \in \llbracket 0, n-1 \rrbracket, \frac{w}{w_0} = \zeta_n^k && (\text{théorème 1.6.2}) \\
 &\iff \exists k \in \llbracket 0, n-1 \rrbracket, w = w_0 \times \zeta_n^k.
 \end{aligned}$$

L'équation $w^n = z$ admet donc exactement les n solutions annoncées. \square

Remarque 1.6.9. Comme pour la racine carrée, on n'écrit pas « $\sqrt[n]{z}$ » ni « $z^{\frac{1}{n}}$ » pour désigner une racine n -ième d'un nombre complexe quelconque, car on ne saurait pas de quelle racine on parle. Ces notations sont réservées aux réels positifs (et aux réels négatifs si n est impair). \blacktriangleleft

Exemple 1.6.10. Cherchons les racines cubique de $z := -2+2i$. Dans l'exemple 1.4.10, on a vu qu'une forme trigonométrique de ce nombre est $z = 2\sqrt{2}e^{\frac{3i\pi}{4}}$. Une racine cubique de z est :

$$w_0 := \left(2\sqrt{2}\right)^{\frac{1}{3}} e^{\frac{3i\pi}{4} \times \frac{1}{3}} = \sqrt{2}e^{i\frac{\pi}{4}}.$$

Les racines cubiques de l'unité sont 1, j et $j^2 = \bar{j}$, où $j := e^{\frac{2i\pi}{3}}$. Par conséquent, les trois racines cubiques de z sont :

$$\begin{aligned}
 w_0 &= \sqrt{2}e^{i\frac{\pi}{4}} && w_1 := w_0 \times j && w_2 := w_0 \times \bar{j} \\
 & && = \sqrt{2}e^{i(\frac{\pi}{4} + \frac{2\pi}{3})} && = \sqrt{2}e^{i(\frac{\pi}{4} - \frac{2\pi}{3})} \\
 & && = \sqrt{2}e^{i\frac{11\pi}{12}} && = \sqrt{2}e^{-i\frac{5\pi}{12}}.
 \end{aligned}$$

1.6.c Équations de degré n

Dans la section 1.3, on a vu que tout nombre complexe non nul admettait exactement deux racines carrées dans \mathbb{C} . Plus généralement, on a vu que tout polynôme de degré 2 à coefficients complexes admettait exactement deux racines dans \mathbb{C} (dans le cas des polynômes dont le discriminant est nul, la racine double compte comme deux racines).

Dans cette section, on a vu plus généralement que tout nombre complexes non nul admettait exactement n racines n -ièmes dans \mathbb{C} . Naturellement, on peut conjecturer que les polynômes de degré n admettent n racines dans \mathbb{C} . Et c'est effectivement le cas : ce résultat porte le nom de théorème fondamental de l'algèbre, ou théorème de d'Alembert–Gauss. On ne démontrera pas ce théorème ici, car il nécessite des outils qui seront étudiés en analyse ultérieurement.

Théorème 1.6.11 (d'Alembert–Gauss, admis). Tout polynôme de degré $n \geq 1$ à coefficients complexes admet exactement n racines dans \mathbb{C} , comptées avec multiplicité (une racine double est comptée 2 fois, une racine triple est comptée trois fois, etc).

Il est remarquable que le simple ajout à \mathbb{R} du nombre i , c'est-à-dire d'une racine carrée de -1 , suffise à ce que toutes les équations polynomiales admettent des solutions. Ainsi, l'ensemble des nombres complexes est le cadre idéal dans lequel étudier les équations polynomiales.

Chapitre 2

Résolution de systèmes linéaires

Vous savez déjà résoudre des systèmes linéaires à 2 ou 3 inconnues. L'objectif de ce chapitre d'apprendre une méthode de résolution systématique des systèmes de p équations linéaires à n inconnues (réelles ou complexes), avec n et p des entiers quelconques.

Dans ce chapitre, le symbole \mathbb{K} désigne l'ensemble \mathbb{R} ou l'ensemble \mathbb{C} .

2.1 Systèmes linéaires

Définition 2.1.1. Un **système linéaire** de p équations à n inconnues x_1, \dots, x_n est un système d'équations de la forme :

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,n}x_n = b_2 \\ \vdots \\ a_{p,1}x_1 + a_{p,2}x_2 + \dots + a_{p,n}x_n = b_p, \end{cases} \quad (2.1)$$

où les $a_{i,j} \in \mathbb{K}$ sont les coefficients du système, et les $b_i \in \mathbb{K}$ forment le second membre. Si tous les coefficients b_i sont nuls, on dit que le système est **homogène**. Résoudre le système (2.1), c'est trouver tous les n -uplets $(x_1, \dots, x_n) \in \mathbb{K}^n$ qui satisfont simultanément toutes les équations de (2.1).

Remarque 2.1.2. Les coefficients $a_{i,j}$ sont repérés par deux indices : le premier est celui de l'équation et le deuxième celui de la variable. Ainsi, lorsqu'on écrit le système en alignant les variables verticalement, le premier indice repère la ligne et le deuxième indice repère la colonne où se situe le coefficient. Cette convention doit être impérativement respectée. ◀

Remarque 2.1.3. Lorsque le nombre d'inconnues est petit, on utilise généralement des lettres différentes plutôt des indices pour nommer les inconnues. Par exemple, lorsque $n = 3$, on appellera les inconnues « x, y, z » plutôt que « x_1, x_2, x_3 ». ◀

Exemple 2.1.4. Considérons le système de 2 équations à 2 inconnues suivant :

$$\begin{cases} x - y = 1 \\ 2x + y = 4. \end{cases}$$

Les coefficients de ce système sont $a_{1,1} = 1$, $a_{1,2} = -1$, $a_{2,1} = 2$ et $a_{2,2} = 1$, et le second membre est $b_1 = 1$ et $b_2 = 4$. Au lycée, vous avez appris deux méthodes pour résoudre un tel système : la méthode par substitutions et la méthode par combinaisons. C'est cette dernière qu'on va utiliser et généraliser pour résoudre n'importe quel système linéaire.

Le principe de la méthode par combinaisons est d'ajouter un multiple bien choisi d'une équation à l'autre, de façon à éliminer une inconnue. Dans cet exemple, en retranchant 2 fois la première

équation à la deuxième, on élimine x dans la deuxième équation :

$$\begin{cases} x - y = 1 \\ 3y = 2. \end{cases}$$

Ce système est alors facile à résoudre : la deuxième équation donne $y = \frac{2}{3}$, on reporte ensuite cette valeur dans la première équation pour obtenir $x = 1 + y = \frac{5}{3}$. L'unique solution de ce système est donc $(x, y) = (\frac{5}{3}, \frac{2}{3})$. ◀

2.2 Opérations élémentaires

Définition 2.2.1. On appelle **opérations élémentaires** sur les lignes d'un système linéaire les opérations suivantes :

1. $L_i \leftrightarrow L_j$: échanger la ligne i avec la ligne j .
2. $L_i \leftarrow \lambda L_i$ ($\lambda \neq 0$) : multiplier la ligne i par une constante λ non nulle.
3. $L_i \leftarrow L_i + \lambda L_j$ ($i \neq j$) : ajouter à la ligne i un multiple d'une autre ligne j .

Remarque 2.2.2. On peut combiner les opérations 2 et 3 pour remplacer la ligne i par une combinaison des lignes i et j . On note cette opération $L_i \leftarrow \lambda L_i + \mu L_j$ ($\lambda \neq 0$ et $i \neq j$). ◀

Le point important est que les opérations élémentaires ne changent pas l'ensemble des solutions d'un système linéaire. On peut donc les utiliser pour transformer un système donné en un système équivalent plus simple à résoudre.

Proposition 2.2.3. Si (S) est un système linéaire et si (S') le système obtenu après une opération élémentaire, alors (S) et (S') sont équivalents. ▶

Remarque 2.2.4. L'exemple ci-dessous est une mauvaise utilisation de cette proposition :

$$\begin{cases} x + y = 1 \\ x - y = 0 \end{cases} \begin{matrix} \Rightarrow \\ \not\Rightarrow \end{matrix} \begin{cases} 2x = 1 \\ 2x = 1 \end{cases} \begin{matrix} L_1 \leftarrow L_1 + L_2 \\ L_2 \leftarrow L_2 + L_1 \end{matrix}$$

Le système de gauche implique le système de droite mais la réciproque est fautive : les systèmes ne sont pas équivalents. En effet, le premier système admet pour unique solution $x = y = \frac{1}{2}$, alors que le second ne comporte plus que l'équation $2x = 1$ et admet donc une infinité de solutions ($x = \frac{1}{2}$ et $y \in \mathbb{R}$ libre).

La proposition ne s'applique pas car on a modifié l'équation 1 à l'aide de l'équation 2, en modifiant simultanément l'équation 2 à l'aide de l'équation 1. Pour que les systèmes soient équivalents, il faut appliquer les opérations élémentaires séquentiellement.

En revanche, lorsque des opérations élémentaires n'interfèrent pas les unes avec les autres (typiquement, utiliser une même ligne pour modifier plusieurs autres lignes), on peut bien sûr les effectuer simultanément pour gagner du temps. ◀

Démonstration de la proposition 2.2.3. La proposition est évidente pour les opérations élémentaires 1 et 2. Démontrons-la pour l'opération 3.

Soit (S) un système linéaire :

$$\left\{ \begin{array}{ll} a_{1,1} x_1 + a_{1,2} x_2 + \cdots + a_{1,n} x_n = b_1 & (L_1) \\ \vdots & \\ a_{i,1} x_1 + a_{i,2} x_2 + \cdots + a_{i,n} x_n = b_i & (L_i) \\ \vdots & \\ a_{j,1} x_1 + a_{j,2} x_2 + \cdots + a_{j,n} x_n = b_j & (L_j) \\ \vdots & \\ a_{p,1} x_1 + a_{p,2} x_2 + \cdots + a_{p,n} x_n = b_p, & (L_p) \end{array} \right. \quad (S)$$

et soit (S') le système linéaire obtenu après l'opération $L_i \leftarrow L_i + \lambda L_j$, c'est-à-dire :

$$\left\{ \begin{array}{l} a_{1,1} x_1 + a_{1,2} x_2 + \cdots + a_{1,n} x_n = b_1 \quad (L_1) \\ \vdots \\ a'_{i,1} x_1 + a'_{i,2} x_2 + \cdots + a'_{i,n} x_n = b'_i \quad (L'_i) \\ \vdots \\ a_{j,1} x_1 + a_{j,2} x_2 + \cdots + a_{j,n} x_n = b_j \quad (L_j) \\ \vdots \\ a_{p,1} x_1 + a_{p,2} x_2 + \cdots + a_{p,n} x_n = b_p, \quad (L_p) \end{array} \right. \quad (S')$$

où $a'_{i,k} := a_{i,k} + \lambda a_{j,k}$ pour tout $k \in \llbracket 1, n \rrbracket$ et $b'_i = b_i + \lambda b_j$.

Pour démontrer que (S) et (S') sont équivalents, on procède par double implication : on démontre dans un premier temps $(S) \Rightarrow (S')$, puis on démontre l'implication réciproque.

Soit $x := (x_1, \dots, x_n)$ une solution de (S) ; montrons que x est aussi solution de (S') . Les systèmes (S) et (S') partagent les mêmes équations (L_k) pour $k \neq i$, donc la seule chose à vérifier est que x est solution de (L'_i) . Puisque x est solution de (S) , il vérifie les équations :

$$\left\{ \begin{array}{l} \sum_{k=1}^n a_{i,k} x_k = b_i \quad (L_i) \\ \sum_{k=1}^n a_{j,k} x_k = b_j \quad (L_j) \end{array} \right.$$

Par conséquent, on a :

$$\sum_{k=1}^n a'_{i,k} x_k = \sum_{k=1}^n (a_{i,k} + \lambda a_{j,k}) x_k = \sum_{k=1}^n a_{i,k} x_k + \lambda \sum_{k=1}^n a_{j,k} x_k = b_i + \lambda b_j = b'_i.$$

Ainsi, on a montré que x est solution de l'équation (L'_i) . Par conséquent, x est solution de (S') , ce qui démontre que $(S) \Rightarrow (S')$. Or, on passe de (S') à (S) par l'opération élémentaire $L'_i \leftarrow L'_i - \lambda L'_j$ donc $(S') \Rightarrow (S)$ d'après ce qu'on vient de montrer. Par conséquent, (S) et (S') sont équivalents. \square

Exemple 2.2.5. Considérons le système suivant :

$$\left\{ \begin{array}{l} x + 3y + z = 0 \\ -x + 2y + 2z = 1 \\ 2x + 4y - z = 1. \end{array} \right.$$

On utilise la 1^{re} équation pour éliminer l'inconnue x dans les 2^e et 3^e équations :

$$\left\{ \begin{array}{l} x + 3y + z = 0 \\ -x + 2y + 2z = 1 \\ 2x + 4y - z = 1 \end{array} \right. \iff \left\{ \begin{array}{l} x + 3y + z = 0 \\ 5y + 3z = 1 \quad L_2 \leftarrow L_2 + L_1 \\ -2y - 3z = 1. \quad L_3 \leftarrow L_3 - 2L_1 \end{array} \right.$$

On utilise ensuite la 2^e équation pour éliminer l'inconnue y dans la 3^e équation :

$$\left\{ \begin{array}{l} x + 3y + z = 0 \\ 5y + 3z = 1 \\ -2y - 3z = 1 \end{array} \right. \iff \left\{ \begin{array}{l} x + 3y + z = 0 \\ 5y + 3z = 1 \\ -9z = 7. \quad L_3 \leftarrow 5L_3 + 2L_2 \end{array} \right.$$

Le système est alors sous une forme plus simple, dite triangulaire. ◀

2.3 Systèmes triangulaires

Définition 2.3.1. On dit qu'un système linéaire est **triangulaire** si :

1. le nombre d'équations est égal au nombre d'inconnues : $p = n$;
2. les coefficients sous la diagonale sont nuls : pour tous $i, j \in \llbracket 1, n \rrbracket$, $a_{i,j} = 0$ si $i > j$.

Un système triangulaire est donc de la forme :

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = b_1 \\ a_{2,2}x_2 + \cdots + a_{2,n}x_n = b_2 \\ \vdots \\ a_{n,n}x_n = b_n. \end{cases}$$

Exemple 2.3.2. On considère le système triangulaire suivant :

$$\begin{cases} -2x + y + z = -1 \\ -y + 3z = 1 \\ 2z = 4. \end{cases} \quad (2.2)$$

Un système triangulaire se résout simplement par remontée :

1. la dernière équation donne la valeur de z ;
2. on reporte cette valeur dans la 2^e équation ce qui permet de trouver la valeur de y ;
3. on reporte les valeurs de y et z dans la 1^{re} équation pour obtenir la valeur de x .

On applique cette procédure au système (2.2) :

$$(2.2) \iff \begin{cases} -2x + y + 2 = -1 \\ -y + 6 = 1 \\ z = 2 \end{cases} \iff \begin{cases} -2x + 5 = -3 \\ y = 5 \\ z = 2 \end{cases} \iff \begin{cases} x = 4 \\ y = 5 \\ z = 2. \end{cases}$$

L'unique solution du système est $(x, y, z) = (4, 5, 2)$. ◀

On peut généraliser la méthode de résolution de l'exemple précédent à tout système triangulaire dont les coefficients diagonaux (les coefficients $a_{i,i}$) sont non nuls.

Théorème 2.3.3. Si tous les coefficients diagonaux d'un système triangulaire sont non nuls, alors le système admet une unique solution.

Démonstration. On procède par récurrence sur la taille du système. Soit H_n la proposition : « tout système linéaire triangulaire de taille n avec des coefficients diagonaux non nuls admet une unique solution ».

- Initialisation $n = 1$: toute équation linéaire de la forme $ax = b$ avec $a \neq 0$ admet pour unique solution $x = \frac{b}{a}$.
- Hérédité : soit $n \geq 1$ tel que H_n est vraie; démontrons H_{n+1} . Considérons un système triangulaire de taille $n + 1$:

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n + a_{1,n+1}x_{n+1} = b_1 \\ a_{2,2}x_2 + \cdots + a_{2,n}x_n + a_{2,n+1}x_{n+1} = b_2 \\ \vdots \\ a_{n,n}x_n + a_{n,n+1}x_{n+1} = b_n \\ a_{n+1,n+1}x_{n+1} = b_{n+1}, \end{cases} \quad (S)$$

tel que $a_{i,i} \neq 0$ pour tout $i \in \llbracket 1, n+1 \rrbracket$. La dernière équation admet pour unique solution $x_{n+1} = \frac{b_{n+1}}{a_{n+1}}$. En reportant la valeur de x_{n+1} dans les lignes 1 à n , le reste du système est équivalent à :

$$\left\{ \begin{array}{l} a_{1,1} x_1 + a_{1,2} x_2 + \cdots + a_{1,n} x_n = b'_1 \\ a_{2,2} x_2 + \cdots + a_{2,n} x_n = b'_2 \\ \vdots \\ a_{n,n} x_n = b'_n, \end{array} \right. \quad (S')$$

où $b'_i := b_i - a_{i,n+1} x_{n+1}$. Ce système est triangulaire de taille n avec des coefficients diagonaux non nuls, donc par hypothèse de récurrence, il admet une unique solution, ce qui finit de prouver H_{n+1} . Ainsi, on a montré que pour tout $n \geq 1$, H_n implique H_{n+1} .

Par le principe de récurrence, la proposition H_n est vraie pour tout $n \geq 1$. \square

2.4 Systèmes échelonnés

Tous les systèmes ne sont pas équivalents à des systèmes triangulaires. On généralise la notion de système triangulaire par la notion de système échelonné. On verra que tout système linéaire peut se mettre sous la forme d'un système échelonné par des opérations élémentaires.

Définition 2.4.1. On dit qu'un système est **échelonné** s'il vérifie les deux propriétés suivantes :

1. Si une ligne de coefficients est entièrement nulle, toutes les lignes suivantes le sont aussi. Ces lignes sont appelées les **équations de compatibilité**.
2. Pour les lignes non nulles, le nombre de coefficients nuls commençant une ligne croît strictement d'une ligne à la suivante. Les premiers coefficients non nuls de chaque ligne sont appelés les **pivots**.

Une autre façon de le dire est que dans un système échelonné, le premier coefficient non nul d'une ligne se décale de plus en plus vers la droite quand on passe d'une équation à la suivante, jusqu'à ce qu'il n'y ait plus d'inconnues.

Remarque 2.4.2. Les équations de compatibilité sont des équations de la forme $0 = b_i$. Elles ne comportent pas d'inconnues, elles ne concernent que le second membre. Ces équations sont des conditions que doit vérifier le second membre pour que le système admette des solutions. Si le second membre ne vérifie pas les équations de compatibilité, le système n'admet pas de solutions et on dit qu'il est **incompatible**. \blacktriangleleft

Les premiers exemples de systèmes échelonnés sont les système triangulaires dont les coefficients diagonaux sont non nuls. Dans ce cas, les coefficients diagonaux sont les pivots. Mais la notion de système échelonné est plus générale comme le montre l'exemple suivant.

Exemple 2.4.3. Le système :

$$\left\{ \begin{array}{l} \boxed{2}x_1 - x_2 - 2x_3 + 2x_5 + 3x_6 = 1 \\ \boxed{1}x_2 - 4x_3 + x_4 - 3x_5 = 2 \\ \boxed{-3}x_4 + x_5 - 3x_6 = 0 \\ \boxed{-2}x_6 = 1, \end{array} \right.$$

est échelonné. En effet, le nombre de coefficients nuls commençant chaque ligne est dans l'ordre : 0, 1, 3 et 5. Les pivots sont les coefficients encadrés. En revanche, le système :

$$\left\{ \begin{array}{l} x_1 + 3x_2 - 5x_6 = 0 \\ -2x_3 + 2x_4 - x_5 + 3x_6 = 1 \\ 5x_3 - x_4 + 2x_6 = 1 \\ x_5 - x_6 = 0, \end{array} \right.$$

n'est pas échelonné car les lignes 2 et 3 débutent par le même nombre de coefficients nuls. ◀

Voyons sur l'exemple suivant comment résoudre un système échelonné qui n'est pas triangulaire.

Exemple 2.4.4. Considérons le système échelonné de 2 équations à 3 inconnues suivant :

$$\begin{cases} 2x - y + 2z = -4 \\ \frac{1}{3}y - z = 2. \end{cases} \quad (2.3)$$

En passant les termes en z dans le second membre, le système devient :

$$\begin{cases} 2x - y = -4 - 2z \\ \frac{1}{3}y = 2 + z. \end{cases}$$

Si on traite z comme un paramètre, ce système d'inconnues x, y est triangulaire avec des coefficients diagonaux non nuls, il admet donc une unique solution quelle que soit la valeur de z . Les inconnues x et y sont dites principales et l'inconnue z est dite secondaire. En résolvant ce système triangulaire par remontée, et on trouve :

$$\begin{cases} x = 1 + \frac{1}{2}z \\ y = 6 + 3z. \end{cases}$$

Ainsi, les inconnus x et y s'expriment en fonction de z qui varie librement dans \mathbb{R} . On dit que z est une variable libre et que x et y sont des variables liées (leur valeur est déterminée par celle de z). Par conséquent, les solutions du système sont tous les triplets de la forme $(1 + \frac{1}{2}z, 6 + 3z, z)$ avec $z \in \mathbb{R}$. L'ensemble des solutions du système s'écrit :

$$\{(1 + \frac{1}{2}z, 6 + 3z, z) : z \in \mathbb{R}\}.$$

Le système (2.3) admet une infinité de solutions, décrite par un paramètre réel. Par exemple, pour $z = 0$, on trouve que $(1, 6, 0)$ est une solution du système; pour $z = -2$, on trouve que $(0, 0, -2)$ est une autre solution du système; etc. ▶

Remarque 2.4.5. Si on rajoute artificiellement l'équation $z = \lambda$, où λ est un paramètre réel, le système (2.3) est équivalent à :

$$\begin{cases} x = 1 + \frac{1}{2}\lambda \\ y = 6 + 3\lambda \\ z = \lambda \end{cases} \quad \lambda \in \mathbb{R}.$$

Géométriquement, on reconnaît l'équation paramétrique de la droite passant par le point $A(1, 6, 0)$ et de vecteur directeur $\vec{u}(\frac{1}{2}, 3, 1)$. Ce n'est pas étonnant : dans l'espace muni d'un repère, les équations $2x - y + 2z = -2$ et $\frac{1}{3}y - z = 2$ sont des équations cartésiennes de deux plans \mathcal{P}_1 et \mathcal{P}_2 . Résoudre le système (2.3) revient à chercher l'intersection de \mathcal{P}_1 et \mathcal{P}_2 . Ces deux plans n'étant pas parallèles, leur intersection est bien une droite. ▶

Définition 2.4.6. Dans un système échelonné :

- les inconnues dont les pivots sont les coefficients sont appelées les **inconnues principales** ou **variables liées**.
- les autres inconnues sont appelées les **inconnues secondaires** ou **variables libres**.
- le nombre d'inconnues principales est appelé le **rang** du système linéaire.

La méthode pour résoudre un système échelonné est la suivante :

1. Vérifier que les équations de compatibilité sont satisfaites. Si ce n'est pas le cas, le système n'a pas de solutions.

2. Passer les inconnues secondaires dans le second membre. Ces inconnues deviennent des paramètres.
3. Le système est alors triangulaire en les inconnues principales, avec des coefficients diagonaux non nuls (les pivots). On le résout par remontée.

On obtient ainsi l'expression des inconnues principales en fonction des inconnues secondaires, ces dernières pouvant prendre n'importe quelle valeur dans \mathbb{K} .

Remarque 2.4.7. Étant donné un système linéaire, il y a plusieurs façons de le transformer en un système échelonné équivalent, et toutes ces façons ne conduisent pas nécessairement aux mêmes inconnues principales et secondaires. En revanche, le nombre d'inconnues principales est toujours le même : le rang ne dépend pas de la façon d'échelonner un système. Cette affirmation sera démontrée au chapitre 4. ◀

Proposition 2.4.8. Pour tout système linéaire de rang r , on a :

- (i) $r \leq \min(n, p)$.
- (ii) le nombre d'inconnues secondaires est $n - r$.
- (iii) le nombre d'équations de compatibilité est $p - r$.
- (iv) si $r = p$ alors le système admet des solutions.
- (v) le système admet une unique solution si et seulement si $r = n$ et les équations de compatibilité sont vérifiées. ▶

Démonstration. Soit (S) un système échelonné de rang r .

(i), (ii), (iii) Le rang est par définition le nombre d'inconnues principales. Puisqu'il y a n inconnues au total, on a $r \leq n$ et $n - r$ est le nombre d'inconnues secondaires. Par ailleurs, le rang est également le nombre de lignes de coefficients non nulles dans (S) . Puisqu'il y a p lignes au total, on a $r \leq p$ et $p - r$ est le nombre d'équations de compatibilité. Enfin, on a établi que $r \leq n$ et $r \leq p$, donc $r \leq \min(n, p)$.

(iv) Si $r = p$, alors il n'y a pas d'équations de compatibilité, donc le système admet des solutions.

(v) Le système admet une solution si et seulement si les équations de compatibilité sont vérifiées. Dans ce cas :

- si $r < n$, alors (S) possède des variables libres, donc une infinité de solutions.
- si $r = n$, alors (S) ne possède pas de variables libres, donc la solution est unique. □

2.5 Algorithme d'élimination de Gauss

Théorème 2.5.1 (admis). Tout système linéaire est équivalent à un système échelonné.

La démonstration repose sur l'**algorithme du pivot de Gauss**, ou algorithme d'élimination de Gauss, qui permet de mettre n'importe quel système linéaire¹ sous forme d'un système échelonné équivalent. Le principe général de cet algorithme est d'effectuer des opérations élémentaires pour éliminer les variables l'une après l'autre, en échangeant des lignes si besoin.

Décrivons le fonctionnement de cet algorithme (cf. algorithme 1 pour le pseudo-code). Considérons un système linéaire :

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n = b_1 & (L_1) \\ a_{2,1}x_1 + a_{2,2}x_2 + \cdots + a_{2,n}x_n = b_2 & (L_1) \\ \vdots \\ a_{p,1}x_1 + a_{p,2}x_2 + \cdots + a_{p,n}x_n = b_p. & (L_p) \end{cases}$$

1. plus généralement, cet algorithme permet de trouver une forme échelonnée en lignes d'une matrice, voir le cours d'algèbre linéaire 2.

Le but est d'éliminer x_1 dans toutes les lignes qui suivent la première. Un problème qui peut se poser est que l'inconnue x_1 n'est pas nécessairement présente dans (L_1) : le coefficient $a_{1,1}$ peut être nul. Dans ce cas, il faut d'abord échanger (L_1) avec l'une des lignes qui suivent.

- Si $a_{1,1} \neq 0$, alors on utilise (L_1) pour éliminer x_1 dans les lignes suivantes. Pour tout $i \geq 2$, on effectue l'opération $L_i \leftarrow L_i - \frac{a_{i,1}}{a_{1,1}} L_1$; on obtient ainsi le système équivalent :

$$\left\{ \begin{array}{l} a_{1,1} x_1 + a_{1,2} x_2 + \cdots + a_{1,n} x_n = b_1 \\ a'_{2,2} x_2 + \cdots + a'_{2,n} x_n = b'_2 \quad L_2 \leftarrow L_2 - \frac{a_{2,1}}{a_{1,1}} L_1 \\ \vdots \\ a'_{p,2} x_2 + \cdots + a'_{p,n} x_n = b'_p \quad L_p \leftarrow L_p - \frac{a_{p,1}}{a_{1,1}} L_1 \end{array} \right.$$

L'inconnue x_1 est principale. On réitère la procédure sur le sous-système formé des équations (L_2) à (L_p) .

- Si $a_{1,1} = 0$, alors on cherche une ligne i pour laquelle $a_{i,1} \neq 0$.
 - Si un tel i existe, alors on échange (L_1) et (L_i) . On peut alors effectuer l'étape d'élimination comme précédemment.
 - Si $a_{i,1} = 0$ pour tout $i \in \llbracket 1, p \rrbracket$, alors l'inconnue x_1 est absente du système : elle est secondaire. On recommence la procédure avec la variable x_2 .

Remarque 2.5.2. L'algorithme laisse de la liberté dans le choix du pivot. Le meilleur choix dépend de qui exécute l'algorithme :

- Pour un être humain, mieux vaut choisir un pivot qui facilite les calculs. Le choix idéal est un pivot qui vaut 1 ou -1 , car ça nous évite d'effectuer des divisions lors de l'étape d'élimination.
- Pour une machine, mieux vaut choisir un pivot dont la valeur est loin de zéro. En effet, lors de l'étape d'élimination, on est amené à diviser par le pivot, ce qui conduit à des erreurs d'arrondis si celui-ci est trop petit. De façon générale, choisir un pivot ayant la plus grande valeur absolue conduit à une meilleure stabilité numérique. ◀

Algorithme 1 : Pivot de Gauss

```

r ← 0          // ligne du dernier pivot et nombre d'inconnues principales trouvées
pour j allant de 1 à n faire
    k ← argmaxi ∈ [r+1, p] |ai,j|          // ligne contenant le plus grand pivot potentiel
    si ak,j ≠ 0 alors                      // ak,j est un pivot et xj est une inconnue principale
        r ← r + 1
        si k > r alors Lk ↔ Lr

        // élimination de xj dans les lignes qui suivent
        si r < n alors pour i allant de r + 1 à n faire
            Li ← Li -  $\frac{a_{i,j}}{a_{r,j}}$  Lr
        fin
    fin
fin
// le système est échelonné et r contient le rang du système

```

2.6 Exemples

Exemple 2.6.1. On considère le système de 3 équations à 3 inconnues suivant :

$$\begin{cases} 8x + 2y - 2z = 9 \\ 3x - y + 2z = 7 \\ x + 2y - 3z = -1. \end{cases} \quad (2.4)$$

On échelonne le système avec la méthode du pivot de Gauss. On commence par échanger les lignes 1 et 3 pour avoir un meilleur pivot.

$$\begin{aligned} \begin{cases} 8x + 2y - 2z = 9 \\ 3x - y + 2z = 7 \\ x + 2y - 3z = -1 \end{cases} &\Leftrightarrow \begin{cases} x + 2y - 3z = -1 & L_1 \leftrightarrow L_3 \\ 3x - y + 2z = 7 \\ 8x + 2y - 2z = 9 & L_1 \leftrightarrow L_3 \end{cases} \\ &\Leftrightarrow \begin{cases} x + 2y - 3z = -1 \\ -7y + 11z = 10 & L_2 \leftarrow L_2 - 3L_1 \\ -14y + 22z = 17 & L_3 \leftarrow L_3 - 8L_1 \end{cases} \\ &\Leftrightarrow \begin{cases} x + 2y - 3z = -1 \\ -7y + 11z = 10 \\ 0 = -3. & L_3 \leftarrow L_3 - 2L_2 \end{cases} \end{aligned}$$

Le système (2.4) est incompatible, il n'admet pas de solutions. ◀

Exemple 2.6.2. On considère le système linéaire de 3 équations à 4 inconnues suivant :

$$\begin{cases} -x + 3z + 3t = 3 \\ x + 2y - 8z - 5t = -2 \\ 3x - 2y - 4z - 7t = -10. \end{cases} \quad (2.5)$$

On échelonne le système avec la méthode du pivot de Gauss :

$$\begin{aligned} \begin{cases} -x + 3z + 3t = 3 \\ x + 2y - 8z - 5t = -2 \\ 3x - 2y - 4z - 7t = -10 \end{cases} &\Leftrightarrow \begin{cases} -x + 3z + 3t = 3 \\ 2y - 5z - 2t = 1 & L_2 \leftarrow L_2 + L_1 \\ -2y + 5z + 2t = -1 & L_3 \leftarrow L_3 + 3L_1 \end{cases} \\ &\Leftrightarrow \begin{cases} -x + 3z + 3t = 3 \\ 2y - 5z - 2t = 1 \\ 0 = 0. & L_4 \leftarrow L_4 + L_2 \end{cases} \end{aligned}$$

Le système est de rang 2 et les équations de compatibilité sont vérifiées, donc le système admet une infinité de solutions. On passe les inconnues secondaires z et t dans le second membre :

$$\begin{aligned} \begin{cases} -x + 3z + 3t = 3 \\ 2y - 5z - 2t = 1 \end{cases} &\Leftrightarrow \begin{cases} -x = 3 - 3z - 3t \\ 2y = 1 + 5z + 2t \end{cases} \\ &\Leftrightarrow \begin{cases} x = -3 + 3z + 3t \\ y = \frac{1}{2} + \frac{5}{2}z + t. \end{cases} \end{aligned}$$

L'ensemble des solutions du système (2.5) est $\{(-3 + 3z + 3t, \frac{1}{2} + \frac{5}{2}z + t, z, t) : z, t \in \mathbb{R}\}$. ◀

Exemple 2.6.3. Soit $m \in \mathbb{R}$ un paramètre. On considère le système linéaire de 3 équations à 3 inconnues suivant :

$$\begin{cases} x + y = 1 \\ 2x + 3y + 3z = m \\ -x - 3y - 6z = 0. \end{cases}$$

On échelonne le système avec la méthode du pivot de Gauss :

$$\begin{cases} x + y = 1 \\ 2x + 3y + 2z = m \\ -x - 4y - 6z = 1 \end{cases} \iff \begin{cases} x + y = 1 \\ y + 2z = m - 2 \\ -3y - 6z = 1 \end{cases} \begin{array}{l} L_2 \leftarrow L_2 - 2L_1 \\ L_3 \leftarrow L_3 + L_1 \end{array} \\ \iff \begin{cases} x + y = 1 \\ y + 2z = m - 2 \\ 0 = 3m - 5 \end{cases} \begin{array}{l} \\ L_3 \leftarrow L_3 + 3L_2 \end{array}$$

Le système est compatible si et seulement si $m = \frac{5}{3}$. Dans ce cas, on passe l'inconnue secondaire z dans le second membre et résout le système par remontée :

$$\begin{cases} x + y = 1 \\ y + 2z = -\frac{1}{3} \end{cases} \iff \begin{cases} x = \frac{4}{3} - 2z \\ y = -\frac{1}{3} + 2z \end{cases}$$

Si $m = \frac{5}{3}$, l'ensemble des solutions est $\{(\frac{4}{3} - 2z, -\frac{1}{3} + 2z, z) : z \in \mathbb{R}\}$. Sinon, il n'y a pas de solutions. \blacktriangleleft

Exemple 2.6.4. Soit $m \in \mathbb{R}$ un paramètre. On considère le système linéaire de 3 équations à 3 inconnues suivant :

$$\begin{cases} 2x - y - z = 0 \\ 3x - y + mz = 0 \\ 2x - 3y + z = 0 \end{cases}$$

On échelonne le système avec la méthode du pivot de Gauss :

$$\begin{cases} 2x - y - z = 0 \\ 3x - y + mz = 0 \\ 2x - 3y + z = 0 \end{cases} \iff \begin{cases} 2x - y - z = 0 \\ y + (2m + 3)z = 0 \\ -2y + 2z = 0 \end{cases} \begin{array}{l} \\ L_2 \leftarrow 2L_2 - 3L_1 \\ L_3 \leftarrow L_3 - L_1 \end{array} \\ \iff \begin{cases} 2x - y - z = 0 \\ y + (2m + 3)z = 0 \\ (4m + 8)z = 0 \end{cases} \begin{array}{l} \\ \\ L_3 \leftarrow L_3 + 2L_2 \end{array}$$

Le système est de rang 2 ou 3 selon que $4m + 8$ est nul ou non. On discute selon la valeur de m :

- si $m \neq -2$, alors le système est triangulaire et admet pour unique solution $(0, 0, 0)$.
- si $m = -2$, alors le système est de rang 2. On passe l'inconnue secondaire z dans le second membre, et on résout le système par remontée :

$$\begin{cases} 2x - y - z = 0 \\ y - z = 0 \end{cases} \iff \begin{cases} x = z \\ y = z \end{cases}$$

L'ensemble des solutions du système est $\{(z, z, z) : z \in \mathbb{R}\}$. \blacktriangleleft

Chapitre 3

Espaces vectoriels

Initialement, les vecteurs ont été définis dans le cadre de la géométrie euclidienne en dimensions 2 et 3, et se sont avérés un outil puissant en physique pour modéliser des concepts comme les forces, la vitesse, l'accélération, etc. Les espaces vectoriels permettent de généraliser les techniques de calcul vectoriel à des espaces de dimension quelconques, mais aussi à des espaces non géométriques à priori (espaces de fonctions, de polynômes, de suites, etc.).

Le but de ce chapitre est d'introduire la notion d'espace vectoriel en partant du calcul vectoriel que vous avez étudié au lycée. On étudiera ensuite des parties remarquables d'un espace vectoriel (droites, plans, sous-espaces de dimension supérieure), ainsi que des opérations sur celles-ci (intersection et somme de sous-espaces). Puis, on verra la notion de famille libre de vecteurs (généralisation de la non colinéarité et de la non coplanarité), ce qui permettra de définir la notion de base d'un espace vectoriel. Pour finir, on définira ce qu'est la dimension d'un espace vectoriel.

3.1 Vecteurs en géométrie euclidienne

Cette section est constituée de rappels de lycée concernant les vecteurs en géométrie euclidienne dans l'espace. Aller directement à la section 3.2 pour la définition des espaces vectoriels.

3.1.a Définition géométrique des vecteurs

On se place dans le plan ou dans l'espace euclidien.

Définition 3.1.1. Un **segment orienté** est une paire de points (A, B) , où A est l'origine du segment orienté et B son extrémité. On le représente géométriquement par un segment fléché allant de l'origine à l'extrémité.

Notons qu'un segments orienté est localisé dans l'espace : des segments orientés sont égaux si et seulement si ils ont la même origine et la même extrémité. Or, on voudrait définir les vecteurs comme des « segments orientés libres », c'est-à-dire non localisés dans l'espace. L'idée est de définir un vecteur comme l'ensemble de tous les segments orientés identiques à leur position près.

Pour ce faire, il faut déterminer les propriétés invariantes d'un segment orienté lorsqu'on change sa position. Si on considère un segment orienté (A, B) et qu'on lui applique une translation pour obtenir le segment orienté (C, D) , on peut montrer (cf. figure 3.1) :

- qu'ils ont la même direction : (AB) et (CD) sont parallèles ;
- qu'ils ont la même longueur : $AB = CD$;
- qu'ils ont le même sens : le quadrilatère $ABDC$ n'est pas croisé.

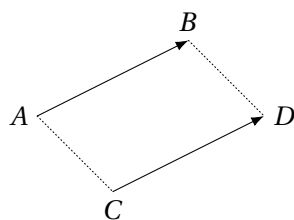


Figure 3.1 – Segments orientés équipollents.

Ces trois propriétés sont conjointement équivalentes au fait que $ABDC$ est un parallélogramme, c'est-à-dire au fait que les diagonales $[AD]$ et $[BC]$ ont le même milieu¹.

Définition 3.1.2. On dit que deux segments orientés (A, B) et (C, D) sont **équipollents** si $ABDC$ est un parallélogramme (potentiellement aplati). On appelle vecteur \overrightarrow{AB} l'ensemble de tous les segments orientés équipollents à (A, B) .

Remarque 3.1.3. L'équipollence est une relation d'équivalence sur les segments orientés et un vecteur est une classe d'équivalence pour cette relation (cf. annexe A). ◀

Le segment orienté (A, B) est appelé un représentant du vecteur \overrightarrow{AB} . Si (C, D) est un segment orienté équipollent à (A, B) , alors (C, D) est aussi un représentant de \overrightarrow{AB} , et on a l'égalité entre les vecteurs \overrightarrow{AB} et \overrightarrow{CD} :

$$\overrightarrow{AB} = \overrightarrow{CD} \stackrel{\text{def.}}{\iff} (A, B) \text{ et } (C, D) \text{ sont équipollents } \stackrel{\text{def.}}{\iff} ABDC \text{ est un parallélogramme.}$$

La longueur du segment orienté (A, B) est appelée la norme du vecteur \overrightarrow{AB} et on la note $\|\overrightarrow{AB}\|$. Elle ne dépend pas du choix du représentant du vecteur, car des segments orientés équipollents ont tous la même longueur.

Définition 3.1.4. Pour tout point A , le vecteur \overrightarrow{AA} est un vecteur particulier appelé **vecteur nul** et noté $\vec{0}$. Il n'a pas de direction, pas de sens et sa norme est nulle.

Les vecteurs ne sont pas que des objets géométriques, ce sont aussi des objets algébriques : on peut effectuer des opérations avec. La première est l'addition vectorielle.

Définition 3.1.5. Pour tous points A , B et C , on définit l'addition vectorielle de \overrightarrow{AB} et \overrightarrow{BC} par la **relation de Chasles** :

$$\overrightarrow{AB} + \overrightarrow{BC} := \overrightarrow{AC}.$$

La relation de Chasles permet d'additionner des vecteurs en mettant bout à bout des représentants. Pour additionner des vecteurs à partir de représentants ayant la même origine, on utilise la règle du parallélogramme (cf. figure 3.2).

Proposition 3.1.6 (règle du parallélogramme). Soient A , B , C et D des points du plan. On a l'égalité vectorielle $\overrightarrow{AB} + \overrightarrow{AC} = \overrightarrow{AD}$ si et seulement si $ABDC$ est un parallélogramme. ◀

L'autre opération que l'on peut effectuer sur les vecteurs est la multiplication par des scalaires, c'est-à-dire des nombres réels.

1. cette définition a l'avantage de s'appliquer même dans les cas dégénérés, par exemple si $A = B$, ou si les segments sont alignés.

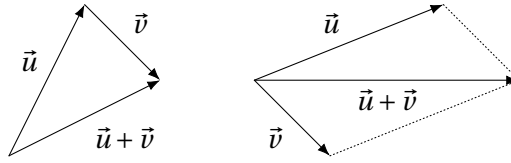


Figure 3.2 – Relation de Chasles et règle du parallélogramme.

Définition 3.1.7. Soit $\lambda \in \mathbb{R}^*$ et soit \vec{u} un vecteur non nul. On définit le vecteur $\lambda \vec{u}$ comme le vecteur :

- de même direction que \vec{u} ;
- de même sens que \vec{u} si $\lambda > 0$ et de sens opposé si $\lambda < 0$;
- de norme $|\lambda| \times \|\vec{u}\|$.

Si $\lambda = 0$ ou si $\vec{u} = \vec{0}$, on définit que $\lambda \vec{u} = \vec{0}$.

On dit que deux vecteurs sont colinéaires si l'un est un multiple de l'autre. Géométriquement, les vecteurs \vec{AB} et \vec{AC} sont colinéaires si et seulement si les points A , B et C sont alignés.

Si $\vec{u}_1, \dots, \vec{u}_n$ sont des vecteurs et si $\lambda_1, \dots, \lambda_n$ sont des scalaires, on peut combiner des multiplications par des scalaires et des additions vectorielles pour former le vecteur $\lambda_1 \vec{u}_1 + \dots + \lambda_n \vec{u}_n$. Une telle expression est appelée une **combinaison linéaire** des vecteurs $\vec{u}_1, \dots, \vec{u}_n$.

On dit que trois vecteurs sont coplanaires si l'un d'eux peut s'écrire comme une combinaison linéaire des deux autres. Géométriquement, les vecteurs \vec{AB} , \vec{AC} et \vec{AD} sont coplanaires si et seulement si les points A , B , C et D sont contenus dans un même plan (cf. figure 3.3).

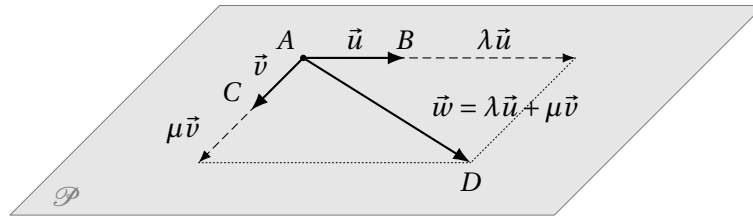


Figure 3.3 – Vecteurs coplanaires.

3.1.b Vecteurs, droites et plans de l'espace

Définition 3.1.8. Un repère de l'espace est un quadruplet $(O, \vec{i}, \vec{j}, \vec{k})$ où O est un point appelé l'origine du repère et $\vec{i}, \vec{j}, \vec{k}$ sont trois vecteurs non coplanaires.

Un repère de l'espace permet d'associer des coordonnées aux points et aux vecteurs de l'espace (cf. figure 3.4).

Définition 3.1.9. On munit l'espace d'un repère $(O, \vec{i}, \vec{j}, \vec{k})$.

- Tout vecteur \vec{u} s'écrit de manière unique comme $\vec{u} = a\vec{i} + b\vec{j} + c\vec{k}$ avec a, b et c des scalaires. Le triplet $(a, b, c) \in \mathbb{R}^3$ s'appelle les coordonnées de \vec{u} dans la base $(\vec{i}, \vec{j}, \vec{k})$.
- Si M est un point de l'espace, on définit ses coordonnées $(x_M, y_M, z_M) \in \mathbb{R}^3$ comme celles du vecteur \vec{OM} dans la base $(\vec{i}, \vec{j}, \vec{k})$:

$$\vec{OM} = x_M \vec{i} + y_M \vec{j} + z_M \vec{k}.$$

Les coordonnées de la somme de deux vecteurs et du produit d'un vecteur par un scalaire se calculent facilement : il suffit d'effectuer les opérations coordonnée par coordonnée.

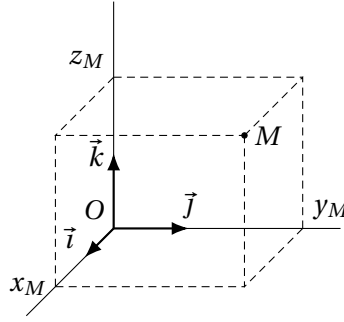


Figure 3.4 – Coordonnées d'un point dans un repère de l'espace.

Proposition 3.1.10. On munit le plan d'un repère. Pour tous vecteurs $\vec{u}(a, b, c)$ et $\vec{v}(a', b', c')$ et pour tout scalaire λ , la somme $\vec{u} + \vec{v}$ a pour coordonnées $(a + a', b + b', c + c')$ et $\lambda \vec{u}$ a pour coordonnées $(\lambda a, \lambda b, \lambda c)$. ◀

Corollaire 3.1.11. On munit l'espace d'un repère. Pour tous points $A(x_A, y_A, z_A)$ et $B(x_B, y_B, z_B)$ de l'espace, le vecteur \overrightarrow{AB} a pour coordonnées $(x_B - x_A, y_B - y_A, z_B - z_A)$. ◀

Toute droite est caractérisée par un point et un vecteur directeur. En effet, si on munit l'espace d'un repère, et qu'on considère \mathcal{D} une droite de vecteur directeur $\vec{u}(a, b, c)$ et passant par le point $A(x_A, y_A, z_A)$, alors pour tout point $M(x, y, z)$ de l'espace, on a :

$$M(x, y, z) \in \mathcal{D} \iff \exists \lambda \in \mathbb{R}, \overrightarrow{AM} = \lambda \vec{u} \iff \exists \lambda \in \mathbb{R}, \begin{cases} x = x_A + \lambda a \\ y = y_A + \lambda b \\ z = z_A + \lambda c. \end{cases}$$

Le système précédent est appelé une **représentation paramétrique** de la droite \mathcal{D} . Ce n'est pas un système à résoudre mais un critère d'appartenance à la droite \mathcal{D} . Chaque valeur de λ fournit les coordonnées d'un point de la droite et chaque point de la droite possède des coordonnées de cette forme pour une certaine valeur de λ (cf. figure 3.5).

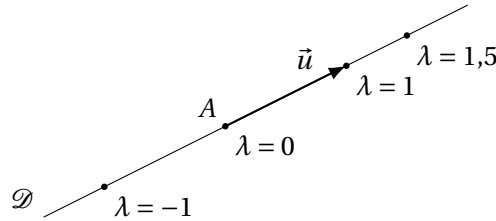


Figure 3.5 – Droite paramétrée.

Les plans admettent également une représentation paramétrique.

Définition 3.1.12. Soit A un point de l'espace et soient \vec{u} et \vec{v} des vecteurs non colinéaires. Soient B et C les points de l'espace tels que $\overrightarrow{AB} := \vec{u}$ et $\overrightarrow{AC} := \vec{v}$. Puisque \vec{u} et \vec{v} ne sont pas colinéaires, les points A , B et C sont non alignés, donc ils définissent un unique plan \mathcal{P} . On dit que \mathcal{P} est le plan passant par A et dirigé par \vec{u} et \vec{v} .

Soit \mathcal{P} le plan dirigé par les vecteurs $\vec{u}(a, b, c)$ et $\vec{v}(a', b', c')$, et passant par le point $A(x_A, y_A, z_A)$. Pour tout point $M(x, y, z)$ de l'espace, on a :

$$M(x, y, z) \in \mathcal{P} \iff \exists (\lambda, \mu) \in \mathbb{R}^2, \overrightarrow{AM} = \lambda \vec{u} + \mu \vec{v} \iff \exists (\lambda, \mu) \in \mathbb{R}^2, \begin{cases} x = x_A + \lambda a + \mu a' \\ y = y_A + \lambda b + \mu b' \\ z = z_A + \lambda c + \mu c'. \end{cases}$$

On obtient ainsi une **représentation paramétrique** du plan \mathcal{P} . Les points d'un plan sont décrits par deux paramètres λ et μ : chaque valeur de λ et de μ fournit les coordonnées d'un point du plan et chaque point du plan possède des coordonnées de cette forme pour certaines valeurs des paramètres λ et μ .

Dans l'espace, un plan est également caractérisé par une équation cartésienne.

Théorème 3.1.13. On munit l'espace d'un repère. Pour tous $a, b, c, d \in \mathbb{R}$ tels que $(a, b, c) \neq (0, 0, 0)$, l'ensemble des points de coordonnées $(x, y, z) \in \mathbb{R}^3$ tels que :

$$ax + by + cz + d = 0.$$

est un plan de l'espace. L'équation $ax + by + cz + d = 0$ est appelée une **équation cartésienne** de ce plan. Réciproquement, tout plan de l'espace est caractérisé par une équation de cette forme.

Puisque l'intersection de deux plans non parallèles est une droite, il est possible de caractériser une droite par un système de deux équations, chaque équation étant l'équation cartésienne d'un plan.

Corollaire 3.1.14. On munit l'espace d'un repère. L'ensemble des points de l'espace dont les coordonnées (x, y, z) sont solution d'un système linéaire compatible de rang 2 :

$$\begin{cases} ax + by + cz = d \\ a'x + b'y + c'z = d', \end{cases}$$

est une droite de l'espace. Réciproquement, toute droite de l'espace est caractérisée par un système linéaire compatible de rang 2. ◀

3.1.c Vecteurs en dimension n

On a vu que l'ensemble des solutions d'un système linéaire à 3 inconnues, lorsque celui-ci est infini, s'interprète géométriquement comme la représentation paramétrique d'une droite ou d'un plan de l'espace (selon le rang du système). On peut se demander ce qu'il en est lorsque le nombre d'inconnues est supérieur à 3.

Exemple 3.1.15. Considérons le système linéaire à 4 inconnues suivant :

$$\begin{aligned} \begin{cases} x + 2y + z - 2t = 1 \\ 2x + 5y + z - 6t = 5 \\ -x - 3y - z + 5t = -5 \end{cases} &\iff \begin{cases} x + 2y + z - 2t = 1 \\ y - z - 2t = 3 & L_2 \leftarrow L_2 - 2L_1 \\ -y + 3t = -4 & L_3 \leftarrow L_3 + L_1 \end{cases} \\ &\iff \begin{cases} x + 2y + z - 2t = 1 \\ y - z - 2t = 3 \\ -z + t = -1 & L_3 \leftarrow L_3 + L_2 \end{cases} \\ &\iff \begin{cases} x = -8 - 5t \\ y = 4 + 3t \\ z = 1 + t \end{cases} \\ &\iff \begin{cases} x = -8 - 5\lambda \\ y = 4 + 3\lambda \\ z = 1 + \lambda \\ t = \lambda \end{cases} \quad \lambda \in \mathbb{R}. \end{aligned}$$

Si on interprète (x, y, z, t) comme les coordonnées d'un point dans un repère d'un espace à 4 (!) dimensions, alors les équations ci-dessus sont une représentation paramétrique de la droite passant par le point $A(-8, 4, 1, 0)$ et de vecteur directeur $\vec{u}(-5, 3, 1, 1)$. Ainsi, l'ensemble des solutions du système d'interprète comme une droite dans un espace à 4 dimensions. ◀

Plus généralement, si un système linéaire à n inconnues est compatible et de rang r , alors l'ensemble de ses solutions est décrit par $n - r$ paramètres, ce qu'on peut interpréter géométriquement comme une représentation paramétrique d'un sous-espace de dimension $n - r$ dans un espace de dimension n . Ainsi, pour étudier théoriquement les systèmes linéaires, il est naturel de définir les notions d'« espace de dimension n » et de « sous-espaces », lorsque $n > 3$.

Si l'idée d'un espace géométrique à n dimensions échappe à notre intuition, il est en revanche très simple de considérer un nombre arbitraire de coordonnées. Plutôt que définir les vecteurs de façon géométrique (c'est-à-dire à partir de segments orientés en géométrie euclidienne), on va les définir de façon algébrique, à partir de leurs coordonnées.

Définition 3.1.16. On définit un vecteur numérique dans un espace de dimension n comme une liste de n nombres réels (x_1, \dots, x_n) , c'est-à-dire comme un élément de \mathbb{R}^n . Pour tous vecteurs numériques $\vec{x} := (x_1, \dots, x_n)$ et $\vec{y} := (y_1, \dots, y_n)$ de \mathbb{R}^n , on définit leur somme ainsi que leur produit par un scalaire $\lambda \in \mathbb{R}$ par :

$$\vec{x} + \vec{y} := (x_1 + y_1, \dots, x_n + y_n), \quad \lambda \vec{x} := (\lambda x_1, \dots, \lambda x_n).$$

Le n -uplet $\vec{0} := (0, \dots, 0)$ est appelé le vecteur nul de \mathbb{R}^n .

Une fois qu'un repère est fixé, les vecteurs géométriques de l'espace s'identifient à des vecteurs numériques de \mathbb{R}^3 via leurs coordonnées. De plus, d'après la proposition 3.1.10, les opérations d'addition vectorielle et de multiplication par un scalaire des vecteurs géométriques correspondent à celles définies pour les vecteurs numériques.

À partir des propriétés de l'addition et de la multiplication dans \mathbb{R} , il est facile de vérifier que l'addition de \mathbb{R}^n vérifie les propriétés suivantes :

- elle est associative : $(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$;
- elle est commutative : $\vec{x} + \vec{y} = \vec{y} + \vec{x}$;
- elle possède un élément neutre : le vecteur nul est tel que $\vec{x} + \vec{0} = \vec{x}$;
- tout vecteur $\vec{x} = (x_1, \dots, x_n)$ possède un opposé $-\vec{x} := (-x_1, \dots, -x_n)$, de sorte que $\vec{x} + (-\vec{x}) = \vec{0}$;

et que la multiplication par un scalaire possède quant à elle les propriétés suivantes :

- $1\vec{x} = \vec{x}$;
- $\lambda(\vec{x} + \vec{y}) = \lambda\vec{x} + \lambda\vec{y}$;
- $\lambda(\mu\vec{x}) = (\lambda\mu)\vec{x}$;
- $(\lambda + \mu)\vec{x} = \lambda\vec{x} + \mu\vec{x}$.

Ces propriétés sont le cœur du calcul vectoriel et vont servir de modèle pour définir les espaces vectoriels.

Remarque 3.1.17. Dans \mathbb{R}^n , il n'y a pas de distinction claire entre points et vecteurs. En effet, un n -uplet (x_1, \dots, x_n) peut être vu aussi bien comme la position d'un point que comme un vecteur. ◀

En algèbre linéaire, lorsque $n = 2$ ou $n = 3$, on adoptera la représentation suivante des vecteurs numériques de \mathbb{R}^n (cf. figure 3.6). On munit le plan ($n = 2$) ou l'espace ($n = 3$) d'un repère orthonormé. Le vecteur nul est représenté par un point : l'origine O du repère. Les vecteurs numériques non nuls (x_1, x_2) ou (x_1, x_2, x_3) sont représentés par des segments orientés d'origine O et d'extrémité le point de coordonnées (x_1, x_2) ou (x_1, x_2, x_3) . Alternativement, un vecteur numérique peut être représenté par un unique point : l'extrémité du segment orienté. C'est la représentation qu'on utilisera de préférence pour les ensembles de vecteurs : on les représente par des ensembles de points plutôt que par des ensembles de segments orientés.

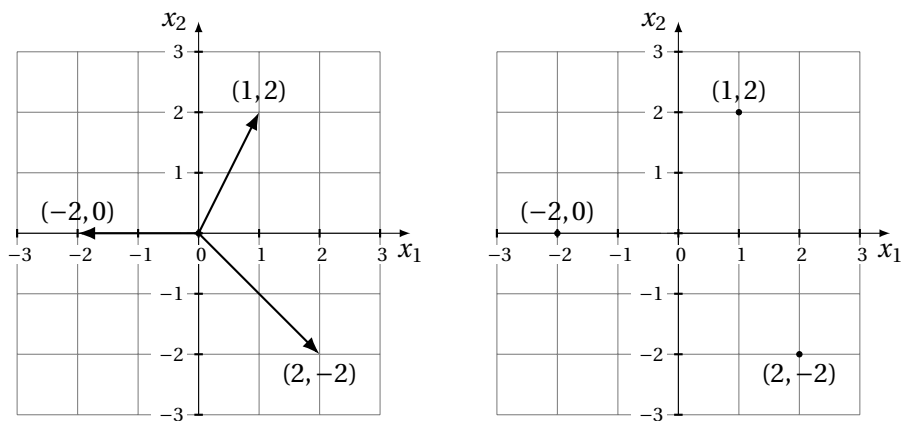


Figure 3.6 – Représentation des éléments de \mathbb{R}^2 comme des segments orientés ou comme des points.

3.2 Structure d'espace vectoriel

3.2.a Définition axiomatique

La définition algébrique des vecteurs comme éléments de \mathbb{R}^n met en lumière les deux opérations fondamentales sur les vecteurs : l'addition vectorielle et la multiplication par un scalaire. Or, on connaît d'autres objets mathématiques qui peuvent être additionnés entre eux et multipliés par des nombres réels, et pour lesquels les mêmes propriétés algébriques de l'addition vectorielle et de la multiplication par un scalaire sont vérifiées.

Par exemple, c'est le cas des fonctions : on peut additionner des fonctions et on peut multiplier une fonction par une constante réelle pour former une nouvelle fonction ; la fonction constante égale à zéro faisant office de « vecteur nul » pour les fonctions. La structure algébrique est la même que pour les vecteurs de \mathbb{R}^n , il est donc légitime de considérer les fonctions comme des sortes de vecteurs.

Remarque 3.2.1. On ne dit pas que les fonctions sont la même chose que les vecteurs géométriques ou les n -uplets de nombres : il y a toutes sortes d'opérations qu'on peut effectuer sur les fonctions (dérivation, intégration, évaluation, etc.) qui n'existent pas pour des vecteurs géométriques. Mais si on s'en tient à l'addition et à la multiplication par des réels, alors les vecteurs géométriques, les n -uplets de nombres réels et les fonctions se comportent de la même façon d'un point de vue algébrique, et c'est en ce sens qu'on les qualifie de « vecteurs ».

L'ensemble des vecteurs géométriques du plan ou de l'espace, l'ensemble des n -uplets de nombres et l'ensemble des fonctions sont ce qu'on appelle des **espaces vectoriels**. Cette notion se forme au cours du XIX^e siècle et consiste en une approche axiomatique des vecteurs. On qualifie d'espace vectoriel n'importe quel ensemble d'objets pouvant être additionnés entre eux et multipliés par des nombres, en satisfaisant certains axiomes (cf. définition 3.2.4). L'important n'est pas la nature des vecteurs (segments orientés, n -uplets, fonctions, etc), mais la structure algébrique (les opérations d'addition vectorielle et de multiplication par un scalaire).

Enfin, on a considéré jusqu'à présent que les nombres par lesquels on pouvait multiplier les vecteurs étaient les nombres réels. Mais d'autres ensembles de nombres peuvent être utilisés, comme par exemple les nombres complexes.

Définition 3.2.2. Dans ce cours, \mathbb{K} désignera l'ensemble \mathbb{R} , ou bien l'ensemble \mathbb{C} . Les éléments de \mathbb{K} sont appelés les **scalaires**.

Remarque 3.2.3. Les résultats de ce cours sont valables plus généralement si \mathbb{K} est un corps (commutatif) quelconque, mais on se limitera aux nombres réels et aux nombres complexes dans les exemples et les exercices.

Définition 3.2.4. Un **espace vectoriel** sur \mathbb{K} (\mathbb{K} -espace vectoriel) est un ensemble E dont les éléments sont appelés des **vecteurs**, qui est muni d'une loi de composition interne notée $+$ appelée **addition vectorielle**, ainsi que d'une loi de composition externe de domaine \mathbb{K} notée \cdot appelée **multiplication par un scalaire** :

$$\begin{aligned} + : E \times E &\longrightarrow E & \cdot : \mathbb{K} \times E &\longrightarrow E \\ (u, v) &\longmapsto u + v, & (\lambda, u) &\longmapsto \lambda \cdot u, \end{aligned}$$

vérifiant les axiomes suivants :

1. Pour tous $u, v, w \in E$, on a $(u + v) + w = u + (v + w)$.
2. Pour tous $u, v \in E$, on a $u + v = v + u$.
3. Il existe un vecteur noté 0_E , appelé **vecteur nul**, tel que pour tout $u \in E$, on a $u + 0_E = u$.
4. Pour tout $u \in E$, il existe un vecteur noté $-u$, appelé **vecteur opposé** de u , tel que $u + (-u) = 0_E$.
5. Pour tout $u \in E$, on a $1 \cdot u = u$.
6. Pour tous $\lambda, \mu \in \mathbb{K}$ et pour tout $u \in E$, on a $\lambda \cdot (\mu \cdot u) = (\lambda \times \mu) \cdot u$.
7. Pour tout $\lambda \in \mathbb{K}$ et pour tous $u, v \in E$, on a $\lambda \cdot (u + v) = (\lambda \cdot u) + (\lambda \cdot v)$.
8. Pour tous $\lambda, \mu \in \mathbb{K}$ et pour tout $u \in E$, on a $(\lambda + \mu) \cdot u = (\lambda \cdot u) + (\mu \cdot u)$.

Remarque 3.2.5. Les axiomes 1 à 4 signifient que $(E, +)$ est un groupe abélien, voir annexe A. ◀

Remarque 3.2.6. En algèbre linéaire, on ne note généralement pas de flèche au-dessus des vecteurs. En géométrie en revanche, on les écrit pour distinguer les vecteurs des points. ▶

Dans un espace vectoriel, il y a unicité du vecteur nul et de l'opposé d'un vecteur. Pour cette raison, on peut parler du vecteur nul d'un espace vectoriel et du vecteur opposé d'un vecteur.

Proposition 3.2.7. Pour tout \mathbb{K} -espace vectoriel E :

- (i) le vecteur nul de E est unique.
- (ii) pour tout $u \in E$, le vecteur opposé de u est unique. ▶

Démonstration. Soit E un \mathbb{K} -espace vectoriel.

(i) Soient 0_E et $0'_E$ deux vecteurs satisfaisant l'axiome 3. D'une part, on a $0_E + 0'_E = 0'_E$ car 0_E est un vecteur nul. D'autre part, on a $0_E + 0'_E = 0_E$ car $0'_E$ est aussi un vecteur nul. Par conséquent, on a l'égalité $0_E = 0'_E$, ce qui prouve l'unicité du vecteur nul.

(ii) Soit $u \in E$ et soient $(-u)$ et $(-u)'$ deux vecteurs satisfaisant l'axiome 4. D'une part, on a :

$$[(-u) + u] + (-u)' = 0_E + (-u)' = (-u)',$$

car $(-u)$ est l'opposé de u . D'autre part, on a :

$$(-u) + [u + (-u)'] = (-u) + 0_E = (-u),$$

car $(-u)'$ est aussi l'opposé de u . Par associativité de l'addition (axiome 1), ces expressions sont égales. Par conséquent, on a $(-u)' = (-u)$, ce qui prouve l'unicité du vecteur opposé de u . ◻

On définit la soustraction vectorielle $u - v$ comme l'addition de u et de $(-v)$. Comme d'habitude, le symbole de la multiplication peut être omis : on note $\lambda u := \lambda \cdot u$. De plus, la multiplication par un scalaire est prioritaire sur l'addition vectorielle. Ainsi, l'expression « $\lambda u + v$ » doit être comprise comme $(\lambda \cdot u) + v$.

Proposition 3.2.8. Soit E un \mathbb{K} -e.v. Pour tout $\lambda \in \mathbb{K}$ et pour tout $u \in E$, on a :

- | | |
|----------------------------|---|
| (i) $0u = 0_E$. | (iii) $(-1)u = -u$. |
| (ii) $\lambda 0_E = 0_E$. | (iv) $\lambda u = 0_E \iff \lambda = 0 \text{ ou } u = 0_E$. |

Démonstration. Soient $\lambda \in \mathbb{K}$ et $u \in E$.

(i) On a d'après l'axiome 8 :

$$0u = (0 + 0)u = 0u + 0u.$$

En ajoutant $-0u$ à chaque membre de cette égalité, on obtient $0_E = 0u$.

(ii) On a d'après l'axiome 7 :

$$\lambda 0_E = \lambda(0_E + 0_E) = \lambda 0_E + \lambda 0_E.$$

En ajoutant $-\lambda 0_E$ à chaque membre, on obtient $0_E = \lambda 0_E$.

(iii) On veut montrer que $(-1) \cdot u$ est l'opposé de u , c'est-à-dire que $u + (-1)u = 0_E$. On a d'après les axiomes 5 et 8 :

$$u + (-1)u = 1u + (-1)u = [1 + (-1)]u = 0u = 0_E.$$

(iv) On procède par double implication. Si $\lambda = 0$ ou $u = 0_E$, alors $\lambda \cdot u = 0_E$ d'après (i) et (ii). Réciproquement, supposons que $\lambda u = 0_E$. On fait une disjonction de cas :

- ou bien $\lambda \neq 0$, alors on multiplie chaque membre de l'égalité $\lambda u = 0_E$ par λ^{-1} , et on obtient :

$$\lambda u = 0_E \implies (\lambda^{-1} \lambda)u = \lambda^{-1} 0_E \implies 1 \cdot u = 0_E \implies u = 0_E.$$

- ou bien $\lambda = 0$.

Par conséquent, on a montré que $\lambda = 0$ ou $u = 0_E$. □

3.2.b Exemples fondamentaux

Exemple 3.2.9. L'ensemble \mathbb{K} est un \mathbb{K} -espace vectoriel, comme on le vérifie facilement : l'addition vectoriel est l'addition de \mathbb{K} , la multiplication par un scalaire est la multiplication de \mathbb{K} et le vecteur nul est le nombre 0.

Plus généralement, pour $n \in \mathbb{N}^*$, on considère l'ensemble \mathbb{K}^n des n -uplets d'éléments de \mathbb{K} . Pour tous $x := (x_1, \dots, x_n) \in \mathbb{K}^n$ et $y := (y_1, \dots, y_n) \in \mathbb{K}^n$, et pour tout $\lambda \in \mathbb{K}$, on définit la somme $x + y$ et le produit λx par :

$$x + y := (x_1 + y_1, \dots, x_n + y_n), \quad \lambda x := (\lambda x_1, \dots, \lambda x_n).$$

On peut vérifier que l'ensemble \mathbb{K}^n muni de ces lois est un \mathbb{K} -espace vectoriel, le vecteur nul de \mathbb{K}^n étant le n -uplet $(0, \dots, 0)$. ◀

Exemple 3.2.10. L'ensemble \mathbb{C} est un \mathbb{R} -espace vectoriel : l'addition est celle de \mathbb{C} et la multiplication par un scalaire est la multiplication d'un nombre complexe par un nombre réel. Le vecteur nul de \mathbb{C} est le nombre 0. Notons que \mathbb{C} est également un \mathbb{C} -espace vectoriel d'après l'exemple précédent. ◀

Exemple 3.2.11. Soit A un ensemble non vide, on considère l'ensemble $\mathcal{F}(A, \mathbb{K})$ des applications définies sur A et à valeurs dans \mathbb{K} . Pour tous $f, g \in \mathcal{F}(A, \mathbb{K})$ et pour tout $\lambda \in \mathbb{K}$, on définit les applications $f + g \in \mathcal{F}(A, \mathbb{K})$ et $\lambda f \in \mathcal{F}(A, \mathbb{K})$ par :

$$\forall a \in A, \quad (f + g)(a) := f(a) + g(a), \quad (\lambda f)(a) := \lambda \times f(a).$$

L'ensemble $\mathcal{F}(A, \mathbb{K})$ muni de ces lois est un \mathbb{K} -espace vectoriel et le vecteur nul de $\mathcal{F}(A, \mathbb{K})$ est l'application constante égale à 0. En particulier :

- l'ensemble $\mathbb{K}^{\mathbb{N}} := \mathcal{F}(\mathbb{N}, \mathbb{K})$ des suites numériques est un espace vectoriel.
- l'ensemble $\mathcal{F}(I, \mathbb{K})$ des fonctions définies sur un même intervalle I est un espace vectoriel. ◀

À partir de maintenant et dans le reste de ce chapitre, on fixe E un \mathbb{K} -espace vectoriel.

3.3 Sous-espaces vectoriels

L'espace euclidien contient des parties remarquables, les droites et les plans, dont l'étude est au cœur de la géométrie euclidienne. De la même façon, un espace vectoriel E contient des parties remarquables qui ont la propriété d'être elles aussi des espaces vectoriels. Ces parties sont appelées des sous-espaces vectoriels de E .

3.3.a Caractérisation des sous-espaces vectoriels

Définition 3.3.1. Soit F une partie de E . On dit que F est un **sous-espace vectoriel** de E si la restriction à F des lois $+$ et \cdot munissent F d'une structure de \mathbb{K} -espace vectoriel.

Exemple 3.3.2. Les ensemble E et $\{0_E\}$ sont des s.e.v.² de E . ◀

Soit F est un s.e.v. de E . Puisque F est un espace vectoriel, il possède un vecteur nul 0_F et tout vecteur de F possède un opposé dans F . Mais F est inclus dans E qui possède son propre vecteur nul 0_E , et chaque vecteur de F possède aussi un vecteur opposé dans E . D'où les questions :

- le vecteur nul de F est-il le même que le vecteur nul de E ?
- si $u \in F$, l'opposé de u dans F est-il égal à son opposé dans E ?

La réponse à ces questions est affirmative.

Proposition 3.3.3. Soit F est un s.e.v. de E .

- (i) Le vecteur nul de F est 0_E .
- (ii) Soit $u \in F$ et soit $-u$ son vecteur opposé dans E . Alors $-u$ est aussi l'opposé de u dans F . ◀

Démonstration. Soit F un s.e.v. de E .

(i) D'après la proposition 3.2.8-(ii) dans le \mathbb{K} -e.v. F , on a $0 \cdot 0_F = 0_F$, et d'après la proposition 3.2.8-(i) dans le \mathbb{K} -e.v. E , on a $0 \cdot 0_F = 0_E$. Par conséquent, $0_F = 0_E$.

(ii) Soit $u \in F$. On note $(-u)_E$ l'opposé de u dans E , et $(-u)_F$ l'opposé de u dans F . D'après la proposition 3.2.8-(iii) appliquée aux \mathbb{K} -e.v. E et F successivement, on a $(-1)u = (-u)_E$ et $(-1)u = (-u)_F$. Donc $(-u)_F = (-u)_E$. ◻

Pour montrer qu'une partie de E est un sous-espace vectoriel, on utilise la caractérisation suivante. Grâce à celle-ci, il n'est pas nécessaire de vérifier les huit axiomes de la définition 3.2.4, il suffit de vérifier seulement deux conditions. En général, pour montrer qu'un ensemble est un espace vectoriel, on montrera que c'est un sous-espace vectoriel d'un espace vectoriel déjà connu (\mathbb{K}^n , espace des suites, espaces des fonctions, etc).

Théorème 3.3.4 (caractérisation des s.e.v.). Une partie F est un sous-espace vectoriel de E si et seulement si les deux propriétés suivantes sont vérifiées :

1. F est non vide.
2. F est stable par combinaisons linéaires : $\forall u, v \in F, \forall \lambda \in \mathbb{K}, u + \lambda v \in F$.

Remarque 3.3.5. En pratique, pour montrer que F est non vide, on vérifie si le vecteur nul appartient à F : si c'est le cas alors F est non vide et il reste à vérifier la stabilité par combinaisons linéaires ; sinon F n'est pas un s.e.v. de E . ◀

2. abréviation de « sous-espace vectoriel ».

Démonstration. Soit $F \subset E$. On procède par double implication.

(\implies) Si F est un s.e.v. de E , alors F est non vide car il contient le vecteur nul, et la stabilité de F par combinaisons linéaires est évidente.

(\impliedby) Supposons que F soit non vide et stable par combinaisons linéaires. Montrons que F est un s.e.v. de E .

• **Affirmation.** La restriction de la loi $+$ à F est une loi de composition interne sur F qui vérifie les axiomes 1 à 4 de la définition 3.2.4.

En effet, si $u, v \in F$, alors par stabilité on a $u + v \in F$, donc $+$ est une loi de composition interne sur F . Les axiomes 1 et 2 sont vraies pour tous les éléments de F , puisqu'ils le sont pour tous les éléments de E et que $F \subset E$.

Montrons que 0_E appartient à F . Soit $u \in F$ (un tel u existe car F est non vide), par stabilité on a $0_E = u + (-1)u \in F$. De plus, le vecteur 0_E est aussi le vecteur nul de F : pour tout $u \in F$, on a $u + 0_E = 0_E + u = u$ car 0_E est le vecteur nul de E et $F \subset E$. Ainsi, l'axiome 3 est vérifié.

Enfin, si $u \in F$ alors $-u = 0_E + (-1)u \in F$ par stabilité, donc l'axiome 4 est vérifié.

• **Affirmation.** La restriction de la loi \cdot à F est une loi de composition externe sur F qui vérifie les axiomes 5 à 8 de la définition 3.2.4.

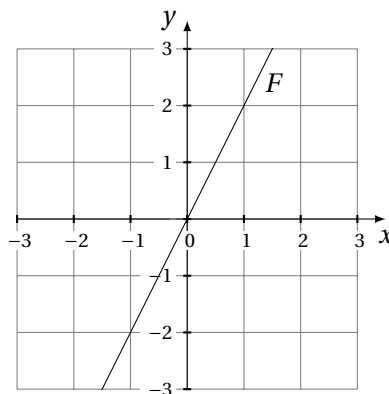
En effet, si $v \in F$ et $\lambda \in \mathbb{K}$ alors on a $\lambda v = 0_E + \lambda v \in F$ par stabilité. Les axiomes 5 à 8 étant vérifiés pour tous les éléments de E , ils le sont également pour tous les éléments de F puisque $F \subset E$.

Bilan. On a montré que la restriction des lois $+$ et \cdot à F le munissait d'une structure de \mathbb{K} -espace vectoriel. \square

Exemple 3.3.6. Dans \mathbb{R}^2 , on considère l'ensemble $F := \{(x, y) \in \mathbb{R}^2 \mid 2x - y = 0\}$. Montrons que F est un sous-espace vectoriel de \mathbb{R}^2 . Premièrement, il est non vide car le vecteur nul $(0, 0)$ appartient à F . Deuxièmement, montrons que F est stable par combinaisons linéaires. Soit $\lambda \in \mathbb{R}$, et soient $u := (x, y) \in F$ et $v := (x', y') \in F$, c'est-à-dire tels que $2x - y = 0$ et $2x' - y' = 0$. Montrons que $u + \lambda v \in F$. Puisque $u + \lambda v = (x + \lambda x', y + \lambda y')$, on calcule :

$$2(x + \lambda x') - (y + \lambda y') = 2x + 2\lambda x' - y - \lambda y' = \underbrace{(2x - y)}_{=0} + \lambda \underbrace{(2x' - y')}_{=0} = 0.$$

Ainsi, $u + \lambda v$ vérifie l'équation définissant F , donc $u + \lambda v \in F$. Par conséquent, F est un sous-espace vectoriel de \mathbb{R}^2 . Géométriquement, F est une droite passant par l'origine du repère. \blacktriangleleft



Remarque 3.3.7. Plus généralement, les droites et les plans passant par l'origine du repère sont des s.e.v. de l'espace \mathbb{R}^3 . En revanche, s'ils ne passent pas par l'origine, ce ne sont pas des s.e.v. puisqu'ils ne contiennent pas le vecteur nul $(0, 0, 0)$. \blacktriangleleft

Exemple 3.3.8. Soit $E = \mathcal{F}(I, \mathbb{R})$ l'espace vectoriel des fonctions définie sur un intervalle ouvert I , alors l'ensemble $F := \mathcal{C}^0(I, \mathbb{R})$ des fonctions continues sur I à valeurs réelles est un sous-espace vectoriel de E . En effet, il est non vide et il est stable par combinaisons linéaires car les sommes de fonctions continues sont continues, et la multiplication d'une fonction continue par une constante est une fonction continue.

On vérifie de même que $\mathcal{C}^k(I, \mathbb{R})$, l'ensemble des fonctions k fois dérivables de dérivée k -ième continue, est un s.e.v. de E . ◀

3.3.b Intersection de sous-espaces

Proposition 3.3.9. Pour tous sous-espaces vectoriels F et G de E , leur intersection $F \cap G$ est un sous-espace vectoriel. ▶

Démonstration. Il suffit de vérifier les deux conditions du théorème 3.3.4. Premièrement, on a $0_E \in F \cap G$ car $0_E \in F$ et $0_E \in G$, donc l'ensemble $F \cap G$ est non vide. Deuxièmement, montrons que $F \cap G$ est stable par combinaisons linéaires. Soient $u, v \in F \cap G$ et $\lambda \in \mathbb{K}$. Puisque F est un s.e.v. de E , il est stable par combinaisons linéaires, donc $u + \lambda v \in F$. Par le même raisonnement, on a aussi $u + \lambda v \in G$, donc $u + \lambda v \in F \cap G$. Par conséquent, $F \cap G$ est un s.e.v. de E . □

Remarque 3.3.10. Ce résultat est valable plus généralement pour des familles quelconques de sous-espaces vectoriels : si $(F_i)_{i \in I}$ est une famille de sous-espace vectoriels indexée par un ensemble non vide I , alors $\bigcap_{i \in I} F_i$ est un sous-espace vectoriel. ◀

Exemple 3.3.11. On considère un système linéaire homogène :

$$\begin{cases} a_{1,1} x_1 + a_{1,2} x_2 + \cdots + a_{1,n} x_n = 0 \\ a_{2,1} x_1 + a_{2,2} x_2 + \cdots + a_{2,n} x_n = 0 \\ \vdots \\ a_{p,1} x_1 + a_{p,2} x_2 + \cdots + a_{p,n} x_n = 0. \end{cases} \quad (S)$$

Soit F l'ensemble des solutions de (S). Alors F est l'intersection de F_1, \dots, F_p où :

$$F_i := \{(x_1, \dots, x_n) \in \mathbb{K}^n \mid a_{i,1} x_1 + \cdots + a_{i,n} x_n = 0\}.$$

On vérifie facilement³ que les F_i sont des s.e.v. de \mathbb{K}^n , donc F est un s.e.v. de \mathbb{K}^n . ◀

3.3.c Sous-espace engendré par une partie

En géométrie euclidienne, on a vu qu'une droite est dirigée par un vecteur non nul et qu'un plan est dirigé par deux vecteurs non colinéaires. Plus généralement, on peut s'intéresser aux sous-espaces engendrés par une famille de vecteurs.

Définition 3.3.12. Soient v_1, \dots, v_p des vecteurs. On dit qu'un vecteur u est une **combinaison linéaire** de v_1, \dots, v_p s'il existe des scalaires $\lambda_1, \dots, \lambda_p$ tels que :

$$u = \lambda_1 v_1 + \cdots + \lambda_p v_p.$$

On note $\text{Vect}(v_1, \dots, v_p)$ l'ensemble de toutes les combinaisons linéaires de v_1, \dots, v_p :

$$\text{Vect}(v_1, \dots, v_p) := \{\lambda_1 v_1 + \cdots + \lambda_p v_p : (\lambda_1, \dots, \lambda_p) \in \mathbb{K}^p\}.$$

Proposition 3.3.13. Pour tous vecteurs v_1, \dots, v_p , l'ensemble $\text{Vect}(v_1, \dots, v_p)$ est un sous-espace vectoriel de E appelé le **sous-espace vectoriel engendré** par v_1, \dots, v_p . ▶

3. on peut le démontrer avec le théorème 3.3.4, mais on verra une autre façon au chapitre 4.

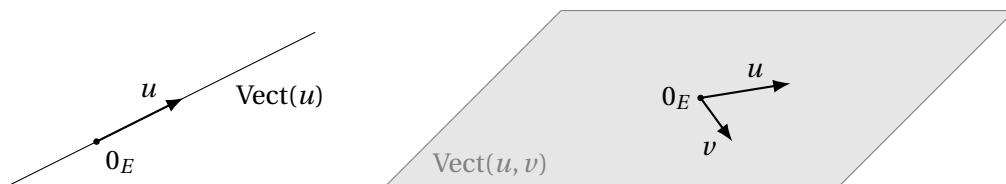


Figure 3.7 – Droite vectorielle et plan vectoriel.

Démonstration. Soient $v_1, \dots, v_p \in E$ et notons $F := \text{Vect}(v_1, \dots, v_p)$. On vérifie les deux conditions du théorème 3.3.4 :

1. F est non vide car il contient le vecteur nul : $0_E = 0v_1 + \dots + 0v_p \in F$.
2. F est stable par combinaisons linéaires : soient $u := \sum_{i=1}^p \lambda_i v_i \in F$ et $u' := \sum_{i=1}^p \lambda'_i v_i \in F$, et soit $\mu \in \mathbb{K}$, alors on a :

$$u + \mu u' = \sum_{i=1}^p \lambda_i v_i + \mu \sum_{i=1}^p \lambda'_i v_i = \sum_{i=1}^p (\lambda_i + \mu \lambda'_i) v_i \in F.$$

Par conséquent, F est un s.e.v. de E . □

Remarque 3.3.14. Plus généralement, si A est une partie de E , on note $\text{Vect}(A)$ l'ensemble de toutes les combinaisons linéaires **finies**⁴ de vecteurs de A . On vérifie que $\text{Vect}(A)$ est un sous-espace vectoriel de E . Il est appelé le sous-espace vectoriel engendré par A , et c'est le plus petit sous-espace vectoriel de E qui contient A : si F est un sous-espace vectoriel de E et si $A \subset F$, alors on a $\text{Vect}(A) \subset F$. ◀

Exemple 3.3.15.

1. Soit $u \in E$, alors le s.e.v. engendré par u est l'ensemble des vecteurs colinéaires à u :

$$\text{Vect}(u) := \{\lambda u : \lambda \in \mathbb{K}\}.$$

On le note également $\mathbb{K}u$. Si $u \neq 0_E$, le s.e.v. $\mathbb{K}u$ est appelé une **droite vectorielle** de E , voir figure 3.7.

2. Soient $u, v \in E$. Le sous-espace engendré par u et v est :

$$\text{Vect}(u, v) := \{\lambda u + \mu v : (\lambda, \mu) \in \mathbb{K}^2\}.$$

Si u et v ne sont pas colinéaires, alors $\text{Vect}(u, v)$ est appelé un **plan vectoriel** de E , voir figure 3.7. ◀

Proposition 3.3.16. Pour tous vecteurs v_1, \dots, v_p, v_{p+1} , si $v_{p+1} \in \text{Vect}(v_1, \dots, v_p)$ alors on a :

$$\text{Vect}(v_1, \dots, v_p, v_{p+1}) = \text{Vect}(v_1, \dots, v_p). \quad \blacktriangleleft$$

Démonstration. Soient v_1, \dots, v_p, v_{p+1} des vecteurs de E tels que $v_{p+1} \in \text{Vect}(v_1, \dots, v_p)$. On a toujours $\text{Vect}(v_1, \dots, v_p) \subset \text{Vect}(v_1, \dots, v_p, v_{p+1})$, montrons l'inclusion réciproque. Par hypothèse, on a l'inclusion $\{v_1, \dots, v_p, v_{p+1}\} \subset \text{Vect}(v_1, \dots, v_p)$, donc $\text{Vect}(v_1, \dots, v_p, v_{p+1}) \subset \text{Vect}(v_1, \dots, v_p)$ d'après la remarque précédente. Par conséquent, $\text{Vect}(v_1, \dots, v_p, v_{p+1}) = \text{Vect}(v_1, \dots, v_p)$. □

4. même si A est infinie, on ne considère que des sommes finies de vecteurs de A . En algèbre linéaire, on ne travaille qu'avec des sommes finies de vecteurs.

3.4 Familles de vecteurs

3.4.a Familles génératrices

Définition 3.4.1. On dit qu'une famille \mathcal{G} de vecteurs est une **famille génératrice** de E si tout vecteur de E peut s'écrire comme une combinaison linéaire de vecteurs de \mathcal{G} , c.-à-d. si $\text{Vect}(\mathcal{G}) = E$.

Exemple 3.4.2. Dans \mathbb{R}^2 , soient $v_1 := (1, 2)$, $v_2 := (0, 1)$ et $v_3 := (-1, 1)$. Alors $\mathcal{G} = (v_1, v_2, v_3)$ est une famille génératrice de \mathbb{R}^2 . En effet, pour tout $(x, y) \in \mathbb{R}^2$, on a :

$$\begin{aligned} (x, y) = \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 &\iff \begin{cases} \lambda_1 - \lambda_3 = x \\ 2\lambda_1 + \lambda_2 + \lambda_3 = y \end{cases} \\ &\iff \begin{cases} \lambda_1 - \lambda_3 = x \\ \lambda_2 + 3\lambda_3 = y - x. \end{cases} \quad L_2 \leftarrow L_2 - 2L_1 \end{aligned}$$

Ce système échelonné est compatible, donc tout vecteur de \mathbb{R}^2 s'écrit comme une combinaison linéaire de v_1, v_2, v_3 . Par conséquent, la famille (v_1, v_2, v_3) est génératrice de \mathbb{R}^2 .

Notons que le système admet une infinité de solutions, donc l'écriture d'un vecteur comme combinaison linéaire de v_1, v_2, v_3 n'est pas unique. ◀

Définition 3.4.3. Si E admet une famille génératrice finie, on dit que E est de dimension finie. Dans le cas contraire, on dit que E est de dimension infinie.

Exemple 3.4.4. L'espace \mathbb{K}^n est de dimension finie, car les vecteurs $e_1 := (1, 0, \dots, 0)$, $e_2 := (0, 1, 0, \dots, 0)$, ..., $e_n = (0, \dots, 0, 1)$ forment une famille génératrice finie. En revanche, l'espace $\mathbb{K}^{\mathbb{N}}$ des suites à valeurs dans \mathbb{K} ou l'espace $\mathcal{F}(I, \mathbb{K})$ des fonctions définies sur un intervalle I sont de dimension infinie (on le justifiera plus tard). ◀

3.4.b Familles libres et familles liées

Dans l'exemple 3.4.2, on voit qu'il y a pas forcément unicité de l'écriture d'un vecteur comme combinaison linéaire des éléments d'une famille génératrice. Dans cet exemple, la raison est que les vecteurs v_1, v_2, v_3 sont coplanaires : il est possible d'exprimer l'un en fonction des deux autres, par exemple $v_3 = 3v_2 - v_1$. Par conséquent, dans une écriture $\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3$, on peut remplacer v_3 par son expression en fonction de v_1 et v_2 pour obtenir :

$$\begin{aligned} \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 &= \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 (3v_2 - v_1) \\ &= (\lambda_1 - \lambda_3) v_1 + (\lambda_2 + 3\lambda_3) v_2 \\ &=: \lambda'_1 v_1 + \lambda'_2 v_2 + 0v_3. \end{aligned}$$

Le fait que les générateurs soient coplanaires empêchent l'unicité de la décomposition, car l'un des générateurs est redondant : v_1 et v_2 suffisent déjà pour engendrer \mathbb{R}^2 , l'ajout de v_3 donne un degré de liberté dans la façon de décomposer un vecteur de \mathbb{R}^2 .

On peut généraliser cette remarque à des familles génératrices \mathcal{G} quelconque : si l'un des générateurs peut s'exprimer à partir des autres, l'écriture d'un vecteur comme combinaison linéaire d'éléments de \mathcal{G} n'est pas unique. Lorsqu'un vecteur d'une famille s'écrit comme une combinaison linéaire des autres, on dit que la famille est liée. Cette notion généralise les notions de vecteurs colinéaires et de vecteurs coplanaires.

Définition 3.4.5. On dit qu'une famille (v_1, \dots, v_p) de vecteurs de E est une **liée** si l'un des vecteurs de la famille peut s'écrire comme une combinaison linéaire des autres :

$$\exists i \in \llbracket 1, p \rrbracket, v_i \in \text{Vect}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_p).$$

Cette définition n'est cependant pas pratique, car elle demande de vérifier pour chaque vecteur s'il est combinaison linéaire des autres. La proposition suivante fournit une caractérisation des familles liées plus utile en pratique.

Proposition 3.4.6. Une famille (v_1, \dots, v_p) est liée si et seulement si il existe une combinaison linéaire nulle de v_1, \dots, v_p dont les coefficients ne sont pas tous nuls :

$$(v_1, \dots, v_p) \text{ est liée} \iff \exists (\lambda_1, \dots, \lambda_p) \in \mathbb{K}^p \setminus \{(0, \dots, 0)\}, \quad \lambda_1 v_1 + \dots + \lambda_p v_p = 0_E.$$

Une telle combinaison est appelée une **relation de liaison** entre les vecteurs v_1, \dots, v_p . ▶

Remarque 3.4.7. Dire que les λ_i ne sont pas tous nuls ne signifie pas qu'ils sont tous non nuls, ça signifie qu'au moins l'un d'eux est non nul. ▶

Démonstration. Soient $v_1, \dots, v_p \in E$. On procède par double implication :

(\implies) Supposons que (v_1, \dots, v_p) est liée. Alors il existe $i \in \llbracket 1, p \rrbracket$ et $\lambda_1, \dots, \lambda_{i-1}, \lambda_{i+1}, \dots, \lambda_p \in \mathbb{K}$ tels que $v_i = \lambda_1 v_1 + \dots + \lambda_{i-1} v_{i-1} + \lambda_{i+1} v_{i+1} + \dots + \lambda_p v_p$. On a alors :

$$\lambda_1 v_1 + \dots + \lambda_{i-1} v_{i-1} + (-1) v_i + \lambda_{i+1} v_{i+1} + \dots + \lambda_p v_p = 0_E.$$

Ainsi, il existe une combinaison linéaire nulle de v_1, \dots, v_p dont les coefficients ne sont pas tous nuls.

(\impliedby) Réciproquement, supposons qu'il existe des scalaires $\lambda_1, \dots, \lambda_p \in \mathbb{K}$ non tous nuls tels que $\lambda_1 v_1 + \dots + \lambda_p v_p = 0_E$. Soit $i \in \llbracket 1, p \rrbracket$ tel que $\lambda_i \neq 0$. Alors on a :

$$v_i = \frac{-\lambda_1}{\lambda_i} v_1 + \dots + \frac{-\lambda_{i-1}}{\lambda_i} v_{i-1} + \frac{-\lambda_{i+1}}{\lambda_i} v_{i+1} + \dots + \frac{-\lambda_p}{\lambda_i} v_p.$$

Le vecteur v_i s'écrit comme une combinaison linéaire des autres. □

Remarque 3.4.8.

1. Toute famille contenant le vecteur nul est liée.
2. Une famille constituée d'un seul vecteur u est liée si et seulement si $u = 0_E$.
3. Une famille de deux vecteurs (u, v) est liée si et seulement si u et v sont colinéaires.
4. Une famille de trois vecteurs (u, v, w) est liée si et seulement si u, v et w sont coplanaires. ▶

On définit qu'une famille de vecteurs est libre si elle n'est pas liée. En prenant les négations dans la proposition 3.4.6, on obtient :

$$\begin{aligned} (v_1, \dots, v_p) \text{ est libre} &\iff \forall (\lambda_1, \dots, \lambda_p) \in \mathbb{K}^p \setminus \{(0, \dots, 0)\}, \quad \lambda_1 v_1 + \dots + \lambda_p v_p \neq 0_E \\ &\iff \forall \lambda_1, \dots, \lambda_p \in \mathbb{K}, \quad (\lambda_1, \dots, \lambda_p) \neq (0, \dots, 0) \implies \lambda_1 v_1 + \dots + \lambda_p v_p \neq 0_E. \end{aligned}$$

En prenant la contraposée de la dernière implication, on obtient la définition de famille libre qu'on utilise en pratique.

Définition 3.4.9. On dit qu'une famille (v_1, \dots, v_p) de vecteurs de E est une **famille libre** si elle n'est pas liée, c'est-à-dire si :

$$\forall \lambda_1, \dots, \lambda_p \in \mathbb{K}, \quad \lambda_1 v_1 + \dots + \lambda_p v_p = 0_E \implies \lambda_1 = \lambda_2 = \dots = \lambda_p = 0.$$

On dit aussi que les vecteurs v_1, \dots, v_p sont **linéairement indépendants**.

Pour montrer qu'une famille (v_1, \dots, v_p) est libre ou liée, on résout l'équation vectorielle :

$$\lambda_1 v_1 + \dots + \lambda_p v_p = 0_E,$$

d'inconnues $\lambda_1, \dots, \lambda_p \in \mathbb{K}$ (si $E = \mathbb{K}^n$, c'est un système linéaire de n équations à p inconnues) :

- si l'unique solution est $(\lambda_1, \dots, \lambda_p) = (0, \dots, 0)$, alors la famille est libre ;
- si l'équation admet des solutions $(\lambda_1, \dots, \lambda_p) \neq (0, \dots, 0)$, alors la famille est liée et on obtient au passage les relations de liaison entre les vecteurs de la famille.

Exemple 3.4.10. Dans \mathbb{R}^4 , soient $v_1 := (1, 1, 0, 2)$, $v_2 := (1, 2, -2, 1)$ et $v_3 := (2, 0, -1, 1)$. Cherchons si v_1 , v_2 et v_3 sont libres. Pour $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$, on résout :

$$\begin{aligned}
 \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = 0_{\mathbb{R}^4} &\iff \begin{cases} \lambda_1 + \lambda_2 + 2\lambda_3 = 0 \\ \lambda_1 + 2\lambda_2 = 0 \\ -2\lambda_2 - \lambda_3 = 0 \\ 2\lambda_1 + \lambda_2 + \lambda_3 = 0 \end{cases} \\
 &\iff \begin{cases} \lambda_1 + \lambda_2 + 2\lambda_3 = 0 \\ \lambda_2 - 2\lambda_3 = 0 & L_2 \leftarrow L_2 - L_1 \\ -2\lambda_2 - \lambda_3 = 0 \\ -\lambda_2 - 3\lambda_3 = 0 & L_4 \leftarrow L_4 - 2L_1 \end{cases} \\
 &\iff \begin{cases} \lambda_1 + \lambda_2 + 2\lambda_3 = 0 \\ \lambda_2 - 2\lambda_3 = 0 \\ -5\lambda_3 = 0 & L_3 \leftarrow L_3 + 2L_2 \\ -5\lambda_3 = 0 & L_4 \leftarrow L_4 + L_2 \end{cases} \\
 &\iff \lambda_1 = \lambda_2 = \lambda_3 = 0.
 \end{aligned}$$

Par conséquent, la famille (v_1, v_2, v_3) est libre. ◀

Exemple 3.4.11. Dans \mathbb{R}^3 , soient $v_1 := (1, 2, -1)$, $v_2 := (0, 1, -1)$ et $v_3 := (-1, 1, -2)$. Cherchons si v_1 , v_2 et v_3 sont libres. Pour $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$, on résout :

$$\begin{aligned}
 \lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = 0_{\mathbb{R}^3} &\iff \begin{cases} \lambda_1 - \lambda_3 = 0 \\ 2\lambda_1 + \lambda_2 + \lambda_3 = 0 \\ -\lambda_1 - \lambda_2 - 2\lambda_3 = 0 \end{cases} \\
 &\iff \begin{cases} \lambda_1 - \lambda_3 = 0 \\ \lambda_2 + 3\lambda_3 = 0 & L_2 \leftarrow L_2 - 2L_1 \\ -\lambda_2 - 3\lambda_3 = 0 & L_3 \leftarrow L_3 + L_1 \end{cases} \\
 &\iff \begin{cases} \lambda_1 - \lambda_3 = 0 \\ \lambda_2 + 3\lambda_3 = 0. \end{cases}
 \end{aligned}$$

Le système est homogène de rang 2 et possède 3 inconnues, donc il admet des solutions (une infinité). Par conséquent, la famille (v_1, v_2, v_3) est liée. Les solutions du système sont $\lambda_1 = \lambda_3$ et $\lambda_2 = -3\lambda_3$ avec $\lambda_3 \in \mathbb{R}$ une variable libre. En prenant $\lambda_3 = 1$ par exemple, on obtient la relation de liaison :

$$v_1 - 3v_2 + v_3 = 0_{\mathbb{R}^3}.$$

Dans ce cas, toutes les relations de liaison entre v_1, v_2, v_3 sont proportionnelles à celles-ci. ◀

Exemple 3.4.12. Dans $E := \mathcal{F}(\mathbb{R}, \mathbb{R})$, on considère la famille (\cos, \sin) . Montrons que c'est une famille libre. Soient $\lambda, \mu \in \mathbb{R}$ tels que $\lambda \cos + \mu \sin = 0_E$, c'est-à-dire :

$$\forall x \in \mathbb{R}, \quad \lambda \cos(x) + \mu \sin(x) = 0.$$

En prenant $x = 0$, on obtient $\lambda = 0$, et pour $x = \frac{\pi}{2}$, on a $\mu = 0$. Par conséquent, les fonctions \cos et \sin sont linéairement indépendantes. ◀

Proposition 3.4.13. Pour tous vecteurs v_1, \dots, v_p de E , la famille (v_1, \dots, v_p) est libre si et seulement si la décomposition de tout vecteur de $\text{Vect}(v_1, \dots, v_p)$ sur les vecteurs v_1, \dots, v_p est unique. ◀

Démonstration. Soient $v_1, \dots, v_p \in E$. On procède par double implication :

(\Rightarrow) Supposons que la famille (v_1, \dots, v_p) est libre. Soit $u \in \text{Vect}(v_1, \dots, v_p)$, et soient deux familles de scalaires $(\lambda_1, \dots, \lambda_p) \in \mathbb{K}^p$ et $(\mu_1, \dots, \mu_p) \in \mathbb{K}^p$ telles que :

$$u = \lambda_1 v_1 + \dots + \lambda_p v_p = \mu_1 v_1 + \dots + \mu_p v_p.$$

Alors $(\lambda_1 - \mu_1)v_1 + \dots + (\lambda_p - \mu_p)v_p = 0_E$. Puisque la famille est libre, les coefficients de cette combinaison sont tous nuls, donc on a $\lambda_i = \mu_i$ pour tout $i \in \llbracket 1, p \rrbracket$. Par conséquent, la décomposition de u sur les vecteurs v_1, \dots, v_p est unique.

(\Leftarrow) Supposons que la décomposition de tout vecteur de $\text{Vect}(v_1, \dots, v_p)$ sur les vecteurs v_1, \dots, v_p est unique. Soient $\lambda_1, \dots, \lambda_p \in \mathbb{K}$ tels que :

$$\lambda_1 v_1 + \dots + \lambda_p v_p = 0_E.$$

Le vecteur nul s'écrit aussi $0_E = 0v_1 + \dots + 0v_p$, donc par unicité de la décomposition, tous les λ_i doivent être nuls. Ainsi, on a montré que la seule combinaison linéaire nulle des v_1, \dots, v_p est celle dont tous les coefficients sont nuls. Par conséquent, la famille est libre. ◻

La proposition suivante donne une condition nécessaire et suffisante pour pouvoir ajouter un vecteur à une famille libre de façon à obtenir une famille encore libre.

Proposition 3.4.14. Soit (v_1, \dots, v_p) une famille libre et soit $u \in E$. Alors :

$$(v_1, \dots, v_p, u) \text{ est libre} \iff u \notin \text{Vect}(v_1, \dots, v_p). \quad \blacktriangleleft$$

Démonstration. On procède par double implication.

(\Rightarrow) Si (v_1, \dots, v_p, u) est libre, alors aucun vecteur de la famille n'est combinaison linéaire des autres, donc en particulier $u \notin \text{Vect}(v_1, \dots, v_p)$.

(\Leftarrow) Supposons que $u \notin \text{Vect}(v_1, \dots, v_p)$. Soient $\lambda_1, \dots, \lambda_p, \lambda \in \mathbb{K}$ tels que :

$$\lambda_1 v_1 + \dots + \lambda_p v_p + \lambda u = 0_E.$$

Alors on a nécessairement $\lambda = 0$ (sinon, u serait combinaison linéaire des v_1, \dots, v_p , contradiction). Ainsi, $\lambda_1 v_1 + \dots + \lambda_p v_p = 0_E$. Puisque la famille (v_1, \dots, v_p) est libre, alors tous les λ_i sont nuls. Par conséquent, la famille (v_1, \dots, v_p, u) est libre. ◻

3.4.c Base d'un espace vectoriel

Définition 3.4.15. Une famille de vecteurs de E qui est à la fois libre et génératrice est appelée une **base** de E .

Théorème 3.4.16. Une famille $\mathcal{B} := (v_1, \dots, v_n)$ est une base de E si et seulement si tout vecteur de E s'écrit de manière unique comme combinaison linéaire des v_i :

$$\forall u \in E, \quad \exists! (\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n, \quad u = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

Les nombres $\lambda_1, \dots, \lambda_n$ sont appelés les **coordonnées** de u dans la base \mathcal{B} .

Démonstration. L'existence de la décomposition est équivalente au fait que la famille est génératrice et l'unicité est équivalente au fait que la famille est libre. ◻

Remarque 3.4.17. Les coordonnées $\lambda_1, \dots, \lambda_n$ d'un vecteur u dans une base \mathcal{B} s'écrivent en général en colonne :

$$\begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \quad \text{ou} \quad \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}.$$

Cette notation sera utilisée en calcul matriciel, voir cours d'algèbre linéaire 2. ◀

Remarque 3.4.18. Les coordonnées d'un vecteur dépendent de la base choisie. Par exemple dans \mathbb{R}^2 , considérons les vecteurs $e_1 := (1, 0)$, $e_2 := (0, 1)$, $u := (1, 1)$ et $v := (2, 1)$. On admet que $\mathcal{B} := (e_1, e_2)$ et $\mathcal{C} := (u, e_2)$ sont des bases de \mathbb{R}^2 (le montrer!). Cherchons les coordonnées de v dans ces bases :

- dans la base \mathcal{B} , on a $v = 2e_1 + 1e_2$ donc les coordonnées de v sont $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$.
- dans la base \mathcal{C} , on a $v = 2u + (-1)e_2$ donc les coordonnées de v sont $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$. ◀

Exemple 3.4.19. Dans \mathbb{K}^n , notons e_i le vecteur dont la i -ème composante est 1 et les autres sont 0, c'est-à-dire :

$$e_1 := (1, 0, \dots, 0), \quad e_2 := (0, 1, 0, \dots, 0), \quad \dots, \quad e_n := (0, \dots, 0, 1).$$

La famille (e_1, e_2, \dots, e_n) est une base de \mathbb{K}^n appelée la **base canonique** de \mathbb{K}^n . En effet, tout vecteur $x := (x_1, \dots, x_n) \in \mathbb{K}^n$ s'écrit de manière unique comme :

$$x = x_1 e_1 + \dots + x_n e_n.$$

Les coordonnées d'un vecteur dans cette base sont ses composantes. ◀

Exemple 3.4.20. La famille $(1, i)$ est une base de \mathbb{C} en tant qu'espace vectoriel sur \mathbb{R} . En effet, tout nombre complexe z s'écrit de manière unique comme $z = a + ib$ avec $a, b \in \mathbb{R}$. Les coordonnées d'un nombre complexe dans cette base sont sa partie réelle et sa partie imaginaire. ◀

Exemple 3.4.21. Dans \mathbb{R}^4 , soit $F := \{(x, y, z, t) \in \mathbb{R}^4 \mid 2x - y + z + 3t = 0\}$. Soit $u := (x, y, z, t) \in \mathbb{R}^4$, alors on a :

$$\begin{aligned} u \in F &\iff 2x - y + z + 3t = 0 \\ &\iff y = 2x + z + 3t \\ &\iff u = (x, 2x + z + 3t, z, t) \\ &\iff u = (x, 2x, 0, 0) + (0, z, z, 0) + (0, 3t, 0, t) \\ &\iff u = x(1, 2, 0, 0) + z(0, 1, 1, 0) + t(0, 3, 0, 1). \end{aligned}$$

Par conséquent, les vecteurs $v_1 := (1, 2, 0, 0)$, $v_2 := (0, 1, 1, 0)$ et $v_3 := (0, 3, 0, 1)$ forment une famille génératrice de F . Vérifions qu'ils sont libres. Pour tous $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$, on a :

$$\lambda v_1 + \lambda_2 v_2 + \lambda_3 v_3 = 0_{\mathbb{R}^4} \iff \begin{cases} \lambda_1 = 0 \\ 2\lambda_1 + \lambda_2 + 3\lambda_3 = 0 \\ \lambda_2 = 0 \\ \lambda_3 = 0. \end{cases} \iff \lambda_1 = \lambda_2 = \lambda_3 = 0.$$

Ainsi, la famille (v_1, v_2, v_3) est libre et génératrice, donc c'est une base de F . ◀

3.5 Théorie de la dimension

Dans cette section, on souhaite définir ce qu'est la dimension d'un espace vectoriel. Prenons des exemples géométriques qu'on connaît bien : les droites et les plans. Intuitivement, une droite est de dimension 1 et un plan est de dimension 2. Une raison est la suivante : une base d'une droite sera formée d'un vecteur tandis qu'une base d'un plan contiendra deux vecteurs. L'idée est de définir plus généralement la dimension d'un espace vectoriel E comme le nombre de vecteurs d'une base de E . Pour que cette définition ait un sens, il faut vérifier deux choses :

1. qu'un espace vectoriel E de dimension finie possède toujours des bases ;
2. que toutes les bases de E aient le même nombre d'éléments.

3.5.a Existence de bases

Dans cette sous-section, on va voir qu'un espace vectoriel de dimension finie admet toujours une base. De plus, en partant d'une famille libre et d'une famille génératrice, il est toujours possible de construire une base en complétant la famille libre avec des vecteurs de la famille génératrice. La démonstration de ce résultat repose sur le lemme suivant qui permet de majorer la taille d'une famille libre par la taille d'une famille génératrice.

Lemme 3.5.1. Si E possède une famille génératrice de n vecteurs, alors toute famille comportant strictement plus de n vecteurs est liée. ◀

Remarque 3.5.2. Par contraposée, si E possède une famille génératrice de n vecteurs, alors toute famille libre possède au plus n vecteurs. Une conséquence de ce lemme est que pour montrer qu'un espace vectoriel est de dimension infinie, il suffit de trouver des familles libres de tailles arbitraires. C'est ainsi qu'on montre que les espaces $\mathbb{K}^{\mathbb{N}}$ et $\mathcal{F}(I, \mathbb{K})$ sont de dimension infinie. ◀

Démonstration. On procède par récurrence sur n . Soit H_n la proposition : « si un espace vectoriel est engendré par n vecteurs, alors toute famille comportant strictement plus de n vecteurs est liée ».

- Initialisation $n = 1$: si E est engendré par un vecteur u , alors tous les vecteurs de E sont colinéaires à u . Par conséquent, toute famille comportant deux vecteurs ou plus est liée.
- Hérédité : soit $n \geq 2$ tel que H_{n-1} est vraie ; démontrons H_n . Soit (g_1, \dots, g_n) une famille génératrice de n vecteurs et soit (v_1, \dots, v_p) une famille de p vecteurs avec $p > n$. On écrit chaque vecteur v_i comme une combinaison linéaire des vecteurs g_j :

$$\begin{cases} v_1 = \lambda_{1,1} g_1 + \dots + \lambda_{1,n} g_n \\ v_2 = \lambda_{2,1} g_1 + \dots + \lambda_{2,n} g_n \\ \vdots \\ v_p = \lambda_{p,1} g_1 + \dots + \lambda_{p,n} g_n. \end{cases}$$

Si $v_1 = 0_E$, alors la famille est liée. Sinon, l'un des vecteurs $\lambda_{1,j} g_j$ est non nul. Quitte à renuméroter les vecteurs, supposons que $\lambda_{1,1} g_1 \neq 0_E$. On utilise le vecteur v_1 pour éliminer g_1 dans la décomposition des autres vecteurs : pour chaque $i \in \llbracket 2, p \rrbracket$, on pose $w_i := v_i - \frac{\lambda_{i,1}}{\lambda_{1,1}} v_1$. Ainsi, w_2, \dots, w_p sont des vecteurs de $F := \text{Vect}(g_2, \dots, g_n)$.

Le s.e.v. F est engendré par $n-1$ vecteurs. Puisque $p-1 > n-1$, la famille (w_2, \dots, w_p) de F possède strictement plus de $n-1$ vecteurs, donc elle est liée par hypothèse de récurrence. Par conséquent, il existe des scalaires μ_2, \dots, μ_p non tous nuls tels que :

$$\mu_2 w_2 + \dots + \mu_p w_p = 0_E,$$

c'est-à-dire :

$$\mu_2 \left(v_2 - \frac{\lambda_{2,1}}{\lambda_{1,1}} v_1 \right) + \dots + \mu_p \left(v_p - \frac{\lambda_{p,1}}{\lambda_{1,1}} v_1 \right) = 0_E.$$

En développant, on obtient une combinaison linéaire nulle de v_1, \dots, v_p dont l'un des coefficients est non nul. Par conséquent, la famille (v_1, \dots, v_p) est liée.

Par le principe de récurrence, H_n est vraie pour tout $n \geq 1$. □

Théorème 3.5.3 (de la base incomplète). Si E est non nul et de dimension finie, alors pour toute famille génératrice \mathcal{G} de E et pour toute famille libre \mathcal{L} de E , on peut compléter \mathcal{L} avec des vecteurs de \mathcal{G} de manière à former une base de E .

Démonstration. Supposons que $E \neq \{0_E\}$. Soient \mathcal{L} une famille libre de E et soit \mathcal{G} une famille génératrice. On ajoute séquentiellement des vecteurs de \mathcal{G} à la famille \mathcal{L} jusqu'à former une base, en utilisant la procédure suivante.

- On initialise $\mathcal{L}_0 := \mathcal{L}$.
- Pour chaque $i \geq 0$:
 - si \mathcal{L}_i est génératrice, alors c'est une base de E . On arrête la procédure et le théorème est démontré.
 - sinon, il existe un vecteur $u \in \mathcal{G}$ qui n'appartient pas à au s.e.v. engendré par les vecteurs de \mathcal{L}_i (sinon \mathcal{L}_i serait génératrice, ce qui est exclu). Notons \mathcal{L}_{i+1} la famille \mathcal{L}_i augmentée du vecteur u . Cette famille est encore libre car \mathcal{L}_i est libre et u n'est pas combinaison linéaire des vecteurs de \mathcal{L}_i (proposition 3.4.14).

On construit ainsi une suite de familles libres dont la longueur croît strictement et qui s'arrête dès qu'on a obtenu une base. Il reste à montrer que cette suite s'arrête nécessairement. Puisque E possède une famille génératrice finie, la taille d'une famille libre est majorée d'après le lemme 3.5.1, donc la suite de familles libres ne peut pas croître indéfiniment. □

Corollaire 3.5.4. Si E est non nul et de dimension finie, alors :

- (i) toute famille libre de E peut être complétée en une base de E .
- (ii) on peut extraire une base de toute famille génératrice de E . ◀

Remarque 3.5.5.

1. Tout espace vectoriel non nul de dimension fini admet une base : il suffit de considérer un vecteur non nul (qui existe par hypothèse), celui-ci forme une famille libre qu'on complète en une base grâce à (i).
2. L'énoncé (ii) est aussi appelé **théorème de la base extraite**. ◀

Démonstration. Supposons que $E \neq \{0_E\}$.

(i) Soit \mathcal{L} une famille libre. On applique le théorème 3.5.3 avec n'importe quelle famille génératrice pour compléter \mathcal{L} en une base de E .

(ii) Soit \mathcal{G} une famille génératrice de E et soit $u \neq 0_E$ un vecteur de \mathcal{G} . Alors u forme une famille libre de E . On applique le théorème 3.5.3 pour compléter cette famille en une base de E avec des vecteurs de \mathcal{G} . On obtient ainsi une base de E formée de vecteurs de \mathcal{G} . □

3.5.b Dimension d'un espace vectoriel

Théorème 3.5.6 (théorème de la dimension). Si E est de dimension finie, alors toutes les bases de E ont le même nombre de vecteurs. Ce nombre est appelé la **dimension** de E sur \mathbb{K} et on le note $\dim_{\mathbb{K}}(E)$, ou simplement $\dim(E)$ s'il n'y a pas d'ambiguïté. Par convention, l'espace nul $\{0_E\}$ est de dimension 0.

Démonstration. Soient \mathcal{B} et \mathcal{B}' des bases de E de longueurs respectives n et n' . La base \mathcal{B}' est une famille génératrice de n' vecteurs, donc d'après le lemme 3.5.1 le nombre de vecteurs de \mathcal{B} ne peut pas dépasser n' , sinon \mathcal{B} serait une famille liée. Par conséquent, on a $n \leq n'$. En inversant les rôles de \mathcal{B} et \mathcal{B}' , on obtient l'inégalité contraire $n' \leq n$, donc $n = n'$. □

Définition 3.5.7. Supposons que E est de dimension finie n . Soit F un sous-espace vectoriel de E .

- Si $\dim(F) = 1$, on dit que F est une **droite vectorielle** de E .
- Si $\dim(F) = 2$, on dit que F est un **plan vectoriel** de E .
- Si $\dim(F) = n - 1$, on dit que F est un **hyperplan vectoriel** de E .

Exemple 3.5.8. Dans \mathbb{R}^4 , soit $F := \{(x, y, z, t) \in \mathbb{R}^4 \mid 2x - y + z + 3t = 0\}$. On a vu dans l'exemple 3.4.21 que les vecteurs $(1, 2, 0, 0)$, $(0, 1, 1, 0)$ et $(0, 3, 0, 1)$ forment une base de F , donc $\dim(F) = 3$. Le sous-espace F est un hyperplan vectoriel de \mathbb{R}^4 . ◀

Remarque 3.5.9. La dimension dépend du corps \mathbb{K} . Par exemple, l'ensemble \mathbb{C} considéré comme un \mathbb{R} -espace vectoriel admet pour base $(1, i)$, donc $\dim_{\mathbb{R}}(\mathbb{C}) = 2$. Mais si on considère \mathbb{C} comme un \mathbb{C} -espace vectoriel, alors $\dim_{\mathbb{C}}(\mathbb{C}) = 1$ (puisque $\dim_{\mathbb{C}}(\mathbb{C}^n) = n$). ◀

Corollaire 3.5.10. Supposons que E est de dimension finie n .

- Toute famille libre a au plus n éléments. Si une famille libre possède n éléments, alors c'est une base de E .
- Toute famille génératrice a au moins n éléments. Si une famille génératrice possède n éléments, c'est une base de E .
- Une famille de n vecteurs est une base si et seulement si elle est libre, si et seulement si elle est génératrice. ▶

Ce corollaire est très utile pour montrer qu'une famille de vecteurs est une base lorsqu'on connaît déjà la dimension d'un espace vectoriel : il suffit de vérifier que la famille est libre et qu'elle contient autant de vecteurs que la dimension.

Démonstration.

(i) La base \mathcal{B} est une famille génératrice de n vecteurs, donc toute famille libre possède au plus n éléments d'après le lemme 3.5.1. Si une famille libre possède exactement n éléments, alors on peut la compléter en une base de E . Or, une base possède n éléments par définition de la dimension, donc cette base est la famille libre elle-même.

(ii) S'il existait une famille génératrice de strictement moins de n éléments, alors on pourrait en extraire une base comportant strictement moins de n éléments; absurde! Par conséquent, toute famille génératrice de E possède au moins n éléments. Si elle possède exactement n éléments, on peut en extraire une base. Or, une base possède n éléments par définition de la dimension, donc cette base est la famille génératrice elle-même.

(iii) C'est une conséquence de (i) et (ii). ◻

Exemple 3.5.11. Montrons que les vecteurs $v_1 := (1, 2)$ et $v_2 := (-1, 1)$ forment une base de \mathbb{R}^2 . On sait que $\dim(\mathbb{R}^2) = 2$, donc il suffit de montrer que la famille $\mathcal{B} := (v_1, v_2)$ est libre. Pour tous $\lambda, \mu \in \mathbb{R}$, on a :

$$\lambda v_1 + \mu v_2 = 0_{\mathbb{R}^2} \iff \begin{cases} \lambda - \mu = 0 \\ 2\lambda + \mu = 0 \end{cases} \iff \begin{cases} \lambda - \mu = 0 \\ 3\mu = 0 \end{cases} \xrightarrow{L_2 \leftarrow L_2 - 2L_1} \iff \lambda = \mu = 0.$$

Par conséquent, la famille \mathcal{B} est une base de \mathbb{R}^2 . Cherchons les coordonnées du vecteur $u := (3, 3)$ dans cette base :

$$\begin{aligned} u = \lambda v_1 + \mu v_2 &\iff \begin{cases} \lambda - \mu = 3 \\ 2\lambda + \mu = 3 \end{cases} \\ &\iff \begin{cases} \lambda - \mu = 3 \\ 3\mu = -3 \end{cases} \xrightarrow{L_2 \leftarrow L_2 - 2L_1} \\ &\iff \begin{cases} \lambda = 2 \\ \mu = -1. \end{cases} \end{aligned}$$

Ainsi, $u = 2v_1 - v_2$, donc les coordonnées de u dans la base \mathcal{B} sont $\begin{pmatrix} 2 \\ -1 \end{pmatrix}$ (cf. figure 3.8). ◀

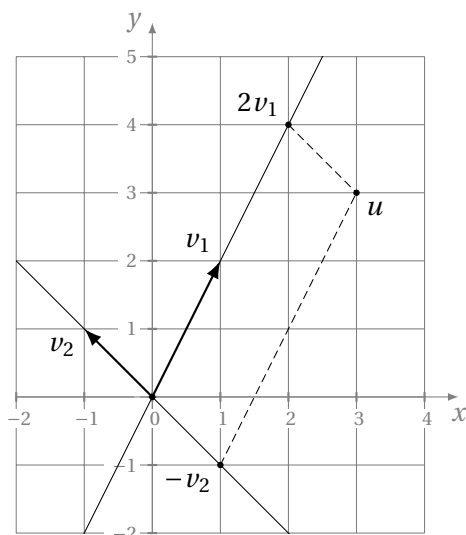


Figure 3.8 – Vecteurs de l'exemple 3.5.11.

Proposition 3.5.12. Supposons que E est de dimension finie. Soit F un sous-espace vectoriel de E , alors :

- (i) F est aussi de dimension finie et $\dim(F) \leq \dim(E)$.
- (ii) $\dim(E) = \dim(F) \iff E = F$. ◀

Pour montrer que deux sous-espaces F et G vectoriels sont égaux, on peut employer la stratégie suivante : on montre d'abord que $F \subset G$, puis on montre que $\dim(F) = \dim(G)$, et la proposition précédente permet de conclure que $F = G$.

Démonstration. Soit F un s.e.v. de E .

(i) Si $F = \{0_E\}$, il n'y a rien à montrer. Sinon, soit $A \subset \mathbb{N}^*$ l'ensemble des tailles des familles libres de F . Alors A est une partie non vide de \mathbb{N}^* car $F \neq \{0_E\}$, et majorée par $\dim(E)$. Par conséquent, A admet un élément maximal $n \leq \dim(E)$.

Soit (v_1, \dots, v_n) une famille libre maximale de F . Montrons qu'elle est aussi génératrice de F . Soit $u \in F$, alors la famille (v_1, \dots, v_n, u) est liée car c'est une famille de vecteurs de F de plus de n éléments. Par conséquent, il existe des scalaires $\lambda_1, \dots, \lambda_n, \lambda \in \mathbb{K}$ non tous nuls tels que :

$$\sum_{i=1}^n \lambda_i v_i + \lambda u = 0_E.$$

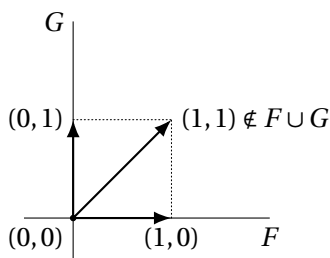
On a nécessairement $\lambda \neq 0$. En effet, si λ était nul, alors on aurait une combinaison linéaire nulle de v_1, \dots, v_n . Cette famille étant libre, tous les λ_i seraient nuls, et la famille (v_1, \dots, v_n, u) serait libre ; contradiction. Puisque $\lambda \neq 0$, on peut écrire $u = \sum_{i=1}^n \frac{-\lambda_i}{\lambda} v_i$. Par conséquent, tout vecteur de F s'écrit comme une combinaison linéaire de v_1, \dots, v_n , c'est donc une famille génératrice de F . Ainsi, F est de dimension finie et $\dim(F) = n \leq \dim(E)$.

(ii) Seule l'implication « \implies » est à démontrer. Supposons que $\dim(F) = \dim(E)$. Si $\dim(E) = 0$, alors $E = F = \{0_E\}$. Sinon, considérons \mathcal{B} une base de F . Alors \mathcal{B} est une famille libre de E de cardinal $\dim(E)$, donc c'est aussi une base de E d'après le corollaire 3.5.10. Par conséquent, \mathcal{B} est une famille génératrice de E , d'où $F = \text{Vect}(\mathcal{B}) = E$. ◻

3.6 Sommes de sous-espaces, sommes directes et supplémentaires

3.6.a Sommes de sous-espaces

Si F et G sont des s.e.v. de E , on a vu que leur intersection est un s.e.v. de E . En revanche, leur réunion $F \cup G$ n'est pas un s.e.v. en général. Par exemple, si $F := \{(x, 0) : x \in \mathbb{R}\}$ et $G := \{(0, y) : y \in \mathbb{R}\}$, alors F et G sont des s.e.v. de \mathbb{R}^2 (ce sont des droites vectorielles) mais leur réunion n'est pas stable par addition : on a $(1, 0) \in F \cup G$ et $(0, 1) \in F \cup G$ mais $(1, 0) + (0, 1) = (1, 1) \notin F \cup G$.



À la place, on peut chercher le plus petit s.e.v. qui contienne $F \cup G$, c'est-à-dire le s.e.v. engendré par $F \cup G$. Un tel s.e.v. doit contenir F et G , ainsi que toutes les sommes de vecteurs de F et de G .

Définition 3.6.1. Soit F et G des sous-espaces vectoriels de E . L'ensemble des sommes $u + v$ où $u \in F$ et $v \in G$ est appelé la **somme** de F et G , et on le note $F + G$:

$$F + G := \{u + v : u \in F \text{ et } v \in G\}.$$

Proposition 3.6.2. Pour tous sous-espaces vectoriels F et G , on a :

- (i) $F + G$ est un sous-espace vectoriel de E .
- (ii) $F + G = \text{Vect}(F \cup G)$. ▶

Démonstration. Le point (i) se démontre en vérifiant les deux conditions du théorème 3.3.4 (exercice!). Pour démontrer (ii), on procède par double inclusion :

(\subset) Soit $u \in F + G$, alors il existe $v \in F$ et $w \in G$ tels que $u = v + w$. Or, $F \subset \text{Vect}(F \cup G)$ et $G \subset \text{Vect}(F \cup G)$, donc v et w appartiennent à $\text{Vect}(F \cup G)$. Puisque $\text{Vect}(F \cup G)$ est un s.e.v., il est stable pour l'addition vectoriel, donc $v + w \in \text{Vect}(F \cup G)$. Ainsi, on a montré que $F + G \subset \text{Vect}(F \cup G)$.

(\supset) On a $F \subset F + G$ (en effet, si $u \in F$, alors $u = u + 0_E \in F + G$) et $G \subset F + G$, donc $F \cup G \subset F + G$. Ainsi, $F + G$ est un s.e.v. qui contient $F \cup G$, donc par définition il contient le s.e.v. engendré par $F \cup G$. Par conséquent, $\text{Vect}(F \cup G) \subset F + G$.

On conclut que $F + G = \text{Vect}(F \cup G)$. □

Proposition 3.6.3. Soient F et G des s.e.v. de E . Si \mathcal{G}_1 est une famille génératrice de F et \mathcal{G}_2 est une famille génératrice de G , alors $\mathcal{G}_1 \cup \mathcal{G}_2$ est une famille génératrice de $F + G$. ▶

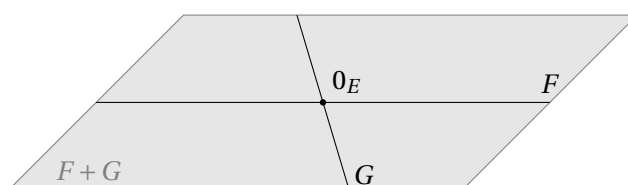


Figure 3.9 – Somme de sous-espaces vectoriels.

Démonstration. Soit $u \in F + G$. Alors il existe $v \in F$ et $w \in G$ tel que $u = v + w$. Puisque F est engendré par la famille \mathcal{G}_1 , il existe v_1, \dots, v_p des vecteurs de \mathcal{G}_1 tels que $v = v_1 + \dots + v_p$. De même, il existe w_1, \dots, w_q des vecteurs de \mathcal{G}_2 tels que $w = w_1 + \dots + w_q$. Donc u s'écrit :

$$u = v + w = v_1 + \dots + v_p + w_1 + \dots + w_q.$$

Ainsi, on a montré que tout vecteur de $F + G$ est une somme de vecteurs de $\mathcal{G}_1 \cup \mathcal{G}_2$. \square

3.6.b Somme directe de sous-espaces vectoriels

Définition 3.6.4. Soient F et G des sous-espaces vectoriels de E . On dit que F et G sont en **somme directe** si tout vecteur de $F + G$ s'écrit de manière unique comme la somme d'un vecteur de F et d'un vecteur de G . Lorsque F et G sont en somme directe, on note $F \oplus G$ leur somme.

Remarque 3.6.5. Les sous-espaces vectoriels $F \oplus G$ et $F + G$ sont les mêmes, le symbole \oplus donne simplement l'information supplémentaire que la somme est directe. \blacktriangleleft

Plus généralement, des sous-espaces vectoriels F_1, \dots, F_p sont en somme directe s'il y a unicité de l'écriture d'un vecteur de $F_1 + \dots + F_p$ comme somme de vecteurs des F_i . On peut voir ça comme une généralisation de la décomposition d'un vecteur sur une base. En effet, si (e_1, \dots, e_n) est une base de E , alors tout vecteur de E s'écrit manière unique comme une combinaison linéaire des e_i , c'est-à-dire comme une somme de vecteurs des droites engendrées par les e_i . Par conséquent, on a la décomposition $E = \mathbb{K}e_1 \oplus \dots \oplus \mathbb{K}e_n$. Une base de E correspond donc à une décomposition de E comme une somme directe de sous-espaces de dimensions 1. Les sommes directes permettent des décompositions plus générales, en utilisant des sous-espaces de dimensions plus grandes (voire infinies).

Proposition 3.6.6. Pour tous sous-espaces vectoriels F et G , leur somme est directe si et seulement si $F \cap G = \{0_E\}$. \blacktriangleleft

Démonstration. Soient F et G des s.e.v. de E . On procède par double implication :

(\implies) Supposons que F et G sont en somme directe. Soit $u \in F \cap G$, alors u s'écrit de deux façons comme somme d'un vecteur de F et d'un vecteur de G : $u = u + 0_E$ et $u = 0_E + u$. Par unicité de la décomposition, on a $u = 0_E$. Par conséquent, $F \cap G = \{0_E\}$.

(\impliedby) Supposons que $F \cap G = \{0_E\}$. Soit $u \in F + G$ et soient $v, v' \in F$ et $w, w' \in G$ tels que $u = v + w$ et $u = v' + w'$. Alors on a $v + w = v' + w'$, d'où

$$\underbrace{v - v'}_{\in F} = \underbrace{w' - w}_{\in G}.$$

Par conséquent, $v - v'$ et $w' - w$ appartiennent à $F \cap G = \{0_E\}$, donc $v - v' = 0_E$ et $w' - w = 0_E$, c'est-à-dire $v = v'$ et $w = w'$. L'écriture de u comme somme d'un vecteur de F et d'un vecteur de G est unique. \square

Remarque 3.6.7. Pour les sommes directes $F_1 \oplus \dots \oplus F_p$ de plusieurs sous-espaces vectoriels, la caractérisation de la proposition 3.6.6 ne se généralise pas telle quelle : les s.e.v. F_1, \dots, F_p sont en somme directe implique que $F_1 \cap \dots \cap F_p = \{0_E\}$, mais la réciproque est fausse si $p \geq 3$. \blacktriangleleft

Proposition 3.6.8. Pour tous sous-espaces vectoriels F et G de dimension finie, si F et G sont en somme directe, alors la concaténation d'une base de F et d'une base de G est une base de $F \oplus G$. \blacktriangleleft

Démonstration. Si F ou G est nul, il n'y a rien à montrer. Supposons que F et G sont non nuls, et considérons $\mathcal{B}_F := (f_1, \dots, f_p)$ une base de F et $\mathcal{B}_G := (g_1, \dots, g_q)$ une base de G . Montrons que la famille $\mathcal{B} := (f_1, \dots, f_p, g_1, \dots, g_q)$ est une base de $F \oplus G$. Si $u \in F \oplus G$, alors u s'écrit de manière unique sous la forme $v + w$ avec $v \in F$ et $w \in G$. De plus, v s'écrit de manière unique comme une combinaison linéaire de vecteurs de \mathcal{B}_F , et w s'écrit de manière unique comme combinaison linéaire de vecteurs de \mathcal{B}_G . Par conséquent, u s'écrit de manière unique comme combinaison linéaire de vecteurs de \mathcal{B} , ce qui prouve que \mathcal{B} est une base de $F \oplus G$. \square

On en déduit que la dimension d'une somme directe de sous-espaces vectoriels est la somme des dimensions des sous-espaces.

Corollaire 3.6.9. Pour tous sous-espaces vectoriels F et G de dimension finie, si F et G sont en somme directe alors on a :

$$\dim(F \oplus G) = \dim(F) + \dim(G).$$

Démonstration. Si F ou G est nul, la formule est triviale. Sinon, en concaténant une base de F et une base de G , on obtient une base de $F \oplus G$, ce qui montre en comptant les vecteurs de chaque base que $\dim(F \oplus G) = \dim(F) + \dim(G)$. \square

Théorème 3.6.10 (formule de Grassmann). Pour tous sous-espaces vectoriels F et G de dimensions finies, leur somme $F + G$ est de dimension finie et on a la formule :

$$\dim(F + G) = \dim(F) + \dim(G) - \dim(F \cap G).$$

Démonstration. Soient F et G des s.e.v. de E . On considère plusieurs cas.

- Si $F \subset G$, alors $F + G = G$ et $F \cap G = F$. Dans ce cas, la formule est triviale (de même si $G \subset F$).
- Si $F \cap G = \{0_E\}$, alors F et G sont en somme directe et la formule de Grassmann se réduit à l'égalité $\dim(F \oplus G) = \dim(F) + \dim(G)$, ce qu'on a déjà montré.
- Supposons que $F \not\subset G$, $G \not\subset F$ et $F \cap G \neq \{0_E\}$. Ces hypothèses impliquent que F et G ne sont pas nuls et que $F \cap G$ est un s.e.v. strict de F et de G .

Soit $\mathcal{B}_{F \cap G} := (e_1, \dots, e_r)$ une base $F \cap G$. On complète $\mathcal{B}_{F \cap G}$ en une base $\mathcal{B}_F := (e_1, \dots, e_r, f_1, \dots, f_p)$ de F . Ensuite, on complète $\mathcal{B}_{F \cap G}$ en une base $\mathcal{B}_G := (e_1, \dots, e_r, g_1, \dots, g_q)$ de G .

Montrons que $\mathcal{B} := (e_1, \dots, e_r, f_1, \dots, f_p, g_1, \dots, g_q)$ est une base de $F + G$. La famille \mathcal{B} est génératrice de $F + G$ car elle contient une famille génératrice de F et une famille génératrice de G . Montrons qu'elle est libre. On considère une combinaison linéaire nulle des vecteurs de \mathcal{B} :

$$\underbrace{\sum_{i=1}^r \lambda_i e_i}_{=:u} + \underbrace{\sum_{i=1}^p \mu_i f_i}_{=:v} + \underbrace{\sum_{i=1}^q \nu_i g_i}_{=:w} = 0_E.$$

On a $u + v = -w \in G$, mais aussi $u + v \in F$ car u et v sont des vecteurs de F . Ainsi, $u + v \in F \cap G$, donc $u + v$ est une combinaison linéaire de e_1, \dots, e_r , ce qui implique que les μ_i sont tous nuls. Par conséquent, on a $u + w = 0_E$. Par liberté de $(e_1, \dots, e_r, g_1, \dots, g_q)$, les coefficients λ_i et ν_i sont tous nuls, ce qui finit de prouver que \mathcal{B} est une famille libre. Ainsi, on a montré que \mathcal{B} est une base de $F + G$. En comptant le nombre de vecteurs des bases \mathcal{B} , \mathcal{B}_F , \mathcal{B}_G et $\mathcal{B}_{F \cap G}$, on obtient la formule de Grassmann :

$$\dim(F + G) = p + q + r = (p + r) + (q + r) - r = \dim(F) + \dim(G) - \dim(F \cap G).$$

Corollaire 3.6.11. Pour tous sous-espaces vectoriels F et G de dimensions finies, on a :

$$\dim(F + G) \leq \dim(F) + \dim(G),$$

avec égalité si et seulement si F et G sont en somme directe.

3.6.c Supplémentaires d'un sous-espace vectoriel

Définition 3.6.12. On dit que deux sous-espaces vectoriels F et G sont **supplémentaires** dans E s'ils sont en somme directe et si leur somme est égale à E , c'est-à-dire si $E = F \oplus G$.

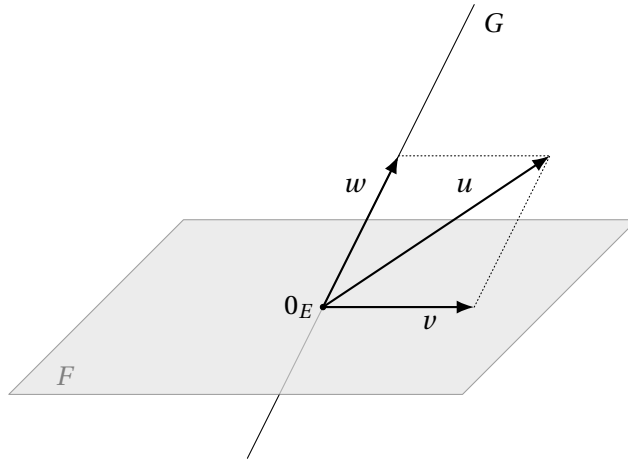


Figure 3.10 – Un plan et une droite supplémentaires dans \mathbb{R}^3 . Tout vecteur u de \mathbb{R}^3 se décompose de manière unique comme la somme d'un vecteur du plan F et d'un vecteur de la droite G .

Des s.e.v. F et G sont supplémentaires dans E si et seulement si tout vecteur de E s'écrit de manière unique comme la somme d'un vecteur de F et d'un vecteur de G :

$$E = F \oplus G \iff \forall u \in E, \exists ! (v, w) \in F \times G, u = v + w.$$

En effet, l'existence d'une décomposition de tout vecteur de E comme somme d'un vecteur de F et d'un vecteur de G est équivalente à dire que $E = F + G$, et l'unicité d'une telle décomposition signifie que la somme est directe.

Exemple 3.6.13. Dans l'espace vectoriel $\mathcal{F}(\mathbb{R}, \mathbb{R})$, l'ensemble \mathcal{P} des fonctions paires et l'ensemble \mathcal{I} des fonctions impaires sont supplémentaires. En effet, soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une application. Posons g et h les applications définies sur \mathbb{R} par :

$$g(x) := \frac{f(x) + f(-x)}{2}, \quad h(x) := \frac{f(x) - f(-x)}{2}.$$

On vérifie facilement que g est une fonction paire, que h est une fonction impaire et que $f = g + h$. Par conséquent, on a montré que $\mathcal{F}(\mathbb{R}, \mathbb{R}) = \mathcal{P} + \mathcal{I}$. Enfin, montrons que la somme est directe en calculant $\mathcal{P} \cap \mathcal{I}$. Soit $f \in \mathcal{P} \cap \mathcal{I}$ et soit $x \in \mathbb{R}$, alors on a $f(-x) = f(x)$ car f est paire et $f(-x) = -f(x)$ car f est impaire. Ainsi, $f(x) = -f(x)$, donc $f(x) = 0$. Par conséquent f est identiquement nulle, d'où $\mathcal{P} \cap \mathcal{I} = \{0_{\mathcal{F}(\mathbb{R}, \mathbb{R})}\}$. On a montré que $\mathcal{F}(\mathbb{R}, \mathbb{R}) = \mathcal{P} \oplus \mathcal{I}$, donc \mathcal{P} et \mathcal{I} sont supplémentaires. ◀

Théorème 3.6.14. Si E est de dimension finie, alors tout sous-espace vectoriel F admet un supplémentaire dans E .

Démonstration. Soit F un s.e.v. de E . Si F est égal à $\{0_E\}$ ou à E , il n'y a rien à montrer. Supposons que F un s.e.v. propre de E . Notons n la dimension de E et p la dimension de F . Soit (v_1, \dots, v_p) une base de F . Par le théorème de la base incomplète, soient v_{p+1}, \dots, v_n des vecteurs de E tels que (v_1, \dots, v_n) est une base de E et posons $S := \text{Vect}(v_{p+1}, \dots, v_n)$. Puisque (v_1, \dots, v_n) est une base de E , alors tout vecteur de E s'écrit de manière unique comme la somme d'un vecteur de F et d'un vecteur de S , donc F et S sont supplémentaires. ◻

Remarque 3.6.15. Il n'y a pas unicité du supplémentaire d'un s.e.v. de E . Par exemple, si F un plan vectoriel dans \mathbb{R}^3 , alors toute droite vectorielle non contenue dans F est un supplémentaire de F . ◀

Proposition 3.6.16. Supposons que E est de dimension finie. Pour tous sous-espaces vectoriels F et G de E , les trois assertions suivantes sont équivalentes :

- (i) F et G sont supplémentaires dans E .
- (ii) $\dim(E) = \dim(F) + \dim(G)$ et $F \cap G = \{0_E\}$.
- (iii) $\dim(E) = \dim(F) + \dim(G)$ et $E = F + G$. ◀

Démonstration. Soient F et G des s.e.v. de E .

(i) \implies (ii) Si F et G sont supplémentaires, alors $E = F \oplus G$ donc $\dim(E) = \dim(F) + \dim(G)$. De plus, F et G sont en somme directe, donc $F \cap G = \{0_E\}$.

(ii) \implies (iii) D'après la formule de Grassmann, $\dim(F + G) = \dim(F) + \dim(G) - \dim(F \cap G)$. Par hypothèse, $\dim(F) + \dim(G) = \dim(E)$ et $\dim(F \cap G) = 0$, donc $\dim(F + G) = \dim(E)$. Puisque $F + G$ est un s.e.v. de même dimension que E , on a $F + G = E$.

(iii) \implies (i) D'après la formule de Grassmann, on a $\dim(F \cap G) = \dim(F) + \dim(G) - \dim(F + G) = \dim(E) - \dim(E) = 0$ par hypothèse, donc $F \cap G = \{0_E\}$. Ainsi, on a $E = F + G$ et $F \cap G = \{0_E\}$, donc F et G sont supplémentaires dans E . ◻

Remarque 3.6.17. Plus généralement, on a l'équivalence entre les assertions suivantes :

- (i) $E = F_1 \oplus \cdots \oplus F_p$.
- (ii) $E = F_1 + \cdots + F_p$ et $\dim(E) = \dim(F_1) + \cdots + \dim(F_p)$. ◀

Chapitre 4

Applications linéaires

On considère E et F des \mathbb{K} -espaces vectoriels.

4.1 Définition et premiers exemples

Dans ce chapitre, on s'intéresse à des applications entre espaces vectoriels qui ont la propriété d'être compatibles avec les opérations vectorielles : elles transforment les sommes dans l'espace de départ en sommes dans l'espace d'arrivée, de même pour la multiplication par un scalaire. Ces applications sont dites *linéaires* et vous en avez déjà croisées.

4.1.a Définition des applications linéaires

Définition 4.1.1. On dit qu'une application $f: E \rightarrow F$ est \mathbb{K} -**linéaire** si :

1. pour tous $x, y \in E$, on a $f(x + y) = f(x) + f(y)$;
2. pour tout $\lambda \in \mathbb{K}$ et pour tout $x \in E$, on a $f(\lambda x) = \lambda f(x)$.

On note $\mathcal{L}_{\mathbb{K}}(E, F)$ l'ensemble des applications \mathbb{K} -linéaires de E dans F , ou simplement $\mathcal{L}(E, F)$ s'il n'y a pas d'ambiguïté.

Remarque 4.1.2. La linéarité implique que l'image du vecteur nul de E est le vecteur nul de F . En effet, on a $f(0_E) = f(0 \cdot 0_E) = 0 \cdot f(0_E) = 0_F$. On utilise souvent la contraposée de cette proposition : si $f(0_E) \neq 0_F$ alors f ne peut pas être linéaire. ◀

Exemple 4.1.3. Voyons quelques exemples élémentaires d'applications linéaires.

1. Les fonctions linéaires $f: \mathbb{K} \rightarrow \mathbb{K}$ définies pour tout $x \in \mathbb{K}$ par $f(x) := \lambda x$, avec $\lambda \in \mathbb{K}$, sont des applications linéaires.
2. L'application identité $\text{id}_E: E \rightarrow E$ définie pour tout $x \in E$ par $\text{id}_E(x) := x$ est linéaire.
3. L'application nulle $0_{\mathcal{L}(E, F)}: E \rightarrow F$ définie pour tout $x \in E$ par $0_{\mathcal{L}(E, F)}(x) := 0_F$ est linéaire. ◀

Pour montrer qu'une application est linéaire, on pourra utiliser la caractérisation suivante qui permet de vérifier les deux conditions de la définition d'un seul coup.

Proposition 4.1.4. Une application $f: E \rightarrow F$ est linéaire si et seulement si pour tous $x, y \in E$ et pour tout $\lambda \in \mathbb{K}$, on a $f(x + \lambda y) = f(x) + \lambda f(y)$. ▶

Démonstration. La condition est évidemment nécessaire, montrons qu'elle est suffisante. Soit f une application de E dans F telle que :

$$\forall x, y \in E, \forall \lambda \in \mathbb{K}, \quad f(x + \lambda y) = f(x) + \lambda f(y). \quad (*)$$

Montrons que f est linéaire en vérifiant les deux conditions de la définition. Premièrement, en appliquant (*) avec $\lambda = 1$, on obtient que pour tous $x, y \in E$, on a $f(x + y) = f(x) + f(y)$.

Deuxièmement, en appliquant (*) avec $x = y = 0_E$ et $\lambda = -1$, on a $f(0_E) = f(0_E) - f(0_E) = 0_F$. Ainsi, en appliquant (*) avec $x = 0_E$, on obtient que pour tout $y \in E$ et pour tout $\lambda \in \mathbb{K}$, on a :

$$f(\lambda y) = f(0_E) + \lambda f(y) = 0_F + \lambda f(y) = \lambda f(y).$$

Par conséquent, f est une application linéaire. □

Définition 4.1.5 (vocabulaire sur les applications linéaires).

- Une application linéaire de E dans lui-même est appelée un **endomorphisme** de E . L'ensemble des endomorphismes de E se note $\mathcal{L}(E)$.
- Une application linéaire bijective est appelée un **isomorphisme**. S'il existe un isomorphisme entre E et F , on dit que E et F sont isomorphes.
- Une application linéaire à valeurs scalaires est appelée une **forme linéaire** sur E .

Exemple 4.1.6. L'application $f: \mathbb{R}^2 \rightarrow \mathbb{C}$ définie pour tout $(x, y) \in \mathbb{R}^2$ par $f(x, y) := x + iy$ est \mathbb{R} -linéaire. De plus, elle est bijective par existence et unicité de la forme algébrique de tout nombre complexe. Par conséquent, f est un isomorphisme : les \mathbb{R} -espaces vectoriels \mathbb{R}^2 et \mathbb{C} sont isomorphes. ◀

4.1.b Exemples d'applications linéaires

Commençons par voir à quoi ressemblent les applications linéaires entre les espaces \mathbb{K}^n et \mathbb{K}^p .

Proposition 4.1.7. Une application $f: \mathbb{K}^n \rightarrow \mathbb{K}$ est une forme linéaire si et seulement si il existe des scalaires $a_1, \dots, a_n \in \mathbb{K}$ tels que pour tout $x := (x_1, \dots, x_n) \in \mathbb{K}^n$, on ait $f(x) = a_1 x_1 + \dots + a_n x_n$. ◀

Démonstration. Soit $f \in \mathcal{L}(\mathbb{K}^n, \mathbb{K})$ et soit (e_1, \dots, e_n) la base canonique de \mathbb{K}^n . Pour tout $i \in \llbracket 1, n \rrbracket$, posons $a_i := f(e_i) \in \mathbb{K}$. Alors pour tout $(x_1, \dots, x_n) \in \mathbb{K}^n$, on a par linéarité de f :

$$f(x_1, \dots, x_n) = f(x_1 e_1 + \dots + x_n e_n) = x_1 f(e_1) + \dots + x_n f(e_n) = a_1 x_1 + \dots + a_n x_n.$$

Réciproquement, montrons que les applications de cette forme sont linéaires en utilisant la proposition 4.1.4. Soient $x := (x_1, \dots, x_n) \in \mathbb{K}^n$, $y := (y_1, \dots, y_n) \in \mathbb{K}^n$ et $\lambda \in \mathbb{K}$, alors on a :

$$f(x + \lambda y) = \sum_{j=1}^n a_j (x_j + \lambda y_j) = \sum_{j=1}^n a_j x_j + \lambda \sum_{j=1}^n a_j y_j = f(x) + \lambda f(y).$$

Par conséquent, l'application f est linéaire. □

Le lemme suivant montre qu'on peut construire une application linéaire à valeurs dans \mathbb{K}^p en combinant plusieurs formes linéaires dans un vecteur.

Lemme 4.1.8. Soient f_1, \dots, f_p des applications de E dans \mathbb{K} et soit $f: E \rightarrow \mathbb{K}^p$ l'application définie pour tout $x \in E$ par $f(x) := (f_1(x), \dots, f_p(x))$. Si les applications f_1, \dots, f_p sont des formes linéaires sur E , alors f est une application linéaire. ◀

Démonstration. Soient $x, y \in E$ et $\lambda \in \mathbb{K}$, on a :

$$\begin{aligned} f(x + \lambda y) &= (f_1(x + \lambda y), \dots, f_p(x + \lambda y)) \\ &= (f_1(x) + \lambda f_1(y), \dots, f_p(x) + \lambda f_p(y)) && \text{(linéarité des } f_j) \\ &= (f_1(x), \dots, f_p(x)) + \lambda (f_1(y), \dots, f_p(y)) \\ &= f(x) + \lambda f(y). \end{aligned}$$

D'après la proposition 4.1.4, f est linéaire. □

Exemple 4.1.9. L'application $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ définie par $f(x, y) := (2x - y, x + y, -x)$ est linéaire. En effet, les applications $f_1(x, y) := 2x - y$, $f_2(x, y) := x + y$ et $f_3(x, y) := -x$ sont des formes linéaires sur \mathbb{R}^2 , donc f est une application linéaire. ◀

Proposition 4.1.10. Toute application $f: \mathbb{K}^n \rightarrow \mathbb{K}^p$ définie pour tout $x := (x_1, \dots, x_n) \in \mathbb{K}^n$ par :

$$f(x) := \left(\sum_{j=1}^n a_{1,j} x_j, \sum_{j=1}^n a_{2,j} x_j, \dots, \sum_{j=1}^n a_{p,j} x_j \right),$$

avec $a_{i,j} \in \mathbb{K}$ pour tous $i \in \llbracket 1, p \rrbracket$ et $j \in \llbracket 1, n \rrbracket$, est une application linéaire. ▶

Remarque 4.1.11. La réciproque est vraie : les applications linéaires de \mathbb{K}^n dans \mathbb{K}^p sont toutes de cette forme, voir le cours d'algèbre linéaire 2. ▶

Démonstration. On peut le vérifier avec la proposition 4.1.4, mais on va plutôt utiliser la proposition 4.1.7 et le lemme 4.1.8. En effet, l'application f est de la forme $f(x) = (f_1(x), \dots, f_p(x))$ avec $f_i(x) := a_{i,1}x_1 + \dots + a_{i,n}x_n$. Les applications f_i sont des formes linéaires sur \mathbb{K}^n , donc f est une application linéaire. ◻

Enfin, voyons des exemples d'applications linéaires sur des espaces vectoriels de fonctions.

Exemple 4.1.12. Soit I un intervalle ouvert de \mathbb{R} , soit $E := \mathcal{C}^1(I, \mathbb{R})$ l'espace vectoriel des fonctions dérivables sur I de dérivée continue, et soit $F := \mathcal{C}^0(I, \mathbb{R})$ l'espace vectoriel des fonctions continues sur I . L'application $\mathcal{D}: E \rightarrow F$ qui à une fonction $u \in E$ associe sa dérivée $u' \in F$ est une application linéaire. En effet, pour tous $u, v \in E$ et pour tout $\lambda \in \mathbb{R}$, on a :

$$\mathcal{D}(u + \lambda v) = (u + \lambda v)' = u' + \lambda v' = \mathcal{D}(u) + \lambda \mathcal{D}(v). \quad \blacktriangleleft$$

Exemple 4.1.13. Soient $a, b \in \mathbb{R}$ tels que $a < b$ et soit $E := \mathcal{C}^0([a, b], \mathbb{R})$ l'espace vectoriel des fonctions continues sur $[a, b]$. On considère l'application $\mathcal{J}: E \rightarrow \mathbb{R}$ définie pour tout $u \in E$ par :

$$\mathcal{J}(u) := \int_a^b u(t) dt.$$

Alors \mathcal{J} est une forme linéaire sur E . En effet, on a pour tous $u, v \in E$ et pour tout $\lambda \in \mathbb{R}$:

$$\mathcal{J}(u + \lambda v) = \int_a^b (u(t) + \lambda v(t)) dt = \int_a^b u(t) dt + \lambda \int_a^b v(t) dt = \mathcal{J}(u) + \lambda \mathcal{J}(v). \quad \blacktriangleleft$$

4.2 Noyau, image et équations linéaires

Exemple 4.2.1. Considérons un système linéaire de p équations à n inconnues :

$$\begin{cases} a_{1,1} x_1 + a_{1,2} x_2 + \dots + a_{1,n} x_n = b_1 \\ a_{2,1} x_1 + a_{2,2} x_2 + \dots + a_{2,n} x_n = b_2 \\ \vdots \\ a_{p,1} x_1 + a_{p,2} x_2 + \dots + a_{p,n} x_n = b_p. \end{cases}$$

On peut voir ce système comme une seule équation dont l'inconnue est un vecteur de \mathbb{K}^n . En effet, si on pose $x := (x_1, \dots, x_n) \in \mathbb{K}^n$, $b := (b_1, \dots, b_p) \in \mathbb{K}^p$, et $f: \mathbb{K}^n \rightarrow \mathbb{K}^p$ l'application linéaire définie par :

$$f(x) := \left(\sum_{j=1}^n a_{1,j} x_j, \sum_{j=1}^n a_{2,j} x_j, \dots, \sum_{j=1}^n a_{p,j} x_j \right),$$

alors le système est équivalent à l'équation $f(x) = b$. Les équations de ce type sont dites *linéaires*. ◀

Définition 4.2.2. Une **équation linéaire** d'inconnue $x \in E$ est une équation de la forme :

$$f(x) = b,$$

où $f: E \rightarrow F$ est une application linéaire et b est un vecteur de F . Si $b = 0_F$, on dit que l'équation linéaire est **homogène**.

L'existence de solutions d'une équation linéaire revient à étudier l'*image* de l'application f : il existe des solutions si et seulement si b est dans l'image de f . L'unicité de la solution est équivalent à l'injectivité de l'application f :

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

On verra que la linéarité de f permet de se ramener à l'étude des antécédents de 0_F . L'ensemble des antécédents de 0_F par f est appelé le *noyau* de f : c'est l'ensemble des solutions de l'équation $f(x) = 0_F$, appelée *équation linéaire homogène* associée.

Ainsi, l'étude des solutions d'une équation linéaire nous conduit à étudier le noyau et l'image d'une application linéaire.

4.2.a Noyau et image d'une application linéaire

Définition 4.2.3. Soit $f: E \rightarrow F$ une application linéaire.

- On appelle **noyau** de f l'ensemble $\ker f := \{x \in E \mid f(x) = 0_F\}$.
- On appelle **image** de f l'ensemble $\operatorname{Im} f := \{f(x) : x \in E\}$.

Proposition 4.2.4. Soit $f: E \rightarrow F$ une application linéaire.

- L'image directe par f d'un sous-espace vectoriel de E est un sous-espace vectoriel de F .
- La préimage par f d'un sous-espace vectoriel de F est un sous-espace vectoriel de E . ◀

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

(i) Soit A un s.e.v. de E , montrons que $f(A)$ est un s.e.v. de F . L'ensemble $f(A)$ est non vide car A est non vide, montrons que $f(A)$ est stable par combinaison linéaire. Soient $x, y \in E$ et soit $\lambda \in \mathbb{K}$, il faut montrer que $f(x) + \lambda f(y) \in f(A)$. Par linéarité de f , on a $f(x) + \lambda f(y) = f(x + \lambda y)$. Or, $x + \lambda y \in A$ car A est un s.e.v. donc $f(x) + \lambda f(y) \in f(A)$.

(ii) Soit B un s.e.v. de F , montrons que $f^{-1}(B)$ est un s.e.v. de E . L'ensemble $f^{-1}(B)$ est non vide car il contient 0_E (en effet, $f(0_E) = 0_F \in B$). Soient $x, y \in f^{-1}(B)$ et $\lambda \in \mathbb{K}$, montrons que $x + \lambda y \in f^{-1}(B)$. Par linéarité de f , on a $f(x + \lambda y) = f(x) + \lambda f(y)$. Or, $f(x)$ et $f(y)$ sont des vecteurs de B par hypothèse, et B est un s.e.v. de F , donc $f(x) + \lambda f(y) \in B$. Par conséquent, on a montré que $x + \lambda y \in f^{-1}(B)$. ◻

Corollaire 4.2.5. Pour toute application linéaire $f: E \rightarrow F$, son noyau est un sous-espace vectoriel de E et son image est un sous-espace vectoriel de F . ◀

Démonstration. Soit $f: E \rightarrow F$ une application linéaire. Le noyau de f est la préimage de $\{0_F\}$, donc c'est un s.e.v. de E . L'image de f est l'image directe de E par f , donc c'est un s.e.v. de F . ◻

Exemple 4.2.6. On considère un système linéaire homogène :

$$\begin{cases} a_{1,1} x_1 + a_{1,2} x_2 + \cdots + a_{1,n} x_n = 0 \\ a_{2,1} x_1 + a_{2,2} x_2 + \cdots + a_{2,n} x_n = 0 \\ \vdots \\ a_{p,1} x_1 + a_{p,2} x_2 + \cdots + a_{p,n} x_n = 0. \end{cases} \quad (S)$$

Soit F l'ensemble des solutions de (S). Alors F est le noyau de l'application linéaire $f: \mathbb{K}^n \rightarrow \mathbb{K}^p$ définie pour tout $x := (x_1, \dots, x_n) \in \mathbb{K}^n$ par :

$$f(x) := \left(\sum_{j=1}^n a_{1,j} x_j, \sum_{j=1}^n a_{2,j} x_j, \dots, \sum_{j=1}^n a_{p,j} x_j \right).$$

Par conséquent, F est un s.e.v. de \mathbb{K}^n . Résoudre un système linéaire homogène de p équations à n inconnues est équivalent à déterminer le noyau d'une application linéaire de \mathbb{K}^n dans \mathbb{K}^p . ◀

Connaitre le noyau d'une application linéaire nous permet de déterminer les antécédents d'un vecteur par cette application : la différence entre deux antécédents d'un même vecteur est un élément du noyau.

Lemme 4.2.7. Pour tous $x, x' \in E$, on a $f(x) = f(x')$ si et seulement si $x - x' \in \ker f$. ▶

Démonstration. Soient $x, x' \in E$, alors on a :

$$\begin{aligned} f(x) = f(x') &\iff f(x) - f(x') = 0_F \\ &\iff f(x - x') = 0_F && \text{(linéarité de } f) \\ &\iff x - x' \in \ker f. \end{aligned}$$

□

La connaissance du noyau permet de savoir si une application linéaire est injective.

Proposition 4.2.8. Soit $f \in \mathcal{L}(E, F)$, alors :

- (i) f est injective si et seulement si $\ker f = \{0_E\}$.
- (ii) f est surjective si et seulement si $\text{Im } f = F$. ▶

Démonstration. La proposition (ii) est la définition de la surjectivité (cf. annexe A). Montrons la proposition (i) par double implication.

(\implies) Supposons que f est injective et montrons que $\ker f = \{0_E\}$. Soit $x \in \ker f$, c'est-à-dire tel que $f(x) = 0_F$. Puisque $f(0_E) = 0_F$ également, alors on a $x = 0_E$ par injectivité de f . Par conséquent, $\ker f = \{0_E\}$.

(\impliedby) Supposons que $\ker f = \{0_E\}$ et montrons que f est injective. Soient x et x' des vecteurs de E tels que $f(x) = f(x')$, alors d'après le lemme 4.2.7, on a :

$$f(x) = f(x') \iff x - x' \in \ker f.$$

Puisque $\ker f = \{0_E\}$, on a $x - x' = 0_E$, donc $x = x'$. Par conséquent, f est injective. □

4.2.b Structure des solutions des équations linéaires

Pour comprendre la structure des solutions d'une équation linéaire, regardons un exemple qu'on connaît bien : les système linéaires.

Exemple 4.2.9. Reprenons le système (2.5) de l'exemple 2.6.2. Après avoir échelonné le système, on a obtenu :

$$\begin{cases} -x & + 3z + 3t = 3 \\ x + 2y - 8z - 5t = -2 \\ 3x - 2y - 4z - 7t = -10. \end{cases} \iff \begin{cases} x = -3 + 3z + 3t \\ y = \frac{1}{2} + \frac{5}{2}z + t. \end{cases}$$

Si on ajoute les équations $z = \lambda_1$ et $t = \lambda_2$ avec $\lambda_1, \lambda_2 \in \mathbb{R}$ des paramètres libres, alors les solutions s'écrivent :

$$\begin{cases} x = -3 + 3\lambda_1 + 3\lambda_2 \\ y = \frac{1}{2} + \frac{5}{2}\lambda_1 + \lambda_2 \\ z = 0 + \lambda_1 \\ t = 0 + \lambda_2 \end{cases} \quad \lambda_1, \lambda_2 \in \mathbb{R}.$$

Autrement dit, les solutions du système s'écrivent $(x, y, z, t) = (-3, \frac{1}{2}, 0, 0) + \lambda_1 z_1 + \lambda_2 z_2$ avec $z_1, z_2 \in \mathbb{R}^4$ les vecteurs définis par $z_1 = (3, \frac{5}{2}, 1, 0)$ et $z_2 = (3, 1, 0, 1)$. La structure des solutions est la suivante :

- le vecteur $(-3, \frac{1}{2}, 0, 0)$ est une solution particulière du système (c'est le cas $\lambda_1 = 0$ et $\lambda_2 = 0$) ;
- les vecteurs z_1 et z_2 forment une base de l'ensemble des solutions du système homogène associé (on rappelle que l'ensemble des solutions d'un système homogène est un s.e.v.). Autrement dit, (z_1, z_2) est une base du noyau de l'application linéaire associé au système. ◀

Pour résumer, les solutions d'un système linéaires s'écrivent comme la somme d'une solution particulière du système et d'un élément quelconque du noyau de l'application linéaire associée au système. Ce fait est valable plus généralement pour n'importe quelle équation linéaire.

Théorème 4.2.10. Pour $f \in \mathcal{L}(E, F)$ et $b \in F$, on considère l'équation linéaire d'inconnue $x \in E$:

$$f(x) = b. \quad (4.1)$$

- Si $b \notin \text{Im } f$, alors l'équation (4.1) n'a pas de solution.
- Si $b \in \text{Im } f$, soit $x_0 \in E$ une solution particulière de (4.1). L'ensemble des solutions de (4.1) est :

$$\{x_0 + u : u \in \ker f\}.$$

Par définition, les éléments de $\ker f$ sont les solutions de $f(x) = 0_F$. L'équation $f(x) = 0_F$ est appelée l'équation linéaire homogène associée à (4.1). Ce théorème affirme donc que la solution générale d'une équation linéaire s'obtient en faisant la somme d'une solution particulière avec la solution générale de l'équation homogène associée.

Démonstration. Soit $f : E \rightarrow F$ une application linéaire et soit $b \in F$.

- Si $b \notin \text{Im } f$, alors par définition l'équation (4.1) n'a pas de solution.
- Si $b \in \text{Im } f$, soit $x_0 \in E$ tel que $f(x_0) = b$. Alors pour tout $x \in E$, on a :

$$f(x) = b \iff f(x) = f(x_0) \iff x - x_0 \in \ker f \iff \exists u \in \ker f, x = x_0 + u.$$

Par conséquent, l'ensemble des solutions de (4.1) est $\{x_0 + u : u \in \ker f\}$. ◻

Exemple 4.2.11 (un théorème vu en terminale). Soit I un intervalle de \mathbb{R} . Pour $a \in \mathbb{R}$ et b une fonction continue sur I , on considère l'équation différentielle linéaire d'ordre 1 :

$$\forall t \in I, \quad y'(t) = ay(t) + b(t). \quad (\mathcal{E})$$

Notons $E := \mathcal{C}^1(I, \mathbb{R})$ l'espace vectoriel des fonctions continument dérivables sur I , et $F := \mathcal{C}^0(I, \mathbb{R})$ l'espace vectoriel des fonctions continues sur I . Soit $f : E \rightarrow F$ l'application définie pour tout $y \in E$ par $f(y) := y' - ay$. On vérifie facilement que f est linéaire et que l'équation (\mathcal{E}) se réécrit $f(y) = b$. D'après le théorème 4.2.10, la solution générale de (\mathcal{E}) est $y = y_0 + u$ où y_0 est une solution particulière de (\mathcal{E}) et u est la solution générale de l'équation différentielle homogène associée :

$$\forall t \in I, \quad y'(t) = ay(t). \quad (\mathcal{E}_h)$$

La solution générale de cette équation est $u(t) := \lambda e^{at}$ où λ est un nombre réel quelconque. Autrement dit, les solutions de (\mathcal{E}_h) forment un s.e.v. de E de dimension 1 et la fonction $x \mapsto e^{ax}$ est une base de ce sous-espace. Par conséquent, si y_0 est une solution particulière de (\mathcal{E}) , alors les solutions de (\mathcal{E}) sont les fonctions :

$$\begin{aligned} y_\lambda : I &\longrightarrow \mathbb{R} \\ t &\longmapsto y_0(t) + \lambda e^{at}, \end{aligned}$$

avec $\lambda \in \mathbb{R}$. ◀

4.3 Applications linéaires et dimension

4.3.a Image d'une famille libre, d'une famille génératrice, d'une base

Lemme 4.3.1. Soit $f: E \rightarrow F$ une application linéaire.

- (i) Les images d'une famille génératrice de E est une famille génératrice de $\text{Im } f$.
- (ii) Si f est injective, alors les images d'une famille libre de vecteurs de E sont une famille libre de vecteur de F .
- (iii) Si f est bijective, alors l'image d'une base de E est une base de F . ▶

Remarque 4.3.2. En particulier, si (e_1, \dots, e_n) est une base de E , alors $\text{Im } f = \text{Vect}(f(e_1), \dots, f(e_n))$. On peut ensuite extraire une base de $\text{Im } f$ de la famille des $f(e_i)$. ▶

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

(i) Soit \mathcal{G} une famille génératrice de E . Soit $y \in \text{Im } f$ et soit $x \in E$ tel que $f(x) = y$. Puisque \mathcal{G} engendre E , il existe v_1, \dots, v_p des vecteurs de \mathcal{G} et des scalaires $\lambda_1, \dots, \lambda_p \in \mathbb{K}$ tels que :

$$x = \lambda_1 v_1 + \dots + \lambda_p v_p.$$

Par linéarité de f , on a $y = f(x) = \lambda_1 f(v_1) + \dots + \lambda_p f(v_p)$. Par conséquent, tout vecteur de $\text{Im } f$ est une combinaison linéaire d'images de vecteur de \mathcal{G} .

(ii) Supposons que f est injective. Soit $\mathcal{L} := (v_1, \dots, v_p)$ une famille libre de E , montrons que la famille $(f(v_1), \dots, f(v_p))$ est libre dans F . Soient $\lambda_1, \dots, \lambda_p \in \mathbb{K}$, on a :

$$\lambda_1 f(v_1) + \dots + \lambda_p f(v_p) = 0_E \implies f(\lambda_1 v_1 + \dots + \lambda_p v_p) = 0_E \implies \lambda_1 v_1 + \dots + \lambda_p v_p \in \ker f.$$

Puisque f est injective, on a $\ker f = \{0_E\}$, donc $\lambda_1 v_1 + \dots + \lambda_p v_p = 0_E$. Puisque \mathcal{L} est une famille libre, les scalaires $\lambda_1, \dots, \lambda_p$ sont tous nuls.

(iii) Supposons que f est bijective. Soit $\mathcal{B} := (e_1, \dots, e_n)$ une base de E . La famille \mathcal{B} est libre et f est injective, donc $(f(e_1), \dots, f(e_n))$ est une famille libre d'après (ii). De plus, \mathcal{B} engendre E , donc $(f(e_1), \dots, f(e_n))$ engendre $\text{Im } f = F$ (surjectivité de f) d'après (i). Par conséquent, $(f(e_1), \dots, f(e_n))$ est une base de F . □

Proposition 4.3.3. Supposons que E est de dimension finie. Si F est isomorphe à E , alors F est aussi de dimension finie et $\dim(E) = \dim(F)$. ▶

Démonstration. Supposons que E est de dimension finie n . Si $n = 0$, alors $E = \{0_E\}$ et $F = \{0_F\}$ car f est bijective, donc $\dim(F) = 0$. Sinon, soit (e_1, \dots, e_n) une base de E et soit $f: E \rightarrow F$ un isomorphisme. Alors $(f(e_1), \dots, f(e_n))$ est une base de F d'après le lemme 4.3.1, donc $\dim(F) = n$. □

4.3.b Rang d'une application linéaire

Définition 4.3.4. Soit $f: E \rightarrow F$ une application linéaire. Si $\text{Im } f$ est de dimension finie, on appelle **rang** de f , et on note $\text{rg}(f)$, la dimension de $\text{Im } f$.

Proposition 4.3.5. Soit $f: E \rightarrow F$ une application linéaire.

- (i) Si E est de dimension finie, alors $\text{Im } f$ est de dimension finie et $\text{rg}(f) \leq \dim(E)$.
- (ii) Si F est de dimension finie, alors $\text{Im } f$ est de dimension finie et $\text{rg}(f) \leq \dim(F)$. ▶

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

(i) Supposons que E est de dimension finie n . Si $n = 0$, il n'y a rien à montrer. Sinon, soit (e_1, \dots, e_n) une base de E . Alors $(f(e_1), \dots, f(e_n))$ est une famille génératrice de $\text{Im } f$, donc $\text{rg}(f) \leq n$.

(ii) $\text{Im } f$ est un s.e.v. de F , donc il est de dimension finie si F l'est, et sa dimension est inférieure à celle de F . □

Le noyau et l'image de f ne sont pas des sous-espaces vectoriels du même espace : le noyau est un sous-espace de l'ensemble de départ alors que l'image est un sous-espace de l'ensemble d'arrivée. Leurs dimensions sont malgré tout liées par la relation suivante, appelée **théorème du rang**.

Théorème 4.3.6 (théorème du rang). Si E est de dimension finie, alors pour toute application linéaire $f: E \rightarrow F$, son noyau et son image sont de dimensions finies et on a :

$$\dim(E) = \dim(\ker f) + \operatorname{rg}(f).$$

Exemple 4.3.7 (rang d'un système linéaire). Au chapitre 2, on a défini le rang d'un système linéaire échelonné comme le nombre d'inconnues principales. Dans ce chapitre, on a défini la notion de rang d'une application linéaire. Voyons en quoi ces deux définitions coïncident.

On a vu à l'exemple 4.2.1 qu'un système linéaire de p équations à n inconnues peut s'écrire comme une équation linéaire $f(x) = b$ avec f une application linéaire de \mathbb{K}^n dans \mathbb{K}^p . Si on échelonne le système, on sait que le nombre n d'inconnues est égal à la somme du nombre d'inconnues principales (le rang du système) et du nombre d'inconnues secondaires. Or, le nombre d'inconnues secondaires, c'est-à-dire le nombre de variables libres, est égal la dimension du noyau de f . Par conséquent, le nombre d'inconnues principales est égal à $n - \dim(\ker f)$, c'est-à-dire à $\operatorname{rg}(f)$ d'après le théorème du rang. Ainsi, le rang du système est égal au rang de l'application f .

On peut maintenant justifier la remarque 2.4.7 : si (S) et (S') sont des systèmes équivalents, alors les applications linéaires associées à chaque système ont le même noyau (les systèmes ont les mêmes solutions), donc ils ont le même rang d'après le théorème du rang. ◀

Démonstration. Soit $f: E \rightarrow F$ une application linéaire et notons $p := \dim(\ker f)$. On traite à part les cas $p = 0$ et $p = n$.

- Si $\ker f = E$, alors f est l'application nulle donc $\operatorname{rg}(f) = 0$ et il n'y a rien à montrer.
- Si $\ker f = \{0_E\}$, alors f est injective. D'après le lemme 4.3.1 l'image d'une base de E est une base de $\operatorname{Im} f$, donc $\dim(E) = \dim(\operatorname{Im} f)$ et le théorème est démontré.

Supposons maintenant que $\ker f$ est différent de $\{0_E\}$ ($p \neq 0$) et de E ($p \neq n$). Soit (e_1, \dots, e_p) une base de $\ker f$. Par le théorème de la base incomplète, il existe des vecteurs $e_{p+1}, \dots, e_n \in E$ tels que $(e_1, \dots, e_p, e_{p+1}, \dots, e_n)$ soit une base de E . On va montrer que les vecteurs $f(e_{p+1}), \dots, f(e_n)$ forment une base de l'image de f .

1. Montrons que $(f(e_{p+1}), \dots, f(e_n))$ est une famille génératrice de $\operatorname{Im} f$. En effet, on a :

$$\begin{aligned} \operatorname{Im} f &= \operatorname{Vect}(f(e_1), \dots, f(e_p), f(e_{p+1}), \dots, f(e_n)) \\ &= \operatorname{Vect}(0_E, \dots, 0_E, f(e_{p+1}), \dots, f(e_n)) \\ &= \operatorname{Vect}(f(e_{p+1}), \dots, f(e_n)). \end{aligned}$$

2. Montrons que $(f(e_{p+1}), \dots, f(e_n))$ est une famille libre. Soient $\lambda_{p+1}, \dots, \lambda_n \in \mathbb{K}$ tels que :

$$\lambda_{p+1}f(e_{p+1}) + \dots + \lambda_n f(e_n) = 0_F.$$

Par linéarité, on a $f(\lambda_{p+1}e_{p+1} + \dots + \lambda_n e_n) = 0_F$, donc $\lambda_{p+1}e_{p+1} + \dots + \lambda_n e_n \in \ker f$. Donc il existe $\lambda_1, \dots, \lambda_p \in \mathbb{K}$ tels que $\lambda_{p+1}e_{p+1} + \dots + \lambda_n e_n = \lambda_1 e_1 + \dots + \lambda_p e_p$, d'où :

$$(-\lambda_1)e_1 + \dots + (-\lambda_p)e_p + \lambda_{p+1}e_{p+1} + \dots + \lambda_n e_n = 0_E.$$

Puisque (e_1, \dots, e_n) est une base de E , tous les λ_i sont nuls, donc $(f(e_{p+1}), \dots, f(e_n))$ est libre. Par conséquent, $(f(e_{p+1}), \dots, f(e_n))$ est une base de $\operatorname{Im} f$, donc en comptant le nombre de vecteurs de cette base :

$$\operatorname{rg}(f) = \dim(\operatorname{Im} f) = n - p = \dim(E) - \dim(\ker f). \quad \square$$

Le théorème permet de calculer la dimension de l'image d'une application linéaire f à partir de celle de son noyau, ce qui s'avère utile pour trouver une base de l'image. En effet, il est facile de déterminer une famille génératrice de $\text{Im } f$ (il suffit de prendre l'image par f d'une base de E d'après le lemme 4.3.1), il suffit ensuite d'en extraire une sous-famille libre comportant autant de vecteurs que le rang.

Exemple 4.3.8 (examen 2022-2023, session 1). Considérons l'application linéaire $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ définie pour tout $(x, y, z) \in \mathbb{R}^3$ par :

$$f(x, y, z) := (x - y + z, 2x + y + 3z, x + 2y + 2z).$$

Déterminons des bases du noyau et de l'image de f .

• **Base du noyau.** Soit $(x, y, z) \in \mathbb{R}^3$, on a :

$$\begin{aligned} (x, y, z) \in \ker f &\iff f(x, y, z) = (0, 0, 0) \\ &\iff \begin{cases} x - y + z = 0 \\ 2x + y + 3z = 0 \\ x + 2y + 2z = 0 \end{cases} \\ &\iff \begin{cases} x - y + z = 0 \\ 3y + z = 0 & L_2 \leftarrow L_2 - 2L_1 \\ 3y + z = 0 & L_3 \leftarrow L_3 - L_1 \end{cases} \\ &\iff \begin{cases} x - y + z = 0 \\ 3y + z = 0 \\ 0 = 0. & L_3 \leftarrow L_3 - L_2 \end{cases} \end{aligned}$$

Le système est de rang $2 < 3$, donc il admet une infinité de solutions dépendant d'une variable libre. Choisissons y comme variable libre, alors :

$$(x, y, z) \in \ker f \iff \begin{cases} x = 4y \\ z = -3y \end{cases} \iff (x, y, z) = (4y, y, -3y) \iff (x, y, z) \in \text{Vect}((4, 1, -3)).$$

Ainsi, $\ker f$ est de dimension 1 et une base de $\ker f$ est le vecteur $(4, 1, -3)$.

• **Base de l'image.** Notons (e_1, e_2, e_3) la base canonique de \mathbb{R}^3 . Alors $(f(e_1), f(e_2), f(e_3))$ est une famille génératrice de l'image de f :

$$f(e_1) = f(1, 0, 0) = (1, 2, 1), \quad f(e_2) = f(0, 1, 0) = (-1, 1, 2), \quad f(e_3) = f(0, 0, 1) = (1, 3, 2).$$

Ainsi, $\text{Im } f = \text{Vect}((1, 2, 1), (-1, 1, 2), (1, 3, 2))$. Or, d'après le théorème du rang, le rang de f est donné par $\text{rg}(f) = \dim(\mathbb{R}^3) - \dim(\ker f) = 3 - 1 = 2$, donc $\text{Im } f$ est de dimension 2. Il suffit de choisir une sous-famille de 2 vecteurs libres parmi les générateurs de $\text{Im } f$ pour former une base, par exemple $(1, 2, 1)$ et $(-1, 1, 2)$. Une base de $\text{Im } f$ est $((1, 2, 1), (-1, 1, 2))$. ◀

Corollaire 4.3.9. Supposons que E et F sont de même dimension finie. Pour toute application linéaire $f: E \rightarrow F$, on a l'équivalence :

$$f \text{ est injective} \iff f \text{ est surjective} \iff f \text{ est bijective.} \quad \blacktriangleleft$$

Démonstration. Il suffit de démontrer la première équivalence.

$$\begin{aligned} f \text{ est injective} &\iff \ker f = \{0_E\} && \text{(proposition 4.2.8)} \\ &\iff \dim(E) = \text{rg}(f) && \text{(théorème du rang)} \\ &\iff \dim(F) = \text{rg}(f) && (\dim(E) = \dim(F)) \\ &\iff F = \text{Im } f && \text{(proposition 3.5.12)} \\ &\iff f \text{ est surjective.} && \text{(proposition 4.2.8)} \quad \square \end{aligned}$$

Remarque 4.3.10.

1. Attention, ce corollaire ne dit pas que l'injectivité et la surjectivité sont équivalentes en général, mais seulement pour les applications linéaires entre des espaces vectoriels de même dimension finie.
2. En dimension finie, pour montrer qu'une application linéaire est un isomorphisme, il suffit donc de vérifier que les espaces de départ et d'arrivée ont la même dimension et que l'application est injective (en calculant son noyau par exemple). ◀

4.3.c Caractérisation d'une application linéaire

Lorsque E est de dimension finie, une application linéaire définie sur E est caractérisée par les valeurs qu'elle prend sur une base de E : si on connaît l'image des vecteurs d'une base de E , alors on connaît l'image de tout vecteur.

Théorème 4.3.11. Supposons que E est de dimension finie $n \geq 1$. Soit $\mathcal{B} := (e_1, \dots, e_n)$ une base de E . Pour toute famille (v_1, \dots, v_n) de vecteurs de F , il existe une unique application linéaire $f: E \rightarrow F$ telle que $f(e_i) = v_i$ pour tout $i \in \llbracket 1, n \rrbracket$.

Démonstration. Soit $\mathcal{B} := (e_1, \dots, e_n)$ une base de E et soit (v_1, \dots, v_n) une famille de vecteurs de F . On procède par analyse-synthèse.

Analyse. Soit $f: E \rightarrow F$ une application linéaire telle que $f(e_i) = v_i$. Soit $x \in E$ et soient $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ ses coordonnées dans la base \mathcal{B} . Par linéarité, on a :

$$f(x) = f(\lambda_1 e_1 + \dots + \lambda_n e_n) = \lambda_1 f(e_1) + \dots + \lambda_n f(e_n) = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

Ceci démontre l'unicité de f (sous réserve d'existence).

Synthèse. Soit $f: E \rightarrow F$ l'application définie pour tout $x := \lambda_1 e_1 + \dots + \lambda_n e_n \in E$ par :

$$f(x) := \lambda_1 v_1 + \dots + \lambda_n v_n.$$

Par définition, on a $f(e_i) = v_i$ pour tout $i \in \llbracket 1, n \rrbracket$. Montrons que f est linéaire. Pour tous vecteurs $x := \lambda_1 e_1 + \dots + \lambda_n e_n \in E$ et $y := \lambda'_1 e_1 + \dots + \lambda'_n e_n \in E$, et pour tout scalaire $\mu \in \mathbb{K}$, on a :

$$\begin{aligned} f(x + \mu y) &= f\left(\sum_{i=1}^n (\lambda_i + \mu \lambda'_i) e_i\right) \\ &= \sum_{i=1}^n (\lambda_i + \mu \lambda'_i) v_i \\ &= \sum_{i=1}^n \lambda_i v_i + \mu \sum_{i=1}^n \lambda'_i v_i \\ &= f(x) + \mu f(y). \end{aligned}$$

Par conséquent, l'application f est linéaire. Ainsi, on a montré l'existence d'une application linéaire satisfaisant $f(e_i) = v_i$ pour tout $i \in \llbracket 1, n \rrbracket$. ◻

Corollaire 4.3.12. Des espaces vectoriels de dimensions finies sont isomorphes si et seulement si ils ont la même dimension. ◀

Démonstration. Si E est de dimension finie et que F est isomorphe à E , alors ils ont la même dimension d'après la proposition 4.3.3. Réciproquement, supposons que E et F ont la même dimension n . Si $n = 0$, alors $E = \{0_E\}$ et $F = \{0_F\}$ sont trivialement isomorphes. Sinon, soit $\mathcal{B} := (e_1, \dots, e_n)$ une base de E et soit $\mathcal{B}' := (e'_1, \dots, e'_n)$ une base de F . On considère l'application linéaire $f: E \rightarrow F$ telle que $f(e_i) = e'_i$ pour tout $i \in \llbracket 1, n \rrbracket$ (cette application existe d'après le théorème 4.3.11).

Montrons que f est injective. Soit $x \in \ker f$, montrons que $x = 0_E$. Si $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$ sont les coordonnées de x dans la base \mathcal{B} , alors on a par linéarité de f :

$$f(x) = f(\lambda_1 e_1 + \dots + \lambda_n e_n) = \lambda_1 e'_1 + \dots + \lambda_n e'_n.$$

Par conséquent, les coordonnées de $f(x)$ dans la base \mathcal{B}' sont égales à celles de x dans la base \mathcal{B} . Ainsi, on a :

$$f(x) = 0_F \implies \forall i \in \llbracket 1, n \rrbracket, \lambda_i = 0 \implies x = 0_E,$$

donc f est injective. Puisque E et F ont la même dimension, f est bijective d'après le corollaire 4.3.9. Par conséquent, E et F sont isomorphes. \square

Remarque 4.3.13. Si E est dimension finie n , alors E est isomorphe à \mathbb{K}^n via l'application qui à tout vecteur de E associe ses coordonnées dans une base. Ainsi, on connaît à un isomorphisme près tous les \mathbb{K} -espaces vectoriels de dimension finie. Étudier l'espace vectoriel \mathbb{K}^n et les applications linéaires de \mathbb{K}^n dans \mathbb{K}^p permet donc de comprendre tous les espaces vectoriels de dimension n et toutes les applications linéaires d'un espace de dimension n dans un espace de dimension p . \blacktriangleleft

Annexe A

Notions générales

A.1 Théorie des ensembles

Depuis le début du XX^e siècle, on considère que les mathématiques sont fondées sur les notions d'ensembles et de relation d'appartenance. Intuitivement, un **ensemble** est une *collection non ordonnée d'objets uniques*, au sens suivant :

- **non ordonnée** : on ne tient pas compte de l'ordre dans lequel apparaissent les éléments de la collection. Par exemple, les ensembles $\{a, b, c\}$ et $\{b, c, a\}$ sont identiques.
- **d'objets uniques** : on ne tient pas compte du nombre de fois qu'un objet apparaît dans la collection. Par exemple, les ensembles $\{a, a, a, b, b\}$ et $\{a, b\}$ sont identiques.

Cependant, si on accepte n'importe quelle collection comme étant un ensemble, on aboutit à des paradoxes (le plus connu étant le paradoxe de Russel). Pour éviter ça, la théorie des ensembles doit reposer sur une liste d'axiomes qui précisent quelles sont les opérations permises pour construire des ensembles acceptables. La théorie axiomatique des ensembles la plus connue et la plus largement utilisée est celle de Zermelo et Fraenkel (théorie ZF), à laquelle on ajoute généralement l'axiome du choix (théorie ZFC).

La suite de cette section ne constitue en aucun cas un exposé d'une théorie axiomatique des ensembles, mais une description des notions et opérations ensemblistes les plus courantes pour la pratique des mathématiques.

Égalité et sous-ensembles

On écrit « $x \in E$ » pour dire que l'objet x est un élément de l'ensemble E . Un ensemble est caractérisée uniquement par ses éléments.

Définition A.1.1. Soient E et F des ensembles. On dit que E et F sont égaux s'ils ont les mêmes éléments :

$$E = F \stackrel{\text{def.}}{\iff} \forall x, (x \in E \iff x \in F).$$

Il existe un ensemble particulier qui ne possède aucun élément. On l'appelle l'ensemble vide et on le note \emptyset . Il est unique au vu de la définition précédente.

Définition A.1.2. Soient E et F des ensembles. On dit que F est inclus dans E , ou encore que F est un sous-ensemble ou une partie de E , et on note $F \subset E$, si tout élément de F est aussi un élément de E :

$$F \subset E \stackrel{\text{def.}}{\iff} \forall x, (x \in F \implies x \in E).$$

On note $\mathcal{P}(E)$ l'ensemble des parties de E .

Remarque A.1.3. Deux ensembles sont égaux si et seulement si ils sont inclus l'un dans l'autre :

$$E = F \iff E \subset F \text{ et } F \subset E.$$



Exemple A.1.4. Si $E := \{1;2;3\}$, alors l'ensemble des parties de E est :

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

Modes de définition

Il existe plusieurs façons de définir un ensemble :

- **Ensemble défini en extension** : si l'ensemble est fini, on peut le définir en listant explicitement ses éléments. Par exemple, l'ensemble $\{1,2,3\}$ est défini en extension, ses éléments sont 1, 2 et 3.
- **Ensemble défini en compréhension** : on définit un ensemble par une propriété qui caractérise ses éléments. Soient E un ensemble et P un prédicat sur E , c'est-à-dire que pour chaque $x \in E$, $P(x)$ est une assertion vraie ou fausse. L'ensemble des éléments $x \in E$ pour lesquels $P(x)$ est vraie se note $\{x \in E \mid P(x)\}$. Par exemple, un intervalle $[a,b]$ est défini en compréhension par $[a,b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}$.
- **Ensemble défini par image directe** : on définit un ensemble par une formule décrivant ses éléments. Soient E et F des ensembles et soit $f(x) \in F$ une formule dépendant d'un paramètre $x \in E$. L'ensemble des toutes les valeurs prises par $f(x)$ lorsque x décrit l'ensemble E se note $\{f(x) : x \in E\}$. Par exemple, l'ensemble des carrés parfaits peut se noter $\{n^2 : n \in \mathbb{N}\}$.

Remarque A.1.5. Un ensemble défini par image directe est en réalité un ensemble défini en compréhension :

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

Il faut garder à l'esprit que derrière une définition par image directe se cache toujours un quantificateur existentiel.

Opérations ensemblistes

On rappelle la définition des opérations ensemblistes de base.

Définition A.1.6. Soient E et F des ensembles.

- On appelle réunion de E et F , et on note $E \cup F$, l'ensemble des éléments appartenant à E ou à F :

$$x \in E \cup F \stackrel{\text{def.}}{\iff} x \in E \text{ ou } x \in F.$$

- On appelle intersection de E et F , et on note $E \cap F$, l'ensemble des éléments appartenant à la fois à E et à F :

$$x \in E \cap F \stackrel{\text{def.}}{\iff} x \in E \text{ et } x \in F.$$

- On appelle différence de E et F , et on note $E \setminus F$, l'ensemble des éléments de E qui ne sont pas des éléments de F :

$$x \in E \setminus F \stackrel{\text{def.}}{\iff} x \in E \text{ et } x \notin F$$

- Si F est une partie de E , on appelle complémentaire de F dans E l'ensemble $F^c := E \setminus F$.

Produit cartésien

Définition A.1.7. Soient x et y deux objets. Le couple (x,y) est l'objet formé de x et de y dans cet ordre. L'élément x est la première composante du couple et l'élément y est la deuxième composante. Deux couples sont égaux s'ils sont égaux composante par composante :

$$(x,y) = (x',y') \stackrel{\text{def.}}{\iff} x = x' \text{ et } y = y'.$$

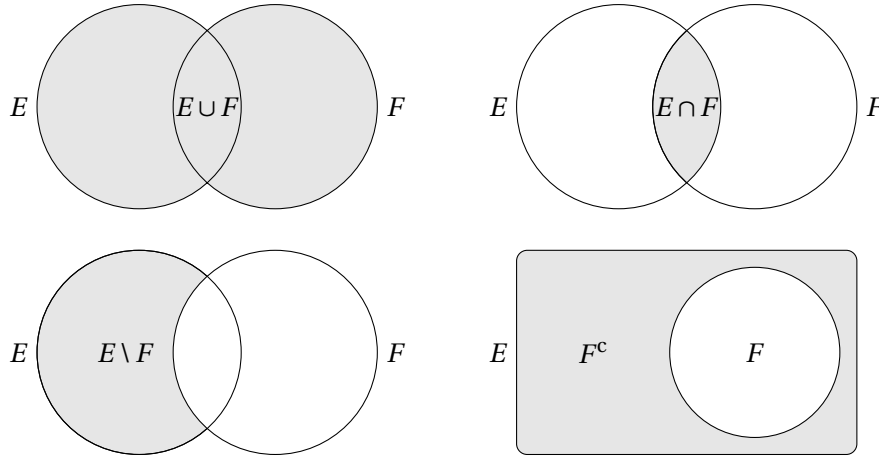


Figure A.1 – Diagramme de Venn des opérations ensemblistes.

Remarque A.1.8. Le couple (x, y) ne doit pas être confondu avec la paire $\{x, y\}$. Dans un couple, l'ordre des composantes est important : le couple (x, y) n'est pas égal au couple (y, x) , sauf si $x = y$. Dans une paire, l'ordre des éléments n'a pas d'importance : $\{x, y\} = \{y, x\}$. ◀

Définition A.1.9. Soient E et F des ensembles. On appelle **produit cartésien** de E et F , et on note $E \times F$, l'ensemble de tous les couples (x, y) avec $x \in E$ et $y \in F$:

$$E \times F := \{(x, y) : x \in E \text{ et } y \in F\}.$$

Lorsque $E = F$, on note $E^2 := E \times E$.

L'idée de couple se généralise aux listes de 3 éléments (triplets), de 4 éléments (quadruplets), etc.

Définition A.1.10. Soient x_1, \dots, x_n des objets. Le n -uplet (x_1, \dots, x_n) est l'objet formé de la liste des x_1, \dots, x_n dans cet ordre. Pour tout $i \in \llbracket 1, n \rrbracket$, l'élément x_i est appelé la i -ième composante du n -uplet. Deux n -uplets sont égaux s'ils sont égaux composante par composante :

$$(x_1, \dots, x_n) = (x'_1, \dots, x'_n) \stackrel{\text{def.}}{\iff} \forall i \in \llbracket 1, n \rrbracket, x_i = x'_i.$$

On définit ainsi la notion de produit cartésien d'un nombre quelconque d'ensembles.

Définition A.1.11. Soit $n \geq 2$ et soient E_1, \dots, E_n des ensembles. On appelle **produit cartésien** des ensembles E_1, \dots, E_n , et on note $E_1 \times \dots \times E_n$, l'ensemble de tous les n -uplets (x_1, \dots, x_n) avec $x_i \in E_i$ pour tout $i \in \llbracket 1, n \rrbracket$:

$$E_1 \times \dots \times E_n := \{(x_1, \dots, x_n) : x_1 \in E_1, \dots, x_n \in E_n\}.$$

Lorsque tous les E_i sont égaux à un même ensemble E , on note $E^n := \underbrace{E \times \dots \times E}_{n \text{ fois}}$.

A.2 Relations binaires

Définition A.2.1. Soient E et F des ensembles. On appelle **relation binaire** entre E et F un prédicat sur $E \times F$. Si \mathcal{R} est une relation entre E et F , on note $x\mathcal{R}y$ le fait que $x \in E$ et $y \in F$ soient en relation pour \mathcal{R} . Lorsque $E = F$, on dit que la relation est interne et on parle simplement de relation binaire sur E .

Exemple A.2.2.

1. L'égalité est une relation binaire sur n'importe quel ensemble E .
2. \leq est une relation binaire sur \mathbb{R} .
3. \perp est une relation binaire sur les droites du plan. ◀

Relations d'équivalence

Définition A.2.3. Soit E un ensemble. On dit qu'une relation binaire \sim sur E est une **relation d'équivalence** si elle vérifie les trois propriétés suivantes :

1. **Réflexivité** : $\forall x \in E, x \sim x$.
2. **Symétrie** : $\forall x, y \in E, x \sim y \iff y \sim x$.
3. **Transitivité** : $\forall x, y, z \in E, (x \sim y \text{ et } y \sim z) \implies x \sim z$.

Exemple A.2.4.

1. Soit $n \in \mathbb{N}$. Sur \mathbb{Z} , la relation de congruence modulo n : « $x \sim y \iff x \equiv y \pmod{n}$ », est une relation d'équivalence.
2. Soit E l'ensemble des droites du plan. La relation de parallélisme : « $d \sim d' \iff d \parallel d'$ » est une relation d'équivalence sur E .
3. Soit E l'ensemble des segments orientés du plan, c'est-à-dire l'ensemble des bipoints (A, B) où A est l'origine du segment orienté et B son extrémité. La relation d'équipollence :

$$(A, B) \sim (C, D) \iff ABDC \text{ est un parallélogramme,}$$

est une relation d'équivalence sur E . ◀

Définition A.2.5. Soit E un ensemble et soit \sim une relation d'équivalence sur E . Pour tout $x \in E$, on appelle **classe d'équivalence** de x , et on note $[x]$, l'ensemble de tous les éléments de E qui sont équivalents à x :

$$[x] := \{y \in E \mid y \sim x\}.$$

Tout élément d'une classe d'équivalence est appelé un **représentant** de cette classe.

Définition A.2.6. On appelle **ensemble quotient** de E par \sim , et on note E/\sim l'ensemble de toutes les classes d'équivalence pour la relation \sim :

$$E/\sim := \{[x] : x \in E\}.$$

Remarque A.2.7. Pour tous $x, y \in E$, on a $x \sim y \iff [x] = [y]$. On interprète E/\sim comme l'ensemble E dans lequel les éléments équivalents ont été identifiés : la relation d'équivalence \sim dans E devient la relation d'égalité dans E/\sim . ◀

Exemple A.2.8.

1. On note $\mathbb{Z}/n\mathbb{Z}$ le quotient de \mathbb{Z} pour la relation de congruence modulo n . La classe d'équivalence d'un entier a est l'ensemble $[a] := \{a + kn : k \in \mathbb{Z}\}$. Le quotient $\mathbb{Z}/n\mathbb{Z}$ est constituée de n classes d'équivalences :

$$\mathbb{Z}/n\mathbb{Z} = \{[0], [1], \dots, [n-1]\}.$$

2. Si E est l'ensemble des droites du plan, alors la classe d'équivalence d'une droite d est l'ensemble de toutes les droites parallèles à d , c'est-à-dire les droites de même direction que d . Le quotient de E pour cette relation est l'ensemble des directions du plan.
3. Si E est l'ensemble des segments orientés du plan, alors la classe d'équivalence d'un segment orienté (A, B) est l'ensemble de tous les segments orientés équipollents à (A, B) , c'est-à-dire le vecteur \overrightarrow{AB} . Le quotient de E pour la relation d'équipollence est l'ensemble des vecteurs du plan. ◀

A.3 Applications

Définition A.3.1. Soient E et F des ensembles. Une **application** $f: E \rightarrow F$ est une relation binaire entre E et F qui à tout élément $x \in E$ associe un unique élément $y \in F$. L'élément y est appelé l'**image** de x par f , et on note $y = f(x)$. Les éléments de E ayant pour image y sont appelés les **antécédents** de y par f . L'ensemble des applications de E dans F est noté $\mathcal{F}(E, F)$.

Une application est entièrement caractérisée par ses images : deux applications sont égales si elles associent les mêmes images aux mêmes entrées.

Définition A.3.2. Soient E et F des ensembles et soient $f, g \in \mathcal{F}(E, F)$. On dit que les applications f et g sont égales si pour tout $x \in E$, $f(x) = g(x)$.

Définition A.3.3. Soient E, F et G des ensembles et soient $f: E \rightarrow F$ et $g: F \rightarrow G$ des applications. On définit la **composée** de f par g , et on note $g \circ f$, l'application de E dans G définie par :

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

Définition A.3.4. Soit $f: E \rightarrow F$ et soient $A \subset E$ et $B \subset F$.

- L'**image directe** de A par f est la partie de F définie par :

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

C'est l'ensemble des images par f des éléments de A .

- L'**image réciproque** de B par f est la partie de E définie par :

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

C'est l'ensemble des antécédents par f des éléments de B .

Exemple A.3.5. Soit $f: \mathbb{R} \rightarrow \mathbb{R}$ la fonction carrée $f(x) := x^2$.

- L'image directe de $[-2, 0]$ est $f([-2, 0]) = [0, 4]$.
- L'image réciproque de $[0, 1]$ est $f^{-1}([0, 1]) = [-1, 1]$. ◀

Définition A.3.6. Soit $f: E \rightarrow F$ une application.

- On dit que f est **surjective** si tout élément de F admet un antécédent par f :

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

- On dit que f est **injective** si tout élément de F a au plus un antécédent par f :

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

- On dit que f est **bijjective** si elle est à la fois injective et surjective, c'est-à-dire si tout élément de F admet un unique antécédent par f :

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

Remarque A.3.7.

1. La surjectivité de f peut se réécrire avec l'égalité ensembliste $f(E) = F$.
2. Une application f définie sur E est toujours surjective sur son image $f(E)$. ◀

Définition A.3.8. Soit $f: E \rightarrow F$ une application bijective. On appelle **application réciproque** de f l'application $f^{-1}: F \rightarrow E$ qui à tout $y \in F$ associe son unique antécédent par f .

Exemple A.3.9. L'injectivité ou la surjectivité d'une application dépend de son expression, mais aussi des ensembles E et F .

1. La fonction $f: \mathbb{R} \rightarrow \mathbb{R}$ définie par $f(x) := x^2$ n'est pas injective car $f(1) = f(-1)$ et n'est pas surjective car -1 n'a pas d'antécédents dans \mathbb{R} .
2. La fonction $g: \mathbb{R}_+ \rightarrow \mathbb{R}$ définie par $g(x) := x^2$ est injective mais n'est pas surjective.
3. La fonction $h: \mathbb{R} \rightarrow \mathbb{R}_+$ définie par $h(x) := x^2$ est surjective mais n'est pas injective.
4. La fonction $k: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ définie par $k(x) := x^2$ est injective et surjective, donc bijective. Son application réciproque est la fonction $k^{-1}: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ définie par $k^{-1}(y) := \sqrt{y}$. ◀

A.4 Structures algébriques

Pour terminer ce chapitre de notions à connaître, passons en revue quelques structures algébriques courantes. Vous étudierez ces notions plus en détail dans la suite de vos études.

Lois de composition interne

Définition A.4.1. Soit E un ensemble. Une **loi de composition interne** sur E est une application $*$: $E \times E \rightarrow E$. L'image d'un couple (x, y) par cette application est notée $x * y$.

Exemple A.4.2.

1. L'addition et la multiplication sont des lois de composition interne sur \mathbb{N} .
2. La soustraction n'est pas une loi de composition interne sur \mathbb{N} , car $1 - 2 \notin \mathbb{N}$ par exemple. En revanche, c'est une loi de composition interne sur \mathbb{Z} .
3. Soit E un ensemble non vide, la composition d'applications \circ est une loi de composition interne dans $\mathcal{F}(E, E)$. ◀

Définition A.4.3. Soit $*$ une loi de composition interne sur un ensemble E .

- On dit que $*$ est **associative** si pour tous $x, y, z \in E$, on a $(x * y) * z = x * (y * z)$.
- On dit que $*$ est **commutative** si pour tous $x, y \in E$, on a $x * y = y * x$.
- Soit $e \in E$, on dit que e est un **élément neutre** pour $*$ si pour tout $x \in E$, on a $x * e = e * x = x$. On peut montrer que e est unique, lorsqu'il existe.
- Si $*$ possède un élément neutre e , alors on dit que $x' \in E$ est **symétrique** de $x \in E$ si :

$$x * x' = x' * x = e.$$

Exemple A.4.4.

1. Dans \mathbb{R} , l'addition est associative, commutative, admet pour élément neutre 0 et le symétrique d'un nombre x est son opposé $-x$.
2. Dans \mathbb{R} , la multiplication est associative, commutative, admet pour élément neutre 1, et le symétrique de tout nombre $x \neq 0$ est son inverse $\frac{1}{x}$.
3. Soit E un ensemble. Dans $\mathcal{F}(E, E)$, la composition d'applications est associative mais pas commutative en général, elle admet pour élément neutre la fonction identité $\text{id}_E: x \mapsto x$, et le symétrique d'une application bijective f est son application réciproque f^{-1} . ◀

Groupes

Définition A.4.5. Un **groupe** est un ensemble G muni d'une loi de composition interne $*$ telle que :

1. $*$ est associative ;
2. $*$ possède un élément neutre ;
3. tout élément de G possède un symétrique.

Si de plus la loi $*$ est commutative, on dit que le groupe est commutatif (ou abélien).

Exemple A.4.6.

1. Les ensembles \mathbb{Z} , \mathbb{Q} , \mathbb{R} et \mathbb{C} munis de l'addition sont des groupes commutatifs.
2. Les ensembles $\{-1; 1\}$, \mathbb{Q}^* , \mathbb{R}^* , \mathbb{C}^* munis de la multiplication sont des groupes commutatifs.
3. L'ensemble des applications bijectives de E dans E est un groupe (non commutatif en général) pour la composition. ◀

Anneaux

Un anneau est un ensemble muni d'une structure algébrique qui ressemble à celle de \mathbb{Z} : on peut y faire des additions, des soustractions et des multiplications, et ces opérations ont les propriétés de calcul auxquelles on est habitué.

Définition A.4.7. Un **anneau** est un ensemble A muni de deux lois de composition interne notées $+$ et \times telles que :

1. $(A, +)$ est un groupe commutatif, d'élément neutre noté 0_A ;
2. la loi \times est associative et possède un élément neutre noté 1_A ;
3. pour tous $x, y, z \in A$, on a $x \times (y + z) = (x \times y) + (x \times z)$ et on a $(x + y) \times z = (x \times z) + (y \times z)$.

Si de plus la loi \times est commutative, on dit que A est un anneau commutatif. En pratique, le symbole de la loi multiplicative peut être omis (c.-à-d. on note $xy := x \times y$).

Exemple A.4.8.

1. \mathbb{Z} est un anneau commutatif (pour les opérations d'addition et de multiplication usuelles).
2. $\mathbb{Z}/n\mathbb{Z}$ (les entiers modulo n) est un anneau commutatif.
3. Si A est un anneau commutatif, alors l'ensemble $A[X]$ des polynômes à coefficients dans A est un anneau commutatif.
4. Dans le cours d'algèbre linéaire 2, vous verrez que l'ensemble $\mathcal{M}_n(\mathbb{K})$ des matrices carrées de taille n à coefficients réels ou complexes est un anneau non commutatif. ◀

Corps


Un corps est un anneau commutatif dans lequel on peut effectuer des divisions : tout élément non nul est inversible pour la multiplication.

Définition A.4.9. Un **corps** est ensemble K muni de deux lois de composition interne notées $+$ et \times telles que :

1. $(K, +, \times)$ est un anneau ;
2. $(K \setminus \{0_K\}, \times)$ est un groupe commutatif.

Remarque A.4.10. D'autres auteur·rices définissent un corps comme étant un anneau dont tous les éléments non nuls sont inversibles, sans exiger que la multiplication soit commutative. Ce qu'on appelle ici « corps » est alors appelé « corps commutatif » suivant cette autre convention. En ce qui concerne la théorie des espaces vectoriels exposée dans ce cours, il est essentiel que le corps des scalaires soit commutatif. ◀

Exemple A.4.11.

1. \mathbb{Q} , \mathbb{R} et \mathbb{C} sont des corps (pour les opérations d'addition et de multiplication usuelles).
2. $\mathbb{Z}/p\mathbb{Z}$ (les entiers modulo p) est un corps si et seulement si p est premier.
3. Si K est un corps, alors l'ensemble $K(X)$ des fractions rationnelles, c'est-à-dire les quotients de polynômes à coefficients dans K , est un corps. 

Annexe B

Construction de \mathbb{C}

B.1 Objectif

Notre but dans cette annexe est de démontrer le théorème 1.2.1 en donnant une construction de l'ensemble des nombres complexes. Entendons-nous sur la signification de cette phrase : construire les nombres complexes, c'est donner un exemple d'ensemble \mathbf{C} , muni d'une addition $+$ et d'une multiplication \times pour lesquelles $(\mathbf{C}, +, \times)$ est un corps (cf. annexe A), de telle sorte que :

1. \mathbf{C} contienne un sous-corps \mathbf{R} qui soit isomorphe à \mathbb{R} ;
2. \mathbf{C} contienne un élément \mathbf{i} tel que $\mathbf{i}^2 = -\mathbf{1}$, où $\mathbf{1}$ est l'élément neutre de la multiplication de \mathbf{C} ;
3. tout élément de \mathbf{C} s'écrive sous la forme $\mathbf{a} + \mathbf{i} \times \mathbf{b}$ avec $\mathbf{a}, \mathbf{b} \in \mathbf{R}$.

Expliquons le sens intuitif de la propriété 1. Dire que \mathbf{C} contient un sous-corps \mathbf{R} isomorphe à \mathbb{R} signifie que \mathbf{C} contient une copie de \mathbb{R} et que les opérations $+$ et \times de \mathbf{C} prolongent celles de \mathbf{R} .

Plus précisément, donnons la définition de *sous-corps*. On pourra faire un parallèle entre les notions de corps/sous-corps et celles d'espace vectoriel/sous-espace vectoriel (cf. chapitre 3).

Définition B.1.1. Soient $(K, +, \times)$ un corps et L une partie de K . On dit que L est un sous-corps de K si L est stable pour les opérations $+$ et \times , et si la restriction à L des opérations $+$ et \times munissent L d'une structure de corps.

Donnons ensuite la définition de *corps isomorphes*. Cette définition repose sur la notion de morphisme de corps : ce sont des applications entre deux corps qui conservent la structure de corps. Encore une fois, un parallèle peut être fait avec la notion d'application linéaire (morphisme d'espaces vectoriels).

Définition B.1.2. Soient K et L des corps. On dit qu'une application $f: K \rightarrow L$ est un morphisme de corps si :

1. $f(1_K) = 1_L$.
2. $\forall x, y \in K, f(x + y) = f(x) + f(y)$.
3. $\forall x, y \in K, f(x \times y) = f(x) \times f(y)$.

Si de plus f est bijective, on dit que f est un isomorphisme de corps et que les corps K et L sont isomorphes.

Dire que deux corps sont isomorphes signifie qu'ils sont « les mêmes » en tant que corps. L'isomorphisme est un dictionnaire permettant de passer de K à L : à tout élément de K correspond un unique élément de L et inversement, et cette correspondance préserve les relations algébriques entre les éléments (une addition dans K devient une addition dans L , de même pour la multiplication).

Il existe plusieurs constructions des nombres complexes, c'est-à-dire qu'il y a plusieurs façons de définir un ensemble \mathbf{C} satisfaisant les conditions précédentes. Celle qu'on présente ci-après est basée sur l'interprétation des nombres complexes comme les points d'un plan.

B.2 Structure de corps

On sait que les nombres complexes sont en bijection avec les points du plan : à tout nombre complexe $a+ib$, on associe le point du plan de coordonnées (a, b) dans un repère orthonormé. Pourquoi ne pas définir les nombres complexes directement comme des couples de réels ?

On pose $\mathbf{C} := \mathbb{R}^2$, le couple $(a, b) \in \mathbb{R}^2$ s'interprétant comme le nombre complexe « $a+ib$ ». Ainsi, les composantes du couple (a, b) sont respectivement la partie réelle et la partie imaginaire du nombre complexe. Or, on sait que l'addition et la multiplication des nombres complexes s'expriment de la façon suivante à partir des parties réelles et imaginaires :

$$\begin{aligned}\Re(z+w) &= \Re(z) + \Re(w), & \Re(z \times w) &= \Re(z)\Re(w) - \Im(z)\Im(w), \\ \Im(z+w) &= \Im(z) + \Im(w), & \Im(z \times w) &= \Re(z)\Im(w) + \Im(z)\Re(w).\end{aligned}$$

L'idée est d'utiliser ces expressions comme définitions de l'addition et de la multiplication dans \mathbf{C} .

Définition B.2.1. Sur \mathbf{C} , on définit deux lois de composition interne $+$ et \times par :

$$\forall (a, b), (a', b') \in \mathbf{C}, \quad (a, b) + (a', b') := (a + a', b + b'), \quad (a, b) \times (a', b') := (aa' - bb', ab' + a'b).$$

Montrons que \mathbf{C} muni de ces opérations est effectivement un corps.

Lemme B.2.2. L'addition de \mathbf{C} est associative, commutative, admet pour élément neutre $\mathbf{0} := (0, 0)$ et tout élément $\mathbf{z} := (a, b) \in \mathbf{C}$ admet pour opposé l'élément $-\mathbf{z} := (-a, -b)$. ◀

Démonstration. Soient $\mathbf{z} := (a, b)$, $\mathbf{z}' := (a', b')$ et $\mathbf{z}'' := (a'', b'')$ des éléments de \mathbf{C} .

- Montrons que la multiplication est associative :

$$\begin{aligned}(\mathbf{z} + \mathbf{z}') + \mathbf{z}'' &= (a + a', b + b') + (a'', b'') \\ &= (a + a' + a'', b + b' + b'') \\ &= (a, b) + (a' + a'', b' + b'') \\ &= \mathbf{z} + (\mathbf{z}' + \mathbf{z}'').\end{aligned}$$

- Montrons que l'addition est commutative :

$$\begin{aligned}\mathbf{z} + \mathbf{z}' &= (a, b) + (a', b') \\ &= (a + a', b + b') \\ &= (a' + a, b' + b) \\ &= (a', b') + (a, b) \\ &= \mathbf{z}' + \mathbf{z}.\end{aligned}$$

- Montrons que $\mathbf{0}$ est neutre pour l'addition :

$$\mathbf{0} + \mathbf{z} = (0, 0) + (a, b) = (0 + a, 0 + b) = (a, b) = \mathbf{z}.$$

Puisque l'addition est commutative, on a $\mathbf{z} + \mathbf{0} = \mathbf{0} + \mathbf{z} = \mathbf{z}$.

- Montrons l'opposé de $\mathbf{z} := (a, b)$ est $-\mathbf{z} := (-a, -b)$:

$$\mathbf{z} + (-\mathbf{z}) = (a, b) + (-a, -b) = (a + (-a), b + (-b)) = (0, 0) = \mathbf{0}.$$

Puisque l'addition est commutative, on a $(-\mathbf{z}) + \mathbf{z} = \mathbf{z} + (-\mathbf{z}) = \mathbf{0}$. ◻

Lemme B.2.3. La multiplication de \mathbf{C} est associative, commutative, distributive par rapport à l'addition et admet pour élément neutre $\mathbf{1} := (1, 0)$. ◀

Démonstration. Soient $\mathbf{z} := (a, b)$, $\mathbf{z}' := (a', b')$ et $\mathbf{z}'' := (a'', b'')$ des éléments de \mathbf{C} .

- Montrons que la multiplication est associative. On a d'une part :

$$\begin{aligned} (\mathbf{z} \times \mathbf{z}') \times \mathbf{z}'' &= (aa' - bb', ab' + a'b) \times (a'', b'') \\ &= ((aa' - bb')a'' - (ab' + a'b)b'', (aa' - bb')b'' + a''(ab' + a'b)) \\ &= (aa'a'' - a''bb' - ab'b'' - a'b b'', aa'b'' - bb'b'' + aa''b' + a'a''b). \end{aligned}$$

D'autre part, on a :

$$\begin{aligned} \mathbf{z} \times (\mathbf{z}' \times \mathbf{z}'') &= (a, b) \times (a'a'' - b'b'', a'b'' + a''b') \\ &= (a(a'a'' - b'b'') - b(a'b'' + a''b'), a(a'b'' + a''b') + (a'a'' - b'b'')b) \\ &= (aa'a'' - ab'b'' - a'b b'' - a''bb', aa'b'' + aa''b' + a'a''b - bb'b''). \end{aligned}$$

Les calculs précédents montrent que $(\mathbf{z} \times \mathbf{z}') \times \mathbf{z}'' = \mathbf{z} \times (\mathbf{z}' \times \mathbf{z}'')$, ce qui prouve l'associativité.

- Montrons que la multiplication est commutative :

$$\begin{aligned} \mathbf{z} \times \mathbf{z}' &= (a, b) \times (a', b') \\ &= (aa' - bb', ab' + a'b) \\ &= (a'a - b'b, a'b + ab') \\ &= (a', b') \times (a, b) \\ &= \mathbf{z}' \times \mathbf{z}. \end{aligned}$$

- Montrons que la multiplication est distributive par rapport à l'addition :

$$\begin{aligned} \mathbf{z} \times (\mathbf{z}' + \mathbf{z}'') &= (a, b) \times (a' + a'', b' + b'') \\ &= (a(a' + a'') - b(b' + b''), a(b' + b'') + (a' + a'')b) \\ &= (aa' + aa'' - bb' - bb'', ab' + ab'' + a'b + a''b) \\ &= (aa' - bb', ab' + a'b) + (aa'' - bb'', ab'' + a''b) \\ &= (\mathbf{z} \times \mathbf{z}') + (\mathbf{z} \times \mathbf{z}''). \end{aligned}$$

Puisque la multiplication est commutative, on a aussi $(\mathbf{z}' + \mathbf{z}'') \times \mathbf{z} = (\mathbf{z}' \times \mathbf{z}) + (\mathbf{z}'' \times \mathbf{z})$.

- Montrons que $\mathbf{1}$ est neutre pour la multiplication :

$$\mathbf{1} \times \mathbf{z} = (1, 0) \times (a, b) = (1a - 0b, 1b + 0a) = (a, b) = \mathbf{z}.$$

Puisque la multiplication est commutative, on a $\mathbf{z} \times \mathbf{1} = \mathbf{1} \times \mathbf{z} = \mathbf{z}$. □

Proposition B.2.4. L'ensemble \mathbf{C} muni des opérations $+$ et \times est un corps. ◀

Démonstration. D'après les lemmes B.2.2 et B.2.3, $(\mathbf{C}, +, \times)$ est un anneau commutatif. Il reste à montrer que tout élément non nul de \mathbf{C} est inversible. Soient $\mathbf{z} := (a, b) \in \mathbf{C} \setminus \{\mathbf{0}\}$ et $\mathbf{z}' := (\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2})$. Montrons que \mathbf{z}' est l'inverse de \mathbf{z} :

$$\mathbf{z} \times \mathbf{z}' = \left(a \frac{a}{a^2+b^2} - b \frac{-b}{a^2+b^2}, a \frac{-b}{a^2+b^2} + \frac{a}{a^2+b^2} b \right) = \left(\frac{a^2+b^2}{a^2+b^2}, \frac{-ab+ab}{a^2+b^2} \right) = (1, 0) = \mathbf{1}.$$

Puisque la multiplication est commutative, on a $\mathbf{z}' \times \mathbf{z} = \mathbf{z} \times \mathbf{z}' = \mathbf{1}$, donc \mathbf{z}' est l'inverse de \mathbf{z} . Par conséquent, tout élément non nul de \mathbf{C} est inversible, donc \mathbf{C} est un corps. □

B.3 Sous-corps isomorphe à \mathbb{R}

On sait que les nombres réels sont les nombres complexes de partie imaginaire nulle. L'idée est donc d'identifier les nombres réels aux éléments de \mathbf{C} de la forme $(a, 0)$ avec $a \in \mathbb{R}$. Pour cela, considérons l'application $f: \mathbb{R} \rightarrow \mathbf{C}$ qui à tout réel a associe l'élément $(a, 0)$ de \mathbf{C} . En utilisant cette application, on va montrer que $f(\mathbb{R})$ est un sous-corps de \mathbf{C} isomorphe à \mathbb{R} .

Lemme B.3.1. L'application f vérifie :

- (i) $f(1) = \mathbf{1}$.
- (ii) $\forall a, b \in \mathbb{R}, f(a + b) = f(a) + f(b)$.
- (iii) $\forall a, b \in \mathbb{R}, f(a \times b) = f(a) \times f(b)$.
- (iv) f est injective. ◀

Remarque B.3.2. Les propriétés (i), (ii) et (iii) signifient que f est un morphisme de corps. ◀

Démonstration. Soient $a, b \in \mathbb{R}$.

- (i) $f(1) = (1, 0) = \mathbf{1}$.
- (ii) $f(a + b) = (a + b, 0) = (a, 0) + (b, 0) = f(a) + f(b)$.
- (iii) $f(a) \times f(b) = (a, 0) \times (b, 0) = (ab - 0, 0a + 0b) = (ab, 0) = f(a \times b)$.
- (iv) $f(a) = f(b) \iff (a, 0) = (b, 0) \iff a = b$. □

Proposition B.3.3. L'ensemble $f(\mathbb{R})$ est un sous-corps de \mathbf{C} , isomorphe à \mathbb{R} . ◀

Démonstration. Montrons que $f(\mathbb{R})$ est un sous-corps de $(\mathbf{C}, +, \times)$.

• **Affirmation :** $f(\mathbb{R})$ est stable pour l'addition et la multiplication.

Soient $\mathbf{a}, \mathbf{b} \in f(\mathbb{R})$, alors il existe $a, b \in \mathbb{R}$ tels que $f(a) = \mathbf{a}$ et $f(b) = \mathbf{b}$. Ainsi, on a $\mathbf{a} + \mathbf{b} = f(a) + f(b) = f(a + b) \in f(\mathbb{R})$ et $\mathbf{a} \times \mathbf{b} = f(a) \times f(b) = f(a \times b) \in f(\mathbb{R})$. Donc $f(\mathbb{R})$ est stable par addition et par multiplication.

• **Affirmation :** $(f(\mathbb{R}), +)$ est un groupe commutatif.

L'associativité et la commutativité de $+$ sont automatiquement vérifiées dans $f(\mathbb{R})$ car elles le sont dans \mathbf{C} et $f(\mathbb{R})$ est une partie de \mathbf{C} . Montrons que $\mathbf{0} \in f(\mathbb{R})$. Puisque f est un morphisme de corps, on a $f(0) = f(0 + 0) = f(0) + f(0)$. En retranchant $f(0)$ à chaque membre, on obtient $\mathbf{0} = f(0) \in f(\mathbb{R})$. Enfin, montrons que l'opposé d'un élément de $f(\mathbb{R})$ est un élément de $f(\mathbb{R})$. Soit $a \in \mathbb{R}$, alors en utilisant une fois de plus que f est un morphisme de corps, on a $\mathbf{0} = f(0) = f(a - a) = f(a) + f(-a)$, d'où $f(-a) = -f(a)$. Ainsi, l'opposé de $f(a)$ est $f(-a) \in f(\mathbb{R})$.

• **Affirmation :** $(f(\mathbb{R}), +, \times)$ est un anneau commutatif.

On sait déjà que $(f(\mathbb{R}), +)$ est un groupe commutatif. De plus, la multiplication est associative, multiplicative et distributive par rapport à l'addition dans $f(\mathbb{R})$, car c'est le cas dans \mathbf{C} et $f(\mathbb{R})$ est une partie de \mathbf{C} . Enfin, on a $\mathbf{1} \in f(\mathbb{R})$ car $\mathbf{1} = f(1)$.

• **Affirmation :** Tout élément non nul de $f(\mathbb{R})$ est inversible dans $f(\mathbb{R})$.

Soit $a \in \mathbb{R}$ tel que $f(a) \neq \mathbf{0}$. Puisque f est injective, on a $a \neq 0$ (le seul antécédent de $\mathbf{0}$ est 0). Par conséquent, a est inversible dans \mathbb{R} . Montrons que $f(a^{-1})$ est l'inverse de $f(a)$:

$$f(a) \times f(a^{-1}) = f(a \times a^{-1}) = f(1) = \mathbf{1}.$$

Par conséquent, l'inverse de $f(a)$ est $f(a^{-1}) \in f(\mathbb{R})$.

Bilan. L'ensemble $f(\mathbb{R})$ est un sous-corps de $(\mathbf{C}, +, \times)$. De plus, l'application f est un isomorphisme de corps entre \mathbb{R} et $f(\mathbb{R})$, car f est un morphisme de corps injectif d'après le lemme B.3.1 et f est surjective sur son image $f(\mathbb{R})$, par définition. □

Remarque B.3.4. On vient de démontrer dans un cas particulier le résultat général suivant : si K et L sont des corps et si $f: K \rightarrow L$ est un morphisme de corps, alors $f(K)$ est un sous-corps de L . ◀

Dans la suite, on note \mathbf{R} l'ensemble $f(\mathbb{R}) := \{(a, 0) : a \in \mathbb{R}\}$. L'ensemble \mathbf{R} est l'ensemble des éléments de \mathbf{C} dont la partie imaginaire est nulle, ce qu'on interprète comme des nombres réels.

B.4 Unité imaginaire et forme algébrique

Montrons que \mathbf{C} contient un élément jouant le rôle de l'unité imaginaire. Étant donné notre définition de \mathbf{C} , les couples $(0, 1)$ et $(0, -1)$ devraient convenir.

Proposition B.4.1. L'élément $\mathbf{i} := (0, 1) \in \mathbf{C}$ vérifie l'égalité $\mathbf{i}^2 = -\mathbf{1}$. ◀

Remarque B.4.2. On aurait très bien pu choisir le couple $(0, -1)$ comme définition de \mathbf{i} . ◀

Démonstration. C'est une simple vérification :

$$\begin{aligned}\mathbf{i}^2 &= (0, 1) \times (0, 1) \\ &= ((0 \times 0 - 1 \times 1), (0 \times 1 + 1 \times 0)) \\ &= (-1, 0) \\ &= -\mathbf{1}.\end{aligned}$$

□

Avant d'énoncer le résultat d'existence et d'unicité de la forme algébrique des éléments de \mathbf{C} , il faut voir comment agissent les éléments de \mathbf{R} par multiplication. On s'attend à ce que multiplier par un réel λ résulte en la multiplication de la partie réelle et de la partie imaginaire par λ , et c'est effectivement le cas.

Lemme B.4.3. Pour tout $(\lambda, 0) \in \mathbf{R}$ et pour tout $(a, b) \in \mathbf{C}$, on a $(\lambda, 0) \times (a, b) = (\lambda a, \lambda b)$. ◀

Démonstration. Simple vérification : $(\lambda, 0) \times (a, b) = (\lambda a - 0b, \lambda b + 0a) = (\lambda a, \lambda b)$. □

Pour finir, montrons l'existence de la forme algébrique des éléments de \mathbf{C} .

Proposition B.4.4. Tout élément de $\mathbf{z} \in \mathbf{C}$ s'écrit $\mathbf{z} = \mathbf{a} + \mathbf{i} \times \mathbf{b}$ avec $\mathbf{a}, \mathbf{b} \in \mathbf{R}$. ◀

Démonstration. Soit $\mathbf{z} := (a, b) \in \mathbf{C}$ et soient $\mathbf{a} := (a, 0) \in \mathbf{R}$ et $\mathbf{b} := (b, 0) \in \mathbf{R}$. Alors \mathbf{z} s'écrit :

$$\begin{aligned}\mathbf{z} &= (a, b) \\ &= (a, 0) + (0, b) \\ &= (a, 0) + (0, 1) \times (b, 0) && \text{(lemme B.4.3)} \\ &= \mathbf{a} + \mathbf{i} \times \mathbf{b}.\end{aligned}$$

□

En combinant les propositions B.2.4, B.3.3, B.4.1 et B.4.4, on a démontré que l'ensemble \mathbf{C} est un corps répondant à la définition de \mathbb{C} .

Théorème B.4.5. L'ensemble \mathbf{C} muni des opérations $+$ et \times (cf. définition B.2.1) est un corps qui vérifie les propriétés suivantes.

- (i) \mathbf{C} contient un sous-corps \mathbf{R} , isomorphe à \mathbb{R} .
- (ii) Il existe un élément $\mathbf{i} \in \mathbf{C}$ tel que $\mathbf{i}^2 = -\mathbf{1}$.
- (iii) Tout élément de \mathbf{C} s'écrit sous la forme $\mathbf{a} + \mathbf{i} \times \mathbf{b}$ avec $\mathbf{a}, \mathbf{b} \in \mathbf{R}$.

Index

addition vectoriel, 48
affixe, 10
algorithme du pivot de Gauss, 37
anneau, 85
antécédent, 83
application, 83
 bijective, 83
 injective, 83
 linéaire, 68
 réciproque, 83
 surjective, 83
argument, 23
 principal, 23
associativité, 84

base, 57
 canonique, 58

cas irréductible, 7
cercle trigonométrique, 19
classe d'équivalence, 82
combinaison linéaire, 52
commutativité, 84
composition, 83
congruence, 19
conjugué, 13
coordonnées
 dans une base, 57
 polaires, 23
corps, 9, 85

dimension, 60
droite vectorielle, 53, 61

élément neutre, 84
endomorphisme, 69
ensemble, 79
 quotient, 82
équation cartésienne
 d'un plan, 45
équations de compatibilité, 35
équation linéaire, 71
 homogène, 71
équipollence, 42
espace vectoriel, 48

exponentielle
 d'un nombre complexe, 24
 d'un nombre imaginaire pur, 21

factorisation de l'angle moitié, 27
famille
 génératrice, 54
 libre, 55
 liée, 54
forme algébrique, 9
forme linéaire, 69
forme trigonométrique, 23
formule(s)
 d'Euler, 25
 de duplication, 26
 de factorisation, 28
 de Grassmann, 65
 de Moivre, 26
 du binôme de Newton, 12

groupe, 85

hyperplan vectoriel, 61

image, 83
 d'une application linéaire, 71
 directe, 83
 réciproque, 83
inconnues
 principales, 36
 secondaires, 36
inégalité triangulaire, 15
inverse, 12
isomorphisme, 69

linéarisation, 25
loi de composition interne, 84

méthode de Cardan, 6
module, 13
multiplication par un scalaire, 48

nombre
 complexe, 8, 9
 imaginaire, 8
 imaginaire pur, 9

- noyau d'une application linéaire, 71
- opérations élémentaires, 32
- partie imaginaire, 9
- partie réelle, 9
- pivots, 35
- plan complexe, 10
- plan vectoriel, 53, 61
- point image, 10
- produit cartésien, 81
- racine
 - carrée, 16
 - de l'unité, 28
 - double, 17, 18
 - n -ième, 29
- radical, 9
- rang
 - d'un système linéaire, 36
 - d'une application linéaire, 74
- règle du parallélogramme, 42
- relation
 - binaire, 81
 - d'équivalence, 82
- de liaison, 55
- relation de Chasles, 42
- représentant, 82
- représentation paramétrique
 - d'un plan, 45
 - d'une droite, 44
- scalaire, 47
- segment orienté, 41
- somme
 - de s.e.v., 63
 - directe, 64
- sous-espace vectoriel, 50
 - engendré, 52
- supplémentaire, 65
- symétrique, 84
- système linéaire, 31
 - échelonné, 35
 - homogène, 31
 - incompatible, 35
 - triangulaire, 34
- théorème
 - de d'Alembert–Gauss, 30
 - de la base extraite, 60
 - de la base incomplète, 60
 - de la dimension, 60
 - du rang, 75
- unité imaginaire, 9
- variables
 - libres, *voir* inconnues secondaires
 - liées, *voir* inconnues principales
- vecteur(s), 48
 - linéairement indépendants, 55
 - nul, 42, 48
 - opposé, 48