Cours de mathématiques
Partie III – Algèbre
MPSI 4

Alain TROESCH

Version du:

18 juin 2014

Table des matières

1 Équations

Introduction

Note Historique 1.0.1

- Le développement des mathématiques a souvent été motivé par la nécessité ou la volonté de pouvoir procéder à certains calculs, liés souvent à des problèmes concrets. Les premières traces certifiées de calculs sont localisées en Mésopotamie ou en Egypte, lié à des calculs de surfaces de champs ou de quantité de matériel nécessaire.
- L'outil a souvent été développé avant la théorie, par tâtonnement, de façon parfois un peu empirique (tables pour les calculs de fraction en Égypte antique, table des sinus en Inde vers 500, résolutions par approximations...)

Nous adoptons la même démarche, en présentant dans ce chapitre un certain nombre d'outils, en éludant dans un premier temps les fondements théoriques.

I Equations, ou les origines de l'algèbre

Note Historique 1.1.1

La recherche de solutions d'équation n'est pas un problème récent :

- À Babylone et en Égypte (2e millénaire avant J.-C.), on trouve déjà trace de résolutions de problèmes se ramenant à des équations de degré 2. Ces équations sur des exemples, mais la méthode exposée sur ces exemples est déjà celle utilisée actuellement.
- Vers 300 après J.-C., Diophante formalisation la notion d'équations. Il s'intéresse en particulier à la recherche de solutions rationnelles d'équations à coefficients rationnels, ce qu'on appelle actuellement des *équations diophantiennes*.
- Vers 800 après J.-C., le mathématicien arabe Al Khwarizmi écrit le premier traité de résolution systématique des équations de degré 2. Le titre de ce traité est *kitabu al-mukhtasar fi hisabi al-jabr wa'l-muqabalah*, soit, à peu près : *Abrégé du calcul par la réduction et la comparaison*
  - Le terme arabe « Al-jabr » signifie « par réduction » (ie « en se ramenant à des situations-type par manipulation des termes »). Il a donné naissance au terme « algèbre ».
  - Le nom même d'Al Khwarizmi (qui propose une méthode systématique de résolution) a été occidentalisé en « algorisme » puis en « algorithme »
  - L'importance du traité de Al Khwarizmi tient moins du contenu lui même (il ne propose rien de fondamentalement neuf), que d'une part à la synthèse qu'il propose, par son origine géographique, entre les mathématiques grecques classiques et les mathématiques indiennes calculatoires et pragmatiques, dans la continuité des mathématiques Babyloniennes, et d'autre part à sa large diffusion en Europe jusque vers les années 1200.
  - C'est également grâce à Al Khwarizmi que les chiffres arabes (en fait dérivés du système de numération indienne) arrivent en occident.

I.1 Equations linéaires

La résolution de l'équation linéaire $ax = b$ ne pose bien sûr de problème à personne dans $\mathbb{R}$ ou $\mathbb{C}$. Cependant :

- si a dépend d'un paramètre, n'oubliez pas la discussion sur le paramètre pour avoir la condition $a \neq 0$xs.
- La bonne condition portant sur a n'est pas tant « $a \neq 0$ » que « a inversible ». Plus précisément, le fait que a puisse se simplifier (si b s'écrit sous la forme ac) définit la notion d'« élément régulier » pour la loi de composition étudiée (voir chapitre suivant). On a alors :
  - Le fait que $a \neq 0$ est une condition nécessaire pour que a soit régulier, mais pas suffisante : par exemple, si A et B sont des matrices, et $A \neq 0$, on peut avoir $AB = 0$ sans que $B = 0$ (songez aux résolutions de systèmes d'équations homogènes, B étant dans ce cas une matrice colonne constituée des inconnues ; si A est non nul, mais n'est pas inversible, il existe une solution B non nulle de l'équation...)
  - Le fait que a soit inversible est une condition suffisante pour que a soit régulier, mais pas nécessaire. Par exemple, dans $\mathbb{Z}$, les éléments inversibles sont 1 et -1, les éléments réguliers sont tous les éléments non nuls. L'argument précédent justifie que dans $\mathcal{M}_n$, les éléments réguliers sont exactement les matrices inversibles (pour ramener l'argument précédent au cas de matrices carrées, contruisez la matrice carrée B comme juxtaposition de colonnes toutes égales à une solution non nulle du système $AX = 0$).

  Ainsi, la régularité vient se placer entre la non nullité et l'inversibilité, pouvant être distinct des deux notions simultanément (cas de la régularité dans $\mathbb{R}[X]$ par exemple)

  Évidemment, dans $\mathbb{R}$ ou $\mathbb{C}$, ou dans d'autres cas où l'inversibilité équivaut à la non nullité, les trois conditions sont confondues ; mais il convient de bien les distinguer dans le cas général.

Ainsi, dans le cas général on peut énoncer :

Proposition 1.1.2 (Résolution d'une équation de degré 1)

Soit a et b deux nombres réels ou complexes, ou plus généralement des éléments d'un « anneau » (voir un chapitre ultérieur). Alors,

- *Si a est régulier, et qu'il existe c tel que $b = ac$, alors l'équation $ax = b$ admet une unique solution $x = c$.*
- *En particulier, si a est inversible, l'équation $ax = b$ admet une unique solution $x = a^{-1}b$.*

Si nous ne sommes pas dans une de ces situations, la recherche des solutions est plus délicate, et peut nécessiter des techniques propres à la situation considérée (par exemple la résolution d'un système d'équations, s'écrivant matriciellement $AX = B$).

I.2 Equations de degré 2

Note Historique 1.1.3

Les babyloniens connaissaient déjà la méthode de résolution des équations de degré 2 quasiment deux millénaires avant J.-C. Pour preuve le texte suivant, exposant une méthode de calcul dans une situation précise. Analysez ce texte afin de vous rendre compte qu'il s'agit très exactement de la méthode utilisée encore maintenant. Attention, les calculs sont exprimés en base 60 (pensez heures/minutes/secondes), et il n'est pas toujours très bien précisé si on parle des unités ou des soixantièmes.

> *« J'ai additionné la surface et le côté de mon carré : 45.*
> *Tu poseras 1, la wasitum (=unité). Tu fractionneras la moitié de 1 (: 30). Tu multiplieras 30 et 30 (: 15). Tu ajouteras 15 à 45 : 1. 1 en est la racine carrée. Tu soustrairas le 30, que tu as multiplié, de 1 (: 30). 30 est le côté du carré. »*

La résolution de l'équation de degré 2 passe par un procédé dont la présence dans de nombreuses méthodes montre l'importance en soi :

Proposition 1.1.4 (Mise sous forme canonique)

Toute expression de degré 2 à coefficients dans $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ peut être mise sous la forme $\alpha((x+\beta)^2 - \delta)$, où $(\alpha, \beta, \delta) \in \mathbb{C}$.
Plus précisément, étant donné a, b et c réels ou complexes, avec $a \neq 0$,

$$ax^2 + bx + c = a\left[\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a^2}\right]$$

Définition 1.1.5 (discriminant)

La quantité $b^2 - 4ac$ est appelée discriminant du trinôme $ax^2 + bx + c$, et est souvent noté Δ.

Corollaire 1.1.6 (résolution des équations de degré 2)

Soit $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, et $a, b, c \in \mathbb{K}$. Notons $\mathcal{R}_{\mathbb{K}}(\Delta)$ l'ensemble des racines carrées de $\Delta = b^2 - 4ac$ dans $\mathbb{K}$. Alors l'ensemble des solutions de l'équation $ax^2 + bx + c = 0$ dans $\mathbb{K}$ est :

$$\mathcal{S}_{\mathbb{K}} = \left\{\frac{-b + r}{2a}, r \in \mathcal{R}_{\mathbb{K}}(\Delta)\right\}.$$

On explicite le cas des équations à coefficients réels.

Proposition 1.1.7 (Résolution des équations de degré 2 à coefficients réels)

Soit a, b, c trois réels tels que $a \neq 0$, et soit $\Delta = b^2 - 4ac$. Alors :

- *si $\Delta > 0$, l'équation $ax^2 + bx + c = 0$ admet exactement 2 racines réelles :*

$$x_1 = \frac{-b - \sqrt{\Delta}}{2a} \qquad et \qquad x_2 = \frac{-b + \sqrt{\Delta}}{2a};$$

- *si $\Delta = 0$, l'équation $ax^2 + bx + c = 0$, admet une unique racine (double) :*

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

- *si $\Delta < 0$, l'équation $ax^2 + bx + c = 0$ n'admet pas de racine réelle. Elle admet deux racines complexes conjuguées :*

$$x_1 = \frac{-b - \mathrm{i}\sqrt{-\Delta}}{2a} \qquad et \qquad x_2 = \frac{-b + \mathrm{i}\sqrt{-\Delta}}{2a}.$$

On remarquera au passage que puisque tout nombre complexe différent de 0 a exactement deux racines complexes (opposées l'une de l'autre), toute équation du second degré dans $\mathbb{C}$ admet exactement deux racines distinctes si $\Delta \neq 0$ et une seule si $\Delta = 0$. Ces racines peuvent être déterminées par les méthode de recherche des racines carrées étudiées dans le cours sur les nombres complexes.
Au passage, on a obtenu :

Proposition 1.1.8 (Factorisation d'un trinôme)

Soit x_1 et x_2 les deux racines de l'équation $ax^2 + bx + c = 0$ (avec éventuellement $x_1 = x_2$ si le discriminant est nul). Alors :

$$\forall x \in \mathbb{R}, \quad ax^2 + bx + c = a(x - x_1)(x - x_2).$$

Les racines x_1 et x_2 ne sont en général pas conjuguées. Cette propriété est d'ailleurs une caractérisation des équations à coefficients réel, comme corollaire de la proposition suivante.

Proposition 1.1.9 (Somme et produit des racines)

Soit a, b, c des complexes, et x_1 et x_2 les deux racines (complexes) de $ax^2 + bx + c = 0$ (en posant $x_1 = x_2$ si $\Delta = 0$). Alors :

$$x_1 + x_2 = -\frac{b}{a} \qquad \text{et} \qquad x_1 x_2 = \frac{c}{a}.$$

Réciproquement, si $x_1 + x_2 = \beta$ et $x_1 x_2 = \gamma$, alors x_1 et x_2 sont les deux solutions de l'équation $x^2 - \beta x + \gamma = 0$.

Corollaire 1.1.10 (caractérisation des équations à coefficients réels)

Soit b et c deux complexes. Alors b et c sont tous les deux réels si et seulement si les racines complexes x_1 et x_2 (éventuellement égales si $\Delta = 0$) de l'équation $x^2 + bx + c = 0$ sont soit réelles, soit conjuguées l'une de l'autre.

Remarquez que si $a \neq 0$, l'équation général $ax^2+bx+c = 0$ se ramène toujours à une équation $x^2+b'x+c' = 0$ en divisant par a. Ainsi, si le coefficient a n'est pas égal à 1, on ne peut pas nécessairement conclure que b et c sont réels : on obtient en fait que b et c sont sur la droite réelle de $\mathbb{C}$ engendrée par a, donc peuvent s'écrire sous la forme $b = \beta a$ et $c = \gamma a$, pour des réels β et γ.
La propriété 1.1.9 (qu'on généralisera au cas de sommes et produits des racines de polynômes de degré plus important) fournit une technique de recherche de racines « évidentes ».

Exemples 1.1.11

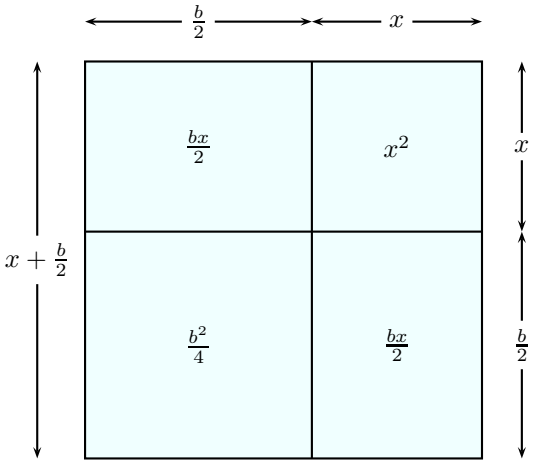
1. (i) Sans calculer le discriminant, étudier l'existence de solutions entières de l'équation $x^2 + 3x - 2 = 0$.
   (ii) De même pour $x^2 + 3x + 2 = 0$.
   (iii) De même pour $x^2 - 24x + 143$.
2. Trouver tous les réels x et y vérifiant $x < y$, $xy = 1$ et $x + y = 3$

Pour terminer ce paragraphe, voici en figure 1.1, à titre de curiosité, une interprétation géométrique de la résolution de l'équation $x^2 + bx = c$ dans le cas où b et c sont positifs.
Cette résolution est extraite de l'ouvrage *Kitabu al-mukhtasar fi hisabi al-jabr wa al-muqabalah* de Al-Khwarizmi, mais était déjà connue des Grecs.

Note Historique 1.1.12

La résolution d'équations de degré plus important a été à l'origine de nombreux développements mathématiques. On peut retracer rapidement l'histoire de ces résolutions :

- En 1545, Jérôme Cardan publie sous son nom une méthode de résolution des équations de degré 3 (formules de Cardan). Il a en fait reçu ces formules du mathématicien italien Tartaglia, sous la promesse (non tenue) du secret. Cette méthode est à la base de l'introduction des nombres complexes.
- En 1545, le mathématicien italien Ferrari propose une méthode de résolution des équations de degré 4, par réduction à une équation de degré 3. Cette méthode est généralisée par Bombelli en 1572.
- Au cours du 18^{e} siècle, Joseph-Louis Lagrange tente de résoudre des équations de degré 5 en introduisant des expressions intermédiaires (les résolvantes de Lagrange), qui dans le cas des équations de degré 3 et 4, font baisser le degré. Mais pour les équations de degré supérieur ou égal à 5, ces résolvantes font au contraire augmenter le degré.
- Paolo Ruffini est le premier en 1799 à affirmer l'impossibilité de résoudre une équation générale de degré 5 par radicaux, mais sa preuve n'est pas correcte.



Aire du grand carré :
$\left(x + \frac{b}{2}\right)^2 = x^2 + \frac{bx}{2} + \frac{bx}{2} + \frac{b^2}{4} = c + \frac{b^2}{4}$

FIGURE 1.1 – Interprétation géométrique de la résolution d'une équation de degré 2

- Le mathématicien norvégien Niels Henrik Abel (mort à 26 ans d'épuisement et de misère) est le premier à donner une preuve rigoureuse de ce théorème en 1824, depuis appelé théorème de Ruffini-Abel (énoncé ci-dessous). Son travail n'est pas reconnu de son vivant, et il n'obtient pas le poste qu'il attendait en Allemagne, et qui aurait pu le sauver de la misère. La communauté scientifique se rend compte de l'importance de son oeuvre mathématique quelques mois après sa mort.
- C'est cependant Évariste Galois (mort quant à lui à 20 ans lors d'un duel) qui propose en 1831 l'approche la plus intéressante et la plus générale, et donne des conditions nécessaires et suffisantes pour que des équations particulières soient résolubles par radicaux. C'est ce travail visionnaire qui est à l'origine de la théorie des groupes et de la théorie des extensions de corps.

Nous utilisons une formulation proche de celle d'Abel :

Théorème 1.1.13 (Ruffini-Abel, 1824)

Il n'existe pas de formule générale exprimant les solutions de l'équation du cinquième degré sous forme de radicaux.

II Inéquations (dans $\mathbb{R}$ uniquement, pas dans $\mathbb{C}$!)

Nous rappelons dans ce paragraphe les règles élémentaires de manipulation des inégalités dans $\mathbb{R}$, et la résolution de certains types classiques d'inéquations (notamment les inéquations de degré 1 et 2).

Proposition 1.2.1 (Opérations sur des inégalités)

Soit a, b, c, d des réels.

- *Si $a \leqslant b$ et $c \leqslant d$ alors $a + c \leqslant b + d$.*
- *Si $a \leqslant b$ alors :*
  - *si $c > 0$, alors $ac \leqslant bc$*
  - *si $c < 0$, alors $ac \geqslant bc$.*
- * *Si $0 \leqslant a \leqslant b$ et $0 \leqslant c \leqslant d$ alors $0 \leqslant ac \leqslant bd$.*
  - * *Pour les autres cas de multiplication : étudier séparément les valeurs absolues et le signe.*
- * *Si $0 < a \leqslant b$, alors $0 < \frac{1}{b} \leqslant \frac{1}{a}$*
  - * *Si $a \leqslant b < 0$, alors $\frac{1}{b} \leqslant \frac{1}{a} < 0$.*
  - * *Si $a < 0 < b$, alors $\frac{1}{a} < 0 < \frac{1}{b}$.*

Avertissement 1.2.2

1. Attention à ne pas soustraire ou à ne pas diviser terme à terme une inégalité :
   - pour la soustraction, multiplier la seconde inégalité par -1 à l'aide de la règle 2 de la proposition ;
   - pour la division, utiliser la règle 4.
2. Ne pas oublier de changer le sens de l'inégalité lorsque vous multipliez ou divisez une inégalité par un scalaire négatif. C'est le cas par exemple pour la résolution de l'inéquation toute simple $ax \leqslant b$ lorsque $a < 0$.

Nous rappelons le résultat suivant :

Proposition 1.2.3 (Résolution des inéquations de degré 2)

Soit $a \neq 0$ et $b, c \in \mathbb{R}$. Alors :

- *si l'équation $ax^2 + bx + c = 0$ n'admet pas de solution réelle, l'ensemble des solutions de l'inéquation $ax^2 + bx + c \leqslant 0$ est :*
  - *l'ensemble vide $\varnothing$ si $a > 0$*
  - *l'ensemble $\mathbb{R}$ entier si $a < 0$*
- *si l'équation $ax^2 + bx + c = 0$ admet deux racines $x_1 \leqslant x_2$ (éventuellement égales), alors l'ensemble des solutions de l'inéquation $ax^2 + bx + c \leqslant 0$ est :*
  - *$[x_1, x_2]$ si $a > 0$*
  - *$]-\infty, x_1] \cup [x_2, +\infty[$ si $a < 0$.*

On résume souvent ce théorème en disant qu'une fonction polynomiale de degré 2 est du signe de son coefficient dominant, sauf entre ses racines (cette condition étant caduque s'il n'y a pas de racine)

Exemple 1.2.4

Résolution de $|2x^2 - 5x + 1| \leqslant 2$.

Dans des situations plus générales, ne pas oublier d'exploiter la monotonie de certaines fonctions.

III Systèmes linéaires de n équations à p inconnues

Voici une petite mise au point sur la résolution systématique des systèmes de n équations linéaires à p inconnues. On peut toujours s'en sortir par la méthode de substitution que vous connaissez généralement bien. Mais pour des gros systèmes, c'est assez lourd à mettre en place. La méthode du pivot de Gauss est à privilégier (sauf dans certaines situations trop symétriques, la méthode du pivot ne permettant pas d'exploiter correctement les symétries du système).

Note Historique 1.3.1 (Algorithme du pivot de Gauss)

Cet algorithme est très ancien, on en trouve des traces au 1er siècle avant JC, en Chine (Chang Ts'ang, chancelier de l'empéreur). Gauss et Jordan abordent au 19e siècle cet algorithme sous un point de vue très différent, en relation non avec la résolution des systèmes linéaires, mais avec la classification des formes quadratiques. Ce n'est que vers 1880 que Frobenius publie plusieurs mémoires faisant un état des lieux de la théorie des matrices, et élucide complètement à l'occasion la théorie des systèmes linéaires à coefficients réels ou complexes.

Nous introduisons ici la méthode du pivot, sans entrer dans les discussions théoriques de la validité générale de cette méthode (nous aurons l'occasion d'aborder ce problème plus tard, mais nous aurons besoin pour cela de plus d'outils matriciels).

III.1 Introduction de la méthode du pivot de Gauss

Globalement, la méthode du pivot consiste à faire des opérations élémentaires sur les lignes du système (échange, multiplication par un scalaire, combinaison linéaire de deux lignes) de sorte à se ramener à un système échelonné (c'est-à-dire dans lequel l'inconnue de plus petit indice d'une équation donnée n'apparaisse plus dans les équations suivantes) : un tel système se résout facilement en partant du bas.
Soit

$$\begin{cases} a_{1,1}x_1 + \cdots + a_{1,p}x_p &= b_1 \\ a_{2,1}x_1 + \cdots + a_{2,p}x_p &= b_2 \\ \cdots & \\ a_{n,1}x_1 + \cdots + a_{n,p}x_p &= b_n \end{cases}$$

un système (dans $\mathbb{R}$ ou $\mathbb{C}$). Ce système peut être traduit par une égalité matricielle $AX = B$, où

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix}, \qquad X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \qquad \text{et} \qquad B = \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix}.$$

Nous adopterons le point de vue matriciel, et même, nous nous contenterons de décrire la méthode de façon purement algorithmique, en décrivant les actions à effectuer sur A et B (ce qui ne modifie pas la matrice X, nous dispensant ainsi de la faire intervenir).
L'intérêt de cette présentation matricielle est d'une part la rapidité apportée par le fait qu'on se dispense de réécrire les variables, et d'autre part une présentation plus claire, du fait de l'alignement obligé des coefficients dans la matrice. Cela supprime une source importante d'erreurs d'inattention.
Dans ce qui suit, A désigne une matrice à n lignes et p colonnes, B une matrice colonne à n lignes, et X la matrice colonne des inconnues, à p lignes.

Définition 1.3.2 (Système homogène associé)

Étant donné le système donné matriciellement par $AX = B$, on appelle système homogène associé au système $AX = B$ le système $AX = 0$.

Définition 1.3.3 (Une solution)

Une solution du système $AX = B$ est un p-uplet $(x_1, \ldots, x_p) \in \mathbb{R}^n$ tel que la matrice colonne X associée vérifie $AX = B$. Ainsi, l'ensemble des solutions de $AX = B$ est un sous-ensemble de $\mathbb{R}^p$.

Théorème 1.3.4 (Ensemble des solutions)

Soit $\mathcal{S}_0$ l'ensemble des solutions de l'équation homogène $AX = 0$, et $X_0 \in \mathbb{R}^n$ une solution particulière de l'équation $AX = B$, s'il en existe une. Alors l'ensemble de toutes les solutions de l'équation $AX = B$ est :

$$\mathcal{S} = \{X_0 + X \mid X \in \mathcal{S}_0\} = X_0 + \mathcal{S}_0.$$

Proposition/Définition 1.3.5 (Structure de l'ensemble des solutions)

1. L'ensemble $\mathcal{S}_0$ des solutions de l'équation homogène est stable par combinaison linéaire et contient le vecteur nul : c'est un *sous-espace vectoriel* de $\mathbb{R}^p$
2. L'ensemble $\mathcal{S}$, s'il est non vide, est donc le translaté d'un sous-espace vectoriel de $\mathbb{R}^p$: un tel sous-ensemble de $\mathbb{R}^p$ est appelé *sous-espace affine* de $\mathbb{R}^p$.

III.2 Recherche d'une solution particulière par la méthode du pivot

Étant donné une matrice dont les lignes sont désignées par $L_1, \dots, L_n$, les opérations admissibles pour le pivot sont les trois opérations suivantes :

- : l'échange des lignes L_i et L_j de la matrice, codée par $L_i \leftrightarrow L_j$
- : la multiplication d'une ligne L_i par un scalaire (réel ou complexe) *non nul* λ, codée par $L_i \leftarrow \lambda L_i$
- : l'ajout à une ligne donnée L_i d'une autre ligne L_j éventuellement multipliée par un scalaire λ (le résultat remplaçant la ligne L_i). cette opération est codée par $L_i \leftarrow L_i + \lambda L_j$.

La combinaison des deux dernières règles amène la règle suivante (avec les notations évidente), souvent bien pratique :

$$L_i \leftarrow \lambda L_i + \mu L_j, \text{ si } \lambda \neq 0.$$

Avertissement 1.3.6

Les différentes opérations s'effectue successivement (même si on les note ensemble dans la même étape) : on ne peut pas effectuer des opérations simultanées. Ainsi, si on a dans la même étape deux opérations $L_1 \leftarrow L_1 + L_2$ et $L_2 \leftarrow L_1 + L_2$, cela signifie que la seconde est effectuée avec la ligne L_1 obtenue à l'issue de la première opération, et non avec la ligne L_1 initiale.

Pour la recherche d'une solution particulière, on effectue des opérations sur les lignes de la matrice A, en effectuant les mêmes simultanément sur les lignes de la matrice colonne B. Pour ne pas s'embrouiller dans ces opérations, on adopte souvent la présentation suivante, permettant d'aligner les lignes de A et celles de B :

$$\left(\begin{array}{ccc|c} a_{1,1} & \cdots & a_{1,p} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{n,1} & \cdots & a_{n,p} & b_n \end{array}\right)$$

Ainsi, effectuer des opérations simultanément sur les lignes de A et de B revient à effectuer des opérations sur cette matrice obtenue par juxtaposition de A et B.

Méthode 1.3.7 (Algorithme du pivot de Gauss, ou méthode de Jordan Gauss)

1. On cherche la première colonne non nulle de la matrice A.
2. Sur cette colonne, on effectue un choix de pivot : n'importe quel coefficient non nul de la colonne convient, mais on a intérêt à choisir un pivot donnant le moins de calculs possible. Il y a trois critères pour cela :
   - Le pivot lui-même doit être facile à inverser. L'idéal est un pivot égal à 1.
   - Les autres coefficients de la ligne du pivot doivent être « simples », de préférence des entiers.
   - Plus il y a de zéros sur la ligne contenant le pivot, moins il y aura de calculs !
3. On fait un échange de lignes pour ramener le pivot choisi sur la première ligne.
4. On annule tous les coefficients situées sous le pivot à l'aide d'opérations élémentaires $L_i \leftarrow L_i + \lambda L_1$, ou bien pour éviter d'introduire des fractions, $L_i \leftarrow \alpha L_i + \beta L_1$, avec $\alpha \neq 0$
5. On recommence récursivement en considérant la sous-matrice obtenue en supprimant la première ligne et la colonne du pivot ainsi que celles qui précèdent : ces lignes et colonnes supprimées ne seront plus modifiées jusqu'à la fin de l'algorithme)

À l'issue de cette méthode, il reste à la place de A une matrice échelonnée :

Définition 1.3.8 (Matrice échelonnée)

Soit m et n deux entiers non nuls, et $M = (a_{i,j})_{\substack{1\leqslant i\leqslant m\\ 1\leqslant j\leqslant n}}$ dans $\mathcal{M}_{m,n}(\mathbb{K})$. On dit que M est une matrice échelonnée s'il existe un entier $k \in [\![1, m]\!]$ et une suite croissante $j_1 < j_2 < \cdots < j_k$ d'éléments de $[\![1, n]\!]$ tels que :

(i) $\forall i \in [\![1, k]\!]$, $a_{i,j_i} \neq 0$;

(ii) $\forall i \in [\![1, k]\!]$, $\forall j \in [\![1, j_i - 1]\!]$, $a_{i,j} = 0$;

(iii) $\forall i \in [\![k + 1, m]\!]$, $\forall j \in [\![1, n]\!]$, $a_{i,j} = 0$

Autrement dit, les lignes nulles sont regroupées au bas de la matrice (lignes $k + 1$ à m), les autres lignes sont classées suivant la position de leur premier élément non nul, ces positions étant deux à deux distinctes. Une matrice échelonnée admet donc la représentation suivante :

$$M = \begin{pmatrix} 0 & \cdots & 0 & a_{1,j_1} & \bullet & & \cdots & & \bullet \\ 0 & \cdots & \cdots & 0 & a_{2,j_2} & \bullet & & \cdots & \bullet \\ \vdots & & & & & & & & \vdots \\ 0 & & \cdots & \cdots & & 0 & a_{k,j_k} & \bullet & \cdots & \bullet \\ 0 & & & \cdots & & & \cdots & & & 0 \\ \vdots & & & & & & & & & \vdots \\ 0 & & & \cdots & & & \cdots & & & 0 \end{pmatrix},$$

les coefficients indiqués d'un $\bullet$ étant quelconques.

Définition 1.3.9 (réduite de Gauss)

La matrice échelonnée A' obtenue à l'aide de la méthode du pivot appliqué à la matrice A s'appelle *réduite de Gauss de la matrice A*. Il n'y a pas unicité d'une résuite de Gauss.

Le plus simple pour bien comprendre la méthode est de voir un exemple explicite.

Exemple 1.3.10

Recherche d'une réduite de Gauss A', et de la matrice B' associée lorsque :

$$A = \begin{pmatrix} 1 & -4 & -2 & 3 & 2 \\ 2 & 2 & 1 & 0 & 1 \\ -1 & 0 & 0 & 1 & 0 \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Théorème 1.3.11 (Conservation de l'équation, admis provisoirement)

Soit $AX = B$ un système d'équations linéaires, et A' et B' les matrices obtenues par la méthode du pivot de Gauss appliqué à A en opérant simultanément sur B. Alors les solutions de l'équation $AX = B$ sont les mêmes que les solutions de l'équation $A'X = B'$.

On est donc ramené à l'étude d'équations du type $AX = B$ lorsque A est échelonnée. Une solution particulière est facile à déterminer en partant de la dernière équation et en remontant petit à petit.

Méthode 1.3.12 (recherche d'une solution particulière d'un système échelonné)

Soit A une matrice échelonnée, et $AX = B$ un système.

1. S'il existe dans ce système une ligne du type $0 = b_i$, avec b_i non nul, alors le système n'admet pas de solution.

2. Sinon, dans chaque équation non triviale, on isole la variable de plus petit indice (on revient à une notation non matricielle).
3. En prenant les équations en sens inverse, cela permet d'exprimer chacune de ces variables (non encore déterminées à ce moment, du fait de l'ordre de lecture des équations) en fonction des constantes b_i et de certaines variables ; les variables auxquelles on n'a pas encore attribué de valeur lors d'une étape précédente sont posées quelconques (par exemple nulles), et cela détermine alors la variable qu'on a isolé.

Exemple 1.3.13

Recherche d'une solution particulière dans le cas de l'exemple 1.3.10.

III.3 Recherche de la solution générale de l'équation homogène associée

À ce stade, l'essentiel du travail a été fait : on réutilise les résultats des calculs effectués dans la phase précédente. La recherche des solutions de l'équation homogène associée $AX = 0$ équivaut à la recherche des solutions du système homogène échelonné $A'X = 0$.

Méthode 1.3.14 (Ensemble des solutions d'un système homogène échelonné)

- On reprend la démarche précédente, en isolant dans chaque ligne du système échelonné $A'X = 0$ la variable de plus petit indice, en partant de la fin.
- à chaque étape, les variables déjà exprimées sont remplacées par leur expression en fonction des paramètres, les variables non encore exprimées sont prises en paramètres.
- À l'issue de ce calcul, on obtient la solution générale sous la forme $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$, vecteur dans lequel on peut remplacer certaines coordonnées (celles qui correspondent aux pivots utilisés) par leur paramétrage en fonction des autres.
- Ainsi, en notant $x_{i_1}, \ldots, x_{i_k}$ les inconnues prises comme paramètre, l'expression générale de la solution étant linéaire en ces variables, on peut exprimer la solution générale sous la forme suivante :

$$X = x_{i_1} e_1 + \cdots + x_{i_k} e_k,$$

où les e_i sont des vecteurs colonnes à p lignes. On dit alors que $(e_1, \ldots, e_n)$ est une base de l'espace des solutions (selon la terminologie utilisée dans la théorie des espaces vectoriels) et on écrit alors l'ensemble des solutions sous la forme suivante :

$$\mathcal{S}_0 = \text{Vect}(e_1, \cdots, e_k)$$

Notation 1.3.15 (Vect$(e_1, \ldots, e_k)$)

Étant donné des vecteurs $e_1, \ldots, e_k$ de $\mathbb{R}^p$, on note $\text{Vect}(e_1, \ldots, e_k)$ l'ensemble de tous les éléments de $\mathbb{R}^p$ s'écrivant comme combinaison linéaire des vecteurs $e_1, \ldots, e_k$:

$$\text{Vect}(e_1, \ldots, e_k) = \{\lambda_1 e_1 + \cdots + \lambda_k e_k \mid (\lambda_1, \ldots, \lambda_k) \in \mathbb{R}^k\}.$$

Il s'agit du *sous-espace vectoriel engendré par la famille* $(e_1, \ldots, e_k)$. Cette notion se généralisera à tout espace vectoriel.

Dire que $(e_1, \ldots, e_k)$ est une base de $\text{Vect}(e_1, \ldots, e_k)$ signifie que tout élément de $\text{Vect}(e_1, \ldots, e_k)$ s'écrit *de*

façon unique comme combinaison linéaire des e_i. Dans le contexte étudié, cela provient tout simplement du caractère minimal du choix des paramètres.

Exemple 1.3.16

Recherche de la solution générale de l'équation homogène, puis de l'équation non homogène dans le cas de l'exemple 1.3.10.

Remarque 1.3.17 (Un mot sur le terme « algorithme »)

La méthode décrite ci-dessus est une méthode algorithmique dans la mesure où elle peut être implémentée facilement sur un ordinateur : on a une démarche toute tracée pour arriver au bout du calcul, ne nécessitant ni reflexion ni astuce. Le seul point où l'intelligence humaine peut surpasser la puissance calculatoire de l'ordinateur réside dans le choix judicieux du pivot. Mais même avec un mauvais choix, on parvient au bout du calcul.

2 Structures algébriques

Note Historique 2.0.18

Il est fréquent de trouver des propriétés communes dans des situations qui au départ semblent totalement sans rapport. Une des grandes découvertes des mathématiques du 19e siècle a été de réussir à unifier ces problèmes en apparence distincts, en faisant ressortir de ces différents problèmes des structures ensemblistes et opératoires ayant des propriétés similaires.
C'est Évariste Galois le premier à mettre en avant ces études de structure à l'occasion de ses travaux visant à étudier la résolubilité des équations polynomiales par radicaux. Il y parle de groupes de permutations des solutions d'une équation, et est amené à étudier des propriétés de certains sous-ensembles de ces groupes de permutations. C'est lui qui introduit la terminologie de « groupe », même si la formalisation précise de cette notion est beaucoup plus tardive.
Le groupe des permutations d'un ensemble avait déjà été étudié auparavant par Lagrange (mais sans en faire ressortir cette structure bien particulière de groupe). Il a notamment établi à cette occasion un résultat important, généralisé plus tard pour tout groupe sous le nom de « théorème de Lagrange ».

La notion de structure algébrique repose de façon essentielle sur la notion de loi de composition (c'est-à-dire d'opération définie sur un ensemble, comme l'addition ou la multiplication) et sur les différentes propriétés que ces lois de composition peuvent vérifier. Nous commençons donc notre étude par l'examen de ces propriétés, après avoir défini de façon précise ce qu'est une loi de composition.

I Lois de composition

I.1 Définitions

Dans ce qui suit, E est un ensemble quelconque.

Définition 2.1.1 (Lois de composition)

On distingue deux types de lois de compositions (opérations), suivant que la loi décrit une opération entre deux éléments de l'ensemble E, ou entre un élément de E et un élément d'un ensemble externe Ω, appelé domaine d'opérateur.

- Une *loi de composition interne* est une application de $\varphi : E \times E$ dans E, souvent notée de façon opérationnelle plutôt que fonctionnelle (par exemple $x + y$ au lieu de $\varphi(x, y)$ pour désigner une addition)
- Une *loi de composition externe à gauche* sur E, d'ensemble d'opérateurs Ω, est une application de $\Omega \times E$ dans E, également notée de façon opérationnelle le plus souvent (par exemple $\lambda \cdot x$ au lieu de $\varphi(\lambda, x)$)

- De même, une *loi de composition externe à droite* sur E d'ensemble d'opérateurs Ω est une application $E \times \Omega \longrightarrow E$.

Exemples 2.1.2

- Les lois $+$ et $\times$ sont des lois de composition interne sur $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$ ou $\mathbb{C}$.
- La loi $+$ est une loi de composition interne sur $\mathbb{R}^n$ ou $\mathbb{C}^n$.
- $(\lambda, X) \mapsto \lambda X$ (multiplication d'un vecteur par un scalaire) est une loi de composition externe sur $\mathbb{R}^n$ (ou $\mathbb{C}^n$), d'ensemble d'opérateurs $\mathbb{R}$ (ou $\mathbb{C}$)
- De même pour la multiplication des polynômes par des scalaires.
- La composition $\circ$ définit une loi de composition interne sur E^E.
- Le produit scalaire sur $\mathbb{R}^n$ n'est pas une loi de composition (interne ou externe), car le résultat de l'opération n'est pas un élément de $\mathbb{R}^n$.

Nous nous limitons ici à l'étude des lois de composition interne, les seules qui interviennent dans les définitions des structures de groupe, d'anneaux et de corps. Nous reparlerons de lois de composition externes lorsque nous introduirons les espaces vectoriels.

I.2 Propriétés d'une loi de composition

Soit E un ensemble, muni d'une loi de composition interne que nous noterons $\star$. Nous étudions ici quelques propriétés pouvant être vérifiées par la loi $\star$

Définition 2.1.3 (Associativité, commutativité)

- On dit que $\star$ est *associative* ssi : $\forall (x, y, z) \in E^3, \quad (x \star y) \star z = x \star (y \star z)$
- On dit que $\star$ est *commutative* ssi : $\forall (x, y) \in E^2, \quad x \star y = y \star x$.

Ainsi, lorsque E est muni d'une loi associative, on peut effectuer les opérations dans l'ordre que l'on veut, à condition de respecter la position respective des éléments les uns par rapport aux autres. Si la loi est commutative, on peut échanger la position respective des éléments (mais pas nécessairement faire les opérations dans l'ordre qu'on veut si la loi n'est pas associative).

Exemples 2.1.4 (Lois commutatives, associatives)

1. Les lois $+$ et $\times$ définies sur $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$ et $\mathbb{C}$ sont associatives et commutatives.
2. Le produit matriciel définit une loi associative sur $\mathcal{M}_n(\mathbb{R})$ (ensemble des matrices carrées d'ordre n), mais pas commutative.
3. La composition définit une loi associative sur E^E mais pas commutative.
4. La soustraction dans $\mathbb{Z}$ est non associative et non commutative.
5. La loi définie sur $\mathbb{R}$ par $(a, b) \mapsto (a + b)^2$ est commutative mais non associative.

Notation 2.1.5 (Suppression des parenthèses)

Lorsque $\star$ est associative, nous nous permettons d'ommettre le parenthésage, en notant $x \star y \star z$ au lieu de $(x \star y) \star z$ ou $x \star (y \star z)$, la propriété d'associativité levant toute ambiguité sur l'interprétation de cette expression.

Avertissement 2.1.6

Attention à toujours bien indiquer le parenthésage lorsque la loi n'est pas associative, ou lorsque plusieurs lois sont en jeu sans qu'il n'ait été défini de façon explicite de relation de priorité sur les opérations.

Convention 2.1.7 (Commutativité d'une loi d'addition, usage)

Nous réserverons la notation additive (signe opératoire $+$) pour des lois de composition commutatives. Cela n'empêche pas en revanche de considérer des lois commutative notées multiplicativement.

Nous voyons maintenant des propriétés liées à l'existence de certains éléments particuliers de E.

Définition 2.1.8 (Élément neutre)

Soit e un élément de E.

1. On dit que e est un *élément neutre à droite* pour la loi $\star$ ssi : $\forall x \in E,\ x \star e = x$.
2. On dit que e est un *élément neutre à gauche* pour la loi $\star$ ssi : $\forall x \in E,\ e \star x = x$.
3. On dit que e est un *élément neutre* pour la loi $\star$ s'il est à la fois neutre à gauche et neutre à droite.

Pour une loi commutative, e est neutre ssi e est neutre à droite ssi e est neutre à gauche.

Exemple 2.1.9 (Éléments neutres)

1. 0 est élément neutre pour $+$ dans $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$. C'est le seul élément neutre pour $+$.
2. 1 est élément neutre pour $\times$ dans $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$. C'est le seul élément neutre pour $\times$.
3. I_n est élément neutre pour $\times$ sur $\mathcal{M}_n(\mathbb{R})$, 0_n est élément neutre pour $+$ sur $\mathcal{M}_n(\mathbb{R})$.
4. id_E est élément neutre pour $\circ$ sur E^E.
5. Sur un ensemble E de cardinal supérieur ou égal à 2, la loi $(x, y) \mapsto y$ admet plusieurs neutres à gauche (tout $x \in E$ est neutre à gauche). En revanche, il n'y a pas de neutre à droite.

Une loi ne peut pas admettre plusieurs éléments neutres, comme le montre la propriété suivante.

Proposition 2.1.10 (Unicité du neutre)

- *Si E admet pour $\star$ un neutre à gauche e_g et un neutre à droite e_d, alors $e_g = e_d$, et cet élément est alors un élément neutre pour $\star$.*
- *L'élément neutre, s'il existe, est unique ; dans ce cas, il s'agit aussi de l'unique neutre à gauche, et de l'unique neutre à droite.*

Notation 2.1.11 (0_E, 1_E)

- On note généralement 0_E (ou 0 s'il n'y a pas de risque d'ambiguité) le neutre (s'il existe) d'une loi notée additivement $+$.
- On note généralement 1_E (ou 1 s'il n'y a pas de risque d'ambiguité) le neutre (s'il existe) d'une loi notée multiplicativement $\times$.

Définition 2.1.12 (Élément symétrique)

Supposons que E admet un élément neutre e pour la loi $\star$. Soit $x \in E$.

- On dit que ${}^s x$ est un symétrique à gauche de x pour la loi $\star$ si ${}^s x \star x = e$.
- On dit que x^s est un symétrique à droite de x pour la loi $\star$ si $x \star x^s = e$.
- On dit que $\overline{x}$ est un symétrique de x pour la loi $\star$ si et seulement si $\overline{x}$ est un symétrique à droite et à gauche de x.
- On dit que x est symétrisable (*resp.* symétrisable à gauche, *resp.* symétrisable à droite) si x admet au moins un symétrique (*resp.* un symétrique à gauche, *resp.* un symétrique à droite).

Terminologie 2.1.13 (Opposé, inverse)

- Dans le cas d'une loi notée additivement, on parle plutôt d'opposé, et en cas d'unicité, on note $-x$ l'opposé de x.
- Dans le cas d'une loi notée multiplicativement (on omet dans ce cas souvent d'écrire le signe), on parle plutôt d'inversibilité (inversibilité à droite, à gauche), et en cas d'unicité, on note x^{-1} l'inverse de x.

Proposition 2.1.14 (Unicité du symétrique)

Si $\star$ est associative, alors, en cas d'existence, le symétrique est unique.

Exemples 2.1.15

1. Dans $\mathbb{N}$ seul 0 admet un opposé pour $+$.
2. Dans $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{Q}$, $\mathbb{C}$, tout élément admet un opposé pour $+$.
3. Dans $\mathbb{N}$ seul 1 admet un inverse, dans $\mathbb{Z}$, seuls 1 et -1 admettent un inverse. Dans $\mathbb{R}$, $\mathbb{Q}$ et $\mathbb{C}$ tous les éléments non nuls admettent un inverse.
4. Dans E^E muni de $\circ$, les éléments symétrisables à gauche sont les injections, les éléments symétrisables à droite sont les surjections, les éléments symétrisables sont les bijections. Une injection non surjective admet plusieurs symétriques à gauche ; une surjection non injective admet plusieurs symétriques à droite.

Proposition 2.1.16 (Symétrique de $x \star y$)

Supposons $\star$ associative. Soit $(x, y) \in E^2$. Si x et y sont symétrisables, de symétriques x^s et y^s, alors $x \star y$ est symétrisable de symétrique $y^s \star x^s$. Notez l'inversion !

Traduisons pour une loi multiplicative : si x et y sont inversibles, d'inverses x^{-1} et y^{-1}, alors xy aussi, d'inverse $y^{-1}x^{-1}$.

Définition 2.1.17 (Élément absobant)

Soit $x \in E$.

- On dit que x est un élément absorbant à gauche pour $\star$ ssi : $\forall y \in E, \quad x \star y = x$.
- On dit que x est absorbant à droite pour $\star$ ssi : $\forall y \in E$, $y \star x = x$.
- On dit que x est absorbant s'il est à la fois absorbant à gauche et à droite.

Exemples 2.1.18

1. 0 est absorbant pour $\times$ dans $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$.
2. Pour la loi $(x, y) \mapsto y$, tout élément y de E est absorbant à droite. Il n'y a pas d'élément absorbant à gauche si E est de cardinal au moins 2.

Définition 2.1.19 (Élément régulier ou simplifiable)

- Un élément x est dit régulier (ou simplifiable) à gauche ssi :

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

- Un élément x est dit régulier (ou simplifiable) à droite ssi :

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

- Un élément x est dit régulier (ou simplifiable) s'il est à la fois régulier à gauche et à droite.

Proposition 2.1.20 (Régularité des éléments symétrisables)

Supposons que E soit muni d'une loi $\star$ associative.

- *Soit x un élément admettant un symétrique à gauche. Alors x est régulier à gauche.*
- *Soit x un élément admettant un symétrique à droite. Alors x est régulier à droite.*
- *Soit x un élément admettant un symétrique. Alors x est régulier.*

Ainsi, le fait de pouvoir simplifier une égalité par un réel ou complexe non nul x ne vient pas tant de la non nullité que de l'inversibilité de x. Par exemple, la non nullité n'est pas un critère suffisant de régularité dans $\mathcal{M}_n(\mathbb{R})$: il est nécessaire d'avoir l'inversibilité de la matrice que l'on veut simplifier.
Il convient toutefois de noter que la condition d'inversibilité, si elle est suffisante, n'est en général pas nécessaire.

Exemple 2.1.21

Dans l'ensemble $\mathbb{R}[X]$ des polynômes à coefficients réels muni du produit, tout polynôme non nul est régulier, alors que seuls les polynômes constants non nuls sont inversibles.

I.3 Ensembles munies de plusieurs lois

Soit E un ensemble muni de deux lois de composition $\star$ et $\diamond$.

Définition 2.1.22 (Distributivité)

- On dit que la loi $\star$ est distributive à gauche sur $\diamond$ ssi : $\forall (x, y, z) \in E^3,\quad x \star (y \diamond z) = (x \star y) \diamond (x \star z)$.
- On dit que la loi $\star$ est distibutive à droite sur $\diamond$ ssi : $\forall (x, y, z) \in E^3,\quad (y \diamond z) \star x = (y \star x) \diamond (z \star x)$.
- On dit que la loi $\star$ est distributive sur $\diamond$ ssi elle est distributive à droite et à gauche.

Exemples 2.1.23

1. La loi $\times$ est distributive sur $+$ dans $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$, $\mathcal{M}_n(\mathbb{R})$...
2. La loi $\cap$ est distributive sur $\cup$ sur $\mathcal{P}(X)$. Inversement, la loi $\cup$ est distributive sur $\cap$.

I.4 Stabilité

Définition 2.1.24

Soit E un ensemble muni d'une loi $\star$ et $F \subset E$ un sous-ensemble de E. On dit que F est stable par $\star$, si la restriction de la loi de E à $F \times F$ peut se corestreindre à F, autrement dit si :

$$\forall (x, y) \in F^2,\ x \star y \in F.$$

Dans ce cas, la loi de E se restreint en une loi $\star_F : F \times F \to F$, appelée *loi induite sur F par $\star$*.

II Structures

II.1 Généralités

Définition 2.2.1

- Une structure de « truc » est la donnée d'un certain nombre d'axiomes (définissant ce qu'on appelle un « truc ») portant sur un ensemble fini de lois de composition (internes et/ou externes).
- On dit que E est muni d'une structure de truc ssi E est muni d'un nombre fini de lois de composition vérifiant les axiomes de structure de truc.

Exemples 2.2.2

1. Une structure de magma se définit comme la donnée d'une loi de composition, et un ensemble vide d'axiomes. Ainsi, tout ensemble E muni d'une loi de composition est muni d'une structure de magma.
2. Une structure de monoïde se définit comme la donnée d'une loi de composition, et de deux axiomes : l'associativité de la loi et l'existence d'un élément neutre. Par exemple $(\mathbb{N}, +)$ est muni d'une structure de monoïde (on dit plus simplement que $(\mathbb{N}, +)$ est un monoïde). L'ensemble des mots sur un alphabet $\mathcal{A}$, muni de l'opération de concaténation est aussi un monoïde (appelé monoïde libre sur l'alphabet $\mathcal{A}$). Contrairement à $\mathbb{N}$, le monoïde libre n'est pas commutatif.
3. Ainsi, la structure de monoïde est plus riche que celle de magma : tout monoïde est aussi un magma ; un monoïde peut être défini comme un magma dont la loi est associative et possède un élément neutre.
4. Une structure de groupe est une structure de monoïde à laquelle on rajoute l'axiome d'existence de symétriques. Par exemple $(\mathbb{Z}, +)$ est un groupe, mais pas $(\mathbb{N}, +)$.

Définition 2.2.3 (Stucture induite)

Soit E un ensemble muni d'une structure de truc, et F un sous-ensemble de E. Si F est stable pour chacune des lois de E, l'ensemble F muni des lois induites sur F par les lois de E est appelée structure induite sur F par la structure de E.

Avertissement 2.2.4

En général, F ne peut pas être muni d'une structure de truc, mais seulement d'une structure moins riche, certains des axiomes de la structure de truc pouvant ne pas être préservée par restriction.

Exemple 2.2.5

$(\mathbb{N}, +)$ est la structure induite sur $\mathbb{N}$ par la structure de groupe additif de $(\mathbb{Z}, +)$. En revanche, $(\mathbb{N}, +)$ n'est pas un groupe. On a perdu l'existence des opposés par restriction.

Définition 2.2.6 (Sous-truc)

Soit E un ensemble muni d'une structure de truc et F un sous-ensemble de E. On dit que F est un sous-truc de E si F est stable par les lois de E, si F contient les neutres imposés de E, et si les lois induites sur F par les lois de E vérifient les axiomes de la structure de truc.

Nous verrons comment traduire de façon effective cette notion dans le cas de sous-groupes et sous-anneaux.

II.2 Morphismes

Lorsqu'on dispose d'une structure de truc, on est souvent amené à considérer des applications entre ensembles munis de la structure de truc. Cependant seules nous intéressent les applications compatibles dans un certain sens avec la structure de truc : les autres ne sont pas pertinentes dans le contexte (si on a à s'en servir, c'est qu'on sort de la structure de truc, et que la structure de truc n'est plus le contexte adapté).

Définition 2.2.7 (Homomorphisme)

Soit E et F deux ensembles munis d'une structure de truc, E étant muni des lois de composition interne $(\star_1, \dots, \star_n)$ et F des lois $(\diamond_1, \dots, \diamond_n)$. On dit qu'une application $f : E \longrightarrow F$ est un homomorphisme de truc (ou plus simplement un morphisme de truc) ssi :

- L'application f est compatible avec les lois :

$$\forall k \in [\![1, n]\!], \quad \forall (x, y) \in E^2, \quad f(x \star_k y) = f(x) \diamond_k f(y).$$

- Si l'existence du neutre e_i pour la loi $\star_i$ est imposée dans les axiomes (et donc le neutre f_i pour la loi $\diamond_i$ existe aussi), f doit être compatible avec le neutre : $f(e_i) = f_i$.

On peut avoir à rajouter certaines propriétés liées à la structure étudiée. On peut aussi ajouter l'existence d'un homomorphisme nul (ne vérifiant pas la compatibilité avec les neutres non additifs), afin d'obtenir une structure intéressante sur l'ensemble des morphismes.

Pour chaque structure étudiée, nous redéfinirons de façon précise la notion d'homomorphisme associée, si celle-ci est à connaître. Nous donnons une propriété générale, dont la démonstration dans le cadre général nous dispensera des démonstrations au cas par cas.

Proposition 2.2.8 (Composition d'homomorphismes)

Soit $f : E \longrightarrow F$ et $g : F \longrightarrow G$ deux morphismes de truc. Alors $g \circ f$ est un morphisme de truc.

Nous définissons alors :

Terminologie 2.2.9

- Un isomorphisme de truc est un homomorphisme de truc bijectif.
- Un endomorphisme de truc est un homomorphisme de truc de E dans lui-même.
- Un automorphisme de truc est un endomorphisme qui est également un isomorphisme.

Proposition 2.2.10

Si $f : E \longrightarrow F$ est un isomorphisme de truc, alors f^{-1} est un isomorphisme de truc.

Ainsi, la réciproque d'un isomorphisme est bijective (ça, ce n'est pas une surprise), et, ce qui est moins évident, c'est aussi un homomorphisme de truc.

II.3 Catégories (HP)

La notion de structure et de morphisme associé permet de définir la notion de catégorie. Grossièrement, une catégorie est la donnée :
- d'une classe d'objets ;
- de flèches entre ces objets ;
- d'une règle de composition entre les flèches.

Par exemple, la catégorie des monoïdes est la catégorie dont les objets sont les monoïdes, les flèches sont les homomorphismes de monoïdes, et la composition des flèches correspond à la composition usuelle des homomorphismes (la composée de deux homomophismes étant encore un homomorphisme, donc une flèche de la catégorie). On définit de même la catégorie des groupes, ou la catégorie des anneaux, ou encore la catégorie des corps.

Cette notion de catégorie nous permet de travailler dans un certain contexte. Se donner une catégorie permet de se concentrer sur un certain type d'objets, et un certain type d'applications, et de les étudier d'un point de vue formel.

La notion de catégorie dépasse largement le cadre de l'étude des structures algébriques, car si les structures algébriques fournissent des catégories, de nombreuses catégories sont issues d'autres contextes, comme :
- la catégorie des ensembles, les morphismes étant toutes les applications ;
- la catégorie des ensembles ordonnés, les morphismes étant les applications croissantes
- la catégorie des espaces topologiques, les morphismes étant les applications continues
- ou encore, la catégorie des catégories, les morphismes étant les foncteurs de C dans D, associant à chaque objet de C un objet de D, et à chaque flèche de C une flèche de D, tout en respectant un certain nombre de règles de compatibilité.
- ou encore, des catégories de foncteurs entre deux catégories, les objets étant cette fois des foncteurs, et les flèches étant des « transformations naturelles » entre foncteurs...

III Groupes

III.1 Axiomatique de la structure groupes

Définition 2.3.1 (Groupe)

Soit G un ensemble. On dit que G est muni d'une structure de groupe si G est muni d'une loi de composition $\star$ telle que :
- $\star$ est associative ;
- il existe un élément neutre e pour la loi $\star$;
- tout élément x admet un symétrique x^s.

En vertu de résultats antérieurs, on peut énoncer :

Proposition 2.3.2 (Unicité du neutre et des symétriques)

Soit $(G, \star)$ un groupe. Alors :
- *G admet un unique élément neutre pour $\star$*
- *Pour tout $x \in G$, il existe un unique symétrique x^s de x.*

Corollaire 2.3.3 (régularité des éléments d'un groupe)

Tous les éléments d'un groupe sont réguliers pour la loi du groupe.

Définition 2.3.4 (Groupe abélien ou commutatif)

On dit qu'un groupe $(G, \star)$ est abélien (ou commutatif) si la loi de G est commutative.

Notation 2.3.5 (Notation additive, notation multiplicative)

La loi d'un groupe est le plus souvent notée additivement (signe $+$) ou multiplicativement (signe $\times$, parfois remplacé par un point, voire omis, comme dans $\mathbb{R}$). La notation additive est réservée au cas de groupes abéliens. Nous avons alors les notations suivantes pour désigner des itérées de la loi de composition sur un même élément x :

- loi multiplicative : $x \times \cdots \times x$ (avec n occurrences) est noté x^n ;
  le neutre est noté 1 ;
  par convention, $x^0 = 1$;
- loi additive : $x + \cdots + x$ (avec n occurrences) est noté $n \cdot x$ ou nx ;
  le neutre est noté 0 ;
  par convention $0x = 0$.

Une définition plus rigoureuse par récurrence pourrait être donnée pour ces itérées.

Notation 2.3.6 (Simplifications d'écriture pour la notation additive)

Soit $(G, +)$ un groupe commutatif. Comme mentionné plus haut, l'opposé d'un élément x est noté $-x$. On note alors $x - y$ au lieu de $x + (-y)$. On a alors les règles suivantes :

- $\forall (x, y, z) \in G^3, \quad x - (y + z) = x - y - z$
- $\forall (x, y, z) \in G^3, \quad x - (y - z) = x - y + z.$

En vertu des définitions générales données dans le paragraphe précédent, nous donnons la définition suivante :

Définition 2.3.7 (Homomorphisme de groupe, HP)

Soit $(G, \star)$ et $(H, \diamond)$ deux groupes.

- On dit qu'une application $f : G \longrightarrow H$ est un homomorphisme de groupe si pour tout $(x, y) \in G$, $f(x \star y) = f(x) \diamond f(y)$.
  On note $\mathrm{Hom}(G, H)$ l'ensemble des homomorphismes de G dans H.
- Si $(G, \star) = (H, \diamond)$, on dit que f est un endomorphisme de $(G, \star)$.
- Un homomorphisme bijectif est appelé isomorphisme ; en vertu de ce qui précède, la réciproque d'un isomorphisme est un isomorphisme.
- Un endomorphisme bijectif est appelé automorphisme ; en vertu de ce qui précède, la réciproque d'un automorphisme est un automorphisme.
  On note $\mathrm{Aut}(G)$ l'ensemble des automophismes de G.

III.2 Exemples importants

Exemples 2.3.8

1. $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$ sont des groupes commutatifs notés additivement.
2. $(\mathbb{Q}^*, \times)$, $(\mathbb{R}^*, \times)$, $(\mathbb{C}^*, \times)$, $(\mathbb{Q}_+^*, \times)$, $(\mathbb{R}_+^*, \times)$ sont des groupes commutatifs notés multiplicativement.
3. $(\mathbb{N}, +)$, $(\mathbb{Q}, \times)$, $(\mathbb{Z}, \times)$, « $(\mathbb{R}_-^*, \times)$ » sont-ils des groupes ?
4. $(\mathbb{U}, \times)$ et $(\mathbb{U}_n, \times)$ sont des groupes.

5. Pour $n \geqslant 2$, $(\mathbb{Z}/n\mathbb{Z}, +)$ est un groupe.
6. $((\mathbb{Z}/n\mathbb{Z}) \setminus \{0\}, \times)$ est-il en général un groupe ?
7. Étant donné X un ensemble, $(\mathfrak{S}_X, \circ)$, l'ensemble des permutations de X est un groupe pour la loi définie par la composition.
8. $\exp : (\mathbb{R}, +) \longrightarrow (\mathbb{R}_+^*, \times)$ est un homomorphisme de groupes. C'est même un isomorphisme.
9. Sa réciproque est donc aussi un isomorphisme de groupes : $\ln : (\mathbb{R}_+^*, \times) \longrightarrow (\mathbb{R}, +)$.
10. L'application $x \mapsto \mathrm{e}^{\mathrm{i}\,x}$ est un morphisme de groupes de $(\mathbb{R}, +)$ à $(\mathbb{U}, \times)$.
11. L'application $z \mapsto \mathrm{e}^z$ est un morphisme de groupes surjectif (mais non injectif) de $(\mathbb{C}, +)$ sur $(\mathbb{C}^*, \times)$.
12. Soit $n \geqslant 2$ et pour tout $k \in [\![0, n-1]\!]$, $\omega_k = \mathrm{e}^{\mathrm{i}\,\frac{2k\pi}{n}}$. Alors, étant donné $k \in [\![0, n]\!]$:
$$\begin{array}{rccc} f : & (\mathbb{Z}, +) & \longrightarrow & (\mathbb{U}_n, \times) \\ & \ell & \mapsto & \omega_k^\ell \end{array}$$
est un homomorphisme de groupe. Il est surjectif si $k = 1$, et plus généralement si k est premier avec n (d'après le théorème de Bézout).
13. Puisque $\omega_k^n = 1$, le morphisme précédent « passe au quotient » et définit un homomorphisme de groupe :
$$\begin{array}{rccc} f : & (\mathbb{Z}/n\mathbb{Z}, +) & \longrightarrow & (\mathbb{U}_n, \times) \\ & \ell & \mapsto & \omega_k^\ell \end{array}$$
Cet homomorphisme est un isomophisme si $k = 1$, et plus généralement si k est premier avec n.

III.3 Sous-groupes

Toujours en suivant les définitions plus générales, nous donnons la définition suivante :

Définition 2.3.9 (Sous-groupe)

Soit $(G, \star)$ un groupe. Un sous-ensemble H de G est appelé *sous-groupe de G* si H est stable pour la loi de G et si la loi induite définit sur H une structure de groupe.

Remarquez qu'on n'a pas donné l'appartenance du neutre à H dans la définition, celle-ci étant automatique en vertu de :

Proposition 2.3.10 (Appartenance de l'élément neutre à H)

Soit H un sous-groupe de G. Alors l'élément neutre e de G est dans H et est l'élément neutre du groupe H.

Dans la pratique, pour vérifier que H est un sous-groupe de G on utilise le résultat suivant, ou sa version compactée :

Théorème 2.3.11 (Caractérisation des sous-groupes)

Un sous-ensemble H d'un groupe $(G, \star)$ (de neutre e_G) est un sous-groupe de G si et seulement si :

(i) H est non vide (on vérifie par exemple $e_G \in H$) ;

(ii) H est stable pour $\star$: $\forall (x, y) \in H, \quad x \star y \in H$,

(iii) H est stable par prise de symétrique : $\forall x \in H, \quad x^s \in H$.

On peut rassembler les deux dernières propriétés en une seule vérification :

Théorème 2.3.12 (Caractérisation des sous-groupes, version compactée)
Un sous-ensemble H d'un groupe $(G, \star)$ (de neutre e_G) est un sous-groupe de G si et seulement si :
(i) H est non vide (on vérifie par exemple $e_G \in H$) ;
(ii) $\forall (x, y) \in H^2, \quad x \star y^s \in H$,

On traduit cette dernière propriété dans les deux cas les plus fréquents :
- pour un sous-groupe d'un groupe additif, la vérification de stabilité à faire est donc :

$$\forall (x, y) \in H^2, \ x - y \in H;$$

- pour un sous-groupe d'un groupe multiplicatif, la vérification de stabilité à faire est donc :

$$\forall (x, y) \in H^2, \quad xy^{-1} \in H.$$

De façon évidente, étant donné G un groupe, d'élément neutre e, $\{e\}$ et G sont des sous-groupes de G.

Définition 2.3.13 (Sous-groupe propre)
Un sous-groupe propre de G est un sous-groupe de G distinct de $\{e\}$ et de G.

Proposition 2.3.14 (Transitivité de la notion de sous-groupe)
Soit G un groupe. Alors si H est un sous-groupe de G et K un sous-groupe de H, alors K est un sous-groupe de G.

Proposition 2.3.15 (Intersection de sous-groupes)
Soit G un groupe, et $(H_i)_{i \in I}$ une famille de sous-groupes de G. Alors $\bigcap_{i \in I} H_i$ est un sous-groupe de G.

Théorème 2.3.16
Les sous-groupes de $\mathbb{Z}$ sont exactement les $n\mathbb{Z}$, $n \in \mathbb{N}$.

Les sous-groupes de $\mathbb{Z}$ sont donc des groupes monogènes, dans le sens suivant :

Proposition/Définition 2.3.17 (Sous-groupe monogène, HP)
Soit G un groupe multiplicatif, et $x \in G$. Soit

$$H = \{x^n, \quad n \in \mathbb{Z}\}$$

(dans cet ensemble, certains éléments peuvent être égaux).
Alors H est un sous-groupe de G. On dit que H est un *sous-groupe monogène* de G.

On verra aussi en exercice le résultat classique suivant

Proposition 2.3.18 (caractérisation des sous-groupes additifs de $\mathbb{R}$, HP)
Les sous-groupes de $(\mathbb{R}, +)$ sont soit égaux à $a\mathbb{Z}$, $a \in \mathbb{R}_+$, soit denses dans $\mathbb{R}$.

III.4 Congruences modulo un sous-groupe

Soit G un groupe (multiplicatif) et H un sous-groupe de G.

Définition 2.3.19 (Congruence modulo H, HP)

- On définit la relation de congruence à droite modulo H sur G par :

$$\forall (x,y) \in G, \quad x \equiv_d y \mod H \iff xy^{-1} \in H.$$

- On définit la relation de congruence à gauche modulo H sur G par :

$$\forall (x,y) \in G, \quad x \equiv_g y \mod H \iff x^{-1}y \in H.$$

Dans le cas de groupes abéliens, les deux notions coïncident évidemment.

Exemple 2.3.20

Si $G = \mathbb{Z}$ et $H = n\mathbb{Z}$, cela correspond (en notation additive cette fois) à la relation de congruence modulo n.

Proposition 2.3.21

Les relations de congruence à gauche et à droite modulo H sont des relations d'équivalence.

Définition 2.3.22 (Classes à droite et à gauche modulo H, HP)

- Les classes d'équivalence à droite modulo H sont appelées classes à droite modulo H. Il s'agit des ensembles Ha, $a \in G$, qui ne sont pas des groupes (sauf si $a \in H$).
- Les classes d'équivalence à gauche modulo H sont appelées classes à gauche modulo H. Il s'agit des ensembles aH, $a \in G$, qui ne sont pas des groupes (sauf si $a \in H$).

On en déduit une des propriétés les plus importantes des cardinaux des groupes finis.

Définition 2.3.23 (Ordre d'un groupe fini, HP)

Soit G un groupe fini. Son ordre est son cardinal.

Lemme 2.3.24 (Cardinaux des classes à gauche et à droite, HP)

Soit G un groupe fini. Pour tout $a \in G$, H, aH et Ha ont même cardinal.

Théorème 2.3.25 (Lagrange, HP)

Soit G un groupe fini, et H un sous-groupe de G. Alors l'ordre de H divise l'ordre de G.

Remarquez aussi que pour que les deux relations de congruence coïncident il faut et il suffit que pour tout $a \in G$, $aH = Ha$, ou encore $aHa^{-1} = H$. Cela motive la définition suivante, très importante dans l'étude des tours de composition, à la base de la théorie de Galois. C'est sous cette condition en effet que la loi quotient de la loi G par la relation de congruence munira le quotient d'une structure de groupe.

Proposition/Définition 2.3.26 (Sous-groupe distingué, HP)

On dit que H est un sous-groupe distingué (ou normal) de G si l'une des deux propriétés équivalentes est satisfaite :

(i) $\forall a \in G,\ aH = Ha$

(ii) $\forall a \in G, \forall h \in H,\ aha^{-1} \in H$.

III.5 Ordre d'un élément

Définition 2.3.27 (Ordre d'un élément d'un groupe, HP)

L'ordre d'un élément $x \neq 1$ d'un groupe (multiplicatif) G, dont le neutre est noté 1, est

$$\mathrm{ord}(x) = \min\{n \in \mathbb{N}^* \mid x^n = 1\}$$

Cet ordre peut être $+\infty$ par convention si l'ensemble ci-dessus est vide.

Proposition 2.3.28

Si $\mathrm{ord}(x)$ *est fini, alors les élément* $x^0, \dots, x^{\mathrm{ord}(x)-1}$ *sont deux à deux distincts, et*

$$\{x^n,\ n \in \mathbb{Z}\} = \{x^n,\ n \in [\![0, \mathrm{ord}(x) - 1]\!]\}.$$

On en déduit :

Théorème 2.3.29 (encore Lagrange, HP)

Soit x *un élément d'un groupe fini* G*. Alors l'ordre de* x *divise l'ordre de* G*.*

IV Anneaux et corps

IV.1 Axiomatiques des structures d'anneaux et de corps

Définition 2.4.1 (Anneau)

Soit A un ensemble, muni de deux lois de composition internes (généralement notées $+$ et $\times$). On dit que $(A, +, \times)$ (ou plus simplement A) est un anneau si :

(i) $(A, +)$ est un groupe abélien

(ii) $(A, \times)$ est un monoïde (autrement dit $\times$ est associative et il existe un élément neutre 1 pour $\times$)

(iii) $\times$ est distributive sur $+$

Remarque 2.4.2

Certains ouvrages (notamment anciens) n'imposent pas l'existence de l'élément neutre 1 pour le produit et parlent alors d'*anneau unifère* ou *unitaire* pour ce que nous appelons ici simplement un *anneau*. La convention que nous adoptons concernant l'existence d'un élément neutre est celle généralement adoptée actuellement, et nous suivons en cela le programme officiel de la classe de MPSI.

Exemples 2.4.3

1. $\{0\}$ muni des opérations triviales est un anneau ; ici le neutre pour le produit est 0.
2. $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{Q}$ et $\mathbb{C}$ munis des opérations usuelles sont des anneaux.
3. Pour tout $n \in \mathbb{N}^*$, $\mathbb{Z}/n\mathbb{Z}$ est un anneau. La structure circulaire de ces anneaux explique la terminologie.
4. L'ensemble $\mathbb{R}[X]$ des polynômes à coefficients réels est un anneau. De même pour $\mathbb{Z}[X]$, $\mathbb{Q}[X]$ ou $\mathbb{Z}[X]$.
5. $\mathbb{N}$ n'est pas un anneau.
6. L'ensemble $\mathcal{M}_n(\mathbb{R})$ des matrices carrées est un anneau.
7. L'ensemble $(\mathcal{P}(E), \triangle, \cap)$ est un anneau (anneau de Boole).
8. $(\mathbb{R}^{\mathbb{R}}, +, \circ)$ est-il un anneau ?

Proposition 2.4.4

Si A est un anneau ayant au moins deux éléments, alors $1 \neq 0$.

Définition 2.4.5 (Anneau commutatif)

On dit qu'un anneau $(A, +, \times)$ est commutatif si et seulement si la loi $\times$ est commutative.

Les exemples donnés ci-dessus sont des exemples d'anneaux commutatifs, à l'exception d'un exemple. Lequel ?
Enfin, conformément à la définition générale, nous donnons :

Définition 2.4.6 (Homomorphisme d'anneaux)

Soit A et B deux anneaux. Un homomorphisme d'anneaux de A à B est une application $f : A \longrightarrow B$, soit égale à la fonction nulle, soit vérifiant :

$$\forall (x, y) \in A^2, \quad f(x+y) = f(x) + f(x), \qquad f(xy) = f(x)f(y) \qquad \text{et} \qquad f(1_A) = 1_B.$$

Ainsi, un homomorphisme d'anneaux (à part le morphisme nul un peu particulier) est à la fois un homomophisme du groupe $(A, +)$ et du monoïde $(A, \times)$.
Un corps est un anneau vérifiant une condition supplémentaire :

Définition 2.4.7 (Corps)

Soit K un ensemble muni de deux lois $+$ et $\times$. On dit que $(K, +, \times)$ (ou plus simplement K) est un corps si K est un anneau commutatif dans lequel tout élément non nul est inversible

Ainsi K est un corps si et seulement si $(K, +)$ et $(K, \times)$ sont des groupes commutatifs.

Remarque 2.4.8

- Conformément au programme, nous adoptons la convention stipulant que tout corps doit être commutatif. Là encore, les ouvrages anciens n'imposent pas cette condition. Il est d'usage actuellement de suivre la terminologie anglaise pour le cas où la commutativité n'est pas assurée, et d'appeler *anneau à divisions* un ensemble admettant cette structure non commutative (*division ring* en anglais)

- Dans le cas des corps finis, les deux notions coïncident, d'après le théorème de Wedderburn, stipulant que « tout corps fini est commutatif », ce qui, avec notre terminologie, se réexprime : « tout anneau à division fini est un corps. ».

Exemples 2.4.9

1. $\mathbb{R}$, $\mathbb{Q}$ et $\mathbb{C}$ sont des corps.
2. Le « corps » des quaternions n'est pas un corps dans le sens défini ci-dessus, mais seulement un anneau à divisions.
3. $\mathbb{Z}$ n'est pas un corps
4. En général $\mathbb{Z}/n\mathbb{Z}$ n'est pas un corps. Par exemple, 2 n'est pas inversible dans $\mathbb{Z}/4\mathbb{Z}$.
5. On peut montrer que $\mathbb{Z}/p\mathbb{Z}$ est un corps si et seulement si p est premier. Ce corps est en général noté $\mathbb{F}_p$ (comme « field », terminologie anglaise pour « corps »). Exemple $\mathbb{F}_2$, corps à 2 éléments.
6. On peut montrer que tout corps fini a un cardinal égal à p^n, où p est un nombre premier et n un entier. On peut également montrer que pour de telles données, il existe (à isomorphisme près) un unique corps à $q = p^n$ éléments, qu'on note $\mathbb{F}_q$. Lorsque $n = 1$, on retrouve les corps $\mathbb{F}_p$ du point précédent.

Définition 2.4.10 (Homomorphisme de corps)

Soit K et L deux corps. Un homomorphisme de corps $f : K \longrightarrow L$ est un homomorphisme des anneaux sous-jacents.

Définition 2.4.11 (Caractéristique d'un corps)

Soit K un corps, d'élément neutre $1_K \neq 0_K$. Soit $H = \{n \cdot 1_K, n \in \mathbb{Z}\}$ le sous-groupe monogène de $(K, +)$ engendré par 1_K.
- Si H est infini, on dit que K est de caractéristique nulle.
- Si H est fini, de cardinal $p \in \mathbb{N}$, on dit que K est de caractéristique p.

Théorème 2.4.12 (Primalité de la caractéristique d'un corps)

Soit K un corps de caractéristique non nulle. Alors sa caractéristique p est un nombre premier.

Remarque 2.4.13

- Un corps fini est toujours de caractéristique non nulle, donc première.
- Il existe des corps infinis de caractéristique p (par exemple le corps des fractions rationnelles à coefficients dans $\mathbb{F}_p$)
- Si K est un corps de caractéristique p, alors pour tout x de K, $px = x + \cdots + x = 0$ (avec p facteurs dans la somme).

Nous démontrerons en exercice le théorème important suivant :

Théorème 2.4.14

Soit K un corps.
- *Si K est de caractéristique nulle, il existe un homomophisme injectif de corps $i : \mathbb{Q} \to K$.*
- *Si K est de caractéristique p, il existe un homomorphisme injectif de corps $i : \mathbb{F}_p \to K$.*

IV.2 Sous-anneaux, sous-corps

Conformément à la définition générale, nous avons :

Définition 2.4.15 (Sous-anneau)

Soit $(A, +, \times)$ un anneau. Un sous-ensemble $B \subset A$ est un sous-anneau de A si et seulement si B est stable pour les lois $+$ et $\times$, et si $1_A \in B$

Remarquez qu'encore une fois, on ne dit rien de l'appartenance de 0_A à B, celle-ci étant ici aussi automatique (puisque $(B, +)$ est un sous-groupe de $(A, +)$). En revanche, l'appartenance de 1_A à B n'est pas automatique, comme le montre l'exemple de $B = \left\{ \begin{pmatrix} \lambda & 0 \\ 0 & 0 \end{pmatrix},\ \lambda \in \mathbb{R} \right\}$, sous-ensemble de $\mathcal{M}_2(\mathbb{R})$, stable pour les lois $+$ et $\times$. Ce n'est pas un sous-anneau au sens que nous en avons donné puisque $I_2 \notin B$. En revanche, les restrictions de $\times$ et $+$ définissent tout de même une structure d'anneau sur B, le neutre étant alors $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$.

Proposition 2.4.16 (Caractérisation des sous-anneaux)

Un sous-ensemble B d'un anneau A est un sous-anneau de A si et seulement si :

(i) $1_A \in B$.

(ii) pour tout $(x, y) \in B$, $x - y \in B$

(iii) pour tout $(x, y) \in B$, $xy \in B$

Exemples 2.4.17

1. $\mathbb{Z}$ est un sous-anneau de $\mathbb{Q}$ qui est un sous-anneau de $\mathbb{R}$ qui est un sous-anneau de $\mathbb{C}$.
2. $\mathbb{Z}/n\mathbb{Z}$ n'a d'autre sous-anneau que lui-même.

Définition 2.4.18 (Sous-corps)

Soit $L \subset K$ un sous-ensemble d'un corps K. On dit que L est un sous-corps de K si L est stable par $+$ et $\times$, $1_K \in L$, et si les lois induites sur L par celles de K le munissent d'une structure de corps.

Remarque 2.4.19

On pourrait remplacer l'hypothèse $1_K \subset L$ par le fait que L contient un élément $x \neq 0_K$.

Proposition 2.4.20 (Caractérisation des sous-corps)

$L \subset K$ est un sous-corps de K si et seulement si :

- $1_K \in L$
- *pour tout* $(x, y) \in L$, $x - y \in L$
- *pour tout* $(x, y) \in L$ *tel que* $y \neq 0$, $xy^{-1} \in L$.

Exemples 2.4.21

$\mathbb{Q}$ est un sous-corps de $\mathbb{R}$, $\mathbb{R}$ est un sous-corps de $\mathbb{C}$.

IV.3 Calculs dans un anneau

Du fait de l'existence d'une addition et d'une multiplication dans un anneau et dans un corps, et des règles d'associativité et de commutativité, tous les calculs que l'on a l'habitude de faire dans $\mathbb{Z}$, $\mathbb{R}$ ou $\mathbb{C}$ peuvent se généraliser à un anneau ou un corps quelconque. Il faut toutefois faire attention que dans un anneau, contrairement à ce qu'il se passe dans $\mathbb{R}$ ou $\mathbb{C}$, tous les éléments ne sont pas inversibles, et que par ailleurs, les calculs nécessitant de permuter l'ordre de certains facteurs multiplicatifs ne peuvent pas être effectués en toute généralité dans un anneau non commutatif. Ainsi :

- Pour des calculs dans un anneau : considérer l'analogie avec $\mathbb{Z}$ (plutôt que $\mathbb{R}$), et se méfier :
  - ∗ des inversions intempestives
  - ∗ des simplifications abusives
  - ∗ des problèmes de commutativité, qui nécessitent parfois l'introduction d'hypothèses supplémentaires, à vérifier scrupuleusement.

  L'analogie avec $\mathbb{Z}$ n'est pas toujours suffisante, puisqu'il peut se produire des situations bien particulières, n'ayant pas lieu dans $\mathbb{Z}$, comme par exemple l'existence de « diviseurs de zéro » (voir un peu plus loin)
- Pour des calculs dans un corps : toute analogie avec $\mathbb{R}$ est permise.

Nous rappelons qu'on peut définir dans un anneau A le produit nx où $n \in \mathbb{Z}$, et $x \in A$ de la manière suivante :

- Si $n = 0$, $nx = 0$
- Si $n > 0$, $nx = x + \cdots + x$ (n facteurs)
- Si $n < 0$, $nx = -|n|x$.

Nous pouvons définir même a^n, pour tout $n \in \mathbb{N}$, et même pour tout $n \in \mathbb{Z}$ si a est inversible.

Nous voyons, outre les règles usuelles découlant des règles d'associativité et de distributivité, deux résultats déjà évoqués dans le cas de $\mathbb{R}$ ou $\mathbb{C}$, et que nous voyons plus généralement dans le cadre d'anneaux, mais qui nécessitent une hypothèse de commutativité.

Théorème 2.4.22 (Factorisation de $a^n - b^n$)

Soit a et b deux éléments d'un anneau A tels que $ab = ba$. Alors pour tout $n \in \mathbb{N}^*$

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

Corollaire 2.4.23 (Factorisation de $1 - a^n$)

Pour tout élément a d'un anneau A,

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

Si $1 - a$ est inversible (condition plus forte que $a \neq 1$), on peut alors écrire :

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

En revanche, évitez d'écrire cela sous forme de fraction lorsqu'on n'est pas dans une structure commutative, et attention à placer l'inverse du bon côté (même si, pour l'expression considérée, ce ne serait pas gênant car les facteurs considérés commutent, même si globalement l'anneau n'est pas commutatif ; mais autant prendre dès maintenant de bonnes habitudes)

Théorème 2.4.24 (Formule du binôme)

Soit a et b deux éléments d'un anneau tels que $ab = ba$. Alors, pour tout $n \in \mathbb{N}$,

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

Attention en revanche au cas où on n'a pas commutativité de a et b : il convient de bien distinguer les deux facteurs ab et ba appraissant dans le développement de $(a+b)(a+b)$ (par exemple, pour $n = 2$) :

$$(a+b)(a+b) = a^2 + ab + ba + b^2 \neq a^2 + 2ab + b^2,$$

si $ab \neq ba$. Cette situation peut se produire notamment dans le cadre du produit matriciel. Il faut être toujours bien vigilant à vérifier l'hypothèse de commutativité $ab = ba$.

IV.4 Éléments inversibles

Un anneau n'étant pas nécessairement commutatif, il convient de distinguer la notion d'inversibilité à droite, inversibilité à gauche. Un inverse est alors à la fois un inverse à gauche et à droite, et en cas d'existence, il est unique, comme nous l'avons montré dans une situation générale. De plus, dans le cas d'un anneau, l'ensemble des éléments inversibles possède une structure particulière.

Théorème 2.4.25 (Groupe des inversibles d'un anneau)

Soit A un anneau. Alors l'ensemble des éléments inversibles de A, généralement noté A^* ou $U(A)$, est stable pour la loi $\times$, et la loi induite munit A^* d'une structure de groupe multiplicatif.

Remarque 2.4.26

Remarquez la cohérence avec les notations déjà rencontrées $\mathbb{R}^*$, $\mathbb{Q}^*$, $\mathbb{C}^*$... mais pas avec $\mathbb{Z}^*$.

Exemples 2.4.27

1. $(\mathcal{M}_n(\mathbb{R}))^* = \mathrm{GL}_n(\mathbb{R})$, ensemble des matrices inversibles, appelé *groupe linaire*
2. L'ensemble des inversibles de $\mathbb{Z}$: $\{-1, 1\}$
3. $(\mathbb{Z}/4\mathbb{Z})^* = \{1, 3\}$
4. $(\mathbb{Z}/p\mathbb{Z})^* = (\mathbb{Z}/p\mathbb{Z}) \setminus \{0\}$, si p est premier. On peut montrer que ce groupe multiplicatif est isomorphe au groupe additif $\mathbb{Z}/(p-1)\mathbb{Z}$.
5. Que dire plus généralement de $(\mathbb{Z}/n\mathbb{Z})^*$?

L'étude de $\mathbb{Z}/4\mathbb{Z}$ amène un résultat peu commun pour qui n'a pas l'habitude de travailler dans des structures algébriques abstraites : $2 \times 2 = 0$. Autrement dit, on a deux éléments a et b non nuls, et vérifiant $ab = 0$. La vieille règle, bien pratique pour résoudre des équations, qui nous dit que si $ab = 0$, alors $a = 0$ ou $b = 0$, ne s'applique donc pas dans ce contexte. Comme elle est bien pratique tout de même, nous allons établir un contexte dans lequel elle est vraie, en définissant une propriété adéquate des anneaux nous permettant de l'utiliser.

Définition 2.4.28 (Diviseurs de zéro, HP)

Soit a un élément d'un anneau A. On dit que a est un diviseur de 0 à gauche si $a \neq 0$ et s'il existe $b \in A$, $b \neq 0$, tel que $ab = 0$. On définit de façon symétrique les diviseurs de zéro à droite.

Remarque 2.4.29

La notion de diviseur de 0 caractérise en fait la non régularité : un élément a non nul d'un anneau est régulier à droite si et seulement s'il n'est pas diviseur de 0.

Définition 2.4.30 (anneau intègre, HP)

Un anneau A est dit intègre s'il est non nul, et s'il n'admet aucun diviseur de zéro (ni à gauche, ni à droite).

En particulier, dans un anneau intègre, toutes les simplifications par des éléments non nuls sont possibles, puisque le seul élément non régulier est 0.

Exemples 2.4.31

1. $\mathbb{Z}$ est intègre, $\mathbb{R}[X]$ est intègre, tout corps est intègre.
2. $\mathcal{M}_n(\mathbb{R})$ n'est pas intègre.
3. À quelle condition sur n, $\mathbb{Z}/n\mathbb{Z}$ est-il intègre ?

IV.5 Idéaux (HP)

La notion de sous-anneau est souvent trop restrictive, et on est souvent amené à considérer une structure moins riche :

Définition 2.4.32 (Idéal d'un anneau commutatif, HP)

Soit A un anneau commutatif, et I un sous-ensemble de A. On dit que I est un idéal si et seulement si I est un sous-groupe du groupe $(A, +)$ et si pour tout $a \in I$ et tout $\lambda \in A$, $\lambda a \in I$.

Ainsi, I est un sous-groupe de $(A, +)$, stable par multiplication par un élément de A.
Nous n'étudierons pas les idéaux cette année, mais nous illuminerons parfois quelques résultats à l'éclat de cette notion, notamment en arithmétique. Nous donnons tout de même quelques exemples importants :

Exemples 2.4.33

1. Pour tout $n \in \mathbb{N}^*$, $n\mathbb{Z}$ est un idéal de $\mathbb{Z}$.
2. L'ensemble des polynômes de $\mathbb{R}[X]$ s'annulant en 0 est un idéal de $\mathbb{R}[X]$. Comment généraliseriez-vous ce résultat ?
3. L'ensemble des polynômes $\{XP(X,Y) + YQ(X,Y), \quad (P,Q) \in \mathbb{R}[X,Y]\}$ est un idéal de l'anneau $\mathbb{R}[X,Y]$ des polynômes à deux indéterminées.
4. Que peut-on dire des idéaux d'un corps ?

Dans les deux premiers exemples, on constate que l'idéal considéré est de la forme $\{\lambda a, \lambda \in A\}$, donc engendré par un unique élément a, par multiplication par les éléments λ de A. Un idéal vérifiant cette propriété est appelé *idéal principal*. Tout idéal n'est pas principal, comme le montre le troisième exemple. Un anneau dont tous les idéaux sont principaux est appelé *anneau principal*. C'est le cas par exemple de $\mathbb{Z}$. Cette définition, qui peut paraître anodine, est à la base d'une généralisation possible de la notion de pgcd et de ppcm à des anneaux autres que $\mathbb{Z}$, comme vous le verrez l'an prochain.

3 Arithmétique des entiers

Note Historique 3.0.34

- L'arithmétique désigne dans un premier temps l'étude des opérations élémentaires entre entiers (arithmétique élémentaire), et les algorithmes permettant de faire ces opérations (algorithmes de multiplication, division euclidienne...). C'est une des disciplines fondamentales des mathématiques dans le sens où, avec la géométrie et le calcul numérique (algébrique), elle constitue le point de départ de toutes les mathématiques.
- En découlent de façon naturelle (et déjà en Grèce antique) l'étude des propriétés de divisibilité, et donc de primalité.
- Plus généralement, l'arithmétique déigne l'étude de problèmes relatifs à des nombres entiers. Ainsi, les problèmes de Diophante, liés à la recherche de solutions entières d'équations relèvent de l'arithmétique. L'exemple le plus célèbre en est certainement le fameux théorème de Fermat-Wiles stipulant que pour tout $n \geqslant 3$, il n'existe pas de triplet (a, b, c) d'entiers naturels non nuls tels que $a^n + b^n = c^n$. Il est instructif d'ailleurs de noter que l'ouvrage dans lequel Pierre de Fermat a écrit en marge sa fameuse note concernant une preuve de ce résultat n'est autre que *Arithmetica* de Diophante.
- L'exemple du théorème de Fermat-Wiles justifie la nécessité de sortir du cadre des entiers pour résoudre des problèmes arithmétiques en apparence simples. Ainsi, l'arithmétique a évolué en diverses branches (théorie algébrique des nombres, théorie analytique des nombres, géométrie algébrique...)
- Les concepts essentiels de l'arithmétique ont également été généralisés dans des contextes différents de celui des entiers. C'est une des motivations de l'introduction de la notion d'anneau et d'idéal. Un exemple que vous aurez l'occasion d'étudier prochainement est l'étude de l'arithmétique des polynômes. Mais cela ne s'arrête pas là !

I Divisibilité, nombres premiers

I.1 Notion de divisibilité

Définition 3.1.1 (Divisibilité, diviseur, multiple)

- Soit a et b deux entiers relatifs, $b \neq 0$. On dit que b *divise* a, et on écrit $b \mid a$, si et seulement s'il existe $q \in \mathbb{Z}$ tel que $a = bq$.
- On dit dans ce cas que b est un *diviseur* de a, et que a est un *multiple* de b.

On note $a \mid b$ pour dire que a divise b.
Ainsi, $2 \mid 4$, $-2 \mid 4$, $2 \mid -4$, et $-2 \mid -4$.

Proposition 3.1.2 (Caractérisation en termes d'idéaux)

Soit a et b deux entiers positifs, $a \neq 0$. Alors $a \mid b$ si et seulement si $b\mathbb{Z} \subset a\mathbb{Z}$.

Définition 3.1.3 (couple d'entiers associés)
On dit que deux entiers a et b sont associés si et seulement si $a \mid b$ et $b \mid a$.

Proposition 3.1.4 (Caractérisation des entiers associés)
Les entiers a et b sont associés si et seulement si il existe $\varepsilon \in \{-1, 1\}$ tel que $a = \varepsilon b$.

Ce résultat peut sembler trivial et sans intérêt. Sa version plus générale, pour un anneau intègre A, est plus intéressante, et affirme que les éléments associés diffèrent d'une constante mutliplicative appartenant au groupe A^* des inversibles de A.

Théorème/Définition 3.1.5 (Théorème de la division euclidienne)
Soit $(a, b) \in \mathbb{Z}^2$, $b \neq 0$.
- *Il existe un unique couple $(q, r) \in \mathbb{Z} \times \mathbb{N}$ tel que :*

$$a = bq + r \qquad et \qquad 0 \leqslant r < |b|.$$

- *L'entier q est appelé quotient de la division euclidienne de a par b.*
- *L'entier r est appelé reste de la division euclidienne de a par b.*

Remarquez que b peut être négatif.

Exemples 3.1.6

1. $27 = 6 \times 4 + 3 = 6 \times 3 + 9 = 6 \times 6 - 3$.
   Ainsi, des identités $a = bq + r$, il y en a beaucoup, mais une seule vérifie la condition imposée sur r. Ici, le quotient de la division de 27 par 6 est 4, et son reste est 3.
2. $27 = (-6) \times (-4) + 3 = (-6) \times (-5) - 3$.
   Sans la valeur absolue dans la condition sur r, c'est la deuxième égalité qui aurait été la bonne. Mais la valeur absolue impose un reste positif. Ainsi, le quotient de la division de 27 par -6 est -4, et le reste est 3
3. $-27 = 6 \times (-4) - 3 = 6 \times (-5) + 3$.
   Ici, on voit que si on change le signe du nombre divisé, le quotient n'est pas simplement l'opposé (attention, cela ne correspond pas à la plupart des implémentations informatiques de la division euclidienne). Ainsi, la première identité ne convient pas. Le quotient de la division euclidienne de -27 par 6 est -5, le reste est 3.
   Remarquez que la situation est la même que pour la partie entière, pour laquelle $\lfloor -x \rfloor \neq -\lfloor x \rfloor$, sauf lorsque x est entier. C'est normale, puisque la partie entière n'est autre que le quotient de la division euclidienne (réelle) par 1.
4. $-27 = (-6) \times 5 + 3$.
   Sans surprise, le quotient de la division euclidienne de -27 par -6 est 5, le reste est 3.

La plupart des propriétés arithmétiques de $\mathbb{Z}$ (pour ne pas dire toutes) découlent de l'existence de cette division euclidienne. On peut définir de façon similaire dans certains anneaux une division euclidienne, la condition sur le reste étant un peu plus dure à exprimer. On parle dans ce cas d'anneau euclidien.
Ainsi, $\mathbb{Z}$ est un anneau euclidien. On verra un peu plus tard que $\mathbb{R}[X]$ est un anneau euclidien, ce qui nous permettra d'établir un certain nombre de propriétés arithmétiques pour les polynômes.

Note Historique 3.1.7

La division euclidienne est appelée ainsi par référence à Euclide qui décrit dans ses éléments le procédé algorithmique de soustractions répétées permettant d'obtenir le quotient et le reste. Cependant, on trouve trace de cette notion à des époques antérieures, notamment en Égypte.

I.2 Congruences

De façon quasi-indissociable de la notion de division euclidienne, nous définissons :

Définition 3.1.8 (Congruences d'entiers)

Soit $n \in \mathbb{N}^*$, et $(a, b) \in \mathbb{Z}^2$. On dit que a et b sont *congrus* modulo n, et on écrit $a \equiv b\ [n]$, si et seulement si n divise $b - a$, ou encore si les divisions euclidiennes de a et b par n ont même reste.

On trouve aussi assez souvent la notation $a \equiv b \mod n$, ou un mélange des 2 : $a \equiv b\ [\mathrm{mod}\ n]$.
Nous rappelons les résultats suivants, que nous avons déjà eu l'occasion de démontrer.

Théorème 3.1.9

La relation de congruence modulo n est une relation d'équivalence.

Théorème 3.1.10

La relation de congruence modulo n est compatible avec le produit et la somme : soit $(a, a', b, b') \in \mathbb{Z}^4$ tels que $a \equiv a'\ [n]$ et $b \equiv b'\ [n]$. Alors $a + b \equiv a' + b'\ [n]$ et $ab \equiv a'b'\ [n]$

En d'autre terme, c'est une congruence sur les monoïdes $(\mathbb{Z}, +)$ et $(\mathbb{Z}, \times)$, au sens vu dans le chapitre sur les ensembles.
Ces règles sont importantes pour pouvoir mener à bien le calcul modulaire de façon efficace : il permet de faire lors d'une succession d'opérations, des réductions modulo n étape par étape, plutôt que de tout calculer dans $\mathbb{N}$ et de réduire à la fin.

Exemples 3.1.11

- Calculer le reste de la division euclidienne de $12 \times 21 \times 28 \times 18 \times 75 \times 23$ par 11.
- Calculer le reste de la division euclidienne de 1685^{1750} par 12.

Cette possibilité de réduire les opérations à chaque étape est également important pour l'implémentation informatique du calcul modulaire, permettant ainsi de travailler avec des entiers plus petit, diminuant de la sorte la complexité des calculs. On peut ainsi, contrairement au cas du calcul dans $\mathbb{Z}$, borner explicitement le temps de calcul des opérations modulo n par un réel dépendant de n mais ne dépendant pas des opérandes.

I.3 Nombres premiers

Nous les avons déjà rencontrés, évidemment. Nous rappelons :

Définition 3.1.12 (Nombres premiers)

Soit $p \in \mathbb{N}^*$. On dit que p est un nombre premier si p admet exactement 2 diviseurs positifs distincts (à savoir 1 et p lui-même)

Remarquez que l'existence de deux diviseurs distincts exclut d'office 1 de l'ensemble des nombres premiers, puisqu'il n'a qu'un diviseur.

Définition 3.1.13 (Nombres composés)

Soit $n \in \mathbb{N}^*$. On dit que n est un nombre composé si n possède au moins 3 diviseurs positifs distincts, ou en d'autres termes, si n possède un diviseur positif distinct de 1 et de n.

Remarque 3.1.14 (La blague du matheux revisitée)

Il y a trois types d'entiers naturels non nuls : les entiers premiers et les entiers composés.
J'en ai oublié un, lequel ?

Proposition 3.1.15

Tout nombre composé admet un diviseur strict premier.

Théorème 3.1.16 (Combien de nombres premiers ?)

Il y a une infinité de nombres premiers.

C'est bien joli tout ça, mais comment faire pour déterminer les nombres premiers (pas trop gros) ? Erathostène, mathématicien, astronome, bibliothécaire en chef d'Alexandrie (excusez du peu), astéroïde et cratère lunaire, répondit à cette question il y a déjà très longtemps, par un procédé d'élimination.

Méthode 3.1.17 (Crible d'Érathostène)

Pour trouver tous les nombres premiers inférieurs ou égaux à n :

1. Écrire tous les nombres entiers de 2 à n.
2. Le plus petit d'eux, à savoir 2, est premier (il n'a pas de diviseur strictement plus petit que lui, autre que 1)
3. Les multiples stricts de 2 ne sont pas premiers, on les barre tous.
4. Parmi les nombres restants (en excluant les nombres premiers précédents, à savoir 2 dans la première étape, et en excluant les nombres barrés), le plus petit est premier (il n'est divisible par aucun nombre premier strictement plus petit que lui et différent de 1, sinon il serait barré). On barre tous ses multiples stricts qui ne peuvent pas être premiers, et on recommence cette étape jusqu'à épuisement de tous les entiers de la liste.

Cet algorithme est très facile à implémenter dans un langage informatique. Il n'est évidemment efficace que pour des petites valeurs de n, mais ne peut pas servir à la recherche de très grands nombres premiers. Notamment, il est à peu près inutilisable pour répondre à la question de savoir si un très grand nombre donné est premier ou non (question cruciale dans certaines situations en rapport avec des cryptages, comme la méthode RSA).

II Arithmétique d'un couple d'entiers

II.1 PGCD et PPCM

Proposition/Définition 3.2.1 (PGCD, PPCM)

Soit a et b deux entiers positifs.

1. Si, l'un au moins des entiers a et b est non nul, l'ensemble $\{d \in \mathbb{N}^* \mid d \text{ divise } a \text{ et } d \text{ divise } b\}$ admet un élément maximum (pour l'ordre usuel). On appelle cet élément maximum « plus grand commun diviseur » de a et b (en abrégé : PGCD), et on le note $a \wedge b$.
2. Si a et b sont non nuls, l'ensemble $\{m \in \mathbb{N}^* \mid a \text{ divise } m \text{ et } b \text{ divise } m\}$ admet un élément maximum. On appelle cet élément maximum « plus petit commun multiple » de a et b (en abrégé : PPCM), et on le note $a \vee b$.

Proposition 3.2.2 (Caractérisation de $a \wedge b$ par les idéaux de $\mathbb{Z}$)

L'idéal $(a \wedge b)\mathbb{Z}$ est le plus petit idéal (au sens de l'inclusion) contenant à la fois a et b, à savoir :

$$(a \wedge b)\mathbb{Z} = a\mathbb{Z} + b\mathbb{Z}.$$

Remarquez qu'on se sert pour prouver ce résultat du fait que tout idéal de $\mathbb{Z}$ s'écrit $\mathbb{Z} \cdot a$, donc que $\mathbb{Z}$ est principal. Le fait que $\mathbb{Z}$ est principal nous assure également que le plus petit idéal contenant a et b est engendré par un élément. C'est une façon de définir le pgcd, comme élément générateur de l'idéal engendré par a et b. Cette définition est valide dans tout anneau principal. Remarquez que le pgcd n'est alors défini qu'à multiplication près par un élément inversible de A.

Proposition 3.2.3 (Caractérisation de $a \vee b$ par les idéaux de $\mathbb{Z}$)

L'idéal $(a \vee b)\mathbb{Z}$ est le plus petit idéal contenant à la fois $a\mathbb{Z}$ et $b\mathbb{Z}$:

$$(a \vee b)\mathbb{Z} = a\mathbb{Z} \cap b\mathbb{Z}.$$

La encore, cette caractérisation permet de généraliser à tout anneau principal la notion de pgcd (défini à élément inversible près).

Proposition 3.2.4 (Caractérisation du PCGD par l'ordre de divisibilité)

Soit a et b deux entiers strictement positifs. Alors

1. • *$a \wedge b$ est le maximum de $\{d \in \mathbb{N}^* \mid d$ divise a et d divise $b\}$ pour l'ordre de divisibilité.*
   • *Autrement dit, si $x \in \mathbb{N}^*$ vérifie $x \mid a$ et $x \mid b$, alors $x \mid a \wedge b$;*
   • *Autrement dit : $a \wedge b = \inf_{(\mathbb{N},|)}(a, b)$*
2. • *$a \vee b$ est le minimum de $\{d \in \mathbb{N}^* \mid d$ divise a et d divise $b\}$ pour l'ordre de divisibilité.*
   • *Autrement dit, si $x \in \mathbb{N}^*$ vérifie $a \mid x$ et $b \mid x$, alors $a \vee b \mid x$.*
   • *Autrement dit : $a \vee b = \sup_{(\mathbb{N},|)}(a, b)$*

On en déduit une des premières règles sur le calcul des pgcd :

Proposition 3.2.5 (Distributivité du produit sur $\wedge$ et $\vee$)

Soit a et b deux entiers naturels, et c un entier naturel non nul.

1. *Si a et b ne sont pas tous les deux nuls, $(a \wedge b) \cdot c = (ac) \wedge (bc)$.*
2. *Si a et b sont non nuls, $(a \vee b) \cdot c = (ac) \vee (bc)$.*

On déduit de la caractérisation par les idéaux l'identité de Bézout :

Théorème 3.2.6 (identité de Bézout, ou théorème de Bachet-Bézout)

1. *Soit a et b deux entiers dont l'un au moins est non nul. Alors il existe des entiers relatifs x et y tels que*
$$ax + by = a \wedge b.$$
2. *Réciproquement, étant donné un entier $d \in \mathbb{N}^*$, s'il existe des entiers relatifs x et y tels que*
$$d = ax + by,$$
alors $a \wedge b \mid d$.

Note Historique 3.2.7

- C'est le nom d'Étienne Bézout, mathématicien français du 18e siècle, qui est le plus souvent associé à ce résultat. C'est pourtant à Claude-Gaspard Bachet de Méziriac que l'on doit la première preuve, parue dans son ouvrage *Problèmes plaisans et délectables qui se font par les nombres*, paru en 1624. Sa preuve est celle que nous présentons ci-dessous (par l'algorithme d'Euclide)
- Qu'a fait Bézout alors pour avoir droit à tous ces honneurs ? Il a généralisé le résultat à d'autres situations, notamment au cas des polynômes.
- Il est intéressant de noter que le fameux ouvrage dans lequel Fermat écrivit dans une marge qu'il savait démontrer ce qu'on appelle aujourd'hui le théorème de Fermat-Wiles est en fait une traduction par Bachet de Mériziac de l'*Arithmétique* de Diophante. Le monde est petit...

La démonstration passant par les idéaux peut se généraliser dans un anneau principal. Elle possède l'inconvénient de ne pas être constructive. Il peut être intéressant de trouver explicitement des entiers x et y assurant l'égalité $ax + by = a \wedge b$. L'algorithme de la division euclidienne itéré permet à la fois de déterminer $a \wedge b$, et d'obtenir une identité de Bézout : il s'agit de l'algorithme d'Euclide. Il est intéressant de noter que cet algorithme est valide à partir du moment où l'on dispose d'une notion de division euclidienne : il peut donc être généralisé à tout anneau euclidien, dans le sens évoqué précédemment. Ainsi, par exemple, il nous permettra d'obtenir des identités de Bézout dans $\mathbb{R}[X]$.

Lemme 3.2.8

Soit a et b deux entiers positifs, $b \neq 0$. Soit r le reste de la division euclidienne de a par b. Alors $a \wedge b = b \wedge r$.

Théorème 3.2.9 (Algorithme d'Euclide)

- *Soit a et b deux entiers positifs, $b \neq 0$. En effectuant la division euclidienne de a par b, puis en effectuant la division euclidienne de b par le reste obtenue, et en continuant ainsi en divisant l'ancien reste par le nouveau reste, on finit par obtenir un reste nul.*
- *Le dernier reste non nul est alors égal au PGCD de a et de b.*
- *L'identité de la division euclidienne permet alors d'écrire, étape par étape, les restes successifs comme combinaison linéaire de a et b à coefficients dans $\mathbb{Z}$. La dernière étape fournit une identité de Bézout.*

Ainsi, en écrivant $r_0 = a$, $r_1 = b$ puis les divisions euclidiennes successives :

$$\begin{cases} r_0 & = r_1 q_2 + r_2 \\ r_1 & = r_2 q_3 + r_3 \\ \vdots & \\ r_{k-2} & = r_{k-1} q_k + r_k \\ r_{k-1} & = r_k q_{k+1} + r_{k+1}, \end{cases}$$

avec $r_2 \neq 0$, $r_3 \neq 0, \dots, r_k \neq 0$ et $r_{k+1} = 0$, on a $r_k = a \wedge b$. De plus, en posant $x_0 = 1$, $x_1 = 0$, $y_0 = 0$, $y_1 = 1$ et pour tout $i \leqslant [\![3, k]\!]$

$$x_i = x_{i-2} - q_i x_{i-1} \qquad \text{et} \qquad y_i = y_{i-2} - q_i y_i,$$

on obtient pour tout $i \in [\![1, n]\!]$,

$$r_i = a x_i + b y_i,$$

donc en particulier pour $i = k$, on obtient une identité de Bézout :

$$a \wedge b = a x_k + b y_k.$$

Exemple 3.2.10

- Trouver à l'aide de l'algorithme d'Euclide le pgcd de 27 et 33, ainsi qu'une identité de Bézout.
- Comment trouver une autre identité de Bézout ?
- À retenir : on n'a pas unicité de la relation de Bézout !

La notion de PGCD de deux entiers peut être généralisée à un plus grand nombre d'entiers :

Définition 3.2.11 (PGCD et PPCM d'un nombre fini d'entiers)

Soit $a_1, \dots, a_n$ des entiers naturels.

1. Si au moins un des a_i est non nul, disons a_1, le PGCD des entiers $a_1, \dots, a_n$ est le plus grand (au sens de l'ordre usuel) des entiers d qui divisent chacun des a_i, $i \in [\![1, n]\!]$.
2. Si les a_i sont tous non nuls, le PPCM des entiers $a_1, \dots, a_n$ est le plus petit (au sens de l'ordre usuel) des entiers m qui sont multiples de chacun des a_i, $i \in [\![1, n]\!]$.

Comme pour le cas de deux entiers, nous obtenons :

Proposition 3.2.12 (caractérisation du PGCD et PPCM par la divisibilité)

- *Les conditions de minimalité définissant le PGCD et le PPCM peuvent également être prises au sens de la divisibilité. Ainsi, d_0 est le pgcd de $(a_1, \dots, a_n)$ si et seulement si pour tout $i \in [\![1, n]\!]$, $d_0 \mid a_i$, et si pour tout $d \in \mathbb{N}^*$,*
$$(\forall i \in [\![1, n]\!],\ d \mid a_i) \Longrightarrow d \mid d_0.$$
- *Un énoncé similaire vaut pour le PPCM.*
- *On note $a_1 \wedge a_2 \wedge \dots \wedge a_n$ le ppcm de $a_1, \dots, a_n$ et $a_1 \vee a_2 \vee \dots \vee a_n$ le pgcd de $a_1, \dots, a_n$.*

Remarquez que cette notation se distingue *a priori* de l'itération du pgcd de deux entiers par l'absence de parenthésage. Mais cette distinction est assez peu important, du fait de la proposition suivante :

Proposition 3.2.13

Soit $a_1, \dots, a_n$ des entiers naturels.

1. *Si $a_1 \neq 0$, alors $a_1 \wedge \dots \wedge a_n = ((a_1 \wedge a_2) \wedge \dots) \wedge a_n$.*
2. *Si les a_i sont non nuls, alors $a_1 \vee \dots \vee a_n = ((a_1 \vee a_2) \vee \dots) \vee a_n$.*

Corollaire 3.2.14 (Associativité de $\wedge$ et $\vee$)

Les lois de composition définies sur $\mathbb{N}^*$ par $\wedge$ et $\vee$ sont associatives (et commutatives).

On peut étendre le théorème de Bachet-Bézout à cette situation :

Théorème 3.2.15 (Relation de Bézout)

Soit $a_1, \dots, a_n$ des entiers naturels non tous nuls. Alors il existe des entiers relatifs $x_1, \dots, x_n$ tels que

$$a_1 \wedge \cdots \wedge a_n = x_1 a_1 + \cdots + x_n a_n.$$

Réciproquement, s'il existe des entiers $x_1, \dots, x_n$ tels que

$$d = x_1 a_1 + \cdots + x_n a_n,$$

alors d est un multiple de $a_1 \wedge \cdots \wedge a_n$.

Méthode 3.2.16

Les coefficients $x_1, \dots, x_n$ peuvent se trouver explicitement, par itération de l'algorithme d'Euclide : on cherche d'abord une relation de Bézout entre $d_1 = a_1 \wedge a_2$, a_1 et a_2, puis entre $d_2 = d_1 \wedge a_3$, d_1 et a_2 ; en substituant à d_1 la première relation trouvée, on obtient une relation de Bézout entre $a_1 \wedge a_2 \wedge a_3$, a_1, a_2 et a_3. On continue alors de la sorte, de proche en proche.

Enfin, toutes les notions introduites dans ce paragraphe peuvent être généralisées à des entiers relatifs quelconques ; le pgcd et le ppcm ne sont alors définis correctement qu'au signe près (c'est le cas général dans un anneau principal, ou le pgcd ne peut être déterminé qu'à un facteur multiplicatif inversible près). Dans le cas de $\mathbb{Z}$, on a un choix privilégié qui consiste à prendre la valeur positive. Le pgcd et les relations de Bézout se trouvent de la même façon, en les cherchant d'abord pour les valeurs absolues, puis en modifiant les signes de façon adéquate.

II.2 Entiers premiers entre eux

Définition 3.2.17 (Entiers premiers entre eux)

Soit a et b deux entiers naturels non tous les deux nuls. On dit que a et b sont premiers entre eux si et seulement si $a \wedge b = 1$, donc si a et b n'ont pas d'autre diviseur positif commun que 1.

Cela peut aussi s'exprimer par la relation $a\mathbb{Z} + b\mathbb{Z} = \mathbb{Z}$.

Note Historique 3.2.18

La première apparition de cette notion est dans le Livre VII des *Éléments* d'Euclide.

Proposition 3.2.19 (Simplification des fractions)

- *Soit a et b deux entiers naturels, $b \neq 0$. Alors $\frac{a}{a \wedge b}$ et $\frac{b}{a \wedge b}$ sont premiers entre eux.*
- *En particulier, il est toujours possible d'écrire un rationnel $\frac{a}{b}$ sous forme irréductible $\frac{a'}{b'}$, c'est-à-dire de sorte que $a' \wedge b' = 1$, en simplifiant par $a \wedge b$.*

On déduit des résultats de la section précédente :

Théorème 3.2.20 (Bézout, ou Bachet-Bézout)

Deux entiers naturels a et b sont premiers entre eux si et seulement s'il existe des entiers relatifs x et y tels que $ax + by = 1$.

En particulier :

Corollaire 3.2.21 (Éléments inversibles de $\mathbb{Z}/n\mathbb{Z}$, HP)

1. *Soit $n \in \mathbb{N}^*$, et $k \in [\![0, n-1]\!]$. La classe de k modulo n est inversible dans $\mathbb{Z}/n\mathbb{Z}$ si et seulement si k et n sont premiers entre eux.*
2. *En particulier, si p est premier, $\mathbb{Z}/p\mathbb{Z}$ et un corps.*

Méthode 3.2.22 (Calcul d'un inverse modulo n)

Soit k premier avec n. Pour calculer l'inverse de k modulo n, c'est-à-dire l'inverse de k dans $\mathbb{Z}/n\mathbb{Z}$, déterminer une relation de Bézout $xk+yn=1$, par l'algorithme d'Euclide. On obtient alors $xk \equiv 1\ [n]$.

Un résultat souvent très utile pour les propriétés de divisibilité :

Lemme 3.2.23 (Lemme ou théorème de Gauss)

Soit a, b et c trois entiers naturels tels que $a \mid bc$ et $a \wedge b = 1$. Alors $a \mid c$.

La première version de ce lemme, moins forte, est due à Euclide, qui s'en sert pour montrer l'unicité de la décomposition primaire. C'est un cas particulier du lemme de Gauss, qui peut être démontré séparément, par exemple par un argument basé sur la division euclidienne, ou encore par utilisation d'une relation de Bézout.

Lemme 3.2.24 (Lemme d'Euclide)

Soit b et c deux entiers naturels non nuls, et p un nombre premier. Alors si p divise bc, alors p divise b ou p divise c.

On peut notamment en tirer une formule reliant de façon simple le PGCD et le PPCM. Cette formule sera rendue encore plus limpide lorsqu'on aura la description par la décomposition primaire du PGCD et du PPCM.

Proposition 3.2.25

Soit a et b deux entiers strictement positifs. Alors $ab = (a \wedge b)(a \vee b)$.

Enfin, nous définissons deux notions sur un nombre quelconque d'entiers, à bien distinguer l'une de l'autre :

Définition 3.2.26 (Nombres premiers entre eux deux à deux)

Soit $a_1, \dots, a_n$ des entiers naturels. On dit que $a_1, \dots, a_n$ sont premiers entre eux deux à deux si deux entiers pris au hasard parmi ces n entiers sont toujours premiers entre eux, c'est-à-dire :

$$\forall (i,j) \in [\![1,n]\!]^2, \quad (i \neq j) \Longrightarrow a_i \wedge a_j = 1.$$

La notion suivante est moins forte :

Définition 3.2.27 (Nombres premiers entre eux dans leur ensemble)

Soit $a_1, \dots, a_n$ des entiers naturels. On dit que $(a_1, \dots, a_n)$ sont premiers entre eux dans leur ensemble si $a_1 \wedge \cdots \wedge a_n = 1$, ou de façon équivalente, s'il existe des entiers $x_1, \dots, x_n$ tels que

$$x_1 a_1 + \cdots + x_n a_n = 1.$$

Par exemple 10, 12 et 15 sont premiers entre eux dans leur ensemble. Vous remarquerez en revanche que deux quelconques d'entre eux ne sont pas premiers entre eux !
La reciproque, en revanche est vraie :

Proposition 3.2.28

Soit $(a_1, \dots, a_n) \in \mathbb{N}^n$. Si $a_1, \dots, a_n$ sont premiers entre eux deux à deux (il suffit même en fait que deux d'entre eux soient premiers entre eux) alors ils sont premiers entre eux dans leur ensemble.

II.3 Fonction indicatrice d'Euler

Arrivé à ce stade, nous ne pouvons nous empêcher de parler de la fonction indicatrice d'Euler, d'une importance capitale en arithmétique.

Définition 3.2.29 (Fonction indicatrice d'Euler, ou fonction phi d'Euler)

La fonction φ d'Euler est la fonction qui à tout n de $\mathbb{N}^*$ associe $\varphi(n)$ le nombre d'entiers de $[\![1, n-1]\!]$ premiers avec n.

En particulier, les résultats précédents amènent :

Proposition 3.2.30 (Cardinal de $(\mathbb{Z}/n\mathbb{Z})^*$, HP)

Soit $n \in \mathbb{N}^*$. Alors $(\mathbb{Z}/n\mathbb{Z})^*$ est de cardinal $\varphi(n)$.

On peut montrer (voir exercices ou problèmes) que φ est multiplicative : $\varphi(ab) = \varphi(a)\varphi(b)$. On peut également calculer facilement $\varphi(p^k)$, pour p premier. On peut en déduire une expression de $\varphi(n)$ pour tout n, à condition de savoir écrire n comme produit de puissances de nombres premiers. Cela fait l'objet de la section suivante.

III Décomposition primaire

III.1 Décomposition primaire et valuations p-adiques

Un théorème incournable de l'arithmétique est bien sûr :

Théorème 3.3.1 (Décomposition primaire)

Tout entier strictement positif n s'écrit de façon unique sous la forme

$$n = p_1 \times \cdots \times p_k,$$

où $p_1 \leqslant \cdots \leqslant p_k$ sont des nombres premiers, ce produit étant éventuellement vide si $n = 1$.

Un nombre premier p pouvant apparaître plusieurs fois dans la décomposition de n, nous définissons :

Définition 3.3.2 (Valuation p-adique)

Soit n un entier et p un entier premier. On appelle valuation p-adique de l'entier n le nombre d'occurrences (éventuellement nul) de l'entier p dans la décomposition primaire de n.

En notant $\mathbb{P}$ l'ensemble des nombres premiers, il vient donc :

Proposition 3.3.3 (Reexpression de la décomposition primaire)

Pour tout $n \in \mathbb{N}^*$

$$n = \prod_{p \in \mathbb{P}} p^{v_p(n)},$$

ce produit ayant un sens, puisque constitué d'un nombre fini de termes non égaux à 1.

Un anneau dans lequel on peut définir une notion d'éléments premiers, et dans lequel on a une propriété d'existence et d'unicité (à facteurs multiplicatifs inversibles près, et à l'ordre près des facteurs) d'une décomposition de ce type est appelé *anneau factoriel*. Ainsi, quitte à multiplier par l'élément inversible -1 pour obtenir la décomposition d'un entier relatif, ce résultat se réexprime en disant que $\mathbb{Z}$ est un anneau factoriel.

Proposition 3.3.4 (Règles sur les valuations)

Soit a et b deux entiers strictement positifs, et p un nombre premier.

1. *On a : $v_p(ab) = v_p(a) + v_p(b)$.*
2. *Si b divise a, on a : $v_p\left(\frac{a}{b}\right) = v_p(b) - v_p(a)$.*

Exemples 3.3.5

1. Déterminer, pour p premier, $v_p((p^2)!)$, et plus généralement $v_p((p^k)!)$, puis plus généralement $v_p(n!)$ (formule de Legendre)
2. Déterminer $v_p\left(\binom{n}{k}\right)$ en fonction de n et k.

On obtient en particulier le :

Lemme 3.3.6

Soit p un nombre premier. Alors pour tout $k \in [\![1, p-1]\!]$, $\binom{p}{k} \equiv 0\ [p]$.

De ce lemme, on tire :

Proposition 3.3.7

Soit a et b deux entiers et p un nombre premier. Alors $(a+b)^p \equiv a^p + b^p\ [p]$.

On en déduit le résultat important suivant :

Théorème 3.3.8 (Petit théorème de Fermat)

- *Soit p un nombre premier, et $x \in \mathbb{Z}$. Alors : $x^p \equiv x\ [p]$.*
- *Si de plus, $x \not\equiv 0\ [p]$, alors $x^{p-1} \equiv 1\ [p]$.*

Reamrquez que ce résultat équivaut à dire que dans le corps $\mathbb{F}_p$, tout élément x non nul vérifie $x^{p-1} = 1$, ce qui est une conséquence immédiate du théorème de Lagrange, puisque $\mathbb{F}_p^*$ est un groupe d'ordre $p-1$ (seul l'élément nul étant non inversible).

Remarque 3.3.9

Le petit de théorème de Fermat est notamment beaucoup utilisé dans les tests de non primalité (avec un ordinateur !). En effet, pour montrer qu'un entier p n'est pas premier, il suffit de trouver un entier

a tel que $a^p \not\equiv a\ [p]$. Ainsi, par exemple, à l'aide d'un ordinateur, on peut trouver facilement, pour $n = \frac{1}{9}(10^{31} - 1)$ (nombre contitué de 31 chiffres 1) que $2^n \not\equiv 2\ [n]$. Ainsi, n n'est pas premier. Trouver une décomposition de n est une autre paire de manches...

III.2 Décomposition primaire, pgcd et ppcm

Étant donné deux nombres a et b, le pgcd et le ppcm de a et b s'obtiennent facilement à l'aide de leur décomposition primaire. Ainsi

$$150 = 2 \times 3 \times 5^2 \qquad \text{et} \qquad 180 = 2^2 \times 3^2 \times 5.$$

Ainsi, $150 \wedge 180 = 2 \times 3 \times 5 = 30$ et $150 \vee 180 = 2^2 \times 3^2 \times 5^2 = 900$.
Plus généralement :

Proposition 3.3.10
Soit a et b deux entiers strictement positifs. Alors, pour tout $p \in \mathbb{P}$,

$$v_p(a \wedge b) = \min(v_p(a), v_p(b)) \qquad \text{et} \qquad v_p(a \vee b) = \max(v_p(a), v_p(b)).$$

III.3 Un peu de cryptographie, HP

Il serait faux de croire que toutes les notions abstraites abordées dans ce chapitre sont réservées à une élite inutile dont les vaines recherches sont loin de toute application pratique. Des nombres premiers et des problèmes de factorisation, il y en a partout dans notre monde. Toute la protection informatique des données se fait par cryptages dont la sécurité repose sur des problèmes de factorisation de grands nombres.
C'est le principe en particulier de la cryptographie à clé publique par la méthode RSA (pour Rivest, Shamir, Adleman, 1977). Le but est de réussir à envoyer des messages représentés par des nombres, de façon à ce que quelqu'un qui l'intercepte ne soit pas en mesure de le déchiffrer. Seul le destinataire du message doit pouvoir le lire ; il doit aussi pouvoir autentifier de façon certaine l'envoyeur.
La modélisation est la suivante : chaque individu i possède une fonction $c_i : \mathbb{N} \to \mathbb{N}$ de chiffrement (qui transforme le message initial en message crypté), et une fonction $d_i : \mathbb{N} \to \mathbb{N}$ de déchiffrement (réciproque de c_i, de sorte qu'appliquer c_i puis d_i à un message renvoie le message initial). Les applications c_i de chiffrement sont connues de tout le monde, mais les d_i ne sont connues que de leur propriétaire. De plus, les d_i ne sont pas trouvables facilement à partir des c_i, de sorte à ce que, malgré le partage des fonctions de chiffrement c_i, personne n'est en mesure de connaître les fonctions de déchiffrement d_i des autres.
Supposons maintenant qu'une personne i envoie un message M à une personne j. Puisqu'il connaît sa fonction d_i et la fonction c_j de l'individu j, il peut envoyer le message crypté $c_j \circ d_i(M)$. Pour décrypter ce message, il faut d'abord appliquer d_j, puis c_i. L'individu j étant le seul à connaître d_j, personne d'autre que lui ne peut assurer la première phase de déchiffrement, ce qui assure que le message ne sera pas lu par quelqu'un d'autre. Par ailleurs, l'individu j connaît c_i et peut donc retrouver le message initial. Par ailleurs, si après avoir appliqué c_i, le message est cohérent, c'est que la personne qui l'a envoyé a effectivement appliqué d_i au message initial. L'individu i étant le seul à connaître d_i, cela assure à l'individu j que le message provient bien de l'individu i.
L'essentiel du fonctionnement d'une telle méthode repose donc sur le choix des fonctions de chiffrement et de déchiffrement, et sur le fait que la connaissance des c_i ne permet pas de retrouver facilement les d_i. C'est là que nos factorisations entrent en jeu.
L'idée initiale est la difficulté de factoriser de très grands nombres, lorsque ceux-ci ont peu de diviseurs, et que ces diviseurs sont gros. Le cas typique est celui d'un nombre n obtenu comme produit de deux grands nombres premiers p et q. Partant de là, l'ordre de $(\mathbb{Z}/n\mathbb{Z})^*$ étant $\varphi(n)$, le théorème de Lagrange

nous donne une généralisation du petit théorème de Fermat : pour tout $a \in \mathbb{N}$, si a est premier avec n, alors $a^{\varphi(n)} \equiv 1\ [n]$.
Or, si $n = pq$, d'après la remarque donnée après la définition de φ, $\varphi(n) = \varphi(p)\varphi(q)$. Il n'est pas dur de voir que si p est premier, $\varphi(p) = p - 1$, donc $\varphi(n) = (p-1)(q-1)$.
Ainsi, on choisit r un entier premier avec $(p-1)(q-1)$, on résout $rs \equiv 1 \mod (p-1)(q-1)$ (cela revient à trouver une relation de Bézout pour le couple $(r, \varphi(n))$, par l'algorithme d'Euclide).
La clé publique sera alors le couple (n, r) (ce que les autres connaîssent), la fonction de chiffrement sera : $c(M) = M^r [\text{mod } n]$ (ce que je note ainsi ici est le reste modulo n de M^r). Cela se calcule rapidement par calcul modulaire, en utilisant la méthode d'exponentiation rapide (voir informatique). La fonction de déchiffrement sera $d(M') = M'^s\ [\text{mod } n]$. On a bien

$$d \circ c(M) = M^{sr}\ [\text{mod } n] = M^{k\varphi(n)+1}\ [\text{mod } n] = M\ [\text{mod } n].$$

Par ailleurs, seuls r et n étant connus, pour trouver s, il faudrait connaître $\varphi(n)$, c'est-à-dire les entiers p et q. Ainsi, trouver s passe par la recherche d'une factorisation du grand nombre n. Ceci est une opération difficile, et si les nombres premiers p et q sont choisis suffisamment grands, ce n'est pas réalisable par ordinateur.

Ainsi, tout le succès d'une telle méthode réside dans la recherche de grands nombres premiers qu'un individu pourra prendre comme point de départ de la construction de son code. Mais les technologies et les méthodes évoluant, on est capable de factoriser des nombres de plus en plus grands. La sécurité d'un code diminue donc avec le temps, et il est nécessaire d'en changer. Par conséquent, la sécurité de la transmission d'informations sensibles est une course à la recherche de grands nombres premiers, d'où l'importance de tests de primalité performants utilisant parfois des résultats pointus d'algèbre. À l'opposé, l'espionnage numérique est une course à la factorisation de grands nombres, course également effectuée par ceux qui veulent se protéger, pour savoir quel niveau de protection ils doivent adopter !

4 Polynômes et fractions rationnelles

Le but de ce chapitre est d'étudier les fonctions polynomiales $x \mapsto a_d x^d + \cdots + a_1 x + a_0$. On s'interesse notamment aux propriétés arithmétiques (produit, somme, divisibilité...) et aux propriétés analytiques (racines, dérivation...)

On se placera dans un cadre plus formel dans le but notamment de généraliser des constructions *a priori* uniquement valables pour des polynômes à coefficients réels (comme la dérivation) à des polynômes à coefficients dans des anneaux plus généraux.

Seuls les polynômes à coefficients dans $\mathbb{R}$ ou $\mathbb{C}$ sont théoriquement au programme. Nous donnerons les définitions formelles plus généralement pour les polynômes à coefficients dans un anneaux commutatif. Ce point de vue a une certaine importance, car c'est lui qui permet d'itérer ensuite la construction pour obtenir les polynômes de plusieurs indéterminés. En effet, $\mathbb{A}$ étant un anneau commutatif, l'ensemble $\mathbb{A}[X]$ des polynômes à coefficients dans $\mathbb{A}$ pourra être muni d'une structure d'anneau commutatif, permettant de considérer l'anneau $\mathbb{A}[X][Y]$ des polynômes de l'indéterminée Y, à coefficients dans $\mathbb{A}[X]$. Un petit réarrangement des termes nous montre que cela donne une définition, telle qu'on pourrait la souhaiter, de l'ensemble $\mathbb{A}[X, Y]$ des polynômes à 2 indéterminées. On peut alors itérer la construction pour obtenir les polynômes à n indéterminées.

Certaines propriétés des anneaux de polynômes seront spécifiques au cas où l'anneau des coefficients est un corps, et on a même parfois besoin de certaines hypothèses supplémentaires. Ainsi, pour certaines propriétés, nous reviendrons aux exigences du programme ($\mathbb{R}$ ou $\mathbb{C}$), en précisant parfois ce qu'il en est dans les autres situations. Il convient de remarquer que dans ce cas, ces propriétés ne sont en général plus satisfaites pour les polynômes à plusieurs indéterminées, puisque $\mathbb{A}[X]$ n'est pas un corps.

I Polynômes à coefficients dans un anneau commutatif

Soit $\mathbb{A}$ un anneau, qu'on supposera commutatif.

I.1 Polynômes formels

Remarque 4.1.1 (Motivation de la définition)

Une fonction polynomiale réelle est entièrement déterminée par la suite de ses coefficients. Les différentes constructions telles la somme, le produit, la dérivation, peuvent s'écrire uniquement sur les coefficients.

La remarque précédente semble justifier de considérer un polynôme comme une suite de coefficients. Seul un nombre fini de ces coefficients doit être non nul.

Définition 4.1.2 (Polynômes formels)

- Un polynôme formel P à coefficients dans $\mathbb{A}$ est une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de $\mathbb{A}$, nulle à partir d'un certain rang.
- Le réel a_k est appelé k-ième coefficient de P, ou coefficient du monôme de degré k de P.
- L'ensemble des polynômes (formels) à coefficients dans $\mathbb{A}$ est noté $\mathbb{A}[X]$.

Exemples 4.1.3

1. $\mathbb{R}[X]$ ou $\mathbb{C}[X]$, polynômes formels à coefficients réels ou complexes ;
2. $\mathbb{Q}[X]$, ensemble des polynômes à ceofficients rationnels ;
3. $\mathbb{Z}[X]$, ensemble des polynômes à coefficients entiers ;
4. $(\mathbb{Z}/n\mathbb{Z})[X]$, polynômes à coefficients dans $\mathbb{Z}/n\mathbb{Z}$, dont un cas particulier important est $\mathbb{F}_p[X]$.

I.2 Opérations arithmétiques sur les polynômes

Les polynômes considérés dans cette section sont à coefficients dans un anneau commutatif.

Définition 4.1.4 (Somme de polynômes de $\mathbb{A}[X]$)

La *somme* de deux polynômes $P = (a_n)_{n\in\mathbb{N}}$ et $Q = (b_n)_{n\in\mathbb{N}}$ de $\mathbb{A}[X]$ est la suite $P+Q = (a_n+b_n)_{n\in\mathbb{N}}$, qui est bien nulle à partir d'un certain rang.

Définition 4.1.5 (Produit d'un polynôme de $\mathbb{A}[X]$ par un élément de $\mathbb{A}$)

Le produit de $P = (a_n)_{n\in\mathbb{N}}$ par $\lambda \in \mathbb{A}$ est $\lambda P = (\lambda \cdot a_n)_{n\in\mathbb{N}}$.

Lorsque $\mathbb{A}$ est un corps, on parle de multiplication par un scalaire.

Définition 4.1.6 (Produit de deux polynômes de $\mathbb{A}[X]$)

Soit $P = (a_n)_{n\in\mathbb{N}}$ et $Q = (b_n)_{n\in\mathbb{N}}$ deux polynômes de $\mathbb{A}[X]$. Soit pour tout $n \in \mathbb{N}$, $c_n = \sum_{k=0}^{n} a_k b_{n-k}$. Alors $(c_n)_{n\in\mathbb{N}}$ est un polynôme. On définit alors $PQ = (c_n)_{n\in\mathbb{N}}$.

Théorème 4.1.7 (Structure d'anneau de $\mathbb{A}[X]$)

La somme et le produit définis ci-dessus munissent $\mathbb{A}[X]$ d'une structure d'anneau commutatif.

Avertissement 4.1.8

Ne généralisez pas trop vite « $\mathbb{A}$ anneau $\Longrightarrow$ $\mathbb{A}[X]$ anneau » en « $\mathbb{K}$ corps $\Longrightarrow$ $\mathbb{K}[X]$ corps ». Cette dernière affirmation est fausse !

Lorsque $\mathbb{K}$ est un corps on peut munir $\mathbb{K}[X]$ d'une structure supplémentaire d'espace vectoriel sur le corps $\mathbb{K}$ (voir chapitre ultérieur). La structure totale (espace vectoriel + anneau) est appelée structure d'algèbre. Dans le cas où $\mathbb{A}$ n'est pas un corps, on peut adapter la définition des espaces vectoriels, en définissant la notion de *module* sur un anneau $\mathbb{A}$: $\mathbb{A}[X]$ est alors muni d'une structure de module sur $\mathbb{A}$.

I.3 Indéterminée formelle

Par commodité, on adopte une notation plus proche de la notation fonctionnelle qu'on connaît pour les fonctions polynomiales ; cette notation est plus facile à manipuler que la définition formelle par les suites.

De ce fait, à partir du moment où nous aurons défini l'indéterminée formelle X (remplaçant la notion de variable pour les fonctions polynomiales), nous n'utiliserons plus la définition formelle des polynômes par les suites.
On rappelle que par définition, l'anneau $\mathbb{A}$ considéré contient un élément neutre pour le produit, noté 1.

Définition 4.1.9 (Indéterminée formelle)
On définit dans $\mathbb{A}[X]$ l'indéterminée formelle X comme étant le polynôme $X = (0, 1, 0, 0, \ldots)$.

Ainsi, X n'est pas une variable (au sens fonctionnel), mais un polynôme bien précis, auquel on donne un nom particulier, et auquel on attribue une notation bien particulière, dont le but est l'analogie avec les fonctions polynomiales.

Avertissement 4.1.10
- En particulier, l'indéterminée formelle X n'étant pas une variable, elle ne doit pas être quantifiée, et ne peut pas être utilisée pour résoudre des équations.
- Un polynôme n'est pas une fonction de l'indéterminée formelle, donc la notation $P(X)$ en lieu et place de P n'est pas de mise. On l'utilise néanmoins dans certaines situations, notamment lorsque plusieurs indéterminées sont en jeu. Cette notation peut être justifiée rigoureusement par la notion de spécialisation qu'on verra un peu plus loin.

Proposition 4.1.11 (Monômes)
Pour tout $n \in \mathbb{N}$, on a $X^n = (\underbrace{0, \ldots, 0}_{n \text{ zéros}}, 1, 0, \ldots)$, le 1 étant donc à l'indice n.

Corollaire 4.1.12 (Expression d'un polynôme à l'aide de l'indéterminée formelle)
Soit $P = (a_n)_{n\in\mathbb{N}}$ un polynôme de $\mathbb{A}[X]$. Alors $P = \sum_{k=0}^{+\infty} a_k X^k$, cette somme ayant un sens puisqu'elle est en fait finie, les a_k étant nuls à partir d'un certain rang.

Encore une fois, il faut bien comprendre ce que signifie cette égalité : il s'agit bien d'une somme de polynômes, et non d'éléments de $\mathbb{A}$ (signification de l'indéterminée).
De la définition même, il vient :

Proposition 4.1.13 (principe d'identification)
Soit P et Q deux polynômes de $\mathbb{K}[X]$. Notons $P = \sum_{k=0}^{+infty} a_k X^k$ et $Q = \sum_{k=0}^{+infty} b_k X^k$, en étant bien conscient que ces sommes sont en fait finies. Alors $P = Q$ si et seulement si pour tout $k \in \mathbb{N}$, $a_k = b_k$.

Les règles de calcul sur les polynômes de $\mathbb{A}[X]$ résultent alors des règles usuelles de calcul dans un anneau découlant des associativités, des commutativités, et des distributivités. La règle élémentaire sur laquelle reposent ces calculs est la suivante :

Lemme 4.1.14 (Produits de monômes)
Soit $(i, j) \in \mathbb{N}^2$. Alors $X^i X^j = X^{i+j}$.

Exemples 4.1.15
1. Calcul de $(2 + 3X + 2X^2)(3X + X^2 + 2X^3)$ dans $\mathbb{F}_5[X]$.
2. Calcul de $(X^2 + X + 2)^7$ dans $\mathbb{F}_7[X]$.

Définition 4.1.16 (Monômes)
Un monôme est un polynôme de la forme aX^n, $a \neq 0$. Le degré d'un tel monôme est n.

Ainsi, tout polynôme est une somme de monômes de degrés deux à deux distincts.

I.4 Dérivation

On sait facilement dériver (au sens analytique) une fonction polynomiale à coefficients réels, $x \mapsto x^n$ se dérivant en $x \mapsto nx^{n-1}$. Cette règle de dérivation peut être vue de façon purement formelle, permettant de généraliser la dérivation des polynômes à un anneau quelconque (dans lequel on ne dispose pas des techniques d'analyse, spécifiques à $\mathbb{R}$)

Définition 4.1.17 (Dérivée formelle d'un polynôme)
Soit $P = \sum_{k=0}^{d} a_k X^k$ un polynôme à coefficients dans un anneau commutatif $\mathbb{A}$. Le *polynôme dérivé* est défini par :

$$P' = \sum_{k=1}^{d} k a_k X^{k-1}.$$

Proposition 4.1.18 (Linéarité de la dérivation)
Soit P, Q deux polynômes de $\mathbb{A}[X]$, et $a \in \mathbb{A}$.

1. $(P+Q)' = P' + Q'$.
2. $(aP)' = aP'$.

La linéarité s'exprime en terme de structures en affirmant que la dérivation est une application linéaire (c'est-à-dire un homomorphisme d'espaces vectoriels) lorsque $\mathbb{A}$ est un corps, ou un homomophisme de $\mathbb{A}$-modules sinon.

Vu que la définition de la dérivation est calquée sur la dérivée analytique des fonctions polynomiales réelles, on a, sans surprise, des règles de dérivation similaires, et notamment :

Proposition 4.1.19 (Dérivée de produits)

1. *Soit P et Q deux polynômes à coefficients dans $\mathbb{A}$. Alors*

$$(PQ)' = P'Q + PQ'.$$

2. *Soit $P_1, \ldots, P_n$ des polynômes à coefficients dans $\mathbb{A}$. Alors*

$$(P_1 \cdots P_n)' = \sum_{i=1}^{n} P_1 \cdots P_{i-1} P_i' P_{i+1} \cdots P_n.$$

3. *(Formule de Leibniz) Soit P et Q deux polynômes à coefficients dans $\mathbb{A}$. Alors*

$$P^{(n)} = \sum_{k=0}^{n} \binom{n}{k} P^{(k)} Q^{(n-k)}.$$

Exemples 4.1.20

1. Que peut-on dire de $(PQ)^{(p)}$ dans $\mathbb{F}_p[X]$, p étant premier ?
2. Justifier que la dérivée de P^k est $kP'P^{k-1}$.

Du dernier exemple, et du fait que tout polynôme est somme de monômes, on déduit notamment que les règles de dérivation de compositions sont aussi valables dans le cadre formel.

I.5 Degré et valuation

Définition 4.1.21 (Degré et valuation)

Soit $P = (a_n)_{n\in\mathbb{N}}$ un polynôme à coefficients dans un anneau commutatif $\mathbb{A}$.

1. Le *degré de* P est $\deg(P) = \max\{n \in \mathbb{N} \mid a_n \neq 0\}$.
   Si P est non nul, cet ensemble est non vide, et majoré. Ainsi, $\deg(P) \in \mathbb{N}$.
   Si $P = 0$, par convention, $\deg(P) = -\infty$.
2. La *valuation de* P est $\mathrm{val}(P) = \min\{n \in \mathbb{N} \mid a_n \neq 0\}$.
   Si P est non nul, cet ensemble est non vide, et minoré. Ainsi, $\mathrm{val}(P) \in \mathbb{N}$.
   Si $P = 0$, par convention, $\mathrm{val}(P) = +\infty$.

On utilise souvent la filtration suivante de $\mathbb{A}[X]$ (une filtration de E est une chaîne d'inclusions d'union totale E)

Notation 4.1.22 (Filtration par les degrés)

Soit $\mathbb{A}$ un anneau et $n \in \mathbb{N}$. On note $\mathbb{A}_n[X]$ l'ensemble des polynômes de degré au plus n.

Proposition 4.1.23

On a évidemment $\mathbb{A}_0[X] \subset \mathbb{A}_1[X] \subset \cdots \subset \mathbb{A}_n[X] \subset \cdots$ *et*

$$\mathbb{A}[X] = \bigcup_{n=0}^{+\infty} \mathbb{A}_n[X].$$

Définition 4.1.24 (Monôme dominant, coefficient dominant, polynôme unitaire)

Soit $P = \sum_{k=0}^{d} a_k X^k$ un polynôme de $\mathbb{A}[X]$, de degré d.

1. Le monôme dominant de P est le monôme $a_d X^d$, donc le monôme de plus haut degré de P.
2. Le coefficient dominant de P est l'élément a_d de $\mathbb{A}$, donc le coefficient du monôme dominant.
3. Le polynôme P est dit unitaire si son coefficient dominant vérifie $a_d = 1_{\mathbb{A}}$.

Proposition 4.1.25 (Degré d'une somme, d'un produit, d'une dérivée)

Soit P et Q deux polynômes de $\mathbb{A}[X]$, et $\lambda \in \mathbb{A}$. Alors :

1. $\deg(P+Q) \leqslant \max(\deg(P), \deg(Q))$.
   Si $\deg(P) \neq \deg(Q)$, *alors* $\deg(P+Q) = \max(\deg(P), \deg(Q))$.
2. *Si $\mathbb{A}$ est intègre, et si $\lambda \neq 0$,* $\deg(\lambda P) = \deg(P)$.
3. *Si $\mathbb{A}$ est intègre (en particulier si $\mathbb{A}$ est un corps) et si P et Q sont non nuls,* $\deg(PQ) = \deg(P) + \deg(Q)$.
4. $\deg(P') \leqslant \deg(P) - 1$.

Exemples 4.1.26

1. Trouver dans $(\mathbb{Z}/6\mathbb{Z})[X]$ un exemple contredisant le point 2.
2. Trouver dans $(\mathbb{Z}/6\mathbb{Z})[X]$ un exemple contredisant le point 3.

Corollaire 4.1.27 (Théorème de permanence de l'intégrité)

Si $\mathbb{A}$ est intègre, alors $\mathbb{A}[X]$ est intègre.

Corollaire 4.1.28 (Propriétés de stabilité)

1. *$\mathbb{A}_n[X]$ est stable par $+$.*
2. *La dérivation $D : \mathbb{A}[X] \longrightarrow \mathbb{A}[X]$ induit un homomophisme de $\mathbb{A}$-modules $D_n : \mathbb{A}_n[X] \longrightarrow \mathbb{A}_{n-1}[X]$*
3. *Si $\mathbb{K}$ est un corps de caractéristique nulle, $D_n : \mathbb{K}_n[X] \longrightarrow \mathbb{K}_{n-1}[X]$ est une surjection. Autrement dit, tout polynôme de $\mathbb{K}_{n-1}[X]$ est primitivable formellement dans $\mathbb{K}_n[X]$.*

Corollaire 4.1.29

$\mathbb{A}_n[X]$ est un sous-groupe de $\mathbb{A}[X]$. Est-ce un sous-anneau ?

Remarque 4.1.30

Si $\mathbb{A}$ est intègre, que dire de $\mathbb{A}[X_1, X_2, \ldots, X_n]$, défini récursivement par $\mathbb{A}[X_1, X_2, \ldots, X_{n-1}][X_n]$? Pourquoi appelle-t-on cette propriété « théorème de permanence » ?

Corollaire 4.1.31 (Intégrité des anneaux usuels de polynômes)

Les anneaux $\mathbb{R}[X]$, $\mathbb{C}[X]$, $\mathbb{F}_p[X]$, $\mathbb{Q}[X]$, $\mathbb{Z}[X]$ sont intègres.

Pour les degrés des dérivées, on peut donner un résultat plus précis dans $\mathbb{R}[X]$ et $\mathbb{C}[X]$:

Proposition 4.1.32 (Degré d'une dérivée dans $\mathbb{K}[X]$, $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$)

Soit $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, et P un polynôme non constant de $\mathbb{K}[X]$. Alors $\deg(P') = \deg(P) - 1$.

Remarque 4.1.33

1. Quelle condition donner au corps $\mathbb{K}$ pour que ce résultat reste vrai dans $\mathbb{K}[X]$?
2. Si cette condition n'est pas satisfaite, quelle condition donner au degré de P, pour que cette condition reste vraie (dans $\mathbb{K}[X]$) pour le polynôme P ?

Corollaire 4.1.34

Soit $\mathbb{K}$ un corps de caractéristique nulle, et soit P et Q deux polynômes de $\mathbb{K}[X]$. Si $P' = Q'$, alors P et Q diffèrent d'une constante additive.

On a pour les valuations des propriétés similaires à celles des degrés :

Proposition 4.1.35

Soit P et Q deux polynômes de $\mathbb{A}[X]$, et $\lambda \in \mathbb{A}$. Alors :

1. $\mathrm{val}(P+Q) \geqslant \min(\mathrm{val}(P), \mathrm{val}(Q))$.
   Si $\mathrm{val}(P) \neq \mathrm{val}(Q)$, *alors* $\mathrm{val}(P+Q) = \min(\mathrm{val}(P), \mathrm{val}(Q))$.
2. *Si $\mathbb{A}$ est intègre et si $\lambda \neq 0$,* $\mathrm{val}(\lambda P) = \mathrm{val}(P)$.
3. *Si $\mathbb{A}$ est intègre et si P et Q sont non nuls,* $\mathrm{val}(PQ) = \mathrm{val}(P) + \mathrm{val}(Q)$.
4. $\mathrm{val}(P') \geqslant \mathrm{val}(P) - 1$.

Encore une fois, on donne un résultat plus précis pour la dérivée dans le cas de $\mathbb{R}$ ou $\mathbb{C}$ ou...

Proposition 4.1.36 (Valuation d'une dérivée dans $\mathbb{K}[X]$, $\mathbb{K} = \mathbb{R}$, ou $\mathbb{C}$)

Soit $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, ou n'importe quel corps de caractéristique nulle. Alors, pour tout polynôme non nul P de $\mathbb{K}[X]$ tel que $\mathrm{val}(P) > 0$, $\mathrm{val}(P') = \mathrm{val}(P) - 1$.

Que se passe-t-il lorsque $\mathrm{val}(P) = 0$?

II Arithmétique dans $\mathbb{K}[X]$

On considère ici des polynômes à coefficients dans un corps $\mathbb{K}$. Vous pouvez considérer $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, mais, sauf mention explicite du contraire, les résultats exposés sont valables dans le cadre plus général d'un corps quelconque.

II.1 Division euclidienne

L'anneau $\mathbb{K}[X]$ est euclidien (c'est-à-dire qu'il y existe une notion de division euclidienne) :

Théorème 4.2.1 (Théorème de la division euclidienne dans $\mathbb{K}[X]$)

Soit $\mathbb{K}$ un corps. Pour tout polynôme A et $B \neq 0$ de $\mathbb{K}[X]$ il existe d'uniques polynômes Q et R tels que :

(i) $A = BQ + R$

(ii) $\deg(R) < \deg(B)$.

Les polynômes Q et R sont appelés respectivement quotient et reste de la division euclidienne de A par B.

Méthode 4.2.2 (Algorithme de la division euclidienne)

- On pose la division euclidienne comme la division des entiers, en disposant A à gauche et B à droite, les monômes étant écrits dans l'ordre décroissant des degrés (donc en marquant d'abord les monômes de plus haut degré).
- On trouve le monôme aX^k tel que aX^kB ait même monôme dominant que A, puis on effectue la différence $A_1 = A - aX^kB$, qui a donc un degré strictement plus petit que A.
- On recommence sur A_1, et on en déduit A_2
- On recommence ainsi jusqu'à obtenir A_k de degré strictement plus petit que B. Alors A_k est le reste recherché, et le quotient est la somme des monômes par lesquels on a multiplié B pour obtenir les A_i successifs.

Cet algorithme peut facilement être implémenté dans un langage informatique dans $\mathbb{R}[X]$ ou $\mathbb{C}[X]$; un polynôme est dans ce cas représenté par la liste de ses coefficients (on revient à la définition formelle des polynômes sous forme d'une suite finie).

Exemple 4.2.3

Division euclidienne de $X^6 + 3X^2 + 1$ par $X^2 + X + 1$.

On verra un peu plus loin une méthode basée sur l'étude des racines pour déterminer rapidement le reste d'une division euclidienne par un polynôme de petit degré dont on connaît les racines.

Remarque 4.2.4

L'algorithme de la division euclidienne peut-il être mené sans restriction dans $\mathbb{A}[X]$ lorsque $\mathbb{A}$ est un anneau commutatif quelconque ? Donner une condition sur le polynôme B pour qu'on puisse effectuer dans $\mathbb{A}[X]$ la division euclidienne d'un polynôme A quelconque par B.

II.2 Idéaux de $\mathbb{K}[X]$

Comme pour l'arithmétique de $\mathbb{Z}$, il est commode de raisonner en terme d'idéaux. Le résultat rendant la situation totalement similaire à celle de $\mathbb{Z}$ est le fait que tous les idéaux de $\mathbb{K}[X]$ sont principaux, donc engendrés par un unique polynôme.

Théorème 4.2.5 (Description des idéaux de $\mathbb{K}[X]$)

Soit $\mathbb{K}$ un corps. Alors $\mathbb{K}[X]$ est un anneau principal. De plus, deux polynômes P et Q engendrent le même idéal si et seulement s'il existe un $\lambda \in \mathbb{K}^*$ tel que $Q = \lambda P$.

On notera (P) l'idéal engendré par un polynôme P.

Remarques 4.2.6

- Souvenez-vous de ce qu'on a dit des idéaux de $\mathbb{R}[X, Y]$: ce résultat peut-il être vrai si $\mathbb{K}$ n'est pas un corps ?
- En fait, on peut montrer que $\mathbb{K}[X]$ est principal si et seulement si $\mathbb{K}$ est un corps.
- Nous avons vu en exercice que tout anneau euclidien est principal. Ce n'est donc ici qu'une conséquence de ce résultat plus général.

II.3 Divisibilité

Soit $\mathbb{K}$ un corps.

Définition 4.2.7 (Divisibilité dans $\mathbb{K}[X]$)

Soit A et B deux polynômes de $\mathbb{K}[X]$. On dit que B divise A s'il existe un polynôme Q tel que $A = BQ$. Inversement, on dit que A est un multiple de B.

Ainsi, B divise A si et seulement si le reste de la division euclidienne de A par B est nul.
Comme dans $\mathbb{Z}$, on a la caractérisation suivante :

Proposition 4.2.8 (Caractérisation en termes d'idéaux)

Soit A et B deux polynômes de $\mathbb{K}[X]$. Alors A divise B si et seulement si $B \in (A)$, ou encore si et seulement si $(B) \subset (A)$.

Comme dans le cadre général, on dit que le couple (A, B) est un couple de polynômes associés si A divise B et B divise A. Il vient alors de la description des idéaux de $\mathbb{A}[X]$ que :

Proposition 4.2.9 (Polynômes associés)

Soit $(A, B) \in \mathbb{K}[X]^2$. Alors (A, B) est un couple de polynômes associés si et seulement s'il existe $\lambda \in \mathbb{K}^*$ tel que $A = \lambda B$.

II.4 PGCD et PPCM

Proposition/Définition 4.2.10 (PGCD de deux polynômes)

Soit $\mathbb{K}$ un corps. Soit A et B deux polynômes de $\mathbb{K}[X]$, dont l'un au moins est non nul et $P \in \mathbb{K}[X]$. Les propositions suivantes sont équivalents :

(i) P divise A et B et est de degré maximal pour cette propriété.

(ii) P divise A et B et tout autre diviseur de A et B est aussi un diviseur de P

(iii) $(P) = (A) + (B)$.

Si ces propriétés sont vérifiées on dit que P est un PGCD de A et B.

Il n'y a pas unicité d'un PGCD de A et B. Plus précisément :

Proposition 4.2.11 (Description des PGCD)

Soit A et B deux polynômes de $\mathbb{K}[X]$, dont l'un au moins est non nul, et P un PGCD de A et B. Alors un polynôme Q est un PGCD de A et B si et seulement s'il existe $\lambda \in \mathbb{K}^*$ tel que $Q = \lambda P$.

Notation 4.2.12 ($A \wedge B$)

En particulier, si A et B sont deux polynômes de $\mathbb{K}[X]$ dont l'un au moins est non nul, il existe un unique PGCD unitaire de A et B. Ce PGCD unitaire est noté $A \wedge B$.

Comme dans le cas de $\mathbb{Z}$, on déduit du troisième point équivalent de la définition l'existence de relations de Bézout.

Proposition 4.2.13 (Relation de Bézout)

Soit A et B deux polynômes de $\mathbb{K}[X]$ dont l'un au moins est non nul.

1. *Il existe des polynômes U et V tels que $AU + BV = A \wedge B$*
2. *Soit $P \in \mathbb{K}[X]$ tel qu'il existe U et V dans $\mathbb{K}[X]$ tels que $AU + BV = P$. Alors P est un multiple de $A \wedge B$.*

Comme dans $\mathbb{Z}$, on peut déterminer un PGCD et une relation de Bézout par l'algorithme d'Euclide étendu, en utilisant le lemme suivant :

Lemme 4.2.14

Soit A et B deux polynômes tels que $B \neq 0$. Soit Q et R le quotient et le reste de la division de A par B. Alors $A \wedge B = B \wedge R$

Méthode 4.2.15 (Calcul d'un PGCD et d'une relation de Bézout)

La méthode est la même que dans $\mathbb{Z}$, par divisions euclidiennes successives, jusqu'à obtenir un reste nul. Le dernier reste non nul est le PGCD, et en combinant les relations de division obtenues, on trouve de la même façon que dans $\mathbb{Z}$ une relation de Bézout (quitte à diviser par un scalaire, pour obtenir le PGCD unitaire)

Exemple 4.2.16

Trouver les PGCD de $X^8 - 1$ et $X^{12} - 1$, et une relation de Bézout.

Comme pour le cas de $\mathbb{Z}$, la définition du PGCD s'étend au cas du PGCD de n polynômes. On obtient alors :

Proposition 4.2.17 (Propriétés du PGCD)

L'opération $\wedge$ est commutative et associative. Par ailleurs, si C est unitaire, $(A \wedge B)C = AC \wedge BC$.

Évidemment, on peut aussi définir les PPCM :

Proposition/Définition 4.2.18 (PPCM de deux polynômes)

Soit $\mathbb{K}$ un corps. Soit A et B deux polynômes non nuls de $\mathbb{K}[X]$, et $P \in \mathbb{K}[X]$. Les propositions suivantes sont équivalents :

(i) A et B divisent P et P est de degré minimal pour cette propriété.

(ii) A et B divisent P et tout autre multiple de A et B est aussi un multiple de P

(iii) $(P) = (A) \cap (B)$.

Si ces propriétés sont vérifiées on dit que P est un PPCM de A et B.

Proposition 4.2.19 (Description des PPCM)

Soit A et B deux polynômes non nuls de $\mathbb{K}[X]$, et P un PPCM de A et B. Alors un polynôme Q est un PPCM de A et B si et seulement s'il existe $\lambda \in \mathbb{K}^*$ tel que $Q = \lambda P$.

Notation 4.2.20 ($A \vee B$)

En particulier, si A et B sont deux polynômes non nuls de $\mathbb{K}[X]$, il existe un unique PPCM unitaire de A et B. Ce PPCM unitaire est noté $A \vee B$.

Exemple 4.2.21

$P = (X + 1)^2$ et $Q = (X + 1)(X - 1)$.

II.5 Polynômes premiers entre eux

Définition 4.2.22

Soit A et B deux polynômes de $\mathbb{K}[X]$. On dit que A et B sont premiers entre eux si $A \wedge B = 1$.

Autrement dit, les seuls diviseurs communs à A et B sont les polynôems constants non nuls.
Plus généralement, on définit comme dans $\mathbb{Z}$ la notion de famille finie de polynômes deux à deux premiers entre eux, ou premiers entre eux dans leur ensemble.
Ici encore, les propriétés valables dans $\mathbb{Z}$ se généralisent :

Théorème 4.2.23 (Théorème de Bézout)

Soit A et B deux polynômes de $\mathbb{K}[X]$. Alors A et B sont premiers entre eux si et seulement s'il existe deux polynômes U et V tels que $AU + BV = 1$.

Exemple 4.2.24

Soit $\lambda \neq \mu$ dans $\mathbb{K}$. Alors les polynômes $X - \lambda$ et $X - \mu$ sont premiers entre eux.

Lemme 4.2.25 (Lemme de Gauss)

Soit A, B et C deux polynômes de $\mathbb{K}[X]$ tels que A divise BC et A et B soient premiers entre eux. Alors A divise C.

Corollaire 4.2.26

Soit A, B et C trois polynômes tels que A et B divisent C et A et B soient premiers entre eux. Alors AB divise C.

Comme dans $\mathbb{Z}$, on a une relation simple entre PPCM et PGCD, à ceci près que comme ces notions sont définies à constante multiplicativre près, il faut faire attention au coefficient dominant :

Proposition 4.2.27 (relation entre PGCD et PPCM)

Soit A et B deux polynômes de coefficients dominants a et b respectivement. Alors

$$ab(A \wedge B)(A \vee B) = AB.$$

II.6 Décomposition en facteurs irréductibles

Définition 4.2.28 (Polynôme irréductible)

Un polynôme non constant P de $\mathbb{K}[X]$ est irréductible si et seulement s'il n'est, à une constante multiplicative non nulle près, divisible que par lui-même et par 1.

Exemples 4.2.29

1. Les polynômes $X - \lambda$ sont irréductibles ($\lambda \in \mathbb{K}$)
2. Dans $\mathbb{R}[X]$, tout polynôme $aX^2 + bX + c$ tel que $\Delta < 0$ est irréductible.
3. Ces polynômes ne sont pas irréductibles dans $\mathbb{C}[X]$.

Lemme 4.2.30

Soit P un polynôme irréductible de $\mathbb{K}[X]$ et A un polynôme, non multiple de P. Alors A et P sont premiers entre eux.

Le lemme de Gauss fournit facilement la généralisation suivante du lemme d'Euclide :

Lemme 4.2.31 (Euclide)

Soit A et B deux polynômes de $\mathbb{K}[X]$ et P un polynôme irréductible. Alors si P divise AB, P divise A ou P divise B.

De façon équivalente, la contraposée fournit :

Corollaire 4.2.32

Soit A et B deux polynômes de $\mathbb{K}[X]$ et P un polynôme irréductible. Alors, si P ne divise ni A ni B, P ne divise pas AB.

Enfin, voici l'analogue du théorème de la décomposition primaire :

Théorème 4.2.33 (Décomposition en facteurs irréductibles)

Soit P un polynôme non nul de $\mathbb{K}[X]$.

1. *Il existe un élément $\lambda \in \mathbb{K}^*$ et des polynômes irréductibles $P_1, \dots, P_k$ tels que*
$$P = \lambda P_1 \cdots P_k.$$
2. *Cette décomposition est unique, à l'ordre près des facteurs, et à multiplication près de chaque facteur (y compris λ) par un élément non nul de $\mathbb{K}$.*
3. *En particulier, si on impose que les P_i soient unitaires, cette décomposition est unique, à l'ordre près des facteurs.*

Nous verrons un peu plus loin la description complète des polynômes irréductibles de $\mathbb{R}[X]$ et de $\mathbb{C}[X]$. Pour cela, il nous faut étudier d'un peu plus près les propriétés liées aux racines d'un polynôme.

III Racines d'un polynôme

Pour pouvoir définir la notion de racine d'un polynôme, il faut d'abord pouvoir « appliquer » un polynôme à un élément de $\mathbb{A}$, donc transformer un polynôme formel en une fonction polynomiale.

III.1 Spécialisation, évaluation

Le lien entre les polynômes formels de $\mathbb{R}$ et les fonctions polynomiales sur $\mathbb{R}$ est assez clair : étant donné un polynôme formel $P = \sum_{k=0}^{d} a_k X^k$ de $\mathbb{R}[X]$, on peut lui associer la fonction polynomiale

$$\tilde{P} : x \mapsto \sum_{k=0}^{d} a_k x^k.$$

La seule condition pour pouvoir faire cela de façon plus générale dans $\mathbb{A}[X]$ est de pouvoir faire dans $\mathbb{A}$ des produits (donc calculer des puissances) et des sommes. Comme $\mathbb{A}$ est un anneau commutatif, cela ne pose pas de problème particulier, et on peut donc définir :

Définition 4.3.1 (Fonction polynomiale associée à un polynôme)

1. Soit $P \in \mathbb{A}[X]$, donné par $P = \sum_{k=0}^{d} a_k X^k$. La fonction polynomiale $\tilde{P} : \mathbb{A} \longrightarrow \mathbb{A}$ associée à P est la fonction définie par :
$$\forall b \in \mathbb{A}, \tilde{P}(b) = \sum_{k=0}^{d} a_k b^k.$$
On rappelle que par convention, $b^0 = 1_{\mathbb{A}}$.
2. L'ensemble des fonctions polynomiales sur $\mathbb{A}$ est l'ensemble :
$$\mathbb{A}[x] = \{\tilde{P} \mid P \in \mathbb{A}[X]\}.$$

Définition 4.3.2 (Évaluation d'un polynôme)
Soit P un polynôme de $\mathbb{A}[X]$. L'évaluation de P en $b \in \mathbb{A}$ est l'élément de $\mathbb{A}$ défini par $\tilde{P}(b)$. Pour simplifier les notations, on désigne souvent cette évaluation plus simplement par $P(b)$.

Proposition 4.3.3 (Respect des structures)
Soit $\mathbb{A}$ un anneau commutatif. L'application $\varphi : \mathbb{A}[X] \longrightarrow \mathbb{A}[x]$ définie par $\varphi(P) = \tilde{P}$ est un homomorphisme d'anneaux.

Intuitivement, il apparaît clair que lorsque $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, on peut identifier les polynômes formels à coefficients dans $\mathbb{K}$ et les fonctions polynomiales sur $\mathbb{K}$. C'est ce que nous exprimons dans le théorème suivant :

Théorème 4.3.4 ($\mathbb{K}[X] \simeq \mathbb{K}[x]$ pour $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$)
Soit $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Soit $\varphi : \mathbb{K}[X] \longrightarrow \mathbb{K}[x]$ définie par $\varphi(P) = \tilde{P}$. Alors φ est un isomophisme d'anneaux.

Remarques 4.3.5

1. En considérant le petit théorème de Fermat, montrer que cette propriété n'est pas vraie pour tous les corps $\mathbb{K}$
2. On montrera un peu plus loin qu'une condition suffisante pour que cette identification soit vraie est que $\mathbb{K}$ soit un corps infini. C'est le cas en particulier lorsque $\mathbb{K}$ est de caractéristique nulle.

Enfin, dans le cas spécifique de $\mathbb{R}$ (seul cas dans lequel on peut considérer la dérivée au sens analytique), on a également, du fait même des définitions :

Proposition 4.3.6
Pour tout polynôme P de $\mathbb{R}[X]$, $\widetilde{P'} = \widetilde{P}'$.

Ainsi, les opérations définies formellement coïncident avec les opérations sur les fonctions polynomiales, y compris la dérivation dans le cas de $\mathbb{R}$.

Il est important de constater que le cadre formel qu'on s'est donné pour définir les polynômes permet d'« appliquer » un polynôme à des éléments qui sortent du cadre initialement fixé. Pour prendre un exemple, étant donné un polynôme $P = \sum_{k=0}^{d} a_k X^k$ de $\mathbb{R}[X]$ et M une matrice carrée à coefficients réels, on peut considérer le polynôme de matrices

$$P(M) = \sum_{k=0}^{k} a_k M^k,$$

où il faut bien prendre garde au fait que M^0 désigne la matrice identité I_n.

Exemple 4.3.7
Soit $P = 2 + 3X + 3X^2$, et $M = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$. Calculer $P(M)$.

On peut formaliser ce type de construction, mais nous resterons assez vague, car nous débordons du programme. Le bon cadre à se fixer est celui donné par la structure de $\mathbb{A}$-algèbre : un ensemble $\mathbb{B}$ est une

$\mathbb{A}$-algèbre si et seulement si $\mathbb{B}$ est un anneau et est muni d'une loi de composition externe d'ensemble d'opérateurs $\mathbb{A}$, avec un certain nombre de propriétés (associativité externe, distributivité de la loi externe sur chacune des lois additives de $\mathbb{A}$ et de $\mathbb{B}$, respect du neutre multiplicatif de $\mathbb{A}$). En gros, ce qu'il faut en retenir, c'est que dans une $\mathbb{A}$-algèbre $\mathbb{B}$, on peut faire, avec des règles de calcul raisonnablement semblables aux situations usuelles, la somme et le produit d'éléments de $\mathbb{B}$ ainsi que le produit d'un élément de $\mathbb{A}$ par un élément de $\mathbb{B}$. En particulier, étant donné un polynôme $P = \sum_{k=0}^{d} a_k X^k$ de $\mathbb{A}[X]$ et $b \in \mathbb{B}$, l'expression suivante a un sens :

$$P(b) = \sum_{k=0}^{d} a_k b^k.$$

Il convient de bien noter que par convention $b^0 = 1_{\mathbb{B}}$.
On parle de *spécialisation* du polynôme P en $b \in \mathbb{B}$.

Exemples 4.3.8

1. L'ensemble des matrices carrées de taille n, à coefficients réels, est une $\mathbb{R}$-algèbre : la situation décrite plus haut est un cas particulier de cette situation plus générale.
2. On utilisera beaucoup en algèbre linéaire des polynômes d'endomorphismes (applications linéaires d'un espace vectoriel dans lui-même), l'ensemble des endomorphismes d'un espace vectoriel E sur $\mathbb{K}$ étant une $\mathbb{K}$-algèbre pour la somme usuelle et le produit défini par la composition. Ainsi, f^n désigne dans ce cas la composition itérée de f, et f^0 désigne la fonction identité id_E.

III.2 Racines et multiplicité

Soit $\mathbb{A}$ un anneau commutatif

Définition 4.3.9 (Racine d'un polynôme)
Soit $P \in \mathbb{A}[X]$ et $a \in \mathbb{A}$. On dit que a est une racine de P si $P(a) = 0$.

Théorème 4.3.10 (Caractérisation des racines par la divisibilité)
Soit $\mathbb{K}$ un corps, $P \in \mathbb{K}[X]$ et $r \in \mathbb{K}$. Alors r est racine de P si et seulement si $X - r$ divise P, donc s'il existe $Q \in \mathbb{K}[X]$ tel que $P = (X - r)Q$.

Remarque 4.3.11
Ce théorème reste valable dans $\mathbb{A}[X]$ pour un anneau commutatif quelconque. Pourquoi ?

Si après factorisation $P = (X - r)Q$, r est encore racine de Q, alors r est « plusieurs fois » racine de P. En comptant le nombre de fois qu'on peut mettre $X - r$ en facteur, on obtient la multiplicité de r :

Définition 4.3.12 (Multiplicité d'une racine)
Soit $P \in \mathbb{K}[X]$ et $r \in \mathbb{K}$. On dit que r est racine d'ordre de multiplicité $k \in \mathbb{N}^*$ si et seulement si $(X-r)^k$ divise P et $(X-r)^{k+1}$ ne divise pas P. Autrement dit, il existe $Q \in \mathbb{R}[X]$ tel que $P = (X-r)^k Q$, avec $Q(r) \neq 0$.

Par convention, on dira que r est racine de multiplicité 0 si r n'est pas racine de P. Une racine de multiplicité 1 est aussi appelée racine simple de P.
Cette mise en facteur maximale de $(X - r)^r$ peut être mise en valeur par la formule de Taylor pour les polynômes :

Théorème 4.3.13 (Formule de Taylor pour les polynômes)

Soit $\mathbb{K}$ un corps de caractéristique nulle, P un polynôme de $\mathbb{K}[X]$, de degré d, et $a \in \mathbb{K}$. Alors,

$$P = \sum_{n=0}^{d} \frac{P^{(n)}(a)}{n!}(X-a)^n.$$

Ainsi, si v est le plus petit indice pour lequel le terme de la somme est non nul, on obtient

$$P = (X-a)^v \sum_{n=v}^{d} \frac{P^{(n)}(a)}{n!}(X-a)^n.$$

Remarques 4.3.14

- Pourquoi supposer que $\mathbb{K}$ est de caractéristique nulle ?
- Le résultat reste vrai dans $\mathbb{A}[X]$, si on suppose qu'il existe $i : \mathbb{Q} \to \mathbb{A}$ un homomorphisme injectif d'anneaux.

Ainsi, l'ordre de multiplicité de r correspond à la valuation de P après changement d'indéterminée $Y = X - r$.
On en déduit de façon immédiate :

Théorème 4.3.15 (Caractérisation de la multiplicité par les dérivées successives)

Soit $\mathbb{K}$ un corps de caractéristique nulle, $P \in \mathbb{K}[X]$ et $a \in \mathbb{K}$. Le réel a est racine d'ordre de multiplicité k de P si et seulement si : $P(a) = P'(a) = \cdots = P^{(k-1)}(a) = 0$ et $P^{(k)}(a) \neq 0$.

Ainsi, il faut toujours garder à l'esprit des deux facettes de la multiplicité des racines : la propriété de divisibilité, et la caractérisation par les dérivées.

Corollaire 4.3.16

Soit $\mathbb{K}$ un corps de caractéristique nulle. Soit $P \in \mathbb{K}[X]$ et $r \in \mathbb{R}$. Si r est racine d'ordre k de P, alors r est racine d'ordre $k-1$ de P'.

III.3 Majoration du nombre de racines

Le corollaire du lemme de Gauss amène :

Théorème 4.3.17

Soit $\mathbb{K}$ un corps. Soit $P \in \mathbb{K}[X]$, et $r_1, \ldots, r_k$ des racines deux à deux distinctes de P, de multiplicités respectives $\alpha_1, \ldots, \alpha_k$. Alors $(X-r_1)^{\alpha_1} \cdots (X-r_k)^{\alpha_k}$ divise P, et $r_1, \ldots, r_k$ ne sont pas racines du quotient.

Corollaire 4.3.18 (Majoration du nombre de racines)

Soit $P \in \mathbb{K}[X]$ de degré n. Alors P admet au plus n racines (comptées avec multiplicité).

Exemple 4.3.19

Soit $\mathbb{K}$ un corps dont $\mathbb{F}_p$ est un sous-corps. Montrer que pour tout $x \in \mathbb{K}$, $x^p = x$ si et seulement si $x \in \mathbb{F}_p$.

Corollaire 4.3.20 (Rigidité des polynômes)

1. *Soit P un polynôme de $\mathbb{K}[X]$ degré au plus n. Alors, si P admet strictement plus de n racines, $P = 0$.*
2. *Soit P un polynôme de $\mathbb{K}[X]$ s'annulant en une infinité de points de $\mathbb{K}$. Alors P est le polynôme nul.*
3. *Si deux polynômes P et Q de $\mathbb{K}_n[X]$ coïncident en strictement plus de n valeurs distinctes. Alors $P = Q$.*
4. *Soit $n \in \mathbb{N}^*$. Étant donnés $x_1, \dots, x_n$ des éléments deux à deux distincts de $\mathbb{K}$ et $y_1, \dots, y_n$ des éléments de $\mathbb{K}$ non nécessairement distincts, il existe au plus un polynôme $P \in \mathbb{K}_{n-1}[X]$ tel que pour tout $i \in [\![1, n]\!]$, $P(x_i) = y_i$. Ainsi, sous réserve d'existence, un polynôme de degré au plus $n-1$ est entièrement déterminé par sa valeur en n points distincts.*

On déduit notamment de cette propriété un résultat annoncé un peu plus haut :

Théorème 4.3.21 (Polynômes formels et fonctions polynomiales)

Soit $\mathbb{K}$ un corps infini. Alors l'application de $\mathbb{K}[X]$ dans $\mathbb{K}[x]$ qui à un polynôme formel associe sa fonction polynomiale est un isomorphisme d'anneaux.

Par ailleurs, le dernier point du corollaire ci-dessus affirme l'unicité sous réserve d'existence d'un polynôme de degré au plus $n-1$ prenant des valeurs données en n points fixés. Il n'est pas dur de construire explicitement un tel polynôme, fournissant ainsi l'existence :

Théorème 4.3.22 (Polynômes d'interpolation de Lagrange)

Soit $\mathbb{K}$ un corps, $n \in \mathbb{N}^*$, $x_1, \dots, x_n$ des éléments distincts de $\mathbb{K}$, et $y_1, \dots, y_n$ des éléments de $\mathbb{K}$. Alors il existe un et un seul polynôme P de $\mathbb{K}_{n-1}[X]$ tel que pour tout $i \in [\![1, n]\!]$, $P(x_i) = y_i$, et ce polynôme est donné explicitement par :

$$P = \sum_{i=1}^{n} y_i \cdot \frac{\prod\limits_{\substack{j \in [\![1,n]\!] \\ j \neq i}} (X - x_j)}{\prod\limits_{\substack{j \in [\![1,n]\!] \\ j \neq i}} (x_i - x_j)}$$

Soit $P_0 = (X - x_1) \dots (X - x_n)$. L'ensemble E des polynômes Q (sans restriction de degré) tels que pour tout $i \in [\![1, n]\!]$, $Q(x_i) = y_i$ est alors décrit par :

$$E = P + (P_0) = \{P + (X - x_1) \dots (X - x_n)R, \quad R \in \mathbb{K}[X].\}.$$

Ces polynômes, appelés polynômes d'interpolation de Lagrange, permettent en particulier d'approcher une fonction réelle f par une fonction polynomiale de degré au plus $n-1$ coïncidant avec f en n points distincts. Cette approximation est cependant dans les faits souvent assez peu efficace. Il existe d'autres polynômes d'interpolation, comme les polynômes d'interpolation de Hermite, de degré strictement plus petit que $2n$, coïncidant avec f en n points, et de courbe tangeante à celle de f en ces points.

Remarque 4.3.23

Quelle est la structure algébrique de l'ensemble E du théorème précédent ?

III.4 Polynômes scindés

Définition 4.3.24 (Polynôme scindé)

Soit $\mathbb{K}$ un corps. On dit qu'un polynôme non nul P de $\mathbb{K}[X]$ est scindé s'il possède autant de racines (comptées avec multiplicité) que son degré, autrement dit si son nombre de racines est maximal.

Théorème 4.3.25 (Factorisation d'un polynôme scindé)

1. *Un polynôme est scindé si et seulement s'il peut se factoriser de la façon suivante :*

$$P = \lambda(X - x_1)(X - x_2)\cdots(X - x_n),$$

où λ est un scalaire non nul (égal au coefficient dominant de P), n est le degré de P, et $x_1, \ldots, x_n$ sont les racines, non nécessairement distinctes, de P.

2. *Si on renomme $y_1, \ldots, y_k$ les racines 2 à 2 distinctes de P, de multiplicités respectives $\alpha_1, \ldots, \alpha_k$, cette factorisation se réécrit :*

$$P = \lambda(X - y_1)^{\alpha_1}\cdots(X - y_k)^{\alpha_k},$$

et on a $\alpha_1 + \cdots + \alpha_k = n$.

Ainsi, un polynôme est scindé si et seulement si sa décomposition en facteurs irréductibles ne fait intervenir que des polynômes irréductibles de degré 1.
Dans $\mathbb{R}[X]$, certaines techniques d'analyse peuvent aider à étudier cette propriété. Ainsi, le théorème de Rolle permet de montrer facilement que :

Proposition 4.3.26

Soit P un polynôme scindé de $\mathbb{R}[X]$, à racines simples. Alors P' est scindé, et ses racines séparent celles de P.

On verra en exercice une propriété plus générale :

Théorème 4.3.27

Soit P un polynôme scindé de $\mathbb{R}[X]$. Alors P' est scindé.

Une propriété importante des polynômes scindés est la possibilité de trouver facilement des relations entre les coefficients et les racines, par développement de la forme factorisée, et par identification des coefficients :

Théorème 4.3.28 (Relations coefficients/racines, ou relations de Viète)

Soit $P = \sum_{k=0}^{n} a_k X^k$ un polynôme de degré n, scindé, de racines (éventuellement non distinctes, apparaissant dans la liste autant de fois que sa multiplicité) $r_1, \ldots, r_n$. Alors pour tout $k \in [\![1, n]\!]$:

$$\sum_{1 \leqslant i_1 < \cdots < i_k \leqslant n} r_{i_1} r_{i_2} \cdots r_{i_k} = (-1)^k \frac{a_{n-k}}{a_n}.$$

IV Polynômes irréductibles dans $\mathbb{C}[X]$ et $\mathbb{R}[X]$

Cette section étudie spécifiquement les polynômes à coefficients complexes ou réels.

IV.1 Factorisations dans $\mathbb{C}[X]$

Nous avons plus ou moins défini $\mathbb{C}$ comme le corps de rupture du polynôme X^2+1, donc le plus petit corps contenant $\mathbb{R}$ dans lequel ce polynôme admet une racine i. Un théorème essentiel, parfois appelé *théorème fondamental de l'algèbre* (c'est dire son importance) est le théorème suivant, que d'Alembert croyait avoir démontré, que Gauss a démontré par différentes méthodes :

Théorème 4.4.1 (d'Alembert-Gauss)
Tout polynôme non constant de $\mathbb{C}[X]$ admet au moins une racine.

Corollaire 4.4.2
Tout polynôme de $\mathbb{C}[X]$ est scindé, donc admet exactement autant de racines (comptées avec multiplicité) que son degré.

Corollaire 4.4.3
Dans $\mathbb{C}[X]$, les seuls polynômes irréductibles sont les polynômes de degré 1, c'est-à-dire les polynômes $aX+b$, $a\neq 0$.

Exemple 4.4.4
Quelles sont les racines et leurs multiplicités du polynôme X^n-1 ? Factoriser ce polynôme en fateurs irréductibles dans $\mathbb{C}[X]$.

Tous les polynômes de $\mathbb{C}[X]$ se factorisant en polynômes non constants de degré minimal, on obtient alors une caractérisation simple de la divisibilité :

Théorème 4.4.5 (Caractérisation de la divisibilité dans $\mathbb{C}[X]$)
Soit P et Q deux polynômes de $\mathbb{C}[X]$. Alors P divise Q si et seulement si toute racine de P est aussi racine de Q, et que sa multiplicité dans Q est supérieure ou égale à sa multiplicité dans P.

IV.2 Facteurs irréductibles dans $\mathbb{R}[X]$

On commence par caractériser les polynômes à coefficients réels parmi les polynômes à coefficients dans $\mathbb{C}$.

Théorème 4.4.6 (Caractérisation des polynômes à coefficients réels)
Soit $P\in\mathbb{C}[X]$. Les propositions suivantes sont équivalentes :

(i) P est à coefficients réels ;

(ii) $P(\mathbb{R})\subset\mathbb{R}$

(iii) pour tout $z\in\mathbb{C}$, $P(\overline{z})=\overline{P(z)}$.

Corollaire 4.4.7 (Racines complexes d'un polynôme réel)
Soit P un polynôme à coefficients réels, et r une racine de P dans $\mathbb{C}$. Si $r\notin\mathbb{R}$, alors $\overline{r}$ est aussi racine de P, et elles ont même multiplicité.

Ainsi, les racines non réelles d'un polynôme à coefficients réels peuvent être groupées en paires de racines conjuguées de même multiplicité.
Le théorème de d'Alembert-Gauss amène alors :

Théorème 4.4.8 (Polynômes irréductibles de $\mathbb{R}[X]$)

1. *Les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré 1 et les polynômes de degré 2 de discriminant strictement négatif.*
2. *Ainsi, tout polynôme $P \in \mathbb{R}[X]$ peut être factorisé en produit de polynômes de $\mathbb{R}[X]$ de degré 1, ou de degré 2, de discriminant strictement négatif.*

Exemple 4.4.9

Factorisation dans $\mathbb{R}[X]$ de $X^n - 1$.

V Fractions rationnelles

La contruction formelle que nous avons donnée des polynômes nous empêche *a priori* de former des quotients de polynômes (donc des fractions rationnelles), comme nous pouvons le faire pour les fonctions polynomiales. En effet, si $\mathbb{K}$ est un corps, les seuls polynômes inversibles sont les polynômes constants non nuls.

Remarque 4.5.1

1. Quels sont les polynômes inversibles de $\mathbb{A}[X]$ lorsque $\mathbb{A}$ est intègre ?
2. Trouver un polynôme inversible et non constant de $(\mathbb{Z}/4\mathbb{Z})[X]$

Une construction similaire à celle permettant de définir $\mathbb{Q}$ à partir de $\mathbb{Z}$ nous permet cependant de définir formellement des quotients de polynômes.

V.1 Définition des fractions rationnelles formelles

Soit, dans tout ce qui suit, $\mathbb{K}$ un corps. On définit sur $\mathbb{K}[X] \times \mathbb{K}[X]^*$ la relation suivante :

$$(P, Q) \sim (R, S) \iff PS = QR,$$

l'égalité étant donnée dans $\mathbb{K}[X]$.

Proposition 4.5.2

La relation ci-dessus est une relation d'équivalence sur $\mathbb{K}[X] \times \mathbb{K}[X]^*$.

Définition 4.5.3 (Fraction rationnelle)

Une fraction rationnelle est une classe d'équivalence de la relation ci-dessus. La classe d'équivalence de (P, Q) sera notée $\dfrac{P}{Q}$. L'ensemble des fractions rationnelles sur le corps $\mathbb{K}$ est noté $\mathbb{K}(X)$.

Ainsi, la relation $PS = QR$ amène assez logiquement l'égalité des fractions rationnelles $\dfrac{P}{Q} = \dfrac{R}{S}$.
Pour définir les lois de composition sur $\mathbb{K}(X)$, on commence par les définir sur $\mathbb{K}[X] \times \mathbb{K}[X]^*$. On définit, pour tout (P_1, Q_1) et (P_2, Q_2) de $\mathbb{K}[X] \times \mathbb{K}[X]^*$:

$$(P_1, Q_1) \times (P_2, Q_2) = (P_1 P_2, Q_1 Q_2) \qquad \text{et} \qquad (P_1, Q_1) + (P_2, Q_2) = (P_1 Q_2 + P_2 Q_1, Q_1 Q_2).$$

Lemme 4.5.4

1. *Les opérations $\times$ et $+$ sont associatives et commutatives*
2. *La relation $\sim$ est une congruence sur les monoïdes $(\mathbb{K}[X]\times\mathbb{K}[X]^*,\times)$ et $(\mathbb{K}[X]\times\mathbb{K}[X]^*,+)$.*

Théorème 4.5.5 (Structure de $\mathbb{K}(X)$)

Les lois $+$ et $\times$ induisent des lois de composition, également notées $+$ et $\times$, sur $\mathbb{K}(X)$. L'ensemble $\mathbb{K}(X)$ muni de ces deux lois est un corps.

Les lois de composition ainsi définies se réécrivent sans surprise :

$$\frac{P_1}{Q_1}\times\frac{P_2}{Q_2}=\frac{P_1P_2}{Q_1Q_2},\qquad \frac{P_1}{Q_1}+\frac{P_2}{Q_2}=\frac{P_1Q_2+P_2Q_1}{Q_1Q_2}\qquad\text{et}\qquad \frac{P_1}{Q}+\frac{P_2}{Q}=\frac{P_1+P_2}{Q}.$$

Définition 4.5.6 (Inclusion canonique de $\mathbb{K}[X]$ dans $\mathbb{K}(X)$)

L'application $P\mapsto\frac{P}{1}$ de $\mathbb{K}[X]$ dans $\mathbb{K}(X)$ est un homomorphisme injectif d'anneaux. La fraction $\frac{P}{1}$ seront désormais identifiée au polynôme P de $\mathbb{K}[X]$.

En particulier, si $P=AB$, alors $B=\frac{P}{A}$.

Proposition 4.5.7 (Simplification)

Soit P et Q deux polynômes et D un diviseur commun à P et Q. Alors $\dfrac{P}{Q}=\dfrac{P/D}{Q/D}$

Proposition/Définition 4.5.8

Soit $F\in\mathbb{K}(X)$ une fraction rationnelle. Il existe un représentant (P,Q), unique à multiplication près par un scalaire non nul, tel que $P\wedge Q=1$ et $F=\frac{P}{Q}$. On dit que $\frac{P}{Q}$ est la forme irréductible de la fraction rationnelle F.

V.2 Degré, racines, pôles

Proposition/Définition 4.5.9 (Degré d'une fraction rationnelle)

Soit $F\in\mathbb{K}(X)$. La quantité $\deg(P)-\deg(Q)$ ne dépend pas de la représentation $\frac{P}{Q}$ choisie de la fraction F. On appelle degré de F et on note $\deg(F)$ la quantité

$$\deg(F)=\deg(P)-\deg(Q)\text{ où }F=\frac{P}{Q}$$

Il s'agit d'un entier relatif, ou de $-\infty$ si $P=0$.

Les degrés des fractions rationnelles vérifient des propriétés semblables aux degrés des polynômes :

Proposition 4.5.10 (Propriétés des degrés)

Soit F, G des élément de $\mathbb{K}(X)$.

- $\deg(F+G)\leqslant\max(\deg(F),\deg(G))$, *avec égalité si* $\deg(F)\neq\deg(G)$.
- $\deg(FG)=\deg(F)+deg(G)$
- $\deg(F^{-1})=-\deg(F)$

Proposition 4.5.11 (Partie entière)

Soit F une fraction rationnelle de $\mathbb{K}(X)$. Il existe un unique polynôme P de $\mathbb{K}[X]$ et une fraction rationnelle G de $\mathbb{K}(X)$ tels que

$$F = P + G \qquad et \qquad \deg(G) < 0.$$

Le polynôme P est appelée partie entière de la fraction rationnelle F.

Définition 4.5.12 (Racine, pôle, multiplicité)

Soit F une fraction rationnelle de $\mathbb{K}(X)$, écrit sous forme irréductible $F = \frac{P}{Q}$.

1. Une racine de F est une racine de P, sa multiplicité est sa multiplicité en tant que racine de P.
2. Un pôle de F est une racine de Q, sa multiplicité est sa multiplicité en tant que racine de Q.

Remarque 4.5.13

Puisque $\frac{P}{Q}$ est irréductible, r ne peut pas être à la fois racine de P et racine de Q.

Exemple 4.5.14

Racines, pôles et leurs multiplicités, de $\dfrac{(X-2)^2}{X^3(X-1)^4}$?

Définition 4.5.15 (Fonction rationnelle associée)

Soit $F = \frac{P}{Q}$ une fraction rationnelle formelle sous forme irréductible, et $\mathcal{P}$ l'ensemble de ses pôles. La fonction rationnelle associée est $\tilde{F} : \mathbb{K} \setminus \mathcal{P} \longrightarrow \mathbb{K}$ définie par

$$\forall x \in \mathbb{K} \setminus \mathcal{P}, \quad F(x) = \frac{P(x)}{Q(x)}.$$

V.3 Décomposition en éléments simples dans $\mathbb{C}(X)$

Toutes les fractions rationnelles de $\mathbb{C}(X)$ peuvent se décomposer comme sommes d'un polynôme et de fractions de type simple, plus précisément de fractions $\dfrac{\lambda}{(X-a)^k}$.

Lemme 4.5.16

Soit F une fraction rationnelle de $\mathbb{C}(X)$, et $r_1, \dots, r_k$ ses pôles, de multiplicités $\alpha_1, \dots, \alpha_k$. Alors, il existe des polynômes $P_1, \dots, P_k$ tels que

$$F = \sum_{i=1}^{k} \frac{P_i}{(X-r_i)^{\alpha_i}}.$$

Théorème 4.5.17 (Décomposition en éléments simples dans $\mathbb{C}(X)$, DÉS)

Soit F une fraction rationnelle de $\mathbb{C}(X)$, et $r_1, \dots, r_k$ ses pôles, de multiplicités $\alpha_1, \dots, \alpha_k$. Alors il existe un unique polynôme E et d'uniques complexes $\lambda_{i,j}$ ($1 \leqslant i \leqslant k$, $1 \leqslant j \leqslant \alpha_i$) tels que

$$F = E + \sum_{i=1}^{k} \sum_{j=1}^{\alpha_i} \frac{\lambda_{i,j}}{(X-r_i)^j}.$$

De plus, le polynôme E est la partie entière de la fraction rationnelle F, donc obtenue en effectuant la division euclidienne de P par Q, où $F = \frac{P}{Q}$

Définition 4.5.18 (Partie polaire)
Avec les notations du théorème précédent, la somme $\sum_{j=1}^{\alpha_i} \frac{\lambda_{i,j}}{(X-r_i)^j}$ est appelée partie polaire de F relativement au pôle r_i.

Exemples 4.5.19

1. Forme de la décomposition en éléments simples de $\frac{X^7}{(X-1)^3(X+1)^4}$.
2. Forme de la décomposition en éléments simples de $\frac{1}{(1+X^2)^3}$

Proposition 4.5.20 (cas d'un pôle simple)
Soit r un pôle simple de $F = \frac{P}{Q}$ (sous forme irréductible), et soit $\hat{Q}$ le polynôme $\frac{Q}{X-r}$. Alors le coefficient λ du terme $\frac{1}{X-\lambda}$ de la DÉS de F est :

$$\lambda = \frac{P(r)}{\hat{Q}(r)} = \frac{P(r)}{Q'(r)}.$$

Un cas important de décomposition en éléments simples est le cas de la fraction rationnelle $\frac{P}{P'}$.

Théorème 4.5.21 (Décomposition en éléments simples de $\frac{P'}{P}$)
Soit P un polynôme non nul de $\mathbb{C}[X]$. Soit $r_1, \dots, r_k$ les racines de P, de muliplicités $\alpha_1, \dots, \alpha_k$. Alors les pôles de $\frac{P'}{P}$ sont $r_1, \dots, r_k$ et sont tous des pôles simples. La DÉS de $\frac{P'}{P}$ est :

$$\frac{P'}{P} = \sum_{i=1}^{k} \frac{\alpha_i}{X-r_i}.$$

Nous avons déjà vu que les racines de la dérivée P' d'un polynôme scindé sont situées entre la racine minimale et la racine maximale de P. De la décomposition de $\frac{P'}{P}$, on déduit une propriété similaire dans $\mathbb{C}[X]$:

Corollaire 4.5.22 (Localisation des racines de P', HP)
Soit P un polynôme de $\mathbb{C}[X]$. Alors les racines de P' sont dans l'enveloppe convexe des racines de P.

V.4 Décomposition en éléments simples dans $\mathbb{R}[X]$

Théorème 4.5.23 (DÉS dans $\mathbb{R}(X)$)
Soit $F = \frac{P}{Q}$ une fraction rationnelle sous forme irréductible, et $Q = Q_1^{\alpha_1} \cdots Q_k^{\alpha_k}$ la décomposition en facteurs irréductibles de Q dans $\mathbb{R}[X]$. Ainsi, les Q_i sont de degré 1 ou 2. Alors il existe un unique polynôme E, et d'uniques polynômes $A_{i,j}$, de degré strictement plus petit que Q_i, tels que

$$F = E + \sum_{i=1}^{k} \sum_{j=1}^{\alpha_k} \frac{A_{i,j}}{Q_i^j}.$$

Remarques 4.5.24

1. Si Q_i est de degré 1, la partie correspondante dans la DÉS dans $\mathbb{R}(X)$ est la même que dans $\mathbb{C}(X)$.
2. Si Q_i est de degré 2, il admet deux racines complexes conjuguées r et $\overline{r}$. La partie correspondante dans la DÉS est obtenue en regroupant les parties polaires relatives à r et $\overline{r}$.
3. On a déjà vu en pratique comment déterminer des DÉS dans des cas simples. On a aussi déjà vu l'intérêt que peuvent avoir ces DÉS, notamment pour le calcul d'intégrales.

5 Espaces vectoriels

La structure d'espace vectoriel est une structure rigide, généralisant le cadre géométrique usuel.
Il s'agit d'une structure algébrique liée à la donnée d'un corps, qui va constituer l'unité de la dimension : la droite réelle représente l'espace vectoriel typique sur $\mathbb{R}$ de dimension 1, alors que le $\mathbb{C}$-espace vectoriel typique de dimension 1 ressemblera au plan complexe, donc à un objet géométrique, qui, en tant que $\mathbb{R}$-espace vectoriel sera en fait de dimension 2.
La rigidité de la structure se traduit par le fait qu'on peut multiplier un élément par un scalaire (un élément du corps de base), ceci de façon injective (sauf pour 0) : ainsi, si un élément x est dans un espace vectoriel E, tous les éléments λx, $\lambda \in \mathbb{K}$ seront aussi dans E, et si $x \neq 0$, l'ensemble des λx, $\lambda \in \mathbb{K}$ « ressemble » à $\mathbb{K}$ (il y a une bijection entre les deux). On parle de la droite engendrée par x. Ainsi un espace vectoriel est une structure droite, qui, dès qu'elle contient un élément non nul, contient tout la droite (au sens du corps $\mathbb{K}$) contenant x.
La rigidité d'un espace vectoriel est même plus forte que cela : plus que la stabilité par multiplication par un scalaire, on a la stabilité par combinaison linéaire (et encore une fois, l'application qui à (λ, μ) de $\mathbb{K}^2$ associe $\lambda x + \mu y$ est bijective, sauf si x et y sont colinéaire). Ainsi, si E contient deux points (non colinéaires), il contient tout un $\mathbb{K}$-plan passant par ces deux points et par l'origine.
Cette structure rigide (plate pourrait-on dire) généralise la situation du plan réel usuel (approximation de la surface localement plate de la Terre sur laquelle nous faisons notre géométrie) ou de l'espace usuel, donc de la géométrie euclidienne classique. Elle ne permet en revanche pas de prendre en compte de façon implicite des phénomènes de courbure intrinsèque (géométrie sphérique définie intrinsèquement sur un objet de dimension 2, sans plongement dans un espace de dimension 3, ou propriétés de courbures de l'espace-temps) : ces structures courbes nécessitent l'introduction d'objets plus complexes (les variétés).

Dans tout le chapitre, on considère un corps $\mathbb{K}$. Vous pouvez considérer que $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ (seul cas à connaître si on respecte le programme), mais sauf mention explicite du contraire, les définitions et les résultats sont valables pour tout corps (commutatif selon notre définition des corps).

I Notion d'espace vectoriel

I.1 Définition

Définition 5.1.1 (Espace vectoriel)

Soit E un ensemble. On dit que E est un espace vectoriel sur $\mathbb{K}$ (en abrégé $\mathbb{K}$-ev) si E est muni de deux lois :
- une loi de composition externe $+ \ : \ E \times E \longrightarrow E$;
- une loi de composition externe $\cdot \ : \ \mathbb{K} \times E \longrightarrow E$;

telles que :

(i) $(E, +)$ soit un groupe abélien

(ii) pour tout $(\lambda, \mu) \in \mathbb{K}^2$, $(x, y) \in E^2$:
- $(\lambda\mu)x = \lambda(\mu x)$ (associativité externe de $\cdot$) ;
- $1_{\mathbb{K}} \cdot x = x$ (compatibilité du neutre multiplicatif de $\mathbb{K}$) ;
- $\lambda(x_1 + x_2) = \lambda x_1 + \lambda x_2$ (distributivité de $\cdot$ sur la loi interne) ;
- $(\lambda + \mu)x = \lambda x + \mu x$ (distributivité de $\cdot$ sur la somme de $\mathbb{K}$).

Propriétés 5.1.2

Soit E un $\mathbb{K}$-ev. Pour tout $x \in E$:

1. $0 \cdot x = 0$, *c'est-à-dire* $0_{\mathbb{K}} \cdot x = 0_E$;
2. $\lambda \cdot 0_E = 0_E$
3. $(-1) \cdot x = -x$.

Terminologie 5.1.3

- Les éléments de E sont appelés *vecteurs*
- Les éléments de $\mathbb{K}$ sont appelés *scalaires*
- Deux éléments x et y de E sont colinéaires s'il existe $(\lambda, \mu) \in \mathbb{K}^2$ tels que $(\lambda, \mu) \neq (0, 0)$ et $\lambda x + \mu y = 0$.

I.2 Combinaisons linéaires

Une propriété cruciale d'un espace vectoriel E est la stabilité par combinaison linéaire : si $(\lambda, \mu) \in \mathbb{K}^2$ et $(x, y) \in \mathrm{E}^2$, alors $\lambda x + \mu y \in E$. La notion de combinaison linéaire étant centrale dans l'étude des espaces vectoriels, nous définissons une notion généralisée de combinaison linéaire.

Définition 5.1.4 (Famille presque nulle)

Soit I un ensemble d'indices, et $(\lambda_i)_i \in I$ une famille d'éléments de $\mathbb{K}$. On dit que la famille $(\lambda_i)_{i \in I}$ est presque nulle, ou qu'elle est à support fini, ou encore que les λ_i sont presque tous nuls, si seul un nombre fini de λ_i est non nul, autrement dit s'il existe $J \subset I$ un sous-ensemble fini de I tel que pour tout $i \in I \setminus J$, $\lambda_i = 0$.

Si I est fini, toute famille est à support fini. Peut-on dire qu'elle est presque nulle ? La terminologie n'est certainement pas très heureuse dans cette situation, puisque toute famille est alors presque nulle même si aucun vecteur n'est nul, mais nous l'utiliserons tout de même.

Définition 5.1.5 (Combinaison linéaire généralisée)

Soit E un $\mathbb{K}$-ev et $(x_i)_{i \in I}$ une famille de vecteurs de E. Une combinaison linéaire des $(x_i)_{i \in I}$ est un vecteur

$$x = \sum_{i \in I} \lambda_i x_i,$$

où $(\lambda_i)_{i \in I}$ est une famille à support fini de scalaires de $\mathbb{K}$.

Ainsi, toute combinaison linéaire sur une famille infinie est une combinaison linaire d'un nombre fini de vecteurs de cette famille.

I.3 Un exemple important : espace de fonctions

Comme nous avons pu nous en rendre compte pour les groupes et les anneaux, on a des critères souvent rapides pour montrer qu'un ensemble est un sous-truc d'un truc plus gros. Nous verrons un peu plus loin

que de la même façon, il est beaucoup plus commode de montrer qu'un ensemble est un sous-espace vectoriel d'un espace vectoriel connu plutôt que de montrer de façon directe qu'il s'agit d'un espace vectoriel, ce qui nécessite beaucoup de petites vérificiations, élémentaires mais fastidieuses. Pour cette raison, il est important de connaître un certain nombre d'espaces vectoriels de référence, qui seront suffisants pour justifier la structure d'espace vectoriel d'autres ensembles dans la plupart des cas rencontrés.

Proposition 5.1.6 (espace vectoriel de référence)

1. *Soit F un ensemble quelconque. Alors l'ensemble de fonctions $\mathbb{K}^F$ est un espace vectoriel sur $\mathbb{K}$.*
2. *Plus généralement, soit E un espace vectoriel sur $\mathbb{K}$ et F un ensemble quelconque. Alors l'ensemble de fonctions E^F est un espace vectoriel sur $\mathbb{K}$.*

On déduit de cet exemple les espaces vectoriels usuels, comme cas particulier de la propriété précédente, suivant l'ensemble F choisi :

Exemples 5.1.7 (Espaces vectoriels à bien connaître)

1. $\mathbb{K}^{\varnothing} = \{0\}$;
2. $\mathbb{K}^{\{1\}} = \mathbb{K}$
3. $\mathbb{K}^{[\![1,n]\!]} = \mathbb{K}^n$;
4. $\mathbb{K}^{\mathbb{N}}$ l'ensemble des suites à valeurs dans $\mathbb{K}$;
5. $\mathbb{K}^{[\![0,n]\!]} = \mathbb{K}_n[X]$ l'espace des polynômes de degré au plus n ;
6. $\mathbb{C} = \mathbb{R}^2$ est un $\mathbb{R}$-ev ;
7. $E_2^{E_1}$ les applications entre deux espaces vectoriels ;

I.4 Produits d'espaces vectoriels

Nous voyons maintenant deux façons de construire des espaces vectoriels à partir d'espaces vectoriels de référence : tout d'abord une construction externe (produit cartésien), puis dans la section suivante, une construction interne (sous-espaces vectoriels).

Proposition 5.1.8 (produit cartésien d'ev)

Soit $E_1, \dots, E_n$ des espaces vectoriels sur un corps $\mathbb{K}$. Alors le produit cartésien $E_1 \times \cdots \times E_n$ est un espace vectoriel sur $\mathbb{K}$, lorsqu'on le munit des lois définies par :

- $\forall \lambda \in \mathbb{K},\ \forall (x_1, \dots, x_n) \in E_1 \times \cdots \times E_n,\quad \lambda \cdot (x_1, \dots, x_n) = (\lambda x_1, \cdots, \lambda x_n)$;
- $\forall (x_1, \dots, x_n), (y_1, \dots, y_n) \in E_1 \times \cdots \times E_n,\quad (x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$.

On retrouve en particulier la structure d'espace vectoriel de $\mathbb{K}^n$, déjà obtenue en considérant $\mathbb{K}^{[\![1,n]\!]}$.

I.5 Sous-espaces vectoriels

Selon les définitions générales sur les structures, on définit :

Définition 5.1.9 (Sous-espace vectoriel, sev)

Soit E un espace vectoriel sur $\mathbb{K}$. Un sous-ensemble $F \subset E$ de E est un sous-espace vectoriel de E si F est stable par les lois $+$ et $\cdot$ et que les lois induites munissent F d'une structure d'espace vectoriel.

Comme dans le cas des groupes et des anneaux, nous disposons d'un critère simple permettant de court-circuiter un certain nombre de vérifications :

Théorème 5.1.10 (Critère de sous-espace vectoriel)

Soit E un $\mathbb{K}$-espace vectoriel. Un ensemble F est un sous-espace vectoriel de E si :

(i) $F \subset E$

(ii) $0 \in F$

(iii) *F est stable par combinaison linéaire, ce qui équivaut à :*

$$\forall x, y \in F^2,\ \forall \lambda \in \mathbb{K},\ \lambda x + y \in F.$$

Exemples 5.1.11

1. Étant donné un espace vectoriel E, le sous-espace vectoriel nul $\{0_E\}$ et le sous-espace vectoriel total E.
2. Étant donné un vecteur X de $\mathbb{R}^2$, la droite $\mathbb{R}X = \{\lambda X,\ \lambda \in \mathbb{R}\}$
3. Étant donné a, b et c trois réels, le plan de $\mathbb{R}^3$ d'équation $ax + by + cy = 0$.
4. $\mathbb{R}[X]$ espace des polynômes
5. $\mathcal{C}(\mathbb{R}, \mathbb{R})$ ensemble des fonctions continues sur $\mathbb{R}$;
6. plus généralement $\mathcal{C}(I, \mathbb{R})$, ensemble des fonctions continues sur un intervalle I ;
7. $\mathcal{C}^n(\mathbb{R}, \mathbb{R})$ ensemble des fonctions de classe $\mathcal{C}^n$ sur $\mathbb{R}$;
8. plus généralement $\mathcal{C}^n(I, \mathbb{R})$, ensemble des fonctions de classe $\mathcal{C}^n$ sur un intervalle I ;
9. Les exemples d'espaces vectoriels de fonctions sont nombreux.

Vous remarquerez dans les premiers exemples les deux points de vue différents pour définir un sous-espace vectoriel : par l'intérieur (sous-espace engendré par un vecteur) ou par l'extérieur (sous-espace défini par une équation sur les coordonnées). On retrouvera souvent ces deux points de vue par la suite.

Définition 5.1.12 (Droite vectorielle, plan vectoriel)

Soit E un $\mathbb{K}$-espace vectoriel. Une droite vectorielle de E est un sous-ensemble D de E tel qu'il existe $x \in E$ non nul tel que $D = \mathbb{K}x$.

Proposition 5.1.13 (Stucture des droites vectorielles)

Les droites vectorielles d'un espace vectoriel E sur $\mathbb{K}$ sont des sous-espaces vectoriels de E.

Proposition 5.1.14

Soit E un $\mathbb{K}$-espace vectoriel, et D_1 et D_2 deux droites vectorielles. Alors soit $D_1 \cap D_2 = \{0_E\}$, soit $D_1 = D_2$.

Corollaire 5.1.15

Soit D une droite vectorielle d'un espace vectoriel E, et $x \in D$. Si $x \neq 0$, alors $D = \mathbb{K}x$.

Proposition 5.1.16 (Sous-espaces vectoriels de $\mathbb{R}^2$)

Les sous-espaces vectoriels de $\mathbb{R}^2$ sont :

- *le sous-espace vectoriel nul ;*

- *les droites vectorielles ;*
- *le sous-espace vectoriel total* $\mathbb{R}^2$.

De même, on montrera par la suite que les sous-espaces vectoriels de $\mathbb{R}^3$ sont exactement l'espace vectoriel nul, les droites vectorielles, les plans vectoriels (plans passant par l'origine) et $\mathbb{R}^3$ tout entier.

Remarque 5.1.17

L'aspect géométrique d'une droite vectorielle dépend du corps de base $\mathbb{K}$:

- Si le corps de base est $\mathbb{R}$, une droite vectorielle a l'aspect d'une droite géométrique usuelle.
- Si $\mathbb{K} = \mathbb{C}$, une droite a l'aspect d'un plan complexe : une droite complexe est donc un objet de dimension géométrique égale à 2.
- Si $\mathbb{K} = \mathbb{F}_p$, alors une droite est constituée d'un nombre fini de points « alignés circulairement » si on peut dire cela ainsi...
- Par exemple, si $\mathbb{K} = \mathbb{F}_2$, une droite est constituée de deux points : il y a dans ce cas autant de droites que de vecteurs non nuls de E, les droites étant les ensembles $\{0, x\}$, $x \neq 0$.

I.6 Intersections de sev

Proposition 5.1.18 (Intersection de sev)

Soit E un $\mathbb{K}$-espace vectoriel, et $(E_i)_{i\in I}$ une famille de sous-espaces vectoriels de E. Alors $\bigcap_{i\in I} E_i$ est un sev de E.

En revanche, l'union de deux sev n'est en général pas un sev.

I.7 Sous-espace vectoriel engendré par un sous-ensemble

Définition 5.1.19 (Sous-espace vectoriel engendré par un sous-ensemble)

Soit E un $\mathbb{K}$-ev, et X un sous-ensemble de E. Le sous-espace vectoriel engendré par X est le plus petit sous-espace vectoriel de E contenant X. Il est noté $\mathrm{Vect}(X)$.

Remarque 5.1.20

Pourquoi un tel espace existe-t-il ?

Si X est exprimée sous forme d'une famille $(x_i)_{i\in I}$, on note $\mathrm{Vect}(X) = \mathrm{Vect}(x_i)_{i\in I}$. Si X est fini et enuméré par exemple $x = \{x_1, \ldots, x_n\}$, on notera $\mathrm{Vect}(X) = \mathrm{Vect}(x_1, \ldots, x_n)$.

Proposition 5.1.21 (Minimalité de Vect(X))

Par définition, tout sous-espace vectoriel de E contenant X contient aussi $\mathrm{Vect}(X)$.

On peut donner une description explicite de $\mathrm{Vect}(X)$ à l'aide de combinaisons linéaires :

Proposition 5.1.22 (Description de Vect(X))

Soit E un $\mathbb{K}$-ev et X un sous-ensemble de E. Alors E est l'ensemble des combinaisons linéaires d'éléments de X.

Ainsi, $x \in \text{Vect}(X)$ si et seulement s'il existe $(x_1, \dots x_n) \in X^n$ et $(\lambda_1, \dots, \lambda_n) \in \mathbb{K}^n$ tels que

$$x = \lambda_1 x_1 + \cdots + \lambda_n x_n.$$

Exemples 5.1.23

1. $\text{Vect}(x) = \{\lambda x, \quad \lambda \in \mathbb{K}\} = \mathbb{K}x$
2. $\text{Vect}(x, y) = \{\lambda x + \mu y, \quad (\lambda, \mu) \in \mathbb{K}^2\}$
   - si $x = y = 0$, $\text{Vect}(x, y) = \{0\}$
   - si x et y sont colinéaires, non tous deux nuls (disons $x \neq 0$), $\text{Vect}(x, y) = \text{Vect}(x)$
   - si x et y ne sont pas colinéaires, $\text{Vect}(x, y)$ est un plan vectoriel.

I.8 Sommes de sev

Définition 5.1.24 (Somme de deux sev)

Soit G un $\mathbb{K}$-ev, et E et F deux sev de G. Alors la somme $E + F$ de E et F est le plus petit sous-espace vectoriel de G, contenant à la fois F et G :

$$E + F = \text{Vect}(E \cup F).$$

Proposition 5.1.25 (Description d'une somme)

Soit G un $\mathbb{K}$-ev, et E et F deux sev de G. Alors :

$$E + F = \{x + y \mid x \in E, y \in F\}.$$

Définition 5.1.26 (Somme d'un nombre fini de sous-espaces)

Plus généralement, on définit la somme des sous-espaces $E_1, \dots, E_n$ par :

$$E_1 + \cdots + E_n = \text{Vect}(E_1 \cup \cdots \cup E_n)$$

Proposition 5.1.27 (Description d'une somme d'un nombre fini de sev)

Soit $E_1, \dots, E_n$ et F des sev de E. Les propriétés suivantes sont équivalentes :

(i) $F = E_1 + \cdots + E_n$

(ii) $F = (((E_1 + E_2) + E_3) + \cdots + E_{n-1}) + E_n$

(iii) $F = \{x_1 + \cdots + x_n \mid (x_1, \dots, x_n) \in E_1 \times \cdots \times E_n\}$.

Exemples 5.1.28

1. Dans $\mathbb{R}^2$: $\text{Vect}((1, 2), (-1, 1)) = \mathbb{R}^2$
2. Soit $\mathcal{E}$ l'ensemble des suites vérifiant une relation de récurrence linéaire de polynôme caractéristique P, ayant n racines deux à deux distinctes $x_1, \dots, x_k$ dans $\mathbb{C}$. Alors, dans $\mathbb{C}^{\mathbb{N}}$:

$$\mathcal{E} = \text{Vect}((x_1^n)_{n\in\mathbb{N}}, (x_2^n)_{n\in\mathbb{N}}, \cdots, (x_k^n)_{n\in\mathbb{N}}).$$

3. Soit f une fonction définie sur $\mathbb{R}$ par une équation différentielle homogène d'ordre 2, dont le polynôme caractéristique admet deux racines rélles distinctes λ et μ. Alors

$$f \in \text{Vect}((x \mapsto \mathrm{e}^{\lambda x}), (x \mapsto \mathrm{e}^{\mu x})).$$

4. Soit E un ev, et $x_1, \dots, x_n$ des éléments de E. Alors :

$$\mathrm{Vect}(x_1, \dots, x_n) = \mathbb{R}x_1 + \cdots + \mathbb{R}x_n.$$

Le dernier exemple peut être vu comme une utilisation de la proposition suivante :

Proposition 5.1.29

Soit $(x_i)_{i\in I}$ et $(x_j)_{j\in J}$ deux familles (pas nécessairement disjointes). Alors :

$$\mathrm{Vect}((x_i)_{i\in I\cup J}) = \mathrm{Vect}((x_i)_{i\in I}) + \mathrm{Vect}((x_j)_{i\in J}).$$

Exemple 5.1.30

$\mathrm{Vect}(x_1, x_2, x_4, x_5) + \mathrm{Vect}(x_3, x_4, x_6) = \mathrm{Vect}(x_1, x_2, x_3, x_4, x_5, x_6)$.

I.9 Sommes directes

Définition 5.1.31 (Somme directe)

Soit E et F deux sev de G. On dit que la somme $E + F$ est *directe*, et on note $E \oplus F$, si $E \cap F = \{0\}$.

Plus généralement, $E_1 + \cdots + E_n$ est directe si $E_1 \oplus E_2$, puis $(E_1 + E_2) \oplus E_3$, etc. On note $\bigoplus_{i=1}^{n} E_i$.

Proposition 5.1.32 (Caractérisation de $\oplus$ par les intersections)

La somme $E_1 + \cdots + E_n$ est directe si et seulement si pour tout $k \in [\![2, n]\!]$,

$$F_k \cap \sum_{i<k} F_i = \{0\}.$$

Proposition 5.1.33 (Caractérisation de $\oplus$ par l'unicité)

La somme $\sum E_i$ de sous-espaces vectoriels de E est directe si et seulement si l'application ci-dessous est injective :

$$\begin{array}{rccc} \varphi & : & E_1 \times \cdots \times E_n & \longrightarrow & E \\ & & (x_1, \dots, x_n) & \longmapsto & x_1 + \cdots + x_n. \end{array}$$

En d'autres termes, $E_1 \oplus \cdots \oplus E_n$ est directe si et seulement tout élément x de $E_1 \oplus \cdots \oplus E_n$ s'écrit de façon unique sous la forme $x = x_1 + \cdots + x_n$, $x_i \in E_i$.

Définition 5.1.34 (Supplémentaire)

Soit E un espace vectoriel, et F et G deux sev de E. On dit que F et G sont supplémentaires dans E si $F \oplus G = E$.

Théorème 5.1.35 (Existence d'un supplémentaire)

Soit E un espace vectoriel quelconque, et F un sev de E. Alors F admet au moins un supplémentaire G.

La démonstration de ce théorème repose sur le lemme de Zorn, donc sur l'axiome du choix. On en verra une démonstration plus élémentaire, indépendante de l'axiome du choix, lorsque E est de dimension finie.

II Familles de vecteurs

Nous rappelons que toute combinaison linéaire d'une famille (finie ou infinie) s'exprime, par définition, comme somme *finie* d'éléments de cette famille, multipliés par des scalaires.

II.1 Familles libres

Proposition/Définition 5.2.1 (famille libre)

Une famille $(x_i)_{i\in I}$ de vecteurs de E est *libre* si une des propriétés équivalentes suivantes est vérifiée :

(i) Pour toute famille $(\lambda_i)_{i\in I}$ de scalaires presque tous nuls : $\sum_{i\in I} \lambda_i x_i = 0 \Longrightarrow \forall i \in I, \lambda_i = 0$;

(ii) Pour tout $x \in \mathrm{Vect}((x_i)_{i\in I})$, il existe une *unique* famille $(\lambda_i)_{i\in I}$ de scalaires presque tous nuls tels que $x = \sum_{i\in I} \lambda_i x_i$

Si de plus, $I = [\![1, n]\!]$, les points (i) et (ii) sont équivalents aux points suivants :

(iii) la somme $\mathbb{R}x_1 \oplus \cdots \oplus \mathbb{R}x_n$ est directe ;

(iv) La fonction $\varphi : \mathbb{K}^n \longrightarrow E$ définie par $\varphi(\lambda_1, \ldots, \lambda_n) = \lambda_1 x_1 + \cdots + \lambda_n x_n$ est injective.

Définition 5.2.2 (famille liée)

Une famille qui n'est pas libre est dite *liée*.

Remarque 5.2.3

Une famille contenant 0 peut-elle être libre ?

Proposition 5.2.4 (Stabilité de la liberté par restriction)

Toute sous-famille d'une famille libre est libre.

La proposition suivante permet de ramener l'étude de la liberté des familles infinies à l'étude de la liberté de familles finies.

Proposition 5.2.5 (Caractérisation de la liberté pour des familles infinies)

Une famille $(x_i)_{i\in I}$ est libre si et seulement si toutes ses sous-familles finies sont libres. Une famille $(x_i)_{i\in\mathbb{N}}$ est libre si et seulement si pour tout $n \in \mathbb{N}$, la famille $(x_1, \ldots, x_n)$ est libre.

Proposition 5.2.6 (Ajout d'un élément à une famille libre)

Soit $(x_i)_{i\in I}$ une famille libre de E et x_j ($j \notin I$) un élément de E. Alors, la famille $(x_i)_{i\in I\cup\{j\}}$, obtenue en ajoutant x_j à la famille libre $(x_i)_{i\in I}$, est libre si et seulement si $x_j \notin \mathrm{Vect}((x_i)_{i\in I})$

Si un tel ajout est impossible, on dira que la famille libre est maximale :

Définition 5.2.7 (Famille libre maximale)

Une famille libre est maximale si et seulement s'il est impossible de lui rajouter un vecteur (quelconque) de E en préservant sa liberté

Proposition 5.2.8 (Caractérisation des sommes directes par la liberté)

Soit $E_1, \dots, E_n$ des sev de E. Alors la somme $E_1 \oplus \cdots \oplus E_n$ est directe si et seulement si tout n-uplet $(x_1, \dots, x_n)$ d'éléments tous non nuls de $E_1 \times \cdots \times E_n$ est une famille libre dans E.

II.2 Familles génératrices

Définition 5.2.9 (familles génératrices)

Une famille $(x_i)_{i\in I}$ de vecteurs de E est une famille *génératrice* de E si l'une des propriétés équivalentes suivantes est satisfaite :

(i) tout $x \in E$ est une combinaison linéaire des x_i, $i \in I$;

(ii) $\mathrm{Vect}((x_i)_{i\in I}) = E$.

Si de plus $I = [\![1, n]\!]$, les points (i) et (ii) sont équivalents à :

(iii) $E = \sum_{i=1}^{n} \mathbb{R}x_i$;

(iv) La fonction $\varphi : \mathbb{K}^n \longrightarrow E$ définie par $\varphi(\lambda_1, \dots, \lambda_n) = \lambda_1 x_1 + \cdots + \lambda_n x_n$ est surjective.

On dit aussi que la famille $(x_i)_{i\in I}$ engendre E.
On obtient des propriétés symétriques à celles des familles libres :

Proposition 5.2.10 (Stabilité des familles génératrices par ajout)

Toute famille contenant une famille génératrice de E est une famille gnératrice de E.

Proposition 5.2.11 (Restriction d'une famille génératrice)

La famille obtenue en retirant un élément x d'une famille génératrice de E est encore génératrice si et seulement si x est une combinaison linéaire des autres vecteurs de la famille.

Lorsque cette situation n'est vérifiée pour aucun élément de la famille, on parle de famille génératrice minimale :

Définition 5.2.12 (famille génératrice minimale)

Une famille génératrice est dite minimale, si et seulement s'il est impossible de lui retirer un élément en préservant son caractère générateur.

II.3 Bases

Définition 5.2.13 (base d'un espace vectoriel)

Soit $(x_i)_{i\in I}$ une famille de vecteurs d'un ev E. On dit que $(x_i)_{i\in I}$ est une *base* de E si elle est une famille à la fois libre et génératrice de E.

Ainsi, $(b_i)_{i\in I}$ est une base de E si et seulement si tout élément x de E s'écrit de façon unique comme combinaison linéaire des éléments b_i : l'existence traduit le caractère générateur, l'unicité traduit la liberté.
Les coefficients de cette combinaison linéaire sont appelés *coordonnées de x dans la base $(b_i)_{i\in I}$*.
Le choix d'une base de E permet donc de définir un système de coordonnées : la donnée d'un vecteur x équivaut à la donnée de ses coordonnées dans une base fixée

Exemple 5.2.14

- Les coordonnées cartésiennes dans $\mathbb{R}^2$ correspondent aux coordonnées dans la base canonique $((1,0),(0,1))$.
- Un autre choix de base fournit d'autres coordonnées, par exemple $(2,3)$ dans la base $((1,0),(1,1))$

Proposition 5.2.15 (Caractérisation des bases par minimalité/ maximalité)

Les propriétés suivantes sont équivalentes :

(i) La famille $(x_i)_{i\in I}$ est une base de E ;

(ii) La famille $(x_i)_{i\in I}$ est une famille génératrice minimale de E ;

(iii) La famille $(x_i)_{i\in I}$ est une famille libre maximale de E.

Exemples 5.2.16 (Exemples importants de bases, à connaître)

- Base canonique de $\mathbb{K}^n$.
- Base canonique $(1, X, X^2, \cdots)$ de $\mathbb{K}[X]$; base canonique de $\mathbb{K}_n[X]$.
- $((X-x_0)^k)_{k\in[\![0,n]\!]}$ est une base de $\mathbb{K}_n[X]$.

Ce dernier exemple a une généralisation importante, qui mérite d'être citée en tant que proposition :

Proposition 5.2.17 (Famille échelonnée en degrés)

Si $(P_0,\dots,P_n)$ est une famille d'éléments de $\mathbb{K}_n[X]$ telle que pour tout $k\in[\![0,n]\!]$, $\deg(P_k)=k$, alors $(P_0,\dots,P_n)$ est une base de $\mathbb{K}_n[X]$.

Avertissement 5.2.18

La réciproque est fausse !

III Espaces vectoriels de dimension finie

Dans tout ce paragraphe $\mathbb{K}$ désigne le corps $\mathbb{R}$ ou $\mathbb{C}$.

III.1 Notion de dimension

Définition 5.3.1 (Espace vectoriel de dimension finie)

Un espace vectoriel E sur $\mathbb{K}$ est dit *de dimension finie* s'il existe une famille génératrice de cardinal fini $(x_i)_{i\in I}$ de E. Une famille génératrice finie est souvent appelé *système de générateurs*.

Proposition 5.3.2

Soit E un espace vectoriel de dimension finie. Alors de toute famille génératrice de E, on peut extraire une famille génératrice finie.

Le premier théorème important est l'existence d'une base de cardinal fini d'un espace vectoriel de dimension finie. On montre un résultat plus fort, disant que toute famille libre peut être vue comme le début d'une base.

Théorème 5.3.3

Soit E un espace vectoriel de dimension finie. Soit L une famille libre de E, et G une famille génératrice de E. Alors on peut compléter L en une base de E par ajout de vecteurs de G.

En particulier, si $(x_i)_{1\leqslant i\leqslant n}$ est génératrice de E, et $(x_i)_{i\in I}$ est libre, pour $I \subset [\![1,n]\!]$, il existe J tel que $I \subset J \subset [\![1,n]\!]$, tel que $(x_j)_{j\in J}$ soit une base de E.

Corollaire 5.3.4 (Cardinal des familles libres dans un en de dimension finie)

Soit E un espace vectoriel de dimension finie. Toute famille libre de E est de cardinal fini. En particulier, toute base est de cardinal fini.

Corollaire 5.3.5 (Théorème de la base extraite)

Soit E un espace vectoriel de dimension finie.

(i) De toute partie génératrice de E on peut extraire une base de E.

(ii) E admet au moins une base.

Corollaire 5.3.6 (Théorème de la base incomplète)

Toute famille libre d'un espace de dimension fini E peut être complétée en une base de E.

Nous établissons maintenant le théorème de la dimension, qui à la base de la théorie de la dimension. Pour le démontrer, nous utilisons le lemme suivant :

Lemme 5.3.7 (Théorème d'échange)

Soit E un espace vectoriel. Soit F une famille de vecteurs de E, et x et y dans E tels que $x \notin \mathrm{Vect}(F)$ et $x \in \mathrm{Vect}(F \cup \{y\})$. Alors $\mathrm{Vect}(F \cup \{x\}) = \mathrm{Vect}(F \cup \{y\})$.

Théorème 5.3.8 (Théorème de la dimension)

Soit E un espace vectoriel de dimension finie. Alors toutes les bases de E sont finies et de même cardinal.

L'importance de ce théorème provient du fait qu'il permet de définir la notion fondamentale suivante :

Définition 5.3.9 (dimension d'un espace vectoriel)

Le cardinal commun de toutes les bases de E est appelé *dimension de E*, et est noté $\dim E$. Si E n'est pas de dimension finie, on dira que E est de dimension infinie.

La notion de dimension est la formalisation de la notion de nombre de degrés de liberté dont on dispose pour construire un objet. Par exemple, pour définir entièrement une suite par une récurrence linéaire homogène d'ordre 3, on dispose de 3 degrés de liberté, à savoir le choix des trois premiers termes de la suite. L'espace des suites vérifiant une telle relation de récurrence est *de facto* de dimension 3.

Exemples 5.3.10

1. Dimension de $\mathbb{K}^n$
2. Dimension de $\mathbb{K}_n[X]$

3. Dimension de l'ensemble des suites vérifiant une récurrence linéaire homogène d'ordre k
4. Dimension de l'ensemble des solutions d'une équation différentielle linéaire homogène d'ordre 1, d'ordre 2 à coefficients constants.
5. Dimension de $\mathbb{C}$ en tant que...
6. Dimension de $\mathcal{M}_{n,m}(\mathbb{R})$, de $\mathcal{M}_n(\mathbb{R})$.

Nous terminons cette section par :

Proposition 5.3.11 (Dimension d'un produit cartésien)

Soit E et F deux espaces vectoriels sur $\mathbb{K}$. Si E et F sont de dimension finie, l'espace vectoriel produit $E \times F$ est de dimension finie, égale à :

$$\dim(E \times F) = \dim(E) + \dim(F).$$

Plus précisément, si $(b_1, \dots, b_m)$ et $(c_1, \dots, c_n)$ sont des bases de E et F, une base de $E \times F$ est $((b_1, 0), \dots, (b_m, 0), (0, c_1), \dots, (0, c_n))$.

III.2 Dimension, liberté et rang

Définition 5.3.12 (Rang d'une famille de vecteurs)

Soit E un espace vectoriel, $k \in \mathbb{N}^*$, et $(x_1, \dots, x_k)$ une famille de vecteurs de E. Le *rang* de la famille $(x_1, \dots, x_k)$ est la dimension de $\mathrm{Vect}(x_1, \dots, x_k)$ (cet espace est de dimension finie, puisque engendré par une famille finie). On note :

$$\mathrm{rg}(x_1, \dots, x_k) = \dim \mathrm{Vect}(x_1, \dots, x_k).$$

Proposition 5.3.13 (Majoration du rang et cas d'égalité)

$\mathrm{rg}(x_1, \dots, x_k) \leqslant k$, *avec égalité si et seulement si la famille $(x_1, \dots, x_k)$ est libre.*

Proposition 5.3.14 (Majoration du cardinal d'une famille libre)

Soit E un espace vectoriel de dimension n. Alors toute famille libre de E est de cardinal au plus n, avec égalité si et seulement si c'est une base.

En particulier, toute famille de strictement plus de n éléments est liée.

Exemple 5.3.15

Montrer que toute matrice $M \in \mathcal{M}_n(\mathbb{R})$ admet un polynôme annulateur, c'est-à-dire un polynôme P tel que $P(M) = 0$. Quelle est la structure de l'ensemble des polynômes annulateurs de M ? Justifier l'existence d'un polynôme annulateur minimal pour la relation de divisibilité.

Proposition 5.3.16 (Minoration du cardinal d'une famille génératrice)

Soit E un espace vectoriel de dimension n. Alors toute famille génératrice de E est de cardinal au moins n, avec égalité si et seulement si c'est une base.

On en déduit en particulier :

Corollaire 5.3.17 (Caractérisation par le cardinal des familles libres maximales)

Soit E un espace vectoriel de dimension n.

- *Une famille libre est maximale dans le sens donné plus haut si et seulement si elle est de cardinal n.*
- *Une famille génératrice est minimale si et seulement si elle est de cardinal n.*

III.3 Dimension de sous-espaces et de sommes

Intuitivement, on ne dispose pas de plus de degré de liberté en restreignant l'espace. C'est ce que traduit le théorème suivant :

Théorème 5.3.18 (Dimension d'un sous-espace)

Soit E un espace vectoriel de dimension finie n. Soit $F \subset E$ un sous-espace vectoriel de E. Alors F est de dimension finie, et $\dim F \leqslant \dim E$. On a égalité si et seulement si $F = E$.

Par ailleurs, étant donné deux sous-espaces vectoriels de E, la dimension de leur somme sera facile à déterminer s'il n'y a pas de redondance (autrement dit, si on a unicité des écritures), ce qui se traduit pas le fait que la somme est directe. Cette absence de redondances correspond aussi intuitivement au cas où la dimension de la somme est maximale par rapport aux sommes des deux sous-espaces, ce que l'on justifiera rigoureusement plus tard.

Théorème 5.3.19 (Dimension d'une somme directe)

Soit E un espace vectoriel, et F et G des sous-espaces vectoriels de dimension finie de E, en somme directe. Alors $F \oplus G$ est de dimension finie, et :

$$\dim F \oplus G = \dim F + \dim G.$$

On en déduit facilement, par récurrence :

Corollaire 5.3.20 (Dimension d'une somme directe de n espaces)

Soit E un espace vectoriel, et $(E_1, \dots, E_n)$ une famille de sous-espaces vectoriels de dimension finie de E, en somme directe. Alors $\bigoplus_{i=1}^{n} E_i$ est de dimension finie, et :

$$\dim \bigoplus_{i=1}^{n} E_i = \sum_{i=1}^{n} \dim E_i.$$

On en déduit en particulier la dimension des supplémentaires, après avoir justifié leur existence de façon plus élémentaire qu'en début de chapitre, dans le cadre de la dimension finie :

Théorème 5.3.21 (Existence et dimension d'un supplémentaire, en dimension finie)

Soit E un espace vectoriel de dimension finie, et F un sous-espace vectoriel de E. Alors il existe un supplémentaire S de F dans E, et :

$$\dim S = \dim E - \dim F.$$

Pour terminer, nous donnons une formule générale exprimant la dimension d'une somme quelconque :

Théorème 5.3.22 (Formule de Grassmann)

Soit E un espace vectoriel, et F et G deux sous-espaces de dimension finie de E. Alors :

$$\dim(F+G) = \dim F + \dim G - \dim F \cap G.$$

Par une récurrence immédiate, on en déduit :

Théorème 5.3.23 (Majoration de la dimension d'une somme)

Soit E un espace vectoriel, et $E_1, \dots, E_n$ des sous-espaces vectoriels de dimension finie de E. Alors

$$\dim(E_1 + \cdots + E_n) \leqslant \dim(E_1) + \cdots + \dim(E_n),$$

avec égalité si et seulement si la somme est directe.

Comme évoqué plus haut, ces deux derniers résultats affirment que moins il y a de redondances dans l'écriture d'une somme, plus la dimension de la somme est importante ; elle est maximale lorsqu'il n'y a aucune redondance, ce qui s'exprime par le fait que la somme est directe.

Remarque 5.3.24

- Comparez cette formule à $|A \cup B| = |A| + |B| - |A \cap B|$
- Peut-on généraliser, en trouvant pour les dimensions une formule analogue à la formule du crible de Poincaré :

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{j=1}^{n} (-1)^{j+1} \sum_{\substack{J \subset [\![1,n]\!] \\ |J|=j}} \left|\bigcap_{i \in J} A_i\right| ?$$

Du cas d'égalité de la formule de Grassmann, on déduit :

Proposition 5.3.25 (Caractérisation des couples de sous-espaces supplémentaires)

Soit E un espace de dimension finie, et F et G deux sous-espaces vectoriels de E. Alors F et G sont supplémentaires l'un de l'autre si et seulement si

$$F \cap G = \{0\} \qquad \textit{et} \qquad \dim F + \dim G = \dim E.$$

Note Historique 5.3.26

Herrmann Günther Grassmann (1809-1877) est le précurseur incompris de toute la théorie de l'algèbre linéaire. Il expose les fondements de cette théorie dans sa thèse *Théorie des flots et des marées*, thèse qui ne sera pas lue par son examinateur, et qui sera publiée uniquement au début du XX-ième siècle. Ce n'est qu'une vingtaine d'année plus tard (vers 1860) que certains mathématiciens se rendent comptent de l'importance des travaux de Grassmann. Entre temps, Grassmann s'est contenté d'un poste d'enseignant en lycée. Détourné de la recherche mathématique, c'est dans des études linguistiques (sanscrit, gotique) qu'il finit par trouver la consécration. Grassmann est également le fondateur du calcul tensoriel.

6 Applications linéaires

Comme à chaque fois qu'on définit une nouvelle structure algébrique, il vient une notion de morphisme, adaptée à cette structure. La catégorie des espaces vectoriels sur un corps $\mathbb{K}$ est ainsi formée d'un ensemble d'objets (les $\mathbb{K}$-ev), et de flèches représentant les morphismes entre espaces vectoriels, c'est-à-dire, selon les définition générales qu'on en a données, les applications respectant les deux lois définissant un espace vectoriel. La propriété définissant ces morphismes se traduit par une propriété de linéarité $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$. Pour cette raison, les morphismes d'espaces vectoriels sont plus fréquemment appelés *applications linéaires*.

Le but de ce chapitre est de donner les propriétés élémentaires des applications linéaires, d'étudier les propriétés de rigidité des applications linéaires (comme pour les polynômes, il suffit en général de connaître l'image d'un petit nombre de vecteurs pour déterminer entièrement une application linéaire), d'étudier certains types d'applications linéaires (endomorphismes, isomorphismes, projecteurs, symétries), de traduire certaines propriétés d'une application linéaire sur son comportement sur une base (caractérisations de l'injectivité, de la surjectivité).
Nous verrons ensuite des résultats spécifiques aux applications linéaires définies sur des espaces de dimension finie, notamment une formule reliant la dimension du noyau et la dimension de l'image (formule du rang).
Pour terminer nous donnons un aperçu de la notion d'espace dual, lié à la notion de forme linéaire et d'hyperplan.

Il est important de remarquer que la notion d'application linéaire en dimension finie est indissociable de la notion de matrice. Nous étudierons ce point de vue plus en détail dans le chapitre suivant.

Dans tout ce qui suit, $\mathbb{K}$ désigne un corps quelconque. Vous pouvez considérer, conformément au programme, que $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, mais, sauf mention explicite du contraire, les résultats donnés sont valables pour tout corps.

I Généralités sur les applications linéaires

I.1 Définitions et propriétés de stabilité

Définition 6.1.1 (Application linéaire, AL)

Une application $f : E \to F$ entre deux $\mathbb{K}$-ev est appelée *application $\mathbb{K}$-linéaire*, ou plus simplement *application linéaire* (en abrégé : AL), si :

$$\forall \lambda \in \mathbb{K},\ \forall x \in E,\quad f(\lambda x) = \lambda f(x) \qquad \text{et} \qquad \forall (x,y) \in E^2,\ f(x+y) = f(x) + f(y).$$

Proposition 6.1.2 (Respect du neutre)
Soit $f : E \to F$ une AL. Alors $f(0_E) = 0_F$.

Proposition 6.1.3 (Caractérisation des AL par respect des CL)
Une application $f : E \to F$ est une application linéaire si et seulement si :

$$\forall \lambda \in \mathbb{K},\ \forall (x, y) \in E^2,\quad f(\lambda x + y) = \lambda f(x) + f(y).$$

Exemples 6.1.4

1. L'application de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par $f(x, y) = 2x + 3y$?
2. L'application de $\mathbb{R}^3$ dans $\mathbb{R}$ définie par $f(x, y) = 3x + 4y + 2z + 1$?
3. L'application de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par $f(x, y) = xy$?
4. La somme $\sum$, de $\mathbb{C}^n$ dans $\mathbb{C}$.
5. L'intégrale, de $\mathcal{C}^0([a, b])$ dans $\mathbb{R}$.
6. L'opérateur de dérivation D de $\mathcal{C}^n$ dans ...
7. L'opérateur de dérivation D de $\mathbb{R}[X]$ dans $\mathbb{R}[X]$.
8. Étant donné une matrice $M \in \mathcal{M}_{n,p}(\mathbb{K})$, et en assimilant $\mathbb{K}^n$ à $\mathcal{M}_{n,1}(\mathbb{K})$, l'application de $\mathbb{K}^p$ dans $\mathbb{K}^n$ qui à une matrice colonne X associé la matrice colonne MX.

Nous verrons dans le chapitre suivant que ce dernier exemple fournit la description générique d'une application linéaire en dimension finie, après choix de bases de E et de F.
Nous obtenons ainsi l'ensemble des flèches entre deux objets de la catégorie des $\mathbb{K}$-espaces vectoriels

Définition 6.1.5 (Ensemble des applications linéaires)
Soit E et F deux $\mathbb{K}$-ev. On note $\mathcal{L}(E, F)$ l'ensemble des applications linéaires de E vers F.

Ces ensembles de flèches possèdent eux-même une structure d'espace vectoriel :

Proposition 6.1.6 (Structure de $\mathcal{L}(E, F)$)
$\mathcal{L}(E, F)$ est un espace vectoriel sur $\mathbb{K}$.

Autrement dit, une combinaison linéaire d'applications linéaires de E vers F est encore une application linéaire.
Étudions maintenant des propriétés liées à la composition.

Proposition 6.1.7 (Composée de deux applications linéaires)
Soit $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$. Alors $g \circ f$ est une application linaire de E vers G.

De manière générale, la composition d'applications à valeurs dans un espace vectoriel est toujours linéaire à gauche, c'est-à-dire $(\lambda f + \mu g) \circ h = \lambda f \circ h + \mu g \circ h$. Si les applications considérés sont linéaires, on obtient aussi la linéarité à droite. On rappelle :

Définition 6.1.8 (Application bilinéaire)

Soit E, F et G trois $\mathbb{K}$-espaces vectoriels, et $\varphi : E \times F \to G$. On dit que φ est bilinéaire si elle est linéaire par rapport à chacune de ses deux variables, l'autre étant fixée, c'est-à-dire si pour tout $(x, x', y, y', \lambda) \in E \times E \times F \times F \times \mathbb{K}$,

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

La propriété énoncée plus haut pour les compositions s'exprime alors de la manière suivante :

Proposition 6.1.9 (Bilinéarité de la composition)

La composition d'applications linéaires est bilinéaire. En termes plus précis, E, F et G étant trois $\mathbb{K}$-ev, l'application Φ de $\mathcal{L}(E, F) \times \mathcal{L}(F, G)$ dans $\mathcal{L}(E, G)$ définie par $\Phi(u, v) = v \circ u$ est une application bilinéaire.

On pourrait de façon similaire définir une notion d'application n-linéaire (à n variables vectorielles). La composition de n AL est alors n-linéaire. On verra dans un chapitre ultérieur comment cette notion d'application multilinéaire est également liée à la notion de déterminant.

I.2 Image et noyau

Dans ce paragraphe, E et F sont deux espaces vectoriels, et $f \in \mathcal{L}(E, F)$. On étudie ici deux sous-espaces liés à une application linéaire : l'image (qui correspond à la notion usuelle d'image) et le noyau.

Définition 6.1.10 (Image et noyau)

1. L'image de f est $\mathrm{Im}(f) = \{y \in F \mid \exists x \in E,\ f(x) = y\} = f(E)$;
2. Le noyau de f est $\mathrm{Ker}(f) = \{x \in E \mid f(x) = 0\} = f^{-1}(\{0\})$

La structure algébrique de l'image et du noyau découle d'un résultat plus général de préservation de la structure par image directe et réciproque :

Lemme 6.1.11 (Structure des images directes et réciproques)

1. *Soit E' un sev de E. Alors $f(E')$ est un sev de F.*
2. *Soit F' un sev de F. Alors $f^{-1}(F')$ est un sev de E.*

En appliquant ce lemme avec $E' = E$ et $F' = \{0\}$, on obtient :

Proposition 6.1.12 (Structure de l'image et du noyau)

$\mathrm{Im}(f)$ *est un sev de* F. $\mathrm{Ker}(f)$ *est un sev de* E.

De façon évidente, l'image mesure le défaut de surjectivité, De façon symétrique, le noyau mesure le défaut d'injectivité : si $\mathrm{Ker}(f)$ n'est pas un singleton, les définitions amènent de façon immédiate la non injectivité de f. La réciproque découle du fait que tout défaut d'injectivité peut être translaté en 0 par linéarité. Nous obtenons :

Théorème 6.1.13 (Caractérisation de la surjectivité et de l'injectivité)

Soit $f \in \mathcal{L}(E, F)$. Alors :

(i) f est surjective ssi $\mathrm{Im}(f) = F$;

(ii) f est injective ssi $\mathrm{Ker}(f) = \{0\}$.

Connaissant une famille génératrice de E (par exemple une base), il n'est pas dur de déterminer l'image de f :

Proposition 6.1.14 (Famille génératrice de Im(f))
Soit $f \in \mathcal{L}(E,F)$ et $(e_i)_{i\in I}$ une famille génératrice de E. Alors $f(e_i)_{i\in I}$ est une famille génératrice de $\mathrm{Im}(f)$ *:*

$$\mathrm{Im}(f) = \mathrm{Vect}(f(e_i), i \in I).$$

En particulier, si f transforme une famille génératrice de E en une famille génératrice de F, alors f est surjective.
Si nous appliquons la propriété précédente à l'application de $\mathbb{K}^p$ dans $\mathbb{K}^n$ définie par $X \mapsto MX$, nous obtenons la description très simple d'une famille génératrice de $\mathrm{Im}(f)$, dans le cas où f est donnée matriciellement :

Corollaire 6.1.15 (Image d'une AL décrite matriciellement)
Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$ et $f : \mathbb{K}^p \to \mathbb{K}^n$ (ces ensembles étant vus comme ensemble de vecteurs colonnes), définie par $f(X) = MX$. Alors l'image de f est engendrée par la famille des colonnes de la matrice M.

Exemple 6.1.16
Décrire l'image de $f : X \in \mathbb{R}^3 \mapsto \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 1 & 2 & 3 \end{pmatrix} X$.

On en déduit :

Méthode 6.1.17 (Déterminer l'image et le noyau d'une AL)
- Pour déterminer l'image d'une AL f :
  - Si f est donnée par une matrice (éventuellement après choix d'une base, voir chapitre suivant), considérer la famille des colonnes de cette matrice, qui engendre $\mathrm{Im}(f)$;
  - Sinon, trouver une famille génératrice de E (par exemple une base) et considérer son image.
- Pour déterminer le noyau d'une AL f :
  - Écrire l'équation $f(x) = 0$.
  - Si nécessaire, décomposer x dans une base $\mathcal{B}$ de E. Grâce à la linéarité de f, l'équation précédente se ramène alors à un système linéaire d'équations portant sur les coordonnées de x dans la base $\mathcal{B}$.

Conformément à la terminologie générale, nous définissons :

Définition 6.1.18 (Isomorphisme)
1. Une application linéaire bijective de E vers F est appelée un *isomorphisme*.
2. On dit que deux espaces vectoriels E et F sont isomorphes s'il existe un isomorphisme $f : E \to F$.

Les résultats généraux sur les structures amènent directement :

Théorème 6.1.19 (Réciproque d'un isomorphisme)
Soit f un isomorphisme entre E et F. Alors f^{-1} est une application linéaire, et donc un isomorphisme de F vers E.

I.3 Endomorphismes

Conformément à la terminologie générale, un endomorphisme est une application linéaire d'un espace dans lui-même :

Définition 6.1.20 (Endomorphisme)

Soit E un espace vectoriel sur $\mathbb{K}$. Une application linéaire de E dans E est appelée *endomorphisme de* E. On note $\mathcal{L}(E)$ l'ensemble $\mathcal{L}(E,E)$ des endomorphismes de E.

Exemples 6.1.21

1. L'identité $\mathrm{Id}_E : x \mapsto x$ de E dans E.
2. L'homothétie (vectorielle) de rapport λ : $\lambda\mathrm{Id}_E : x \mapsto \lambda x$, de E dans E ($\lambda \in \mathbb{K}$).
3. La dérivation formelle $D : \mathbb{R}[X] \to \mathbb{R}[X]$
4. La dérivation analytique $D : \mathcal{C}^?([a,b]) \to \mathcal{C}^?([a,b])$
5. L'application matricielle $X \mapsto MX$ si M est...

La bilinéarité de la composition des applications linéaires permet de définir sur $\mathcal{L}(E)$ une loi de composition $\circ$, distributive sur la somme. On obtient alors la structure de l'ensemble $\mathcal{L}(E)$ des endomorphismes de E :

Proposition 6.1.22 (Structure de $\mathcal{L}(E)$)

L'ensemble $(\mathcal{L}(E), +, \circ)$ est muni d'une structure de $\mathbb{K}$-algèbre.

Nous rappelons ci-dessous la définition d'une $\mathbb{K}$-algèbre :

Définition 6.1.23 ($\mathbb{K}$-algèbre)

Étant donné un corps $\mathbb{K}$, une $\mathbb{K}$-algèbre est un espace vectoriel A sur $\mathbb{K}$, muni d'une seconde loi de composition interne $\times$, compatible avec la loi externe dans le sens suivant :

$$\forall \lambda \in \mathbb{K},\quad \forall (x,y) \in A^2,\quad \lambda \cdot (x \times y) = (\lambda \cdot x) \times y = x \times (\lambda \cdot y),$$

et telle que $(A, +, \times)$ soit un anneau.

La composition ayant les propriétés usuelles d'un produit, on utilise souvent les conventions de notation usuelles pour les produits. En particulier la notation vu désigne l'endomorphisme $v \circ u$ (omission du signe opératoire), et :

Notation 6.1.24 (Composition itérée)

Étant donné un endomorphisme u de E, et un entier $n \in \mathbb{N}$, on désigne par u^n la n-ième composée itérée de u, définie récursivement par $u^0 = Id_E$ et $u^n = u \circ u^{n-1}$ pour $n \in \mathbb{N}^*$.

Remarquez que ces conventions de notation ne sont pas gênantes dans la mesure où dans l'espace vectoriel E, on ne dispose pas d'un produit. L'expression $v(x)u(x)$ n'ayant pas de sens, la notation vu ne peut désigner que la composition. De même pour f^n, l'expression $f(x)^n$ n'ayant pas de sens.

Remarque 6.1.25

1. L'anneau $\mathcal{L}(E)$ est-il commutatif ? À quelle condition nécessaire et suffisante sur la dimension de E l'est-il ?

2. L'anneau $\mathcal{L}(E)$ est-il intègre ?

Cette dernière remarque amène la définition suivante :

Définition 6.1.26 (Endomorphisme nilpotent)
Un endomorphisme nilpotent de E est un endomorphisme n tel qu'il existe $n \in \mathbb{N}$ tel que $u^n = 0_{\mathcal{L}(E)}$. La valeur minimale de n vérifiant cette propriété est appelée *indice de nilpotence de u*.

Exemple 6.1.27

1. $X \mapsto \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} X$
2. La dérivation, vue comme endomorphisme de ...

Ainsi, d'une certaine façon, si u est nilpotent d'incide n, l'endormophisme u « annule » le polynôme X^n. On dira alors que X^n est polynôme annulateur de u. un autre polynôme annulateur est X^{n+1}, ou encore $X^n(X-1)$. La valeur de n étant minimale, X^n est le polynôme unitaire de plus bas degré annulant u : on dira dans ce cas qu'il s'agit du polynôme minimal de u. Nous généralisons ci-dessous ces notions pour des endomorphismes quelconques.

Définition 6.1.28 (Polynôme d'endomorphisme)
Soit $P = \sum_{k=0}^{n} a_k X^k$ un polynôme de $\mathbb{K}[X]$ et u un endomorphisme d'un $\mathbb{K}$-ev E. On définit le polynôme $P(u)$ de l'endomorphisme u par :

$$P(u) = \sum_{k=0}^{n} a_k u^k : \ x \longmapsto \sum_{k=0}^{n} a_k u^k(x)$$

La structure d'algèbre de $\mathcal{L}(E)$ nous permet d'affirmer que $P(u) \in \mathcal{L}(E)$.

Remarquez qu'il s'agit d'un cas particulier de spécialisation d'un polynôme à des éléments d'une algèbre, ainsi que nous l'avons vu dans le chapitre sur les polynômes.
Deux polynômes d'un même endomorphisme commutent, ce qui découle de la commutativité dans $\mathbb{K}[X]$, via le lemme suivant :

Lemme 6.1.29 (Spécialisation d'un produit)
Soit u un endomorphisme de E et P et Q deux polynômes de $\mathbb{K}[X]$. Alors $(PQ)(u) = P(u) \circ Q(u)$.

Proposition 6.1.30 (Commutation des polynômes d'endomorphisme)
Soit u un endormophisme de E et P et Q deux polynômes de $\mathbb{K}[X]$. Alors les deux endomorphismes $P(u)$ et $Q(u)$ commutent :

$$P(u) \circ Q(u) = Q(u) \circ P(u).$$

Cela permet en particulier d'exploiter une factorisation d'un polynôme P pour le calcul de $P(u)$.

Exemple 6.1.31

1. D désignant l'endomorphisme de dérivation de $\mathcal{C}^\infty(RR)$, et P désignant le polynôme $P = (X-1)(X-2)(X-3) = X^3 - 6X^2 + 11X - 6$, on obtient, pour une fonction f de classe $\mathcal{C}^\infty$:

$$f^{(3)} - 6f^{(2)} + 11f' - 6f = (D - \mathrm{Id}) \circ (D - 2\mathrm{Id}) \circ (D - 3\mathrm{Id})(f),$$

   ce qu'on vérifie aisément de façon directe (Id désigne l'identité de $\mathcal{C}^\infty(\mathbb{R})$).
2. Qu'obtient-t-on dans le même contexte que ci-dessus avec $P = (X+1)^n$?

Nous pouvons alors généraliser les notions entrevues lors de l'étude des endomorphismes nilpotents.

Définition 6.1.32 (Polynôme annulateur)

On dit que $P \in \mathbb{K}[X]$ est un polynôme annulateur de $u \in \mathcal{L}(E)$ si $P(u) = 0_{\mathcal{L}(E)}$.

Proposition 6.1.33 (Structure de l'ensemble des polynômes annulateurs)

L'ensemble des polynômes annulateurs de E forme un idéal de $\mathcal{L}(E)$.

L'anneau $\mathbb{K}[X]$ étant principal, on peut alors définir :

Définition 6.1.34 (Polynôme minimal)

Si l'endomorphisme u admet au moins un polynôme annulateur non nul, on définit le polynôme minimal de u comme étant l'unique polynôme unitaire engendrant l'idéal des polynômes annulateurs de u.

Nous verrons plus loin que si E est de dimension finie n, alors $\mathcal{L}(E)$ est également de dimension finie, égale à n^2. Cela se comprend facilement une fois assimilée la corespondance entre applications linéaires et matrices. Admettant provisoirement ce résultat, nous obtenons :

Proposition 6.1.35 (Existence d'un polynôme annulateur)

Soit u un endomorphisme d'un espace E de dimension finie n. Alors u admet un polynôme annulateur non nul, donc un polyôme minimal. De plus, le degré du polynôme minimal est au plus n^2.

La théorie de la réduction des endormophismes (et en particulier le théorème de Cayley-Hamilton que vous verrez l'an prochain) permet d'établir que le polynôme minimal est de degré au plus n. Le théorème de Cayley-Hamilton permet de déterminer de façon explicite un polynôme annulateur d'un endomorphisme dont on connaît une représentation matricielle dans une base, à l'aide des déterminants.

I.4 Automorphisme

Définition 6.1.36 (Automophisme, GL(E))

Un automophisme de E est un endomorphisme bijectif, donc une application linéaire qui est à la fois un endomorphisme et un isomorphisme. On appelle *groupe linéaire de E*, et on note $\mathrm{GL}(E)$, l'ensemble des automorphismes de E

La terminologie est justifiée par le théorème suivant :

Théorème 6.1.37 (Structure de GL(E))
L'ensemble GL(E) *muni de la composition des endomorphismes, est un groupe.*

Comme on le verra, dans le chapitre suivant, si E est de dimension n, il existe un isomophisme de groupe entre GL(E) et le groupe $\mathrm{GL}_n(\mathbb{K})$ des matrices inversibles de taille $n \times n$ à coefficients dans $\mathbb{K}$.

I.5 Projecteurs et symétries

Définition 6.1.38 (Projecteur, symétrie)
1. Soit $p \in \mathcal{L}(E)$. On dit que p est un projecteur (ou une projection) de E ssi $p \circ p = p$.
2. Soit $s \in \mathcal{L}(E)$. On dit que s est une symétrie ssi $s \circ s = \mathrm{id}$.

Ainsi, un projecteur est par définition un endomorphisme dont $X^2 - X$ est polynôme annulateur, et une symétrie est un endomorphisme dont $X^2 - 1$ est polynôme annulateur.

Proposition 6.1.39 (Caractérisation de l'image d'un projecteur)
Soit p un projecteur de E. Alors $x \in \mathrm{Im}(p)$ si et seulement si $p(x) = x$, ce qui se traduit par l'égalité ensembliste :

$$\mathrm{Im}(p) = \mathrm{Ker}(p - \mathrm{Id}_E).$$

Théorème 6.1.40 (Diagonalisation d'un projecteur)
Soit p un projecteur de E. Alors :

$$E = \mathrm{Ker}(p) \oplus \mathrm{Ker}(p - Id_E).$$

Cette propriété exprime le fait que p est « diagonalisable ».
De façon générale, dire qu'un endomorphisme f est diagonalisable signifie que l'espace E peut être décomposé en somme directe d'espaces stables par f sur lesquels p induit une homothétie. Ainsi, f est entièrement déterminé par un ensemble de vecteurs engendrant E et sur lesquels f est simplement une dilatation. En termes plus précis :

Définition 6.1.41 (Endomorphisme diagonalisable)
- Un endomorphisme f de E est diagonalisable s'il existe une base $(b_i)_{i\in I}$ de E et une famille $(\lambda_i)_{i\in I}$ de scalaires tels que pour tout $i \in I$, $f(b_i) = \lambda_i b_i$.
- Les λ_i sont appelées valeurs propres de f.
- Étant donné une valeur propre λ, un vecteur x non nul tel que $f(x) = \lambda x$ est appelé vecteur propre associé à λ
- Si λ est une valeur propre de f, $\mathrm{Ker}(f - \lambda\mathrm{Id})$ est appelé sous-espace propre de f associé à la valeur propre λ.

Ainsi, dire que f est diagonalisable revient à dire que E est somme directe des sous-espaces propres de f. Vous verrez l'année prochaine que cela équivaut en dimension finie à l'existence d'une base (par exemple la base donnée dans l'énoncé de la définition) relativement à laquelle la matrice de f est diagonale (voir chapitre suivant pour la notion de matrice associée à un endomorphisme)

Corollaire 6.1.42

Un projecteur distinct de 0 *ou* Id, *admet exactement deux valeurs propres* 0 *et* 1, *et est diagonalisable.*

Remarque 6.1.43

Les valeurs propres de u sont les racines du polynôme annulateur $X^2 - X$. Ce n'est pas anodin. L'ensemble des valeurs propres est toujours inclus dans l'ensemble des racines d'un polynôme annulateur, et on a l'égalité s'il s'agit du polynôme minimal.

Le théorème précédent est en fait une caractérisation des projecteurs

Théorème 6.1.44 (Caractérisation géométrique des projecteurs)

Soit $p \in \mathcal{L}(E)$.

- *Alors* p *est un projecteur si et seulement s'il existe deux sev* F *et* G *de* E *tels que* $F \oplus G = E$, *et :*

$$\forall u \in F,\ \forall v \in G,\ p(u+v) = u.$$

  Cette dernière identité traduit le fait que p *est la projection géométrique sur* F *parallèlement à* G.
- *Dans ce cas, on a* $F = \mathrm{Im}(p)$ *et* $G = \mathrm{Ker}(p)$.
- *Ainsi, un projecteur est une projection géométrique sur* $\mathrm{Im}(p)$ *parallèlement à* $\mathrm{Ker}(p)$.

Le même travail peut être fait pour les symétries.

Théorème 6.1.45 (Diagonalisation d'une symétrie)

Soit s *une symétrie de* E. *Alors :*

$$E = \mathrm{Ker}(s + \mathrm{Id}_E) \oplus \mathrm{Ker}(s - Id_E).$$

Ce dernier résultat traduit le fait que s est diagonalisable, et si s est distinct de Id_E et $-\mathrm{Id}_E$, alors les valeurs propres de s sont exactement 1 et -1. On peut à nouveau remarquer qu'il s'agit des racines du polynôme annulateur $X^2 - 1$.

Encore une fois, le théorème précédent est une caractérisation des symétries, et donne l'interprétation géométrique des symétries.

Théorème 6.1.46 (Caractérisation géométrique des symétries)

Soit $s \in \mathcal{L}(E)$.

- *Alors* s *est une symétrie si et seulement s'il existe deux sev* F *et* G *de* E *tels que* $F \oplus G = E$, *et :*

$$\forall u \in F,\ \forall v \in G,\ s(u+v) = u - v.$$

  Cette dernière identité traduit le fait que s *est la symétrie géométrique par rapport à* F *parallèlement à* G.
- *Dans ce cas, on a* $F = \mathrm{Im}(s) = \mathrm{Ker}(s - \mathrm{Id}_E)$ *et* $G = \mathrm{Ker}(s + \mathrm{Id}_E)$.
- *Ainsi, une symétrie au sens algébrique* s *est une symétrie géométrique par rapport à* $\mathrm{Ker}(s - \mathrm{Id}_E)$ *(l'ensemble des points fixes), parallèlement à* $\mathrm{Ker}(s + \mathrm{Id}_E)$.

II Applications linéaires et familles de vecteurs

II.1 Détermination d'une application linéaire

Nous commençons par un résultat de rigidité, exprimant le fait que l'image d'un nombre limité de vecteurs de E par une application linéaire u permet de déterminer entièrement l'application linéaire u. En effet, par linéarité, la connaissance de u sur une famille de vecteur amène sa connaissance sur tout l'espace engendré par cette famille. On obtient donc le résultat suivant :

Proposition 6.2.1 (Détermination d'une AL par l'image d'une base)

Étant donné $(b_i)_{i\in I}$ une base de E et $(f_i)_{i\in I}$ une famille quelconque de F, il existe une unique application linéaire $u \in \mathcal{L}(E,F)$ telle que pour tout $i \in I$, $f(b_i) = f_i$.

Ainsi, une application linéaire de E dans F est entièrement déterminée par l'image d'une base de E.

Exemples 6.2.2

1. Déterminer l'expression générale de l'application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^2$ telle que $f(1,0) = (3,2)$ et $f(0,1) = (2,1)$.
2. Montrer que toute application linéaire de $\mathbb{R}^p$ dans $\mathbb{R}^n$ est de la forme $X \mapsto MX$, et décrire M à partir d'une base de $\mathbb{R}^p$.
3. Soit $(b_i)_{i\in I}$ une base de E et $(c_j)_{j\in J}$ une base de F. Alors pour tout $(i,j) \in I \times J$, il existe une unique application linéaire $u_{i,j}$ telle que $u_{i,j}(b_i) = c_j$ et pour tout $k \neq i$, $u_{i,j}(b_k) = 0$.

Proposition 6.2.3 (Base de $\mathcal{L}(E,F)$)

Si E est de dimension finie, la famille $(u_{i,j})_{(i,j)\in I\times J}$ décrite dans l'exemple ci-dessus est une base de $\mathcal{L}(E,F)$.

Corollaire 6.2.4 (Dimension de $\mathcal{L}(E,F)$)

Si E et F sont de dimension finie, alors :

$$\dim \mathcal{L}(E,F) = \dim(E) \times \dim F.$$

Remarques 6.2.5

1. Pensez au cas des applications linaires de $\mathbb{K}^p$ dans $\mathbb{K}^n$, qui sont obtenus, comme on l'a vu plus haut, sous la forme $u_M : X \mapsto MX$, avec $M \in \mathrm{Mat}_{n,p}(\mathbb{K})$. Il n'est pas dur de voir que $M \mapsto u_M$ est un isomorphisme d'espace vectoriel entre $\mathcal{M}_{n,p}(\mathbb{K})$ et $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$. Or, une matrice étant la donnée de $n \times p$ coefficients indépendants, la dimension de $\mathcal{M}_{n,p}(\mathbb{K})$ est np. On retrouve sur cet exemple le résultat précédent.
2. Ce n'est d'ailleurs pas qu'un exemple, car comme on le verra dans le paragraphe suivant, tout $\mathbb{K}$-espace vectoriel E de dimension finie est isomorphe à $\mathbb{K}^n$, où $n = \dim(E)$. *Via* cet isomophisme, la situation décrite ci-dessus est générique.
3. En partant des bases canoniques de $\mathbb{K}^p$ et de $\mathbb{K}^n$, la base décrite ci-dessus de $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ correspond à la base de $\mathcal{M}_{n,p}(\mathbb{K})$ formée des matrices $E_{i,j}$, continuées de coefficients tous nuls, sauf le coefficient en position (i,j), égal à 1. Il s'agit de la *base canonique de* $\mathcal{M}_{n,p}(\mathbb{K})$.
4. L'espace vectoriel $\mathcal{M}_{n,p}(\mathbb{K})$ est isomorphe à $\mathbb{K}^{np}$ par l'isomophisme consistant à réordonner les coefficients en les écrivant en une seule ligne (ou en une seule colonne), en juxtaposant les différentes

lignes les unes à la suite des autres. Via cet isomorphisme, quitte à réordonner les éléments de la base, la base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$ correspond à la base canonique de $\mathbb{K}^{np}$.

II.2 Caractérisations de l'injectivité et de la surjectivité par l'image de bases

L'image d'une base par un endomorphisme déterminant entièrement l'application linéaire, toutes les propriétés telles que l'injectivité et la surjectivité peuvent se voir déjà dans la description de l'image d'une base. Ainsi, ces propriétés peuvent être caractérisées par l'image d'une base.

Proposition 6.2.6 (Caractérisation de l'injectivité par l'image d'une base)

Soit $f \in \mathcal{L}(E,F)$. Les propriétés suivantes sont équivalentes :

(i) f est injective ;

(ii) l'image de toute famille libre de E par f est une famille libre de F ;

(iii) l'image de toute base de E par f est une famille libre de F ;

(iv) il existe une base de E dont l'image par f est une famille libre de F.

Proposition 6.2.7 (Caractérisation de la surjectivité par l'image d'une base)

Soit $f \in \mathcal{L}(E,F)$. Les propriétés suivantes sont équivalentes :

(i) f est surjective ;

(ii) l'image de toute famille génératrice de E par f est une famille génératrice de F ;

(iii) l'image de toute base de E par f est une famille génératrice de F ;

(iv) il existe une base de E dont l'image par f est une famille génératrice de F.

En combinant ces deux caractérisations, nous obtenons :

Proposition 6.2.8 (Caractérisation de la bijectivité par l'image d'une base)

Soit $f \in \mathcal{L}(E,F)$. Les propriétés suivantes sont équivalentes :

(i) f est un isomorphisme ;

(ii) l'image de toute base de E par f est une base de F ;

(iii) il existe une base de S dont l'image par f est une base de F.

On en déduit en particulier :

Corollaire 6.2.9 (Dimension d'espaces isomorphes)

Soit E et F deux espaces isomorphes. Alors si l'un des deux espaces E ou F est de dimension finie, les deux le sont, et $\dim(E) = \dim(F)$.

Corollaire 6.2.10 (Classification à isomorphisme près des espaces de dimension finie)

(i) Tout $\mathbb{K}$-ev E de dimension finie n est isomorphe à $\mathbb{K}^n$.

(ii) Si $n \neq m$, $\mathbb{K}^n$ et $\mathbb{K}^m$ ne sont pas isomorphes

Plus précisément, un isomorphisme tel que dans (i) est obtenu par le choix d'une base de E, et correspond alors à la donnée des coordonnées d'un vecteur x dans cette base.

Le dernier résultat peut se réexprimer en remarquant que la relation d'isomorphisme définit une relation d'équivalence, notée $\simeq$, sur l'ensemble $\mathcal{E}_f$ des $\mathbb{K}$-ev de dimension finie. L'espace quotient est alors :

$$(\mathcal{E}_f/\simeq) = \mathbb{N}.$$

II.3 Recollements

Une généralisation de la détermination d'une application linéaire par l'image d'une base est la possibilité de définir, et ceci de façon unique, une application linéaire à partir de ses restrictions sur une famille $(E_i)_{i\in I}$ de sous-espaces vectoriels tels que $(E_i)_{i\in I}$. Le cas de la détermination à partir de l'image d'une base en découle en considérant la famille $(\text{Vect}(b_i))_{i\in I}$. On énonce :

Théorème 6.2.11
Soit I un ensemble fini. Soit $(E_i)_{i\in I}$ une famille de sous-espaces vectoriels de E tels que $\bigoplus_{i\in I} E_i = E$, et soit pour tout $i \in I$, $u_i \in \mathcal{L}(E_i, F)$. Alors il existe une et une seule application linéaire $u \in \mathcal{L}(E, F)$ telle que pour tout $i \in I$, $u_i = u_{|E_i}$.

Si, sous les mêmes conditions sur les E_i, les u_i sont des endomorphismes des E_i, on peut faire la même chose en considérant les applications $v_i = j_i \circ u_i$, où j_i est l'inclusion de E_i dans E. L'application u obtenue est alors un endomorphisme de E.
En revanche, contrairement au cas précédent où toute application linéaire de $\mathcal{L}(E, F)$ peut être otenue comme recollement d'applcations linéaires de $\mathcal{L}(E_i, F)$, tout endormophisme de E ne peut pas être obtenu comme recollement d'endomorphismes des E_i (le problème provenant de la corestriction faite sur l'espace image). Matriciellement, si E est de dimension finie, après choix d'une base de E obtenue par recollement de bases des E_i, un endomorphisme pouvant s'obtenir ainsi est un endormorphisme dont la matrice sera diagonale par blocs (voir chapitre suivant)
Nous donnons une condition pour que ce soit le cas :

Proposition 6.2.12 (Restriction d'un endomorphisme)
Soit E un $\mathbb{K}$-ev et F un sous-espace vectoriel de E. Soit $u \in \mathcal{L}(E)$. Si F est stable par u (c'est-à-dire si $u(F) \subset F$, ou encore si pour tout $x \in F$, $u(x) \in F$) alors u induit par restriction et corestriction un endomorphisme u_F de F.

Proposition 6.2.13 (Décomposition d'un endomorphisme sur une somme de sev stables)
Soit E un $\mathbb{K}$-ev et $(E_i)_{i\in I}$ des sev en nombre fini, tels que $E = \bigoplus_{i\in I} E_i$. Soit $u \in \mathcal{L}(E)$. Si les E_i sont tous stables par u, et si u_i désigne l'endomorphisme de $\mathcal{L}(E_i)$ induit par u, alors u est le recollement des endomorphismes u_i dans le sens évoqué ci-dessus.

III Applications linéaires en dimension finie

III.1 Rang d'une application linéaire

Définition 6.3.1 (Rang d'une application linéaire)
Soit $u \in \mathcal{L}(E, F)$ une application linéaire. Si $\text{Im}(u)$ est de dimension finie, on définit le rang de u par :

$$\text{rg}(u) = \dim(\text{Im}(u)).$$

Proposition 6.3.2

Soit $(x_i)_{i\in I}$ une famille génératrice de E. Le rang de $\mathrm{Im}(u)$, s'il existe, est égal au rang de la famille $(u(x_i))_{i\in I}$.

Proposition 6.3.3 (Existence du rang en dimension finie)

- *Soit $u \in \mathcal{L}(E,F)$. Si E et/ou F sont de dimension finie alors $\mathrm{Im}(u)$ également, et*

$$\mathrm{rg}(u) \leqslant \dim(E) \qquad \text{et/ou} \qquad \mathrm{rg}(u) \leqslant \dim(F).$$

- *Sous les conditions idoines d'existence :*
  - $\mathrm{rg}(u) = \dim(E)$ *si et seulement si u est injective ;*
  - $\mathrm{rg}(u) = \dim(F)$ *si et seulement si u est surjective.*

On en déduit en particulier :

Théorème 6.3.4 (Caractérisation des isomorphismes en dimension finie)

*Soit E et F deux espaces vectoriels de **même dimension** finie n, et soit $f \in \mathcal{L}(E,F)$. Alors les propositions suivantes sont équivalentes :*

(i) f *est un isomorphisme ;*

(ii) $\mathrm{rg}(f) = n$ *;*

(iii) f *est injective ;*

(iv) f *est surjective.*

En particulier, si E est de dimension finie, un endomorphisme $f \in \mathcal{L}(E)$ est un automorphisme si et seulement si f est injective si et seulement si f est surjective.

Théorème 6.3.5 (Effet d'une composition sur le rang)

Soit $u \in \mathcal{L}(E,F)$ et $v \in \mathcal{L}(F,G)$.

1. $\mathrm{rg}(v\circ u) \leqslant \min(\mathrm{rg}(u), \mathrm{rg}(v))$.
2. *Si v est injective,* $\mathrm{rg}(v\circ u) = \mathrm{rg}(u)$.
3. *Si u est surjective,* $\mathrm{rg}(v\circ u) = \mathrm{rg}(v)$

En particulier :

Corollaire 6.3.6 (Invariance du rang par composition par un isomophisme)

Tout est dit dans le titre, la composition pouvant se faire à droite ou à gauche.

III.2 Théorème du rang

Cette section courte, mais ô combien importante, a pour objet un résultat reliant la dimension de l'image et la dimension du noyau d'une application linaire (théorème du rang). En gros, ce théorème dit que, partant du fait qu'il n'y a pas de perte de dimension entre l'espace initial et l'image lorsque u est injective, tout défaut d'injectivité se traduit par une perte sur l'image égale à la dimension du noyau : la dimension du noyau (donc la dimension d'un ensemble d'éléments tous envoyés sur le même point) correspond à la dimension qu'on perd entre l'espace initial et l'image.
Pour établir ce résultat, nous commençons par étudier les propriétés relatives au noyau et à l'image de restrictions.

Lemme 6.3.7 (Noyau et image d'une restriction)

Soit $u \in \mathcal{L}(E,F)$, et E' un sous-espace de E. Soit $v \in \mathcal{L}(E',F)$ la restriction de u à E'. Alors :

- $\mathrm{Ker}(v) = \mathrm{Ker}(u) \cap E'$
- *Si* $\mathrm{Ker}(u) + E' = E$, $\mathrm{Im}(v) = \mathrm{Im}(u)$.

Corollaire 6.3.8 (Restriction de u à un supplémentaire de $\mathrm{Ker}(u)$)

Soit S un supplémentaire de $\mathrm{Ker}(u)$ *dans E. Alors u induit un isomorphisme de S sur* $\mathrm{Im}(u)$.

Remarquez que ce dernier résultat ne nécessite pas d'hypothèse de finitude. Cependant, si E n'est pas de dimension finie, il est nécessaire de supposer l'axiome du choix, afin d'assurer l'existence du supplémentaire S.
On déduit du corollaire précédent le très important :

Théorème 6.3.9 (Théorème du rang)

Soit E un espace vectoriel de dimension finie, et F un espace vectoriel quelconque. Soit $f \in \mathcal{L}(E,F)$. Alors :

$$\dim \mathrm{Ker}\, f + \mathrm{rg}\, f = \dim E.$$

IV Formes linéaires

Nous terminons ce chapitre par l'étude d'une famille importante d'applications linéaires : les formes linéaires, qui correspondent aux applications linéaires à valeurs dans le corps de base. Les formes linéaires sont intimement liées à la notion d'hyperplan (généralisant la notion de plan en dimension 3). Elles sont aussi à la base des théories de dualité (qui ne sont pas au programme).

IV.1 Formes linéaires, espace dual, hyperplan

Définition 6.4.1 (Forme linéaire)

Une forme linéaire sur un $\mathbb{K}$-espace vectoriel E est une application linéaire de E vers $\mathbb{K}$, donc un élément de $\mathcal{L}(E,\mathbb{K})$.

Exemples 6.4.2

1. $f \mapsto \displaystyle\int_a^b f(t)\,\mathrm{d}t$ sur ...
2. la trace sur $\mathcal{M}_n(\mathbb{R})$.
3. L'évaluation des polynômes en a.

Définition 6.4.3 (Dual)

Soit E un espace vectoriel. On appelle *dual* de E, et on note E^*, l'espace vectoriel $\mathcal{L}(E,\mathbb{K})$ constitué des formes linéaires. Le bidual E^{**} est alors le dual de E^*.

On définit alors les hyperplans de la manière suivante :

Proposition 6.4.4 (Hyperplan)

Soit H un sous-espace vectoriel de E. On dit que H est un hyperplan *de E s'il existe une forme linéaire non nulle $\varphi \in E^*$ telle que $H = \mathrm{Ker}(\varphi)$. L'équation $\varphi(x) = 0$, caractérisant l'appartenance à H, est appelée* équation de H.

Proposition 6.4.5 (Caractérisation des hyperplans en dimension finie)

Si E est de dimension finie n, les hyperplans de E sont exactement les sous-espaces vectoriels de E de dimension $n-1$.

En appelant *codimension du sous-espace F de E* la quantité $\mathrm{codim}_E(F) = \dim(E) - \dim(F)$, les hyperplans sont donc les sous-espaces de E de codimension 1.

Exemples 6.4.6

1. Les plans vectoriels de $\mathbb{R}^3$.
2. Les droites vectorielles de $\mathbb{R}^2$.
3. Les polynômes sans terme constant.

Théorème 6.4.7 (Supplémentaire d'un hyperplan)

- *Soit H un hyperplan d'un espace (non nécessairement de dimension finie). Alors un sous-espace vectoriel S est un supplémentaire de H si et seulement si S est une droite non contenue dans H.*
- *De façon symétrique, soit D une droite de E ; un sous-espace vectoriel S de E est un supplémentaire de E si et seulement si S est un hyperplan ne contenant pas D.*

Malgré l'utilisation de supplémentaires en dimension non nécessairement finie, ce théorème ne nécessite pas l'axiome du choix.

Proposition 6.4.8 (Comparaison de deux équations de H)

Soit H un hyperplan de E, d'équation $\varphi \in \mathrm{E}^*$. Alors pour tout $\psi \in E^*$, $\psi(x) = 0$ est une équation de H si et seulement si $\psi \neq 0$ et $\psi \in \mathrm{Vect}(\varphi)$.

Théorème 6.4.9 (Intersection d'hyperplans)

Soit E un espace de dimension finie n.

1. *L'intersection de m hyperplans de E est un sous-espace vectoriel de dimension au moins $n-m$.*
2. *Réciproquement, tout sous-espace vectoriel F de E de dimension $n-m$ peut s'écrire comme l'intersection de m hyperplans.*

Remarquez qu'en dimension finie, après choix d'une base $(b_1, \cdots, b_n)$, une forme linéaire sera décrite par une expression du type

$$\varphi(x) = \sum_{i=1}^{n} a_i x_i,$$

où les x_i sont les coordonnées de x dans la base $(b_1, \ldots, b_n)$. Ainsi, l'équation d'un hyperplan est de la forme

$$a_1 x_1 + \cdots + a_n x_n = 0.$$

On retrouve les équations usuelles d'une droite dans $\mathbb{R}^2$ ($ax+by=0$) ou d'un plan dans $\mathbb{R}^3$ ($ax+by+cz=0$), les x, y et z correspondant ici aux coordonnées dans la base canonique.

Ce que dit le dernier théorème est alors simplement le fait qu'un sous-espace de dimension $n-m$ (donc de codimension m) peut être décrit par un système de m équations de ce type.

IV.2 Qu'est-ce que le principe de dualité ? (hors-programme)

Nous introduisons ici, sans preuve, quelques notions relatives à la notion de dualité en algèbre linéaire. On suppose que E est de dimension finie. La plupart des notions peuvent se définir en dimension infinie, mais l'identification de E et E^{**}, à la base du principe de dualité, n'est valable qu'en dimension finie.

Soit $\varphi \in E^*$ et $x \in E$. On note $\langle x, \varphi \rangle = \varphi(x)$. Cela définit clairement une application bilinéaire de $E \times E^*$ dans $\mathbb{K}$. En particulier, en fixant x, on obtient une forme linéaire de E^* dans $\mathbb{K}$, donc un élément du bidual E^{**} :

Définition 6.4.10 (Application canonique)
L'application $J : E \mapsto E^{**}$ qui à x associe l'élément de E^{**} défini par $\varphi \mapsto \langle x, \varphi \rangle = \varphi(x)$ est appelée application canonique de E dans son bidual.

Théorème 6.4.11 (Identification de E et E^{**})
Si E est de dimension finie, J est un isomorphisme.

On peut alors identifier l'espace E et son bidual E^{**}. Le principe de dualité est alors que toute propriété de dualité démontrée pour les couples (E, E^*) reste vraie pour les couples (E^*, E).

Il peut être intéressant, de ce fait, de transcrire les objets définis sur E en des objets analogues (leur correspondant par dualité) sur E^*. Nous commençons par les applications linéaires.

Définition 6.4.12 (Transposée d'une AL)
Soit $f \in \mathcal{L}(E, F)$. On appelle transposée de f, et on note ${}^t f$, l'application de $\mathcal{L}(F^*, E^*)$ définie par ${}^t f(\varphi) = \varphi \circ f$.

On peut vérifier que pour tout $x \in E$ et tout $\varphi \in E^*$, $\langle x,\ {}^t f(\varphi) \rangle = \langle f(x), \varphi \rangle$.
On peut également vérifier que, modulo l'isomophisme J, on a ${}^t({}^t f) = f$.

La notation ${}^t f$ n'est pas anodine. Par la notion de base duale exposée ci-dessous, et en anticipant les représentations matricielles des applications linéaires (voir chapitre suivant), la matrice de ${}^t f$ dans la base duale est la transposée de la matrice de f dans la base initiale de E.

Cette base duale, qui est donc l'objet associé par dualité à une base de E se définit de la manière suivante. On suppose ici que E est de dimension finie, et on considère une base $(e_1, \dots, e_n)$ de E.

Définition 6.4.13 (Base duale)
La base duale de $(e_1, \dots, e_n)$ est la famille $(e_1^*, \dots, e_n^*)$ d'éléments de E^*, où e_i^* est l'unique forme linéaire prenant la valeur 1 sur e_i et 0 sur les autres e_j. Il s'agit d'une base de E^*

Remarquez qu'il ne s'agit de rien d'autre que la base de $\mathcal{L}(E, F)$ décrite plus haut, dans le cas où $F = \mathbb{K}$, de base (1). Remarquez également que e_i^* est caractérisé par l'équation :

$$\langle e_j, e_i^* \rangle = \delta_{i,j},$$

où $\delta_{i,j}$ est le symbole de Kronecker, égal à 1 si $i = j$, et 0 sinon.
Là encore, il convient de remarquer que la notion définie est bien une notion duale, dans le sens où la base duale de la base duale s'identifie à la base initiale par l'isomorphisme J.
Enfin, nous définissons les version duales des sous-espaces vectoriels de E.

Définition 6.4.14 (Sous-espace orthogonal)

Soit A un sous-espace vectoriel de E, on appelle orthogonal de H dans E^* le sous-espace vectoriel $A^\circ = \{\varphi \in \mathrm{E}^* \mid \langle x, \varphi\rangle = 0\}$.

L'orthogonalité inverse le sens des inclusions, et échange intersection et union. On vérifie une fois de plus que $A^{\circ\circ} = A$, via J, ce qui nous assure que la notion est la bonne.

Nous ne poussons pas plus l'étude de la dualité, qui pourrait nous mener fort loin, notre but étant uniquement de donner un premier aperçu d'une notion que vous développerez ultérieurement de façon plus approfondie.

7 Matrices

Comme nous l'avons vu dans le chapitre précédent, la notion de matrice est indissociable de celle d'application linéaire. Nous avons déjà constaté au cours d'exemples que toute application linéaire de $\mathbb{K}^p$ dans $\mathbb{K}^n$ peut se représenter matriciellement sous la forme $X \mapsto MX$. Cela reste vrai pour toute application linéaire entre deux espaces vectoriels de dimension finie, à condition d'avoir fixé au préalable une base de chacun de ces espaces : les systèmes de coordonnées définis par ces bases nous ramènent alors au cas de $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$.

Note Historique 7.0.15

La notion de tableau de nombre en soi est vieille (étude de carrés magiques dans l'antiquité). Les règles opératoires, et le rapport avec les applications linéaires est plus récent. C'est Gauss le premier qui met le doigt sur ce rapprochement, à une époque où espaces vectoriels et applications linéaires n'avaient pas encore vu le jour. Le problème qui l'intéressait était de faire des changements de variable dans des formes quadratiques (polynômes de plusieurs variables de degré 2), de sorte à se ramener à des formes simples. Ces changements de variables portant sur plusieurs variables se présentent sous la forme de « substitutions linéaires » dans les termes de Gauss (c'est-à-dire d'applications linéaires). Pour alléger ses notations, il présente ses substitutions linéaires en n'en donnant que les coefficients, rangés dans un tableau... Ce n'est rien d'autre que la représentation matricielle d'une application linéaire. Il fait ensuite remarquer que si on enchaîne deux substitutions, on obtient une règle calculatoire permettant de trouver le tableau associé à la composée : il a découvert le produit matriciel.

Il est important de retenir de cette histoire que l'expression un peu compliquée du produit matriciel n'a pas été donnée au hasard : la définition du produit matriciel a été motivée historiquement par les règles de composition des applications linéaires.

I Calcul matriciel

Dans tout ce qui suit, $\mathbb{K}$ désigne un corps, et n et p deux entiers naturels.

I.1 Définition et motivations

Nous rappelons qu'une application linéaire $f \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ est entièrement déterminée, et ceci sans équivoque, par la donnée de $f(e_1), \ldots, f(e_p)$, où $(e_1, \ldots, e_p)$ est la base canonique de $\mathbb{K}^p$. Or, les $f(e_i)$ sont des vecteurs de $\mathbb{K}^n$, donc déterminés par n coordonnées. En les notant en colonne, on a :

$$f(e_1) = \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{n,1} \end{pmatrix}, \quad f(e_2) = \begin{pmatrix} a_{1,2} \\ \vdots \\ a_{n,2} \end{pmatrix}, \quad \ldots \quad f(e_p) = \begin{pmatrix} a_{1,p} \\ \vdots \\ a_{n,p} \end{pmatrix}.$$

Par conséquent, l'application linéaire f est entièrement déterminée par la donnée d'une famille de scalaires $(a_{i,j})_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]}$ (donc une famille « rectangulaire »). Réciproquement, toute telle famille définit une application linéaire de $\mathbb{K}^p$ dans $\mathbb{K}^n$, comme on l'a vu. Cela motive la définition suivante :

Définition 7.1.1 (Matrice)

Une matrice de taille $n \times p$ (n lignes et p colonnes) à coefficients dans $\mathbb{K}$ est la donnée d'une famille $A = (a_{i,j})_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]}$ d'éléments de $\mathbb{K}$. On utilise la représentation planaire suivante :

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix}.$$

Définition 7.1.2 (Ensemble des matrices)

L'ensemble des matrices de taille $n \times p$ (on dit aussi *de type* (n, p)) à coefficients dans $\mathbb{K}$ est noté $\mathcal{M}_{n,p}(\mathbb{K})$. Si $n = p$, on dit que la matrice est *carrée*, et on note simplement $\mathcal{M}_n(\mathbb{K})$ l'ensemble des matrices carrées de taille n.

Le rapport avec les applications linéaires peut alors se formaliser comme suit :

Définition 7.1.3 (Matrice canoniquement associée à $f \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$)

Soit $f \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$. En reprenant les notations ci-dessus, on définit la matrice $\mathrm{Mat}_{b.c.}(f)$ de $\mathcal{M}_{n,p}(\mathbb{K})$ par :

$$\mathrm{Mat}_{b.c.}(f) = \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix} = \Big([f(e_1)]_{b.c.} \Big| \cdots \Big| [f(e_p)]_{b.c} \Big),$$

cette dernière expression étant une description colonne par colonne de la matrice. Ainsi, la i-ième colonne de $\mathrm{Mat}(f)_{b.c.}$ est constituée des coordonnées du vecteur $f(e_i)$ dans la base canonique de $\mathbb{K}^n$. La matrice $\mathrm{Mat}_{b.c.}(f)$ est appelée matrice canoniquement associée à $f \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$

Remarques 7.1.4

1. Le *b.c.* est là pour indiquer que toutes les coordonnées sont prises dans les bases canoniques respectives de l'espace de départ et de l'espace d'arrivée. On généralisera plus loin cette définition à des bases quelconques.
2. Attention à l'inversion des indices : une application linéaire de $\mathbb{K}^p$ dans $\mathbb{K}^n$ fournit une matrice à n lignes et p colonnes.

Réciproquement, étant donné une matrice $M \in \mathcal{M}_{n,p}(\mathbb{K})$, il existe une unique application linéaire f dont la matrice canoniquement associée soit M : il s'agit de l'unique application linéaire f telle que pour tout $i \in [\![1, p]\!]$, $f(e_i)$ soit égal au vecteur de $\mathbb{K}^n$ dont les coordonnées (dans la base canonique) sont données par la i-ième colonne de M. On dit que f est l'application linéaire de $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ canoniquement associée à la matrice M.

I.2 Combinaisons linéaires de matrices

Étant donné deux applications linéaires f et g de $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$, et un scalaire $\lambda \in \mathbb{K}$, la structure d'espace vectoriel de $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ nous assure que $\lambda f + g$ est encore élément de $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$. Soit $(e_1, \ldots, e_p)$ la base

canonique de $\mathbb{K}^p$. En notant :

$$\mathrm{Mat}_{b.c.}(f) = (a_{i,j})_{i,j} = \Big(f(e_1) \,\Big|\, \cdots \,\Big| f(e_p) \Big) \qquad \text{et} \qquad \mathrm{Mat}_{b.c.}(g) = (b_{i,j})_{i,j} = \Big(g(e_1) \,\Big|\, \cdots \,\Big| g(e_p) \Big),$$

on a alors :

$$\begin{aligned} \mathrm{Mat}_{b.c.}(\lambda f + g) &= \Big((\lambda f + g)(e_1) \,\Big|\, \cdots \,\Big| (\lambda f + g)(e_p) \Big) \\ &= \Big(\lambda f(e_1) + g(e_1) \,\Big|\, \cdots \,\Big| \lambda f(e_p) + g(e_p) \Big) \\ &= (\lambda a_{i,j} + b_{i,j})_{i,j} \end{aligned}$$

Cela motive la définition suivant de la somme et du produit par un scalaire pour des matrices :

Définition 7.1.5 (Somme et multiplication par un scalaire)

La somme et la multiplication par un scalaire sont définis coefficient par coefficient. Plus précisément, soit $n, p \in \mathbb{N}^*$, soit $A = (a_{i,j})_{i,j}$, $B = (b_{i,j})_{i,j}$ deux éléments de $\mathcal{M}_{n,p}(\mathbb{K})$ et soit $\lambda \in \mathbb{K}$. Alors :

- on définit sur $\mathcal{M}_{n,p}(\mathbb{K})$ une loi de composition interne $+$ par :

$$A + B = (a_{i,j} + b_{i,j})_{i,j};$$

- on définit sur $\mathcal{M}_{n,p}(\mathbb{K})$ une loi de composition externe, d'opérateurs dans $\mathbb{K}$, par :

$$\lambda A = (\lambda a_{i,j})_{i,j}.$$

Exemples 7.1.6

1. $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 3 & 5 \\ 6 & 7 & 7 \end{pmatrix}$
2. $3 \cdot \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 6 & 3 \\ 6 & 6 & 9 \end{pmatrix}$
3. $5 \times \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ ...???...

Avertissement 7.1.7

On ne somme que des matrices de même taille (même nombre de lignes et de colonnes) !

Les deux lois (interne et externe) qu'on vient de définir sur $\mathcal{M}_{n,p}(\mathbb{K})$ le munissent d'une structure bien connue :

Proposition 7.1.8 (Structure d'ev de $\mathcal{M}_{n,p}(\mathbb{K})$)

L'ensemble $\mathcal{M}_{n,p}(\mathbb{K})$ est un espace vectoriel, isomorphe à $\mathbb{K}^{np}$. En particuler, $\dim(\mathcal{M}_{n,p}(\mathbb{K})) = np$.

Via l'isomorphisme ci-dessus, la base canonique de $\mathbb{K}^{n,p}$ fournit une base de $\mathcal{M}_{n,p}(\mathbb{K})$.

Proposition/Définition 7.1.9 (Base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$)

Soit pour tout $(i,j) \in [\![1,n]\!] \times [\![1,p]\!]$, $E_{i,j}$ la matrice constituée d'un coefficient 1 en position (i,j) et de 0 partout ailleurs. Alors $(E_{i,j})_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]}$ est une base de $\mathcal{M}_{n,p}(\mathbb{K})$, appelée base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$.

Par ailleurs, notre démarche même motivant la définition de la somme et de la multiplication par un scalaire amène de façon immédiate :

Théorème 7.1.10 (Identification de $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ et $\mathcal{M}_{n,p}(\mathbb{K})$)
L'application $f \mapsto \mathrm{Mat}_{b,c}(f)$ est un isomophisme de l'espace vectoriel $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ sur l'espace vectoriel $\mathcal{M}_{n,p}(\mathbb{K})$.

Remarquez que la base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$ correspond à l'image par cet isomorphisme de la base de $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ associée aux bases canoniques de chacun des espaces $\mathbb{K}^p$ et $\mathbb{K}^n$, telle qu'on l'avait définie dans le chapitre précédent.

I.3 Produit de matrices

Reprenant l'exemple ci-dessus, pour tout $X \in \mathbb{K}^n$ qu'on écrit

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 e_1 + \cdots + x_n e_n,$$

on a, par linéarité de f :

$$f(X) = x_1 f(e_1) + \cdots + x_n f(e_n). \tag{7.1}$$

On souhaite que l'évaluation des applications linéaires coïncide avec le produit matriciel (ici le produit de la matrice canoniquement associée à M et d'une matrice colonne représentant les coordonnées de X). Remarquons d'emblée que la matrice colonne X n'a pas une taille quelconque : elle représente un élément de l'espace de départ $\mathbb{K}^n$, donc elle a n lignes, ce qui correspond aussi au nombre de colonnes de M. Nous définissons alors, avec ces compatibilités de dimension, le produit d'une matrice par une colonne :

Définition 7.1.11 (Produit d'une matrice par une matrice colonne)
Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$ et $X \in \mathcal{M}_{p,1}(\mathbb{K})$.
On note $X = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$ les coordonnées de X, et $M = \left(C_1 \middle| \cdots \middle| C_m \right)$ la décomposition en colonnes de M (ainsi, pour tout $i \in [\![1, m]\!]$, $C_i \in \mathcal{M}_{n,1}$ est la i-ème colonne de M). Le produit $M \cdot X$ est alors la matrice colonne de $\mathcal{M}_{n,1}(\mathbb{K})$ définie par :

$$M \cdot X = x_1 C_1 + \cdots + x_m C_m = \sum_{i=1}^{m} x_i C_i.$$

De la formule 7.1, on déduit immédiatement :

Proposition 7.1.12 (Expression matricielle de $f(X)$)
Soit $f \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$; alors, pour tout $X \in \mathbb{K}^p$ (identifié à la colonne de ses coordonnées), on a :

$$f(X) = \mathrm{Mat}_{b.c.}(f) \cdot X.$$

Remarquez qu'en particulier, si L est une matrice ligne (1 ligne, p colonnes), et C une matrice colonne (p lignes, 1 colonne) on obtient la règle suivante :

$$L \times C = \begin{pmatrix} a_1 & \cdots & a_p \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} = a_1 b_1 + \cdots + a_p b_p = \sum_{k=1}^{p} a_k b_k.$$

En notant $\langle X, Y\rangle$ l'application bilinéaire sur $\mathbb{K}^p$, définie par

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix} \right\rangle = \sum_{k=1}^{p} x_i y_i,$$

et en notant ${}^tL$ la transposée de la ligne L (c'est-à-dire en réarrangeant les coefficients de L sous forme d'une colonne), on obtient :

$$L \times C = \left\langle \begin{pmatrix} a_1 \\ \vdots \\ a_p \end{pmatrix}, \begin{pmatrix} b_1 \\ \vdots \\ b_p \end{pmatrix} \right\rangle = \langle\, {}^tL, C\rangle\,.$$

Vous pouvez noter que si $\mathbb{K} = \mathbb{R}$, l'application bilinéaire $\langle \bullet, \bullet\rangle$ est le produit scalaire canonique de $\mathbb{R}^p$. En étuduant ligne par ligne l'expression du produit MX, on obtient alors la description :

Proposition 7.1.13 (Expression coefficient par coefficient du produit MX)

Soit $M = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathcal{M}_{p,1}(\mathbb{K})$. En notant $L_1, \dots, L_n$ les lignes de la matrice M, on a alors :

$$MX = \left(\begin{array}{c} L_1 \\ \hline \vdots \\ \hline L_n \end{array}\right) X = \begin{pmatrix} L_1 \cdot X \\ \vdots \\ L_n \cdot X \end{pmatrix} = \begin{pmatrix} \langle\, {}^tL_1 X\rangle \\ \vdots \\ \langle\, {}^tL_n, X\rangle \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{p} a_{1,k}x_k \\ \vdots \\ \sum_{k=1}^{p} a_{n,k}x_k \end{pmatrix}$$

On cherche maintenant à définir le produit général de deux matrices. Comme pour le cas de la multiplication par une colonne, on motive la définition par la concordance avec des notions liées aux applications linéaires. Notre but est que le produit matriciel corresponde, comme défini par Gauss, à la composition des applications linéaires. Pour que la composition soit possible, cela nécessite des restrictions sur les tailles des matrices : en effet, la composition $g \circ f$ n'est possible que si l'espace d'arrivée de f est égal à l'espace source de g. Soit n, p, q trois entiers naturels strictement positifs, et $f \in \mathcal{L}(\mathbb{K}^q, \mathbb{K}^p)$ et $g \in \mathcal{L}(K^p, \mathbb{K}^n)$. On cherche alors la description des colonnes de la matrice $\mathrm{Mat}_{b.c.}(g \circ f)$ en déterminant l'image par $g \circ f$ de la base canonique $(e_1, \dots, e_q)$ de $\mathbb{K}^q$. En notant

$$\mathrm{Mat}_{b.c.}(f) = (a_{i,j}) = \left(\, C_1 \;\middle|\; \cdots \;\middle|\; C_q \,\right),$$

et $(f_1, \dots, f_p)$ la base canonique de $\mathbb{K}^p$, on obtient pour tout $i \in [\![1, q]\!]$:

$$g \circ f(e_i) = g\left(\sum_{j=1}^{p} a_{j,i} f_j\right) = \sum_{j=1}^{p} a_{j,i} g(f_j).$$

La définition du produit d'une matrice par une colonne amène alors la description suivante :

$$g \circ f(e_i) = \mathrm{Mat}_{b.c.}(g) \cdot C_i. \tag{7.2}$$

Ainsi, la i-ième colonne de la matrice canoniquement associée à $g \circ f$ est le produit de la matrice de g par la i-ième colonne de la matrice de f. Cela motive la définition suivante.

Définition 7.1.14 (Produit matriciel)

Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,q}(\mathbb{K})$. Alors le produit AB est la matrice de $\mathcal{M}_{n,q}(\mathbb{K})$ dont la décomposition par colonnes est donnée par :

$$AB = \left(A \cdot B_1 \middle| \cdots \middle| A \cdot B_q \right),$$

où $B_1, \ldots, B_q$ sont les colonnes de la matrice B.

Remarque 7.1.15

Il est important de constater que lorsque B est une matrice colonne, la définition ci-dessus coïncide avec celle déjà donnée, ce qui n'était pas évident au départ.

Avertissement 7.1.16

On ne peut pas multiplier n'importe quelles matrices entre elles. Pour pouvoir effectuer le produit AB, il est nécessaire que le nombre de colonnes de A soit égal au nombre de lignes de B. Cette condition équivaut à la compatibilité des espaces vectoriels pour pouvoir effectuer la composition des applications linéaires canoniquement associées.

L'identité (7.2), ayant motivé notre définition, amène immédiatement :

Théorème 7.1.17 (Matrice canoniquement associée à une composition)

Soit $f \in \mathcal{L}(\mathbb{K}^q, \mathbb{K}^p)$ et $g \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$. Alors l'application linéaire $g \circ f$ de $\mathcal{L}(\mathbb{K}^q, \mathbb{K}^n)$ est canoniquement associée à la matrice :

$$\mathrm{Mat}_{b.c.}(g \circ f) = \mathrm{Mat}_{b.c.}(g) \cdot \mathrm{Mat}_{b.c.}(f).$$

Ainsi, le produit des matrices correspond à la composition des applications linéaires.
En utilisant la description de la proposition 7.1.13, on obtient la description coefficient par coefficient du produit matriciel :

Proposition 7.1.18

Soit $A = (a_{i,j}) \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B = (b_{j,k}) \in \mathcal{M}_{p,q}(\mathbb{K})$. Soit $AB = (c_{i,k}) \in \mathcal{M}_{n,q}(\mathbb{K})$ la matrice produit. Alors :

$$\forall i \in [\![1, n]\!],\ \forall k \in [\![1, q]\!],\quad c_{i,k} = \sum_{j=1}^{p} a_{i,j} b_{j,k} = L_i \cdot C_k = \left\langle {}^t L_i, C_k \right\rangle,$$

où L_i est la i-ème ligne de A, et C_k la k-ième colonne de B_k. Ainsi :

$$\left(\begin{array}{c} L_1 \\ \hline \vdots \\ \hline L_n \end{array}\right) \times \left(C_1 \middle| \cdots \middle| C_q \right) = \begin{pmatrix} L_1 \cdot C_1 & \cdots & L_1 \cdot C_q \\ \vdots & L_i \cdot C_k & \vdots \\ L_n \cdot C_1 & \cdots & L_n \cdot C_q \end{pmatrix}$$

Remarque 7.1.19

La description du produit par colonnes est parfois plus efficace que la description coefficient par coefficient, notamment lorsque la matrice de droite possède un grand nombre de 0. On obtient souvent une vision plus immédiate de la matrice, même si formellement le nombre de calcul est exactement le

même. Par ailleurs, pour des arguments sur des matrices de taille générique n, la rédaction est souvent simplifiée par la description par les colonnes.

Exemple 7.1.20

1. Calculer $\begin{pmatrix} 2 & 6 & 12 \\ 2 & -7 & 2 \\ 8 & 5 & 11 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 2 \\ -1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.
2. Quel est l'effet de la multiplication à droite par $C_n = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & \ddots & 1 \\ 1 & 0 & \cdots & \cdots & 0 \end{pmatrix}$?
3. Même question avec la matrice « de Jordan » $J_n = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}$
4. Calculer pour tout $k \in \mathbb{N}$, C_n^k et J_n^k. Que constatez-vous ?

Vue la parfaite symétrie de l'expression obtenue, on obtient une version duale des définitions 7.1.11 et 7.1.14 en inversant le rôle des lignes et des colonnes (cela provient de la description complètement symétrique donnée dans la proposition précédente) :

Proposition 7.1.21 (Expression du produit à l'aide des lignes)

1. *Soit* $A = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \in \mathcal{M}_{1,n}(\mathbb{K})$ *une matrice ligne, et* $B = \left(\begin{array}{c} L_1 \\ \hline \vdots \\ \hline L_n \end{array}\right) \in \mathcal{M}_{n,p}(\mathbb{K})$. *Alors*

$$AB = x_1 L_1 + \cdots + x_n L_n = \sum_{i=1}^{n} x_i L_i.$$

Il s'agit d'une matrice ligne.

2. *Soit* $A = \left(\begin{array}{c} A_1 \\ \hline \vdots \\ \hline A_n \end{array}\right) \in \mathcal{M}_{m,n}(\mathbb{K})$, *dont les lignes sont* $A_1, \ldots, A_m$, *et* $B \in \mathcal{M}_{n,p}(\mathbb{K})$. *Alors*

$$AB = \left(\begin{array}{c} A_1 B \\ \hline \vdots \\ \hline A_n B \end{array}\right).$$

Les règles similaires sur la composition des applications linéaires amènent sans effort :

Proposition 7.1.22 (Propriétés du produit)

Sous reserve de compatibilité des formats des matrices, le produit matriciel est associatif et bilinéaire (donc en particulier distributif)

Avertissement 7.1.23

Le produit matriciel n'est pas commutatif, même lorsque les tailles sont compatibles pour effectuer les opérations dans les deux sens (par exemple pour des matrices carrées).

Exemple 7.1.24

Comparer $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ et $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

Nous terminons cette section par l'étude du produit des éléments de la base canonique.

Proposition 7.1.25 (Produit des éléments de la base canonique)

Soit $(E_{i,j})$ la base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$, $(E'_{j,k})$ la base canonique de $\mathcal{M}_{p,q}(\mathbb{K})$ et $(E''_{i,k})$ la base canonique de $\mathcal{M}_{n,q}(\mathbb{K})$. Soit $(i,j) \in [\![1,n]\!] \times [\![1,p]\!]$, et $(k\ell) \in [\![1,p]\!] \times [\![1,q]\!]$. Alors :

$$E_{i,j} \times E'_{k,\ell} = \delta_{j,k} E''_{i,\ell} = \begin{cases} E''_{i,\ell} & \text{si } j = k \\ 0 & \text{sinon.} \end{cases}$$

I.4 Matrices carrées

La structure d'anneau de $\mathcal{L}(\mathbb{K}^n)$ se transfert à l'ensemble des matrices carrées. Pour le formaliser, il nous faut dans un premier temps définir l'élément neutre, image de l'application identité de $\mathbb{K}^n$.

Proposition/Définition 7.1.26 (Matrice identité)

Soit I_n la matrice de $\mathcal{M}_n(\mathbb{K})$, canoniquement associée à l'endormorphisme $\mathrm{Id}_{\mathbb{K}^n}$, identité de $\mathbb{K}^n$. On a alors :

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

Il s'agit donc de la matrice constituée de 1 sur la diagonale et de 0 partout ailleurs.

Théorème 7.1.27 (Structure de $\mathcal{M}_n(\mathbb{K})$)

L'ensemble $\mathcal{M}_n(\mathbb{K})$, muni de la somme, du produit et de la multiplication par les scalaires, est une $\mathbb{K}$-algèbre. Cette $\mathbb{K}$-algèbre est non commutative dès lors que $n \geqslant 2$.

Évidemment, on a de façon plus générale $I_n \times M = M$ et $N \times I_n = N$, même si M et N ne sont pas carrées (mais de format compatible).

En particulier le théorème précédent affirme que $\mathcal{M}_n(\mathbb{K})$ est un anneau, et toutes les règles de calcul que nous avons développées de façon générale dans les anneaux sont donc valables dans $\mathcal{M}_n(\mathbb{K})$, en particulier :

Théorème 7.1.28 (Factorisation de $A^n - B^n$)

Soit A et B deux éléments de $\mathcal{M}_n(\mathbb{K})$ tels que $AB = BA$. Alors pour tout $n \in \mathbb{N}^*$

$$A^n - B^n = (A - B)\sum_{k=0}^{n-1} A^{n-1-k}B^k.$$

Corollaire 7.1.29 (Factorisation de $I_n - A^n$)

En particulier, pour toute matrice carrée $A \in \mathcal{M}_n(\mathbb{K})$,

$$I_n - A^n = (I_n - A)\sum_{k=0}^{n-1} A^k.$$

Théorème 7.1.30 (Formule du binôme)

Soit A et B deux éléments de $\mathcal{M}_n(\mathbb{K})$ tels que $AB = BA$. Alors, pour tout $n \in \mathbb{N}$,

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

Exemple 7.1.31

1. Déterminer A^n, pour tout $n \in \mathbb{N}$, où $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$
2. Plus généralement, si J_n est définie comme dans l'exemple 7.1.20, calculer $(aI_n + J_n)^k$, pour tout $k \in \mathbb{N}$.

On en déduit une méthode assez efficace, mais pas toujours réalisable, de calcul des puissances.

Remarque 7.1.32

L'anneau $\mathcal{M}_n(\mathbb{K})$ est-il intègre ?

Les notions relatives aux polynômes d'endomorphisme se tranduisent en terme de matrices. En particulier, toute matrice carrée d'ordre n admet un polynôme annulateur de degré inférieur ou égal à n^2 (et même inférieur ou égal à n). De là découle l'existence d'un polynôme minimal.

Exemple 7.1.33

1. Déterminer un polynôme annulateur de $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.
2. Déterminer un polynôme annulateur de C_n
3. Déterminer un polynôme annulateur de J_n. Un polynôme minimal.

Le polynôme annulateur peut être efficace pour la recherche des puissances successives d'une matrice A.

Méthode 7.1.34 (Calcul de A^n à l'aide d'un polynôme annulateur)

- Déterminer un polynôme annulateur P de petit degré de A.
- Chercher le reste R_n de la division euclidienne de X^n par P.
- Évaluer l'égalité de division euclidienne en A, il reste $A^n = R_n(A)$. Ainsi A^n s'exprime comme combinaison linéaire des premières puissances de A.

Exemple 7.1.35

Calculer $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^n$.

I.5 Matrices carrées de type particulier

La matrice I_n a une particularité notable : elle est nulle, sauf sur sa diagonale. Les matrices vérifiant cette propriété jouent un rôle central, notamment dans la théorie de la diagonalisation. En effet, les endomorphismes associés laissent stables les axes définis par les vecteurs de la base, et sont donc faciles à étudier. En particulier les produits de matrices diagonales sont simples à exprimer (et donc aussi les puissances). C'est une des motivations de la théorie de la diagonalisation. Nous présentons d'autres formes de matrices (matrices triangulaires)

Définition 7.1.36 (Matrice diagonale)

Soit D une matrice de $\mathcal{M}_n(\mathbb{K})$. On dit que D est une matrice diagonale si tous ses coefficients sont nuls, à l'exception éventuellement de ses coefficients diagonaux.

Ainsi, une matrice diagonale est de la forme :

$$D = \begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & d_n \end{pmatrix}$$

Définition 7.1.37 (Matrice triangulaire)

Soit T une matrice de $\mathcal{M}_n(\mathbb{K})$. On dit que T est une matrice triangulaire supérieure (resp. inférieure) si tous ses coefficients situés strictement en-dessous (resp. au-dessus) de sa diagonale sont nuls.

Ainsi, une matrice triangulaire supérieure (resp. inférieure) est de la forme :

$$T = \begin{pmatrix} \bullet & \cdots & \cdots & \bullet \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \bullet \end{pmatrix} \quad \left(\text{resp. } T = \begin{pmatrix} \bullet & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \bullet & \cdots & \cdots & \bullet \end{pmatrix}\right)$$

où les $\bullet$ désignent des coefficients quelconques. On définit également les matrices strictement triangulaires supérieures ou inférieures, nulles également sur la diagonale.

Notation 7.1.38

Nous noterons dans ce cours $\mathcal{D}_n(\mathbb{K})$ l'espace des matrices diagonales d'ordre n, $\mathcal{T}_n^+(\mathbb{K})$ l'espace des matrices triangulaires supérieures, $\mathcal{T}_n^-(\mathbb{K})$ l'espace des matrices triangulaires inférieures, $\overline{\mathcal{T}}_n^+(\mathbb{K})$ l'espace des matrices strictement triangulaires supérieures, et $\overline{\mathcal{T}}_n^-(\mathbb{K})$ l'espace des matrices strictement triangulaires inférieures.

Proposition 7.1.39

Nous avons de façon évidente :

$$\mathcal{T}_n^+(\mathbb{K}) = \overline{\mathcal{T}}_n^+(\mathbb{K}) \oplus \mathcal{D}_n(\mathbb{K}) \qquad et \qquad \mathcal{T}_n^-(\mathbb{K}) = \overline{\mathcal{T}}_n^-(\mathbb{K}) \oplus \mathcal{D}_n(\mathbb{K}),$$

ainsi que :

$$\mathcal{M}_n(\mathbb{K}) = \overline{\mathcal{T}}_n^+(\mathbb{K}) \oplus \mathcal{D}_n(\mathbb{K}) \oplus \overline{\mathcal{T}}_n^-(\mathbb{K}) = \overline{\mathcal{T}}_n^+(\mathbb{K}) \oplus \mathcal{T}_n^-(\mathbb{K}) = \mathcal{T}_n^+(\mathbb{K}) \oplus \overline{\mathcal{T}}_n^-(\mathbb{K})$$

Nous avons également des propriétés de stabilité :

Proposition 7.1.40 (Produit de matrices diagonales)

Le produit de deux matrices diagonales de même ordre est encore une matrice diagonale. De façon plus explicite :

$$\begin{pmatrix} c_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & c_n \end{pmatrix} \begin{pmatrix} d_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_n \end{pmatrix} = \begin{pmatrix} c_1 d_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & c_n d_n \end{pmatrix}.$$

En particulier, pour tout $k \in \mathbb{N}$, on a :

$$\begin{pmatrix} d_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_n \end{pmatrix}^k = \begin{pmatrix} d_1^k & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_n^k \end{pmatrix}.$$

Proposition 7.1.41 (Produit de deux matrices triangulaires)

(i) Le produit de deux matrices triangulaires supérieures est une matrice triangulaire supérieure dont les coefficients diagonaux sont les produits deux à deux des coefficients diagonaux des deux matrices.

(ii) En particulier, si l'une des deux matrices est strictement triangulaire supérieure, le produit l'est également.

(iii) De même pour les produits de matrices triangulaires inférieures.

Des règles de stabilité précédentes, on déduit :

Théorème 7.1.42 (Structure des sous-ensembles de matrices pde type particulier)

Les sous-ensembles $\mathcal{D}_n(\mathbb{K})$, $\mathcal{T}_n^+(\mathbb{K})$, $\mathcal{T}_n^-(\mathbb{K})$, $\overline{\mathcal{T}}_n^+(\mathbb{K})$, $\overline{\mathcal{T}}_n^-(\mathbb{K})$ sont des sous-algèbres de $\mathcal{M}_n(\mathbb{K})$.

I.6 Noyau, image, rang d'une matrice

La correspondance entre application linéaire et matrice permet d'étendre aux matrices un certain nombre de notions définies pour les applications linéaires.

Définition 7.1.43 (Image et noyau d'une matrice)

Par définition, l'image et le noyau d'une matrice M sont l'image et le noyau de l'application linéaire canoniquement associée. Plus précisément, soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$ et $f : X \mapsto MX$ l'application de $\mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ canoniquement associé. Alors

- $\mathrm{Im}(M) = \mathrm{Im}(f) = \{Y \in \mathbb{K}^n \simeq \mathcal{M}_{n,1}(\mathbb{K}) \mid \exists X \in \mathbb{K}^p,\ MX = y\}$
- $\mathrm{Ker}(M) = \mathrm{Ker}(f) = \{X \in \mathbb{K}^p \simeq \mathcal{M}_{p,1}(\mathbb{K}) \mid MX = 0\}$.

En particulier, puisque les colonnes de M sont les images par f des éléments de la base canonique, nous obtenons :

Proposition 7.1.44 (Description de l'image d'une matrice)

Soit M une matrice de colonnes $C_1, \ldots, C_p$. Alors

$$\mathrm{Im}(M) = \mathrm{Vect}(C_1, \ldots, C_p).$$

Définition 7.1.45 (Rang d'une matrice)

Soit M une matrice. Son rang est égal au rang de l'application linéaire canoniquement associée, c'est-à-dire :

$$\mathrm{rg}(M) = \mathrm{rg}(f) = \dim(\mathrm{Im}(f)) = \dim(\mathrm{Im}(M)) = \dim(\mathrm{Vect}(C_1, \ldots, C_n)) = \mathrm{rg}(C_1, \ldots, C_n),$$

où les C_i sont les colonnes de M.

Ainsi, le rang d'une matrice est le rang de la famille de ses colonnes. Un exemple important, auquel on se ramène souvent par utilisation du pivot de Gauss est le suivant :

Proposition 7.1.46 (Rang d'une matrice échelonnée)

Soit M une matrice échelonnée (voir chapitre sur les systèmes d'équations linéaires). Alors $\mathrm{rg}(M)$ *est le nombre de lignes non nulles de M.*

I.7 Inverse d'une matrice

Définition 7.1.47 (Matrice inversible)

Une matrice M est dite inversible s'il existe $n \in \mathbb{N}^*$ tel que $M \in \mathcal{M}_n(\mathbb{K})$, et tel que M soit inversible dans l'anneau $\mathcal{M}_n(\mathbb{K})$.

Avertissement 7.1.48

Par définition, une matrice inversible est toujours une matrice carrée.

Proposition 7.1.49 (caractérisation par l'endomorphisme canoniquement associée)

Une matrice $M \in \mathcal{M}_n(\mathbb{K})$ est inversible si et seulement si l'endomorphisme canoniquement associé $f \in \mathcal{C}(\mathbb{K}^n)$ est un automorphisme.

Cette caractérisation explique qu'on se limite aux matrices carrées : en effet un isomorphisme préserve les dimensions, donc une matrice non carrée ne peut pas être canoniquement associée à un isomorphisme. Nous déduisons alors de la caractérisation des isomophismes en dimension finie que :

Proposition 7.1.50

Soit $A \in \mathcal{M}_n(\mathbb{K})$. Pour que A soit inversible, il suffit qu'il existe une matrice $B \in \mathcal{M}_n(\mathbb{K})$ telle que $AB = I_n$ OU $BA = I_n$. Dans ce cas $B = A^{-1}$.

Définition 7.1.51 (Groupe linéaire)

L'ensemble des matrices inversibles de $\mathcal{M}_n(\mathbb{K})$ est noté $\mathrm{GL}_n(\mathbb{K})$, et est appelé *$n$-ième groupe linéaire.*

Proposition 7.1.52 (Structure de $\mathrm{GL}_n(\mathbb{K})$)

$(\mathrm{GL}_n(\mathbb{K}), \times)$ est un groupe (c'est le groupe des inversibles de l'anneau $\mathcal{M}_n(\mathbb{K})$).

Conformément aux propriétés générales concernant l'inverse de produits dans un anneau, on a :

Proposition 7.1.53 (Inverse d'un produit)

Soit A et B deux matrices inversibles de $\mathcal{M}_n(\mathbb{K})$. Alors AB est inversible, et son inverse est $B^{-1}A^{-1}$.

Voici quelques exemples importants de matrices inversibles.

Proposition 7.1.54 (Inverse de matrices triangulaires)

Soit $T \in \mathcal{T}_n^+(\mathbb{K})$ une matrice triangulaire. Alors T est inversible si et seulement si tous ses coefficients diagonaux sont non nuls (donc inversibles), et dans ce cas, T^{-1} est une matrice triangulaire dont les coefficients diagonaux sont les inverses des coefficients diagonaux de T.

En particulier, une matrice diagonale est inversible si et seulement si tous ses coefficients diagonaux sont non nuls, et dans ce cas :

$$D^{-1} = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_n \end{pmatrix}^{-1} = \begin{pmatrix} d_1^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & d_n^{-1} \end{pmatrix}.$$

Les exemples suivants sont importants car ils valident la méthode du pivot de Gauss.

Définition 7.1.55 (Matrices de codage des opérations élémentaires)

Soit $n \in \mathbb{N}^*$. On définit les trois familles suivantes de matrices :

(i) Codage des échanges de lignes (matrice de transposition) :

$$\forall (i,j) \in [\![1,n]\!]^2, \quad i \neq j, \quad E(i,j) = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & & & & \vdots \\ \vdots & \ddots & 0 & \ddots & & 1 & & \vdots \\ \vdots & & \ddots & 1 & \ddots & & & \vdots \\ \vdots & & & \ddots & 1 & \ddots & & \vdots \\ \vdots & & 1 & & \ddots & 0 & \ddots & \vdots \\ \vdots & & & & & \ddots & 1 & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ i \\ \\ \\ \\ j \end{matrix}$$

$$\begin{matrix} & & & i & & & j & & \end{matrix}$$

(ii) Codage de la multiplication d'une ligne par un scalaire :

$$E_i(\lambda) = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & \ddots & & & \vdots \\ \vdots & & \ddots & \lambda & \ddots & & \vdots \\ \vdots & & & \ddots & 1 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ i \\ \\ \\ \\ \end{matrix}$$
$$\begin{matrix} \\ \\ \\ \\ \\ \\ \\ \end{matrix} \quad \begin{matrix} \\ \\ \\ i \\ \\ \\ \\ \end{matrix}$$

(iii) Codage d'une combinaison linéaire (matrice de transvection)

$$E_{i,j}(\lambda) = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \ddots & \ddots & \lambda & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \end{matrix}$$
$$\begin{matrix} \\ \\ \\ \\ \\ \\ \end{matrix} \quad \begin{matrix} \\ \\ \\ \\ \\ i \\ \end{matrix}$$
$$\begin{matrix} \\ \\ \\ \\ \\ \\ \end{matrix} \quad \begin{matrix} \\ \\ \\ \\ \\ j \\ \end{matrix}$$

Proposition 7.1.56 (Interprétation matricielle des opérations élémentaires sur les lignes)
Soit $M \in \mathcal{M}_{m,n}(\mathbb{K})$. Soit $i \neq j$ dans $[\![1, m]\!]$, et $\lambda \in \mathbb{K}$.

(i) La matrice N obtenue de M par l'opération $L_i \leftrightarrow L_j$, est $N = E(i,j) \cdot M$;

(ii) La matrice N obtenue de M par l'opération $L_i \leftarrow L_i + \lambda L_j$ est $N = E_{i,j}(\lambda) \cdot M$;

(iii) La matrice N obtenue de M par l'opération $L_i \leftarrow \lambda L_i$, $\lambda \neq 0$, est $N = E_i(\lambda) \cdot M$.

La validité du pivot de Gauss provient alors du résultat suivant, exprimant que toute opération valide du pivot est réversible :

Proposition 7.1.57 (Inversibilité des matrices de codage des opérations élémentaires)
Pour $i \neq j$, les matrices $E(i,j)$, $E_{i,j}(\lambda)$ sont inversibles, ainsi que $E_i(\lambda)$ lorsque $\lambda \neq 0$.

Théorème 7.1.58 (Validité de la méthode du pivot)
Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$, $B \in \mathcal{M}_{n,1}(\mathbb{K})$, et $P \in \mathrm{GL}_n(\mathbb{K})$. Alors

$$\forall X \in \mathcal{M}_{p,1}(\mathbb{K}), \quad MX = B \iff PMX = PB.$$

Ce théorème exprime la stabilité des solutions d'une équation matricielle par multiplication de l'équation par une matrice inversible.
Au passage, si M lui-même est inversible, ce même théorème affirme qu'on a l'équivalence :

$$MX = Y \iff X = M^{-1}Y.$$

Réciproquement, l'existence et l'unicité d'une solution pour tout second membre Y implique la bijectivité de l'application canoniquement associée à M, donc l'inversibilité de M.
Ainsi, une façon de calculer l'inverse est la résolution d'un système, de second membre Y indéterminé :

Méthode 7.1.59 (Calcul pratique de l'inverse (première méthode))

- Résoudre le système $AX = Y$, où Y est pris en paramètre.
- A est inversible si et seulement si le système admet une et une seule solution, quel que soit le paramètre Y.
- Dans ce cas, écrire le résultat sous forme matricielle $X = BY$.
- B est alors l'inverse de A.

Exemple 7.1.60

Calcul de l'inverse de $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$.

Définition 7.1.61 (Système de Cramer)

Un système de Cramer est un système d'équations linéaires admettant une unique solution.

Ainsi, un système est de Cramer si et seulement si la matrice de ses coefficients est inversible.
Dans le cas de matrices 2×2, on a une méthode plus rapide :

Théorème 7.1.62 (Inverse des matrices 2×2 par la comatrice)

Soit $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{M}_2(\mathbb{K})$. *Alors* M *est inversible si et seulement si* $ad - bc \neq 0$, *et*

$$M^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Définition 7.1.63 (Déterminant d'une matrice 2×2)

La quantité $ad - bc$ est appelée *déterminant* de M, et est noté $\det M$ ou $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$.

Remarque 7.1.64

Il existe une notion de déterminant pour des matrices carrées de taille quelconque n. Nous définirons cette notion générale dans le chapitre suivant. La non nullité du déterminant caractérise alors l'inversibilité de la matrice, et comme dans le cas $n = 2$, il existe une formule générale de l'inverse d'une matrice basée sur la notion de comatrice. Dans des situations concrètes, cette formule est cependant assez peu efficace, sauf pour $n = 2$, et éventuellement $n = 3$ (et encore...)

Une deuxième méthode générale de calcul de l'inverse d'une matrice repose sur l'observation suivante : la méthode du pivot de Gauss consiste en des multiplications à gauche par des matrices codant les opérations élémentaires. Ainsi, si à l'aide d'opérations élémentaires sur les lignes, on parvient à transformer la matrice initiale A en la matrice identité, on aura l'existence d'une matrice inversible P (obtenue en multipliant les matrices de codage) telle que $PA = I_n$. Ainsi, cette matrice P est l'inverse de A. Or, $P = PI_n$, et est donc obtenu en appliquant à la matrice I_n les mêmes opérations sur ses lignes que celles qui ont permis de transformer A en I_n. On en déduit la méthode suivante :

Méthode 7.1.65 (Calcul de l'inverse d'une matrice (seconde méthode))

- Juxtaposer la matrice A et la matrice I_n (séparées d'une barre verticale)
- Effectuer un pivot sur A, en faisant les mêmes opérations sur la matrice I_n, pour obtenir une matrice échelonnée à la place de A
- La matrice A est inversible si et seulement si la matrice échelonnée obtenue est inversible (c'est-à-dire s'il s'agit d'une matrice triangulaire supérieure à ceofficients diagonaux non nuls)
- Dans ce cas, faire un pivot remontant, pour annuler les coefficients au dessus de chaque pivot, et toujours en effectuant les mêmes opérations sur la matrice de droite.
- En normalisant les coefficients diagonaux, on obtient à gauche la matrice identité, et à droite la matrice A^{-1}.

I.8 Rang

Nous avons déjà défini le rang. Une propriété importante, permettant le calcul du rang par la méthode du pivot, est la conservation du rang par multiplication par une matrice inversible. On a même plus que cela :

Théorème 7.1.66 (Conservation de l'image et du noyau)

Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$ une matrice quelconque, et $P \in \mathrm{GL}_n(\mathbb{K})$, $Q \in \mathrm{GL}_p(\mathbb{K})$. Alors :

(i) $\mathrm{Ker}(PM) = \mathrm{Ker}(M)$

(ii) $\mathrm{Im}(MQ) = \mathrm{Im}(M)$.

Ainsi, la multiplication à gauche par une matrice inversible conserve le noyau et la multiplication à droite conserve l'image.

Corollaire 7.1.67 (Conservation du rang)

Soit $M \in \mathcal{M}_{n,p}(\mathbb{K})$ une matrice quelconque, et $P \in \mathrm{GL}_n(\mathbb{K})$, $Q \in \mathrm{GL}_p(\mathbb{K})$. Alors :

$$\mathrm{rg}(PMQ) = \mathrm{rg}(M)$$

Corollaire 7.1.68 (Conservation de l'image et du noyau par opérations élémentaires)

(i) Les opérations élémentaires sur les lignes conservent le noyau

(ii) Les opérations élémentaires sur les colonnes conservent l'image

(iii) Les opérations élémentaires sur les lignes et les colonnes conservent le rang

Méthode 7.1.69 (calcul du rang d'une matrice)

- Effectuer un pivot pour se ramener à une matrice échelonnée
- Le rang est égal au rang de la matrice échelonnée, donc au nombre de ses lignes non nulles
- Si les opérations sur les colonnes semblent plus simples, c'est possible aussi
- On peut même combiner les deux.

Méthode 7.1.70 (Trouver une base de l'image)

- Cette fois, on veut conservation de l'image, on travaille donc sur les colonnes.
- On effectue un pivot sur les colonnes jusqu'à obtenir une matrice échelonnée sur les colonnes, dont les colonnes engendrent l'image.

- Les colonnes non nulles de cette matrice forment une base de l'image.

Méthode 7.1.71 (Trouver un supplémentaire)

- On suppose donné un sous-espace E de $\mathbb{K}^p$ dont on connaît une famille génératrice de n vecteurs.
- On écrit la matrice de cette famille (chaque vecteur est écrit en colonne) : il s'agit de la matrice de l'application linéaire qui envoie les vecteurs de la base canonique de $\mathbb{K}^n$ sur les vecteurs de la famille génératrice. L'image de cette application est E.
- On effectue un pivot en travaillant sur les colonnes, pour conserver l'image.
- À la matrice échelonnée sur les colonnes obtenue au bout, on ajoute les colonnes correspondant aux vecteurs de la base canonique de sorte à obtenir une matrice triangulaire supérieure.
- Les vecteurs ainsi ajoutés forment une base d'un supplémentaire de E dans $\mathbb{K}^p$.

I.9 Transposition

Définition 7.1.72 (Transposée d'une matrice)

Soit $A = (a_{i,j})_{(i,j)\in[\![1,n]\!]\times[\![1,p]\!]} \in \mathcal{M}_{n,p}(A)$. Alors la matrice transposée de A, notée ${}^tA$, est la matrice de $\mathcal{M}_{p,n}\mathbb{K}$ définie par :

$$ {}^tA = (a_{j,i})_{(i,j)\in[\![1,p]\!]\times[\![1,n]\!]}.$$

Ainsi, si $A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix}$, alors ${}^tA = \begin{pmatrix} a_{1,1} & \cdots & a_{n,1} \\ \vdots & & \vdots \\ a_{1,p} & \cdots & a_{n,p} \end{pmatrix}$

On trouve aussi assez souvent la notation A^T.

Exemple 7.1.73

Transposée de $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$?

Proposition 7.1.74 (Transposée d'un produit)

Soit $A \in \mathcal{M}_{m,n}(\mathbb{K})$ et $B \in \mathcal{M}_{n,p}(\mathbb{K})$. Alors :

$$ {}^t(MN) = {}^tN\,{}^tM.$$

Proposition 7.1.75 (Transposée d'une inverse)

Soit A une matrice inversible. Alors ${}^tA$ l'est aussi, et

$$ ({}^tA)^{-1} = {}^t(A^{-1}).$$

Enfin, en effectuant les mêmes opérations sur les lignes de l'une et sur les colonnes de l'autre, et en comparant le rang des matrices échelonnées obtenues, on obtient :

Proposition 7.1.76 (Rang d'une transposée)

On a $\mathrm{rg}({}^tA) = \mathrm{rg}(A)$.

Pour terminer cette section, nous définissons deux types importants de matrices :

Définition 7.1.77 (Matrices symétriques, antisymétriques)

Soit $A \in \mathcal{M}_n(\mathbb{K})$.

1. On dit que A est *symétrique* si $A = {}^tA$.
2. On dit que A est *antisymétrique* si $A = -{}^tA$.

II Écriture d'une AL dans une base

Le but de cette partie est de généraliser à des espaces vectoriels quelconques (de dimension finie) la correspondance entre application linéaire et matrice. Cette correspondance dépendra du choix de bases de l'espace de départ E et de l'espace d'arrivée F. On définit dans un premier temps cette correspondance et ses propriétés, qui permettent une étude matricielle de toute application linéaire en dimension finie. Nous voyons les règles permettant de passer d'une base à une autre.

II.1 Définitions et notations

Soit E et F deux espaces vectoriels de dimensions finies respectives p et n.

Définition 7.2.1 (Coordonnées d'un vecteur)

Soit $\mathcal{C} = (c_1, \dots, c_n)$ une base de F, et $X \in F$. Alors il existe d'uniques scalaires $x_1, \dots, x_n$ tels que $X = x_1c_1 + \cdots + x_nc_n$. On définit la matrice des coordonnées de X dans la base $\mathcal{C}$ par :

$$[X]_{\mathcal{C}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Définition 7.2.2 (Matrice associée à une famille)

Sous les mêmes hypothèses, soit $(X_1, \dots, X_k)$ une famille de vecteurs de F. Alors la matrice de cette famille dans la base $\mathcal{C}$ est :

$$[X_1, \dots, X_k]_{\mathcal{C}} = \left([X_1]_{\mathcal{C}} \middle| \cdots \middle| [X_k]_{\mathcal{C}} \right).$$

Ainsi, il s'agit de la matrice dont la i-ème colonne comporte les coordonnées dans la base $\mathcal{C}$ du vecteur X_i.

Définition 7.2.3 (Matrice d'une AL relativement à des bases)

Soit $f \in \mathcal{L}(E, F)$, et soit $\mathcal{B} = (b_1, \dots, b_p)$ une base de de E et $\mathcal{C} = (c_1, \dots, c_n)$ une base de F. Alors la matrice de f relativement aux bases $\mathcal{B}$ et $\mathcal{C}$ est la matrice $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$ de la famille $(f(b_1), \dots, f(b_p))$ dans la base $\mathcal{C}$:

$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f) = [f(b_1), \dots, f(b_p)]_{\mathcal{C}} = \left([f(b_1)]_{\mathcal{C}} \middle| \cdots \middle| [f(b_p)]_{\mathcal{C}} \right).$$

Ainsi, il s'agit de la matrice de type (n, p) dont la i-ème colonne donne les coordonnées du vecteur $f(b_i)$ dans la base $\mathcal{C}$.

Remarque 7.2.4

Si on prend $E = \mathbb{R}^p$, $F = \mathbb{R}^n$, et pour $\mathcal{B}$ et $\mathcal{C}$ les bases canoniques de ces espaces, on retrouve la matrice $\mathrm{Mat}_{b.c.}(f)$.

Exemple 7.2.5

1. Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par $f(x) = 2x + 3y$. Déterminer $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$ lorsque :
   - $\mathcal{B} = b.c.$, $\mathcal{C} = (1)$
   - $\mathcal{B} = \left(\begin{pmatrix}1\\1\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix}\right)$, $\mathcal{C} = (2)$
2. Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ définie par $f(x, y) = (2x - y, x + 2y)$.
   - $\mathrm{Mat}_{b.c.}(f)$?
   - $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$ lorsque $\mathcal{B} = \mathcal{C} = \left(\begin{pmatrix}1\\-1\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix}\right)$
3. Matrice de l'opérateur de dérivation $D \in \mathcal{L}(\mathbb{R}_n[X])$ relativement à la base canonique (au départ et à l'arrivée)
4. Trouver une base $\mathcal{B}$ de $\mathbb{R}_n[X]$ telle que
$$\mathrm{Mat}_{\mathcal{B},\mathcal{B}}(D) = J_{n+1},$$
où J_{n+1} est la matrice de Jordan d'ordre $n + 1$.
5. Plus généralement, soit E un espace de dimension n, et u un endomorphisme nilpotent, d'indice de nilpotence égal à n. Montrer qu'il existe une base relativement à laquelle la matrice de u est J_n.

Proposition 7.2.6 (Évaluation matricielle d'une AL)

Sous les hypothèses de la définition :
$$\forall X \in E, \quad [f(X)]_{\mathcal{C}} = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f) \cdot [X]_{\mathcal{B}}.$$

Cette formule est même une caractérisation de la matrice de f relativement aux base $\mathcal{B}$ et $\mathcal{C}$:

Proposition 7.2.7 (Caractérisation de $\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$)

Soit f un élément de $\mathcal{L}(E, F)$, $\mathcal{B}$ une base de E et $\mathcal{C}$ une base de F. Alors $M = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f)$ si et seulement si :
$$\forall x \in \mathrm{E}, \quad M[x]_{\mathcal{B}} = [f(x)]_{\mathcal{C}}.$$

La formule suivante donne la compatibilité de l'écriture matricielle avec la composition, ce qui était notre motivation initiale pour la définition du produit matriciel.

Proposition 7.2.8 (Matrice associée à une composition)

Soit E, F et G trois espaces vectoriels de dimension finie, munis respectivement des bases $\mathcal{B}$, $\mathcal{C}$ et $\mathcal{D}$. Soit $f \in \mathcal{L}(E, F)$ et $g \in \mathcal{L}(F, G)$. Alors :
$$\mathrm{Mat}_{\mathcal{B},\mathcal{D}}(g \circ f) = \mathrm{Mat}_{\mathcal{C},\mathcal{D}}(g) \cdot \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f).$$

Cette formule est la clé de l'interprétation matricielle des applications linéaires. C'est en particulier elle qui fournit les formules de changement de base, en composant par l'identité, dont la matrice sera prise relativement à des bases différentes au départ et à l'arrivée.

II.2 Changements de base, matrices équivalentes

Définition 7.2.9

Soit E un espace vectoriel, et $\mathcal{B}_1$ et $\mathcal{B}_2$ deux bases de E. Alors la matrice de passage de la base $\mathcal{B}_1$ à la base $\mathcal{B}_2$ est la matrice :

$$P_{\mathcal{B}_1}^{\mathcal{B}_2} = \mathrm{Mat}_{\mathcal{B}_2,\mathcal{B}_1}(\mathrm{Id}_E) = [\mathcal{B}_2]_{\mathcal{B}_1}.$$

Ainsi, la i-ème colonne de cette matrice est constituée des coordonnées du i-ème vecteur de la base $\mathcal{B}_2$ dans la base $\mathcal{B}_1$. Il s'agit donc de la matrice de la famille des vecteurs de la seconde base $\mathcal{B}_2$ dans la première base $\mathcal{B}_1$.

La formule de composition amène facilement :

Proposition 7.2.10 (Inversibilité des matrices de passage)

Toute matrice de passage $P_{\mathcal{B}_1}^{\mathcal{B}_2}$ est inversible. Son inverse est la matrice de passage $P_{\mathcal{B}_2}^{\mathcal{B}_1}$

Par ailleurs, toute matrice inversible peut être interprétée comme une matrice de passage. Plus précisément :

Proposition 7.2.11

Soit $P \in \mathrm{GL}_n(\mathbb{K})$ une matrice inversible et $\mathcal{B}$ une base d'un espace vectoriel de dimension E. Alors il existe une base $\mathcal{C}$, dont les coordonnées des vecteurs dans la base $\mathcal{B}$ sont les colonnes de P, telle que P soit la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$.

Puisque $[X]_{\mathcal{B}_1} = \mathrm{Mat}_{\mathcal{B}_2,\mathcal{B}_1}(\mathrm{Id})[X]_{\mathcal{B}_2}$, on obtient l'expression de l'effet d'un changement de base sur les coordonnées d'un vecteur :

Proposition 7.2.12 (Effet d'un changement de base sur les coordonnées d'un vecteur)

Soit $X \in E$. Alors $[X]_{\mathcal{B}_1} = P_{\mathcal{B}_1}^{\mathcal{B}_2} \cdot [X]_{\mathcal{B}_2}$.

En composant à gauche et à droite par l'identité, avec bases différentes, on obtient, toujours d'après la formule de composition, la très importante formule de changement de base.

Théorème 7.2.13 (Formule de changement de base)

Soit E un espace vectoriel de dimension finie, muni de deux bases $\mathcal{B}_1$ et $\mathcal{B}_2$, et F un espace vectoriel de dimension finie, muni de deux bases $\mathcal{C}_1$ et $\mathcal{C}_2$. Soit $f \in \mathcal{L}(E,F)$. Alors :

$$\mathrm{Mat}_{\mathcal{B}_2,\mathcal{C}_2}(f) = P_{\mathcal{C}_2}^{\mathcal{C}_1} \cdot \mathrm{Mat}_{\mathcal{B}_1,\mathcal{C}_1}(f) \cdot P_{\mathcal{B}_1}^{\mathcal{B}_2} = (P_{\mathcal{C}_1}^{\mathcal{C}_2})^{-1} \cdot \mathrm{Mat}_{\mathcal{B}_1,\mathcal{C}_1}(f) \cdot P_{\mathcal{B}_1}^{\mathcal{B}_2}$$

Ainsi, cette formule s'écrit $M' = Q^{-1}MP$, où M est la matrice dans les bases initiales, M' la matrice dans les nouvelles bases, P et Q les matrices de passage de la première base vers la seconde, respectivement dans E et dans F.

Exemple 7.2.14

1. Retrouver les matrices des exemples précédents.
2. Donner une relation entre la matrice de D dans la base canonique de $\mathbb{R}_n[X]$ (au départ et à l'arrivée)et la matrice de Jordan J_{n+1}.

On remarque que deux matrices associées à une même application linéaire avec des choix différents de bases, s'obtiennent l'une de l'autre par multiplication à gauche et à droite par des matrices inversibles.

Définition 7.2.15 (Matrices équivalentes)

Soit $(M, N) \in \mathcal{M}_{n,p}(\mathbb{K})$. On dit que M et N sont équivalentes si et seulement s'il existe $P \in \mathrm{GL}_n(\mathbb{K})$ et $Q \in \mathrm{GL}_p(\mathbb{K})$ tels que $N = PMQ$.

Proposition 7.2.16

Cela définit une relation d'équivalence.

Ainsi, on obtient :

Théorème 7.2.17

Soit $f \in \mathcal{L}(E, F)$.

(i) Deux matrices M et N représentant f dans des choix différents de bases sont équivalentes ;

(ii) Réciproquement, si M est la matrice de f relativement à deux bases $\mathcal{B}_1$ et $\mathcal{C}^1$, et si N est équivalente à M, alors il existe deux bases $\mathcal{B}_2$ et $\mathcal{C}_2$ de E et F telles que N soit la matrice de f relativement aux bases $\mathcal{B}_2$ et $\mathcal{C}_2$.

Soit, pour tout $(n, p) \in (\mathbb{N}^*)^2$, et pour tout $r \in [\![0, \min(n, p)]\!]$, soit

$$I_{n,p,r} = \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & & \vdots & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots & & & \vdots \\ \vdots & & \ddots & 1 & 0 & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & \cdots & 0 \end{pmatrix} = \left(\begin{array}{c|c} I_r & 0_{r,p-r} \\ \hline 0_{n-r,r} & 0_{n-r,p-r} \end{array}\right)$$

la matrice de type (n, p), constituée de 1 sur ses r premiers coefficients diagonaux, et de 0 partout ailleurs.

Théorème 7.2.18

Soit E et F des espaces de dimensions respectives p et n. Soit $f \in \mathcal{L}(E, F)$, de rang r. Il existe deux bases $\mathcal{B}$ et $\mathcal{C}$, respectivement de E et de F, telles que

$$\mathrm{Mat}_{\mathcal{B},\mathcal{C}}(f) = I_{n,p,r}.$$

Corollaire 7.2.19

Toute matrice de $\mathcal{M}_{n,p}(\mathbb{K})$ de rang r est équivalente à $I_{n,p,r}$.

Corollaire 7.2.20 (Classification des matrices équivalentes par le rang)

Deux matrices de même format sont équivalentes si et seulement si elles ont même rang.

Ainsi, dans $\mathcal{M}_{n,p}(\mathbb{R})$, les classes d'équivalence, pour la relation d'équivalence des matrices, sont les sous-ensembles de matrices de même rang $r \in [\![0, \min(n, p)]\!]$. En particulier, l'espace quotient est en bijection avec $[\![0, \min(n, p)]\!]$.

Définition 7.2.21 (Matrice extraite)

Soit $A = (a_{i,j})$ une matrice de $\mathcal{M}_{n,p}$. Une matrice B est extraite de A si et seulement s'il existe $1 \leqslant i_1 < \cdots < i_q \leqslant n$ et $1 \leqslant j_1 < \cdots < j_r \leqslant p$ tels que $B = (a_{i_k,j_\ell})_{(k,\ell)\in[\![1,q]\!]\times[\![1,r]\!]}$.

Ainsi, la matrice B est obtenue en ne conservant de A que les lignes d'indice $i_1, \ldots, i_q$ et les colonnes d'indice $j_1, \ldots, j_r$.

Proposition 7.2.22 (Rang d'une matrice extraite)

Soit B une matrice extraite de A. Alors $\mathrm{rg}(B) \leqslant \mathrm{rg}(A)$.

Théorème 7.2.23 (Caractérisation du rang par les matrices carrées extraites)

Le rang de A est l'ordre maximal d'une matrice carrée inversible extraite de A.

II.3 Matrice d'un endomorphisme, matrices semblables

Dan le cas où f est un endomorphisme, il est fréquent de choisir sur E la même base au départ et à l'arrivée (même si ceci n'est en théorie par strictement nécessaire). Dans ce cas, on allège un peu les notations, en notant simplement $\mathcal{M}_{\mathcal{B}}(f)$ au lieu de $\mathcal{M}_{\mathcal{B},\mathcal{B}}(f)$. Par ailleurs, un changement de base sur un endomorphisme s'effectue dans ce cas en faisant le même changement de variable au départ et à l'arrivée. Ainsi :

Théorème 7.2.24 (Changement de base pour un endomorphisme)

Soit $f \in \mathcal{L}(E)$, E étant un espace vectoriel de dimension finie, muni de deux bases $\mathcal{B}_1$ et $\mathcal{B}_2$. Alors :

$$\mathrm{Mat}_{\mathcal{B}_2}(f) = (P_{\mathcal{B}_1}^{\mathcal{B}^2})^{-1}\mathrm{Mat}_{\mathcal{B}_1}(f)P_{\mathcal{B}_1}^{\mathcal{B}^2}.$$

Ainsi, cette relation s'écrit : $M' = P^{-1}MP$, où M est la matrice de f dans l'ancienne base, M' la matrice de f dans la nouvelle base, et P la matrice de passage de l'ancienne base vers la nouvelle base.

Exemple 7.2.25

1. Soit p un projecteur de rang r d'un espace de dimension n. Montrer qu'il existe une base $\mathcal{B}$ de E telle que $\mathrm{Mat}_{\mathcal{B}}(p) = I_{n,n,r}$.
2. Décrire de même une matrice simple représentant une symétrie dans un certain choix de base.
3. Déterminer une base relativement à laquelle la matrice de $f : (x, y) \mapsto (3x + 2y, x)$ est diagonale. En déduire une relation entre cette matrice diagonale et la matrice $\begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix}$.

On dit qu'on a diagonalisé $\begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix}$, ou de façon équivalente, qu'on a diagonalisé f.

La notion de diagonalisation est liée à l'étude des classes d'équivalence de la relation fournie par les changements de base pour les endomorphismes :

Définition 7.2.26 (Matrices semblables)

On dit que deux matrices $A, B \in \mathcal{M}_n(\mathbb{K})$ sont semblables s'il existe une matrice inversible $P \in \mathrm{GL}_n(\mathbb{K})$ telle que $B = P^{-1}AP$.

Proposition 7.2.27 (Relation de similitude)
La relation ainsi définie, appelée relation de similitude, est une relation d'équivalence.

Ainsi :

Corollaire 7.2.28
Les matrices d'un endomorphisme dans différentes bases de E sont semblables.

La classification des matrices semblables est beaucoup plus compliquée que la classification des matrices équivalentes. Cette classification est à l'origine du problème de la réduction des endomorphismes. Le problème de la réduction des endormophismes consiste à trouver un système simple de représentants des classes de similitude, donc de trouver une matrice simple canonique équivalente à une matrice donnée. La diagonalisation des matrices (ou des endomorphismes) est une des branches de ce problème, mais les matrices diagonales ne suffisent pas à donner un système de représentants des classes de similitude. Par ailleurs, il est important de noter que cette notion dépend beaucoup du corps de base. En effet, toutes matrice de $\mathcal{M}_n(\mathbb{C})$ est $\mathbb{C}$-semblable à une matrice triangulaire supérieure (et même à une matrice de forme assez particulière, avec seulement deux diagonales non nulles), donc toute matrice de $\mathcal{M}_n(\mathbb{R})$ aussi ; en revanche elles ne sont pas toutes $\mathbb{R}$-semblables à une matrice triangulaire supérieure. Trouver un système de représentant des classes de $\mathbb{R}$-similitude est un problème plus compliqué que trouver un système de représentant de $\mathbb{C}$-similitude.

La classification des classes de similitude étant loin d'être triviale, une étape importante est la recherche d'invariants de similitude, c'est-à-dire de propriétés ou quantités préservées par similitude. Ces invariants permettent de distinguer certaines matrices non semblables (si elles n'ont pas même invariant) mais sont en général insuffisants pour prouver qu'elles sont semblable (sauf si l'invariant caractérise la similitude).

Nous avons déjà étudié un invariant de similitude : le rang. En effet, deux matrices semblables sont équivalentes. Cet invariant est assez faible, puisqu'il ne permet de classer les matrices de $\mathcal{M}_n(\mathbb{K})$ qu'en $n+1$ catégories.

Un deuxième invariant, encore plus facile à calculer, est la trace.

Définition 7.2.29 (Trace d'une matrice)
Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ une matrice carrée d'ordre n. La trace de A est la somme de ses coefficients diagonaux :

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{i,i}.$$

De façon assez immédiate, on obtient :

Proposition 7.2.30 (Linarité de la trace)
L'application $\mathrm{tr} : \mathcal{M}_n(\mathbb{K}) \longrightarrow \mathbb{K}$ *est une forme linéaire sur* $\mathcal{M}_n(\mathbb{K})$.

Proposition 7.2.31 (Invariance par transposition)
Soit $A \in \mathcal{M}_n(\mathbb{K})$. *Alors* $\mathrm{tr}(A) = \mathrm{tr}({}^tA)$.

Théorème 7.2.32 (Invariance de la trace par commutation interne)
Soit $A \in \mathcal{M}_{n,p}(\mathbb{K})$ et $B \in \mathcal{M}_{p,n}(\mathbb{K})$. Alors $\operatorname{tr}(AB) = \operatorname{tr}(BA)$.

Remarquez que A et B ne sont pas des matrices carrées, mais de format transposé. Ainsi, AB est une matrice carrée d'ordre n alors que BA est une matrice carrée d'ordre p.

Avertissement 7.2.33
Le théorème affirme qu'on peut inverser l'ordre d'un produit dans la trace, lorsque la matrice dont on cherche la trace s'exprime comme produit de 2 termes. Cela ne signifie pas qu'on peut commuter come on veut les termes d'un produit de n termes à l'intérieur de la trace. Ainsi, on peut écrire :

$$\operatorname{tr}(ABC) = \operatorname{tr}(CBA) = \operatorname{tr}(BCA),$$

mais en général, **on n'a pas** :

$$\operatorname{tr}(ABC) = \operatorname{tr}(ACB),$$

l'expression ACB ne pouvant pas se déduire de l'interversion (globale) de 2 termes de ABC.

Exemple 7.2.34
Comparer $\operatorname{tr}\left(\begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}1&1\\0&0\end{pmatrix}\right)$ et $\operatorname{tr}\left(\begin{pmatrix}1&0\\0&0\end{pmatrix}\begin{pmatrix}1&1\\0&0\end{pmatrix}\begin{pmatrix}0&1\\1&0\end{pmatrix}\right)$.

Corollaire 7.2.35 (Invariance de la trace par similitude)
La trace est un invariant de similitude. Autrement dit, si M et N sont semblables, alors $\operatorname{tr}(M) = \operatorname{tr}(N)$.

Cela permet de définir :

Définition 7.2.36 (Trace d'un endomorphisme)
La trace d'un endomorphisme est la valeur commune de la trace des matrices de f relativement à un choix quelconque d'une base.

On en déduit notamment que, comme pour les matrices, la trace est une forme linéaire sur $\mathcal{L}(E)$, et que pour tout $u \in \mathcal{L}(E,F)$ et $v \in \mathcal{L}(F,E)$, on a $\operatorname{tr}(uv) = \operatorname{tr}(vu)$.
En particulier, on a :

Proposition 7.2.37 (Trace d'un projecteur, d'une symétrie)
1. *Soit p un projecteur de E, alors* $\operatorname{tr}(p) = \operatorname{rg}(p)$.
2. *Soit s une symétrie de E. Alors* $\operatorname{tr}(s) = n - 2\operatorname{rg}(s - \mathrm{id})$

Méthode 7.2.38 (Dimension des éléments géométriques d'une projection/symétrie)
- S'assurer que l'endomorphisme considéré est une projection ou une symétrie, en calculant u^2.
- Déterminer la matrice de u dans une base quelconque
- Déterminer la trace de u grâce à cette matrice.
- la propriété précédente donne la dimension des éléments propres $\operatorname{Ker}(u)$ et $\operatorname{Im}(u)$ de u si u est un projecteur, et de $\operatorname{Ker}(u-\mathrm{id})$ et $\operatorname{Ker}(u+\mathrm{id})$ si u est une symétrie (ces deux espaces étant supplémentaires pour une symétrie).

Ces invariants ne représentent qu'un tout petit pas vers la classification des classes de similitude. Vous définirez l'année prochaine les valeurs propres d'une matrice. L'ensemble des valeurs propres d'une matrice est également un invariant de similitude.

Nous donnons sans preuve la description complète d'un système de représentant des classes de similitude dans $\mathcal{M}_n(\mathbb{C})$. Pour cela, nous définissons, pour tout $\ell \in [\![1, n]\!]$ et tout $\lambda \in \mathbb{C}$, la matrice :

$$J_\ell(\lambda) = \lambda I_\ell + J_\ell = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}$$

Dans le cas où $n = 1$, la matrice $J_1(\lambda)$ est réduite à la matrice à une seul coefficient (λ).
Alors, toute matrice de $\mathcal{M}_n(\mathbb{C})$ est semblable, à permutation près des blocs diagonaux, à une unique matrice

$$\begin{pmatrix} J_{\ell_1}(\lambda_1) & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & J_{\ell_k}(\lambda_k) \end{pmatrix}, \qquad \ell_1 + \cdots + \ell_k = n, \quad (\lambda_1, \ldots, \lambda_k) \in \mathbb{C}^k,$$

matrice constituée de blocs diagonaux carrés égaux aux matrices $J_{\ell_i}(\lambda_i)$.
Le problème de la réduction d'un endomorphisme u, ou de façon équivalente, d'une matrice M, est de trouver une base relativement à laquelle la matrice de u est de la forme ci-dessus, ou de façon équivalente, de trouver une matrice P inversible telle que $P^{-1}MP$ soit de la forme ci-dessus.
Si les ℓ_k sont tous égaux à 1, la matrice obtenue est diagonale. Ainsi, les matrices diagonales (à permutation près des facteurs) font partie de la famille décrite ci-dessus. Le problème de la diagonalisation est donc un sous-problème du problème plus vaste de la réduction, et si elle n'aboutit pas, il faudra chercher un représentant de la classe de M sous le format plus général donné ci-dessus. Nous définissons :

Définition 7.2.39 (Endomorphisme diagonalisable, matrice diagonalisable)

(i) Un endomophisme est diagonalisable s'il existe une base dans laquelle sa matrice est diagonale.

(ii) Une matrice est diagonalisable si elle est semblable à une matrice diagonale.

Trouver une telle matrice diagonale (et la matrice de passage associée) s'appelle « diagonaliser » l'endomorphisme ou la matrice. Vous verrez l'année prochaine des techniques efficaces pour diagonaliser un endomorphisme ou une matrice.

De façon plus générale, trouver un représentant sous la forme générale exposée ci-dessus (et la matrice de passage correspondante) s'appelle « réduire » ou « jordaniser » l'endomorphisme ou la matrice.
Pour la culture, le résultat qui est à la base du théorème de jordanisation, disant que toute matrice à coefficients complexes est jordanisable, est le théorème suivant :

Théorème 7.2.40 (Théorème des noyaux itérés)

Soit u un endomorphisme de E de dimension finie, on a les inclusions :

$$\mathrm{Ker}(u^0) \subset \mathrm{Ker}(u^1) \subset \mathrm{Ker}(u^2) \subset \cdots \subset \mathrm{Ker}(u^k) \subset \mathrm{Ker}(u^{k+1}) \subset \cdots .$$

De plus, cette suite est stationnaire, et « s'essouffle », dans le sens où les sauts de dimension forment une suite décroissante, ultimement nulle.

Nous verrons la démonstration de ce théorème en exercice.

III Produit matriciel par blocs

Nous voyons dans cette dernière section qu'on peut effectuer un produit matriciel en groupant les termes par blocs rectangulaires, en utilisant les règles usuelles, c'est-à-dire en remplaçant dans les formules les coefficients par des blocs matriciels. Il faut cependant prendre garde au fait que cela n'est possible que si le découpage en blocs des deux matrices A et B fournit des formats compatibles pour les produits intervenant de la sorte.
Soit $\mathcal{B} = (b_1, \dots, b_p)$ une base de E, $\mathcal{C} = (c_1, \dots, c_n)$ une base de F. Soit $u \in \mathcal{L}(E, F)$ et soit $A = \mathrm{Mat}_{\mathcal{B},\mathcal{C}}(u) = (a_{i,j})$. On définit un découpage par blocs de M par les indices

$$0 = i_0 < i_1 < i_2 < \cdots < i_{q-1} < i_q = n \qquad \text{et} \qquad 0 = j_0 < j_1 < j_2 < \cdots < j_{r-1} < j_r = p.$$

Les blocs sont alors définis par $A_{k,\ell} = (a_{i,j})_{(i,j)\in[\![i_{k-1}-1,i_k]\!]\times[\![j_{\ell-1}+1,j_\ell]\!]}$. La matrice $A_{k,\ell}$ est donc la matrice obtenue par extraction des lignes $i_{k-1}+1$ à i_k et des colonnes $j_{\ell-1}+1$ à j_ℓ de A. On dira que le découpage de A en blocs $(A_{k,\ell})$ est associées aux suites d'indices

$$0 = i_0 < i_1 < i_2 < \cdots < i_{q-1} < i_q = n \qquad \text{et} \qquad 0 = j_0 < j_1 < j_2 < \cdots < j_{r-1} < j_r = p.$$

Par ailleurs, étant donné E, muni de la base $\mathcal{B} = (b_1, \dots, b_p)$, nous définissons la décomposition de E associée à la suite

$$0 = j_0 < j_1 < j_2 < \cdots < j_{r-1} < j_r = p$$

par

$$E = E_1 \oplus \cdots \oplus E_r,$$

où pour tout $k \in [\![1, r]\!]$, $E_k = \mathrm{Vect}(\mathcal{B}_k)$, où $\mathcal{B}_k = (b_{j_{k-1}+1}, \dots, b_{j_k})$.
De même, nous disposons alors d'une décomposition de F associée à la base $\mathcal{C}$, et à la suite

$$0 = i_0 < i_1 < i_2 < \cdots < i_{q-1} < i_q = n,$$

décomposition que nous écrivons $F = F_1 \oplus \cdots \oplus F_q$. La suite (j_i) définit également des tranches adaptées $\mathcal{C}_\ell$ de la base $\mathcal{C}$, comme ci-dessus.
Enfin, étant donnée la décomposition en sommes directes $F = F_1 \oplus \cdots \oplus F_q$, on appelle projection sur le k-ième facteur l'application linéaire de $\mathcal{L}(E, E_k)$ définie par :

$$p_k(x_1 + \cdots + x_q) = x_k.$$

Il s'agit d'un projecteur, et on a de façon évidente,

$$\forall x \in F, \quad x = p_1(x) + \cdots + p_q(x).$$

Théorème 7.3.1 (Interprétation géométrique d'un découpage par blocs)
Avec les notations introduites ci-dessus, pour tout $(k, \ell) \in [\![1, q]\!] \times [\![1, r]\!]$, $A_{k,\ell}$ est la matrice de l'application de $\mathcal{L}(E_\ell, F_k)$, obtenue par projection sur F_k de la restriction de u à E_ℓ relativement aux base $\mathcal{B}_\ell$ et $\mathcal{C}_k$:

$$A_{k,\ell} = \mathrm{Mat}_{\mathcal{B}_\ell,\mathcal{C}_k}(p_k \circ u_{|E_\ell}).$$

De cette interprétation, on déduit le théorème de produit par blocs :

Théorème 7.3.2 (Produit par blocs)
Soit $A \in \mathcal{M}_{n,p}$ et $B \in \mathcal{M}_{p,m}$ deux matrices, et

$$0 = i_0 < i_1 < i_2 < \cdots < i_{q-1} < i_q = n,$$
$$0 = j_0 < j_1 < j_2 < \cdots < j_{r-1} < j_r = p,$$
$$0 = k_0 < k_1 < k_2 < \cdots < k_{s-1} < k_s = m.$$

Les deux premières suites définissent un découpage par blocs $A = (A_{i,j})_{(i,j)\in[\![1,q]\!]\times[\![1,r]\!]}$ *de* A *et les deux dernières définissent un découpage par blocs* $B = (B_{j,k})_{(j,k)\in[\![1,r]\!]\times[\![1,s]\!]}$.
Le produit AB *admet alors une représentation par blocs :*

$$AB = (C_{i,k})_{(i,k)\in[\![1,q]\!]\times[\![1,s]\!]},$$

où pour tout $(i,k) \in [\![1,q]\!]\times[\![1,s]\!]$, $C_{i,k} = \sum_{j=1}^{r} A_{i,j}B_{j,k}$.

En particulier, si on a une représentation diagonale par blocs :

$$A = \begin{pmatrix} A_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & A_k \end{pmatrix},$$

où les A_k sont des matrices carrées, alors pour tout $n \in \mathbb{N}$,

$$A^n = \begin{pmatrix} A_1^n & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & A_k^n \end{pmatrix}.$$

Exemple 7.3.3

Calcul des puissances successives de $\begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$

8 Groupe symétrique et déterminants

Nous rappelons que le groupe symétrique $\mathfrak{S}_n$ ou S_n est le groupe des permutations de $[\![1,n]\!]$, la loi du groupe étant la composition. L'étude algébrique des groupes $\mathfrak{S}_n$ est très intéressante, et aboutit à des résultats aussi importants que la non résolubilité par radicaux des équations de degré au moins 5. Cette étude va bien au-delà des objectifs du programme, qui sont uniquement d'introduire les outils nécessaires pour pouvoir définir correctement les déterminants. Nous nous limiterons donc, avec beaucoup de regrets, à ces objectifs.

I Groupe symétrique

Notre but est d'introduire les outils nécessaires à la définition de la signature d'une permutation, notion utilisée ensuite dans l'étude des déterminants. La signature d'une permutation est défini comme l'unique morphisme de groupes $\varepsilon : \mathfrak{S}_n \mapsto \{-1,1\}$ non constant. Notre objectif est donc de montrer l'existence et l'unicité de ce morphisme.

I.1 Notations liées à des permutations

Nous rencontrerons deux façons de décrire des permutations. La première, la plus naturelle, est d'associer, dans un tableau, à chaque valeur $i \in [\![1,n]\!]$, la valeur de $\sigma(i)$, en rangeant les valeurs de i dans l'ordre croissant. Nous disposerons ces données sous la forme matricielle suivante :

Notation 8.1.1 (Permutation)

Soit $\sigma \in \mathfrak{S}_n$. Nous désignerons explicitement σ par le tableau :

$$\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

Exemple 8.1.2

1. Décrire de la sorte tous les éléments de $\mathfrak{S}_2$ et $\mathfrak{S}_3$.
2. Décrire de la sorte la permutation de $\mathfrak{S}_n$ inversant l'ordre : $\sigma(k) = n+1-k$.

Certaines permutations ont un comportement particulier : elles effectuent une rotation sur un ensemble d'éléments (qu'on peut ranger en cercle), et laissent les autres invariants. On appelle *cycle* ces permutations. Pour rendre le caractère cyclique plus visible sur les notations on adopte pour ces permutations une notation plus adaptée. Nous définissons de façon plus précise :

Définition 8.1.3 (Cycle)

Un cycle est la donnée d'un sous ensemble I de $[\![1,n]\!]$ et d'une permutation σ de $[\![1,n]\!]$ telle qu'il existe $i_1,\dots,i_k$ deux à deux distincts tels que $I=\{i_1,\dots,i_k\}$, et :

(i) $\forall \ell \in [\![1,k-1]\!],\quad \sigma(i_\ell)=i_{\ell+1}$ et $\sigma(i_k)=i_1$

(ii) $\forall i \in [\![1,n]\!]\setminus I,\quad \sigma(i)=i.$

Remarque 8.1.4

- Sauf dans un cas bien particulier, la donnée de I est implicitement fournie par la donnée de σ : il s'agit de l'ensemble des points de $[\![1,n]\!]$ que σ ne laisse pas fixe.
- Le cas problématique pour lequel I ne découle pas de la seule donnée de σ est le cas d'un cycle de longueur 1 (donc I est un singleton). Dans ce cas σ est l'identité, et tous les points de σ sont invariants. La définition que nous donnons d'un cycle (comme étant un couple formé d'un sous-ensemble et d'une permutation) permet de différentier les cycles de longueur 1. Par exemple les cycles (1) et (2) correspondent à la même permutation, mais il ne s'agit pas du même cycle, n'ayant pas même support.
- Dans la suite, nous identifierons les cycles et les permutations qu'ils définissent, en gardant en tête cette différentiation possible des supports dans le cas de cycles de longueur 1.

Exemple 8.1.5

Sont-ce des cycles : $\begin{pmatrix}1&2&3&4&5&6\\1&4&5&3&2&6\end{pmatrix}$? $\begin{pmatrix}1&2&3&4&5&6\\3&4&5&2&6&1\end{pmatrix}$?

Notation 8.1.6 (Cycles)

Avec les conditions de la définition 8.1.3, on note

$$\sigma=(i_1\ i_2\ \dots\ i_k).$$

On appelle support du cycle l'ensemble I. La longueur du cycle est l'entier $k=\mathrm{Card}(I)$.

Cette notation signifie que tout élément de la liste est envoyé sur l'élément suivant, le dernier étant envoyé sur le premier.

Exemple 8.1.7

1. Écrire le cycle de l'exemple 8.1.5 avec les notations 8.1.6
2. Écrire le cycle (1 5 8 2 9 7 3) avec les notations 8.1.1

Avertissement 8.1.8

Attention, l'ordre des éléments $i_1,\dots,i_n$ est important.

Remarque 8.1.9

La représentation d'un cycle sous la forme 8.1.6 est-elle unique ?

On rencontre parfois :

Définition 8.1.10 (Grand cycle)
Un grand cycle de $\mathfrak{S}_n$ est un cycle de support $[\![1, n]\!]$.

Ainsi, un grand cycle effectue une rotation, dans un certain ordre, entre les n éléments de $[\![1, n]\!]$. Un exemple important de grand cycle est :

Définition 8.1.11 (Permutation circulaire)
(i) On appelle permutation circulaire directe de $\mathfrak{S}_n$ le grand cycle $(1\ 2\ \cdots\ n)$.
(ii) On appelle permutation circulaire indirecte de $\mathfrak{S}_n$ le grand cycle $(n \dots\ 2\ 1)$.

Exemple 8.1.12
Écrire les permutations circulaires directes et indirectes avec les notations 8.1.1

Une famille importante de cycles est constituée des cycles de longueur 2. En effet, il s'agit d'une famille génératrice de $\mathfrak{S}_n$, comme on le verra plus tard.

Définition 8.1.13 (Transpositions)
On appelle transposition de $\mathfrak{S}_n$ un cycle de longueur 2 : $\tau = (i\ j)$.

Ainsi, la transposition $\tau = (i\ j)$ est la permutation consistant en l'échange des valeurs i et j.

Exemple 8.1.14
Écrire la transposition $\tau = (2\ 5)$ de $\mathfrak{S}_6$ avec les notations 8.1.1.

I.2 Signature d'une permutation

Lemme 8.1.15
Soit $\sigma \in \mathfrak{S}_n$. Alors σ induit une bijection sur $\mathcal{P}_2([\![1, n]\!])$, l'ensemble des sous-ensembles de cardinal 2 de $[\![1, n]\!]$.

On en déduit :

Lemme 8.1.16
Soit $\sigma \in \mathfrak{S}_n$. On a :

$$\left| \prod_{1 \leqslant i < j \leqslant n} (\sigma(j) - \sigma(i)) \right| = \prod_{1 \leqslant i < j \leqslant n} (j - i)$$

Puisque $\dfrac{\sigma(j) - \sigma(i)}{j - i} = \dfrac{\sigma(i) - \sigma(j)}{i - j}$, cette quantité ne dépend pas de l'ordre respectif entre i et j, mais uniquement de l'ensemble $\{i, j\}$. On peut donc définir une application δ sur $\mathcal{P}_2([\![1, n]\!])$ par :

$$X = \{i, j\} \mapsto \delta(X) = \frac{\sigma(i) - \sigma(j)}{i - j}.$$

Définition 8.1.17 (Signature)

On définit la signature $\varepsilon : \mathfrak{S}_n \longrightarrow \{-1, 1\}$ par :

$$\varepsilon(\sigma) = \frac{\prod\limits_{1 \leqslant i < j \leqslant n} (\sigma(j) - \sigma(i))}{\prod\limits_{1 \leqslant i < j \leqslant n} (j - i)} = \prod_{X \in \mathcal{P}_2(\llbracket 1, n \rrbracket)} \delta(X).$$

Théorème 8.1.18

La signature ε est un morphisme de groupes de $(\mathfrak{S}_n, \circ)$ dans $(\{-1, 1\}, \times)$.

De l'étude du signe de l'expression définissant $\varepsilon(\sigma)$, il découle que celui-ci-dépend du nombre de couples (i, j) avec $i < j$ tel que $s(i) > \sigma(j)$.

Définition 8.1.19 (Inversion)

Soit $\sigma \in \mathfrak{S}_n$. On appelle inversion de σ un couple (i, j) tel que $i < j$ et $\sigma(i) > \sigma(j)$, c'est-à-dire $\delta(\{i, j\}) < 0$.

Ainsi, la signature peut être décrite à l'aide du nombre d'inversions :

Théorème 8.1.20 (Description de la signature par les inversions)

En notant $\mathrm{Inv}(\sigma)$ *le nombre d'inversions de σ, on a :*

$$\varepsilon(\sigma) = (-1)^{\mathrm{Inv}(\sigma)}.$$

Une transposition $(i\ i+1)$ n'admettant que le couple $(i, i+1)$ comme inversion, il vient alors :

Proposition 8.1.21

Notons, pour tout $i \in \llbracket 1, n-1 \rrbracket$, $\tau_i = (i\ i+1)$. On a alors $\varepsilon(\tau_i) = -1$.

Or, toute transposition peut s'écrire à l'aide de ces transpositions entre éléments consécutifs :

Proposition 8.1.22 (Décomposition d'une transposition à l'aide des τ_i)

Soit $1 \leqslant i < j \leqslant n$, et $\tau = (i\ j)$. Alors

$$\tau = \tau_{j-1} \circ \cdots \circ \tau_{i+1} \circ \tau_i \circ \cdots \circ \tau_{j-2} \circ \tau_{j-1}.$$

En utilisant le fait que ε est un morphisme, on en déduit :

Théorème 8.1.23 (Signature d'une transposition)

Soit τ une transposition. Alors $\varepsilon(\tau) = -1$.

Nous avons donc répondu au problème de l'existence d'un morphisme $\varepsilon : \mathfrak{S}_n \to \{-1, 1\}$, prenant la valeur -1 sur les transpositions.
Pour étudier son unicité, nous allons établir que toute permutation s'écrit comme composition de transposition. Ainsi, la valeur d'un morphisme ε sur une permutation est entièrement déterminée par la valeur de ε sur les transpositions. Imposer la valeur sur les transpositions fournit dans ce cas l'unicité.

Théorème 8.1.24 (Caractère générateur des transpositions)
Toute permutation $\sigma \in \mathfrak{S}_n$ est un produit de transpositions.

Les transpositions étant elles mêmes engendrées par les τ_i, on se rend compte que toute permutation s'écrit comme produit des τ_i. Cela ne soit pas nous étonner, on l'avait déjà prouvé en étudiant la correction de certains algorithmes de tri (lesquels ?)

Remarque 8.1.25
Le rapport obtenu entre signature et décomposition en produit de transpositions permet d'affirmer que la parité du nombre de terme de toute décomposition en produits de transpositions d'une permutation donnée est la même.

Terminologie 8.1.26 (Permutation paire, permutation impaire)
On dira qu'une permutation est paire si $\varepsilon(\sigma) = 1$, et impaire si $\varepsilon(\sigma) = -1$.

Définition 8.1.27 (Groupe alterné)
Le groupe alterné $\mathfrak{A}_n$ est le noyau de la signature, c'est-à-dire l'ensemble des permutations paires. C'est un sous-groupe distingué de $\mathfrak{S}_n$.

I.3 Décomposition cyclique d'une permutation

Décomposer une permutation en produit de transpositions, ou compter le nombre d'inversions n'est pas quelque chose d'immédiat. Nous donnons dans ce paragraphe une autre façon, plus rapide, de calculer la signature d'une permutation, en s'aidant du type cyclique de cette permutation.
Soit σ une permutation. On définit sur $[\![1, n]\!]$ la relation : $i \equiv_\sigma j$ si et seulement s'il existe $k \in \mathbb{N}$ tel que $j = \sigma^k(i)$.

Proposition 8.1.28
La relation $\equiv_\sigma$ est une relation d'équivalence.

Soit $x \in [\![1, n]\!]$, et S_x l'unique classe d'équivalence contenant x.

Proposition 8.1.29
Soit k le cardinal de S_x. On a $S_x = \{x, \sigma(x), \dots, \sigma^{k-1}(x)\}$ sont deux à deux distincts, et σ induit sur S_x une permutation de S_x, égale au cycle $(x, \sigma(x), \dots, \sigma^{k-1}(x))$.

Notons C_x le cycle de $\mathfrak{S}_n$ défini par

$$C_x = (x, \sigma(x), \dots, \sigma^{k-1}(x)).$$

Lemme 8.1.30
Soit x et y des éléments d'une même classe d'équivalence modulo $\equiv_\sigma$. Alors $C_x = C_y$.

Théorème 8.1.31 (Décomposition en cycles d'une permutation)

Soit σ une permutation de $\mathfrak{S}_n$. À permutation près des facteurs, il existe une unique décomposition de σ en produits de cycles :

$$\sigma = C_1 \circ \cdots \circ C_k,$$

telle que les supports des cycles forment une partition de $[\![1, n]\!]$

Cette dernière hypothèse signifie que les cycles considérés sont à supports disjoints, et que tout élément de $[\![1, n]\!]$ est dans un cycle, quitte à considérer dans la décomposition des cycles de longueur 1 (en ayant en mémoire la possibilité de différentier leur support.

Remarque 8.1.32

- Les cycles C_i sont les cycles associés à un système de représentants modulo la relation $\equiv_\sigma$.
- L'ordre dans lequel on effectue cette composition importe peu, du fait du lemme qui suit.

Lemme 8.1.33 (Commutation des cycles à supports disjoints)

Soit C_1 et C_2 deux cycles à supports disjoints. Alors $C_1 \circ C_2 = C_2 \circ C_1$.

Méthode 8.1.34 (Comment déterminer la décomposition en produits de cycles)

Soit σ une permutation de $\mathfrak{S}_n$.

- Partir de 1 et suivre ses images successives jusqu'à retomber sur 1. Cela donne le premier cycle
- Recommencer en partant du plus petit élément de $[\![1, n]\!]$ n'appartenant pas au premier cycle trouvé
- Recommencer ainsi jusqu'à épuisement des éléments de $[\![1, n]\!]$.

Exemple 8.1.35

Trouver la décomposition en cycles à supports disjoints de :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 5 & 2 & 8 & 10 & 7 & 9 & 1 & 3 & 4 & 6 \end{pmatrix}$$

Définition 8.1.36 (Support cyclique d'une permutation)

Le support cyclique d'une permutation σ est la partition formée des supports des cycles formant la décomposition en cycles de σ. Ainsi, il s'agit de la partition formée des classes d'équivalence de la relation $\equiv_\sigma$.

Définition 8.1.37 (Type cyclique d'une permutation)

Le type cyclique d'une permutation σ est la suite croissante des tailles des parts du support cyclique. Une telle suite croissante de somme n est appelée partition de l'entier n.

I.4 Cycles et signature

Nous voyons enfin comment utiliser la décomposition en cycles pour déterminer la signature. Pour cela nous remarquons qu'il est facile d'écrire un cycle comme produit de transpositions :

Lemme 8.1.38 (Décomposition d'un cycle en transpositions)
Soit $\{i_1, \dots, i_k\}$ des entiers 2 à 2 distincts de $[\![1, n]\!]$. Alors :

$$(\ i_1 \ i_2 \ \cdots \ i_k \ ) = (i_1 \ i_k) \circ (i_1 \ i_{k-1}) \circ \cdots \circ (i_1 \ i_2).$$

Dans le cas où $k = 1$, le terme de droite est réduit à un produit vide de transpositions (égal à l'identité)

Proposition 8.1.39 (Signature d'un cycle)
Soit C un cycle et $\ell(C)$ sa longueur. Alors

$$\varepsilon(C) = (-1)^{\ell(C)-1}.$$

Théorème 8.1.40 (Détermination de ε par le type cyclique)
Soit σ une permutation de $\mathfrak{S}_n$, et $c(\sigma)$ le nombre de parts dans son support cyclique (ou de façon équivalente dans son type cyclique). Alors :

$$\varepsilon(\sigma) = (-1)^{n-c(\sigma)}.$$

II Déterminants

Dans cette section, nous étudions les déterminants. Nous verrons cette notion tout d'abord vue comme un objet défini sur une famille de vecteurs, puis nous en déduirons une notion de déterminant d'un endomorphisme, puis d'une matrice carrée. Notre but étant en grande partie de caractériser l'inversibilité d'une matrice grâce à cet objet, nous introduisons ensuite des techniques calculatoires efficaces. Nous voyons enfin comment donner une expression de l'inverse d'une matrice à l'aide des déterminants, même si en pratique, la formule obtenue manque d'efficacité.

II.1 Formes multilinéaires

Soit $\mathbb{K}$ un corps.

Définition 8.2.1 (Application multilinéaire)
Soit $E_1, \dots, E_n$ et F des $\mathbb{K}$-esaces vectoriels. Une forme n-linéaire est une application

$$\varphi : E_1 \times \cdots \times E_n \longrightarrow F$$

telle que pour tout $i \in [\![1, n]\!]$, φ soit linéaire par rapport à sa i-ième variable, les autres étant fixées quelconques, donc si pour tout $x_1, \dots, x_n$, x_i' et λ,

$$\varphi(x_1, \dots, x_{i-1}, \lambda x_i + x_i', x_{i+1}, \dots, x_n) = \lambda\varphi(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) + \varphi(x_1, \dots, x_{i-1}, x_i', x_{i+1}, \dots, x_n).$$

une application est une forme multilinéaire si elle est n-linéaire pour un certain $n \geqslant 2$.

Proposition 8.2.2
Soit φ une application n-linéaire. Alors $\varphi(x_1, \dots, x_n) = 0$ dès lors qu'une des variables x_i est nulle.

Définition 8.2.3 (Forme n-linéaire)

Une forme n-linéaire est une application n-linéaire à valeurs dans $\mathbb{K}$.

Exemple 8.2.4

1. Voici des exemples de formes bilinéaires ($n = 2$) :
   - $(x, y) \in (\mathbb{R}^n)^2 \mapsto \langle x, y\rangle$, le produit scalaire canonique
   - $(x, y) \in \mathbb{C}^2 \mapsto xy$
   - $(f, g) \in \mathcal{L}(E) \mapsto f \circ g$
   - $(X, Y) \in (\mathbb{R}^n)^2 \mapsto {}^tYMX$
2. $(x_1, \dots, x_n) \mapsto x_1 \cdots x_n$
3. $(f_1, \dots, f_n) \mapsto \displaystyle\int_0^1 f_1(t) \cdots f_n(t)\ \mathrm{d}t$
4. L'aire (signée) du parallélogramme défini par deux vecteurs x et y
5. Le volume du parallélépipède défini par trois vecteurs.

Théorème 8.2.5 (n-linéarité généralisée)

Soit $\varphi : E_1 \times \cdots \times E_n \to F$ une application n-linéaire. Alors pour tout $(k_1, \dots, k_n) \in \mathbb{N}^*$, pour tout $x_{i,j} \in E_i$ et $\lambda_{i,j} \in \mathbb{K}$, ($i \in [\![1, n]\!]$, $j \in [\![1, k_i]\!]$), on a :

$$\varphi\left(\sum_{i_1=1}^{k_1} \lambda_{1,i_1} a_{1,i_1}, \dots, \sum_{i_n=1}^{k_n} \lambda_{n,i_n} a_{n,i_n}\right) = \sum_{i_1=1}^{k_1} \cdots \sum_{i_n=1}^{k_n} \lambda_{1,i_1} \cdot \lambda_{n,i_n} \varphi(a_{1,i_1}, \dots, a_{n,i_n}).$$

Avertissement 8.2.6

Les indices des sommes doivent être indépendants !

Exemple 8.2.7

Dans le cas de la bilinéarité, on obtient :

$$\varphi\left(\sum_{i=1}^{k} \lambda_i a_i, \sum_{j=1}^{\ell} \mu_j b_j\right) = \sum_{i=1}^{k} \sum_{j=1}^{\ell} \lambda_i \mu_i a_i b_j.$$

Proposition 8.2.8 (L'espace des applications n-linéaires)

L'ensemble, noté $\mathcal{L}(E_1, \dots, E_n; F)$ des applications multilinéaires est un espace vectoriel sur $\mathbb{K}$.

Proposition 8.2.9

On a un isomorphisme d'espaces vectoriels :

$$\mathcal{L}(E_1, \dots, E_n; F) \simeq \mathcal{L}(E_1, \mathcal{L}(E_2, \dots, E_n; F).$$

Par itération, on en déduit un isomorphisme d'espaces vectoriels :

$$\mathcal{L}(E_1, \dots, E_n; F) \simeq \mathcal{L}(E_1, \mathcal{L}(E_2, \mathcal{L}(E_3, \dots \mathcal{L}(E_n, F)))).$$

Notation 8.2.10

Lorsque $E_1 = \cdots = E_n$, on notera plus simplement $\mathcal{L}_n(E;F)$ l'espace des applications n-linéaires de E dans F, et $\mathcal{L}_n(E)$ l'espace des formes n-linéaires

Comme pour les applications linéaires, pour déterminer entièrement une application n-linéaire (et donc en particulier une forme n-linéaire), il suffit d'en connaître l'image sur les vecteurs d'une base.

Proposition 8.2.11 (Détermination d'une application n-linéaire sur une base)

Soit pour tout $i \in [\![1,n]\!]$, $(e_{i,j})_{1\leqslant j\leqslant d_i}$ une base de E_i et pour tout $(j_1,\dots,j_n) \in [\![1,d_1]\!]\times\cdots\times[\![1,d_n]\!]$, $f_{j_1,\dots,j_n}$ un élément de F. Alors il existe une unique application n-linéaire $\varphi : E_1\times\cdots\times E_n \longrightarrow F$ telle que

$$\forall (j_1,\dots,j_n)\in[\![1,d_1]\!]\times\cdots\times[\![1,d_n]\!],\quad \varphi(e_{1,j_1},\dots,e_{n,j_n}) = f_{j_1,\dots,j_n}.$$

Cette application est donnée explicitement par :

$$\varphi\left(\sum_{j_1=1}^{d_1}\lambda_{1,j_1}e_{1,j_1},\dots,\sum_{j_n=1}^{d_n}\lambda_{1,j_n}e_{1,j_n}\right) = \sum_{j_1=1}^{d_1}\cdots\sum_{j_n=1}^{d_n}\lambda_{1,j_1}\dots\lambda_{n,j_n}f_{j_1,\dots,j_n}.$$

En particulier, si $E_1 = \cdots = E_n = E$, muni d'une unique base $(e_1,\dots,e_d)$ de E, et si on se donne des éléments $f_{i_1,\dots,i_n}$ de F, pour $1 \leqslant i_1,\dots,i_n \leqslant d$, il existe une unique application n-linéaire φ de $\mathcal{L}_n(E;F)$ telle que pour tout $(i_1,\dots,i_n)\in[\![1,d]\!]^n$,

$$\varphi(e_{i_1},\dots,e_{i_n}) = f_{i_1,\dots,i_n}.$$

Exemple 8.2.12

Soit $(e_1,\dots,e_n)$ la base canonique de $\mathbb{K}^n$. Une forme bilinéaire φ sur $\mathbb{K}^n$ est entièrement déterminée par la donnée de n^2 scalaires $\varphi(e_i,e_j)$. En notant M la matrice carrée définie par ces scalaires ($M = (\varphi(e_i,e_j))_{1\leqslant i,j\leqslant n}$), on peut vérifier que pour tout $X,Y\in R^n$ (qu'on identifie au vecteur colonne des coordonnées dans la base canonique), on a :

$$\varphi(X,Y) = {}^tXMY.$$

Plus généralement, toute forme bilinéaire sur un espace vectoriel E de dimension finie s'écrit de la sorte après choix d'une base de E (voir un chapitre ultérieur).

II.2 Formes n-linéaires symétriques, antisymétriques, alternées

Nous définissons maintenant deux types important de formes n-linéaires prenant leurs variables dans un même espace.

Définition 8.2.13 (Formes n-linéaires symétriques, alternées)

Soit $\varphi :\in \mathcal{L}_n(E)$ une forme n-linéaire. On dit que :

1. φ est symétrique si elle est invariante par permutation de ses variables : pour tout $\sigma\in\mathfrak{S}_n$,
$$\varphi(x_1,\dots,x_n) = \varphi(x_{\sigma(1)},\dots,x_{\sigma(n)}).$$
2. φ est antisymétrique si pour tout $\sigma\in\mathfrak{S}_n$,
$$\varphi(x_1,\dots,x_n) = \varepsilon(\sigma)\varphi(x_{\sigma(1)},\dots,x_{\sigma(n)}).$$
3. φ est alternée si $\varphi(x_1,\dots,x_n) = 0$ dès lors qu'il existe $i\neq j$ tels que $x_i = x_j$.

Exemple 8.2.14

1. Le produit scalaire canonique est une forme bilinéaire symétrique
2. $(A, B) \mapsto \mathrm{tr}(AB)$ est une forme bilinéaire symétrique
3. $(A, B, C) \mapsto \mathrm{tr}(ABC)$ est-elle symétrique ?
4. L'aire du parallélogramme formé par deux vecteurs de $\mathbb{R}^2$ est une forme bilinéaire alternée.
5. Le volume du parallélépipède formé par trois vecteurs de $\mathbb{R}^3$ est une forme bilinéaire alternée.

Lemme 8.2.15

Soit φ une forme alternée. Alors pour tout $(x_1, \dots, x_n) \in E^n$,

$$\varphi(x_1, \dots, x_i, \dots, x_j, \dots, x_n) = -\varphi(x_1, \dots, x_j, \dots, x_i, \dots, x_n).$$

Réciproquement, si cette condition est satisfaite, alors, si $\mathbb{K}$ n'est pas de caractéristique 2, φ est alternée.

Théorème 8.2.16 (Antisymétrie des formes alternées)

Toute forme n-linéaire alternée est antisymétrique. Si $\mathbb{K}$ n'est pas de caractéristique 2, toute forme antisymétrique est alternée.

Proposition 8.2.17 (Image d'une famille liée par une forme alternée)

Soit $(x_1, \dots, x_n)$ une famille liée, et φ une forme alternée. Alors $\varphi(x_1, \dots, x_n) = 0$.

Nous aurons l'occasion de reparler de formes n-linéaires symétriques (et plus particulièrement pour $n = 2$), lorsque nous étudierons les produits scalaires (la symétrie étant une des 3 propriétés requises pour qu'une forme bilinéaire définisse un produit scalaire). Pour l'heure, nous nous concentrons sur la notion de forme alternée, fortement liée, d'après les exemples, aux problèmes de calculs d'aires et de volumes. Nous allons voir qu'à un scalaire multiplicatif près, en dimension n, la mesure des hypervolumes est la seule forme n-linéaire.

II.3 Déterminant d'une famille de vecteurs

Théorème 8.2.18 (Détermination des formes alternées par l'image d'une base)

Soit E un espace vectoriel de dimension finie, muni d'une base $(e_1, \dots, e_n)$, et $p \in \mathbb{N}^*$.

- *Si $p > n$, la seule forme p-linéaire alternée sur E est la forme nulle.*
- *Si $p \leqslant n$, une forme p-linéaire alternée φ est complètement déterminée, de façon unique, par la donnée de $\varphi(e_{i_1}, \dots, e_{i_p})$, pour $1 \leqslant i_1 < \cdots < i_p \leqslant p$.*

Plus précisément, il s'agit de l'unique forme p-linéaire sur E telle que pour tout $(i_1, \dots, i_p) \in [\![1, n]\!]^p$,

$$\varphi(e_{i_1}, \dots, e_{i_p}) = \begin{cases} 0 & \text{s'il existe } j \neq k \text{ tel que } i_j = i_k \\ \varepsilon(\tau)\varphi(e_{i_{\tau(1)}}, \dots, e_{i_{\tau(p)}}) & \text{sinon,} \end{cases}$$

où $\tau \in \mathfrak{S}_p$ est l'unique permutation de $[\![1, p]\!]$ telle que

$$i_{\tau(1)} < \cdots < i_{\tau(p)}.$$

Corollaire 8.2.19 (Formes n-linéaires d'un espace de dimension n)

Soit E un espace vectoriel de dimension n, et $(e_1, \dots, e_n)$ une base de E.

1. *Il existe une unique forme n-linéaire alternée φ sur E telle que $\varphi(e_1, \dots, e_n) = 1$.*
2. *Cette forme n-linéaire est entièrement décrite sur les vecteurs de la base par :*

$$\begin{cases} \varphi(e_{i_1}, \dots, e_{i_n}) = 0 & \text{s'il existe } j \neq k \text{ tel que } i_j = i_k \\ \varphi(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = \varepsilon(\sigma) & \text{où } \sigma \in \mathfrak{S}_n \end{cases}$$

3. *Toute autre forme n-linéaire alternée sur E est de la forme $\lambda\varphi$, $\lambda \in \mathbb{K}$.*

Remarques 8.2.20

- Si la première condition n'est pas satisfaite, les e_{i_k} sont deux à deux distincts, et en nombre n, donc sont constitués des éléments e_i pris chacun une et une seule fois. Aussi la famille $(e_{i,k})$ peut-elle être vue comme une permutation de la famille (e_i), ce qui permet de l'écrire sous la forme $(e_{\sigma(1)}, \dots, e_{\sigma(n)})$.
- La correspondance entre la permutation σ du corollaire précédent, et la permutation τ qui précède le corollaire est (lorsque $p = n$) : $\sigma = \tau^{-1}$.

Cette unique forme n-linéaire alternée va être notre définition du déterminant (par rapport à une base $\mathcal{B}$).

Définition 8.2.21 (Déterminant d'une famille de vecteurs)

Soit E un espace de dimension n, $\mathcal{B} = (e_1, \dots, e_n)$ une base de E et $(x_1, \dots, x_n)$ une famille de n vecteurs de E. Soit $\det_{\mathcal{B}}$ l'unique forme n-linéaire alternée telle que $\det_{\mathcal{B}}(e_1, \dots, e_n) = 1$. Le *déterminant de la famille $(x_1, \dots, x_n)$ par rapport à $\mathcal{B}$* est le scalaire $\det_{\mathcal{B}}(x_1, \dots, x_n)$.

Ainsi, le déterminant par rapport à $\mathcal{B}$ est l'unique forme n-linéaire alternée prenant la valeur 1 sur la famille $\mathcal{B}$.

Remarque 8.2.22

Si E est un $\mathbb{R}$-espace vectoriel, cette notion est à relier à la notion d'hypervolume relative à une base : le déterminant de n vecteurs par rapport à $\mathcal{B}$ est le volume (orienté) du parallélépipède défini par les n vecteurs, l'unité de volume étant le parallélépipède défini par les vecteurs de la base $\mathcal{B}$.

Exemples 8.2.23

1. Voir l'interprétation géométrique dans $\mathbb{R}^2$ et $\mathbb{R}^3$, muni de la base canonique, puis dans $\mathbb{R}^2$ muni d'une base quelconque.
2. Déterminant par rapport à la base canonique de $\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} \right)$.
3. Déterminant par rapport à la base canonique de $\left(\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \right)$.

Nous obtenons plus généralement la description suivante :

Théorème 8.2.24 (Description du déterminant par les coordonnées)

Soit E un espace vectoriel de dimension n et $\mathcal{B} = (e_1, \dots, e_n)$ une base de E. Soit $(x_1, \dots, x_n)$ une famille d'éléments de E, dont les coordonnées sont :

$$\forall i \in [\![1, n]\!], \quad [x_i]_{\mathcal{B}} = \begin{pmatrix} a_{1,i} \\ \vdots \\ a_{n,i} \end{pmatrix}, \text{ soit: } x_i = \sum_{j=1}^{n} a_{j,i} e_j.$$

On a alors :

$$\det_{\mathcal{B}}(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} = \sum_{\tau \in \mathfrak{S}_n} \varepsilon(\tau) a_{1,\tau(1)} \cdots a_{n,\tau(n)}.$$

Nous rappelons que nous avons obtenu :

Corollaire 8.2.25 (Formes n-linéaires alternées)

Soit E de dimension n, et $\mathcal{B}$ une base de E. Alors l'ensemble des formes n-linéaires alternées est $\mathrm{Vect}(\det_{\mathcal{B}})$.

Ainsi, changer de base ne nous fait pas sortir de la droite $\mathrm{Vect}(\det_{\mathcal{B}})$. Comme par ailleurs, une autre base $\mathcal{B}'$, définira également une droite $\mathrm{Vect}(\det_{\mathcal{B}'})$, aussi égale à l'ensemble des formes n-linéaires alternées, on peut donc affirmer que $\det_{\mathcal{B}}$ et $\det_{\mathcal{B}'}$ diffèrent d'une constante multiplicative. En évaluant en $\mathcal{B}'$, on obtient alors :

Proposition 8.2.26 (Effet d'un changement de base sur le déterminant)

Soit $\mathcal{B}$ et $\mathcal{B}'$ deux bases de E. On a alors :

$$\det_{\mathcal{B}} = \det_{\mathcal{B}}(\mathcal{B}') \times \det_{\mathcal{B}'}.$$

Remarque 8.2.27

En échangeant le rôle de $\mathcal{B}$ et $\mathcal{B}'$, ou en évaluant en $\mathcal{B}$, on obtient la relation suivante :

$$\det_{\mathcal{B}}(\mathcal{B}') = \frac{1}{\det_{\mathcal{B}'}(\mathcal{B})}.$$

Les deux formes $\det_{\mathcal{B}}$ et $\det_{\mathcal{B}'}$ étant non nulles, et ayant déjà vu l'image par une forme alternée d'une famille liée, on obtient la caractérisation suivante :

Proposition 8.2.28 (Caractérisation des bases)

Soit E de dimension fini, muni d'une base $\mathcal{B}$ Une famille $\mathcal{B}'$ est une base de E si et seulement si $\det_{\mathcal{B}}(\mathcal{B}') \neq 0$.

Nous verrons un peu plus loin que cette caratérisation équivaut à l'inversibilité de la matrice de la famille $\mathcal{B}'$ dans la base $\mathcal{B}$, cette matrice étant alors la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{B}'$.

II.4 Orientation d'un espace

Dans ce paragraphe, on suppose que $\mathbb{K} = \mathbb{R}$.

Suivant l'ordre dans lequel on donne les vecteurs d'une famille, un hypervolume, défini par le déterminant, peut être positif ou négatif. Ainsi, étant donnée une base $\mathcal{B} = (b_1, b_2)$ de $\mathbb{R}^2$, $\det_{\mathcal{B}}(b_2, b_1) = -1$. En définissant $\mathcal{B}' = (b_2, b_1)$, on obtient alors $\det_{\mathcal{B}'} = -\det_{\mathcal{B}}$. Ainsi, toutes les aires vont avoir un signe inversé par rapport à $\mathcal{B}'$, comparées aux aires par rapport à $\mathcal{B}$. C'est cette propriété des signes qui va définir la notion d'orientation de l'espace. Grossièrement, deux choix de bases donnant des déterminants différant uniquement d'une constante, soit les déterminants associés à ces deux bases prennent des valeurs qui sont toujours de même signe, soit ils prennent systématiquement des valeurs de signe opposé. On va donc pouvoir classer de la sorte les bases en deux groupes. On dira que ces deux groupes définissent deux orientations différentes de E, et que le choix d'un de ces deux groupes (ou d'un représentant, c'est-à-dire d'une base) définit une orientation de E. Parler d'orientation directe ou indirecte relève alors de la convention.
Nous explicitons un peu l'argument ci-dessus.
Soit E un $\mathbb{R}$-ev de dimension finie. Soit $\mathcal{B}(E)$ l'ensemble des bases de E, et soit $\mathcal{R}$ la relation définie sur $\mathcal{B}(E)$ par :

$$\forall (\mathcal{B}, \mathcal{B}') \in \mathcal{B}(E)^2, \quad \mathcal{B}\mathcal{R}\mathcal{B}' \iff \det_{\mathcal{B}}(\mathcal{B}') > 0.$$

Proposition 8.2.29

La relation $\mathcal{R}$ est une relation d'équivalence. Pour cette relation d'équivalence, il existe précisément deux classes d'équivalence.

Définition 8.2.30 (Orientation de E)

Une orientation de E est une classe d'équivalence de E modulo $\mathcal{R}$. Orienter E signifie choisir l'une de ces classes d'équivalence, par exemple par le choix d'un représentant (donc d'une base).

Il y a donc deux orientations de E, donc deux choix possibles d'orientation.

Définition 8.2.31 (Orientation directe, indirecte)

Soit E, muni d'une base de référence $\mathcal{B}$. On dira qu'une orientation définie par $\mathcal{B}'$ est :
- directe (par rapport à $\mathcal{B}$) si $\det_{\mathcal{B}}(\mathcal{B}') > 0$ (donc $\mathcal{B}\mathcal{R}\mathcal{B}'$)
- indirecte sinon.

Remarque 8.2.32

La notion d'orientation directe ou indirecte relève de la convention : elle ne peut se définir dans l'absolue, et nécessite la donnée d'une base de référence. Dans certaines situations, il existe une base de référence naturelle. Ainsi, dans le cas de $\mathbb{R}^n$, la base canonique constituera en général la base directe de référence. Par ailleurs, les espaces vectoriels (y compris $\mathbb{R}^n$) sont définis indépendamment de toute représentation planaire, spatiale ou autre. Ainsi, définir une orientation directe par certaines caractéristiques d'une représentation planaire ou spatiale relève également de la convention, mais s'avère bien pratique, en physique notamment. Remarquez qu'une telle convention pour $\mathbb{R}^2$ est inversée si on regarde le plan par l'autre coté ! De même pour les conventions dans $\mathbb{R}^3$, si on plonge l'espace $\mathbb{R}^3$ dans $\mathbb{R}^4$, et qu'on le regarde par deux côtés différents.

Proposition 8.2.33 (Effet d'une permutation des vecteurs sur l'orientation)

Soit $\mathcal{B} = (b_1, \dots, b_n)$ une base de E, et $\sigma \in \mathfrak{S}_n$. Soit $\mathcal{B}' = (b_{\sigma(1)}, \dots, b_{\sigma(n)})$ la base obtenue de $\mathcal{B}$ par permutation de ses éléments. Alors $\mathcal{B}$ et $\mathcal{B}'$ définissent la même orientation si et seulement si $\varepsilon(\sigma) = 1$, donc si $\sigma \in \mathfrak{A}_n$.

En particulier, l'échange de deux vecteurs modifie l'orientation, puisqu'une transposition est impaire.

II.5 Déterminant d'un endomorphisme

On aimerait définir la notion de déterminant d'un endomorphisme indépendamment du choix d'une base. L'idée naturelle qui vient à l'esprit pourrait être se se fixer une base $\mathcal{B} = (b_1, \dots, b_n)$ et de définir le déterminant d'un endomorphisme u comme le déterminant de $(u(b_1), \dots, u(b_n))$ par rapport à $\mathcal{B}$. Cette définition est correcte (elle sera donnée en propriété), mais a l'inconvénient d'introduire une base $\mathcal{B}$, dont nous aimerions nous affranchir. Prenant une autre base, par exemple $(\lambda b_1, \dots, \lambda b_n)$, nous allons alors multiplier le déterminant de la famille par λ^n, mais également le déterminant de la base (toujours relativement à la base initiale). Pour s'affranchir de la notion de base, on peut constater que cela reste vrai pour toute forme n-linéaire alternée (puisqu'elles différent du détemrinant d'une constante multiplicative seulement). Cela nous motive la définition suivante :

Lemme 8.2.34

Soit $u \in \mathcal{L}(E)$. Soit, pour toute forme n-linéaire alternée φ sur E, φ_u définie par

$$\varphi_u(x_1, \dots, x_n) = \varphi(u(x_1), \dots, u(x_n)).$$

Alors φ_u est une forme n-linéaire alternée.

Proposition/Définition 8.2.35 (Déterminant d'un endomorphisme)

Avec les notations du lemme précédent, il existe un unique scalaire $\det(u)$ tel que pour toute forme n-linéaire alternée φ, on ait :

$$\varphi_u = \det(u) \cdot \varphi.$$

Ce scalaire $\det(u)$ est appelé déterminant de u.

On a alors la propriété attendue :

Proposition 8.2.36 (Caractérisation du déterminant par l'image d'une base)

Soit $\mathcal{B} = (b_1, \dots, b_n)$ une base de E. Alors

$$\det(u) = \det_{\mathcal{B}}(u(b_1), \dots, u(b_n)).$$

Corollaire 8.2.37 (Déterminant de l'identité)

On a $\det(\mathrm{id}) = 1$.

Proposition 8.2.38

Soit n la dimension de E. Soit $u \in \mathcal{L}(E)$ et $\lambda \in \mathbb{K}$. On a $\det(\lambda u) = \lambda^n \det(u)$.

Théorème 8.2.39 (Déterminant d'une composée)

Soit u et v dans $\mathcal{L}(E)$. Alors $\det(v \circ u) = \det(v)\det(u)$.

Enfin, notre but étant de caractériser les matrices inversibles (donc les automorphismes), on obtient :

Théorème 8.2.40 (Caractérisation des automophismes par le déterminant)

Soit $u \in \mathcal{L}(E)$. Alors u est un automorphime de E si et seulement si $\det(u) \neq 0$.

II.6 Déterminant d'une matrice carrée

Comme souvent, passer des endomorphismes aux matrices est assez automatique, en utilisant la correspondance usuelle entre des endomorphismes et leur matrice dans un choix de base. La seule difficulté peut résider dans l'indépendance vis-à-vis du choix effectué des bases.

Définition 8.2.41 (Déterminant d'une matrice)
Soit $A \in \mathcal{M}_n(\mathbb{K})$, et $f \in \mathcal{L}(\mathbb{K}^n)$ canonique associé. Le déterminant de M, noté $\det(M)$, est par définition égal au déterminant de f : $\det(M) = \det(f)$.

On obtient alors directement des définitions :

Proposition 8.2.42 (Caractérisation du déterminant de M par les colonnes)
Soit M une matrice de $\mathcal{M}_n(\mathbb{K})$, de colonnes $C_1, \dots, C_n$. Alors

$$\det(M) = \det_{b.c.}(C_1, \dots, C_n).$$

Les résultats obtenus pour les endomorphismes se transfèrent alors de façon immédiate aux matrices :

Corollaire 8.2.43 (Déterminant de I_n)
On a $\det(I_n) = 1$.

Proposition 8.2.44 (Effet de la multiplication par un scalaire)
Soit $A \in \mathcal{M}_n(\mathbb{K})$ et $\lambda \in \mathbb{K}$. On a $\det(\lambda A) = \lambda^n \det(A)$.

Théorème 8.2.45 (Déterminant d'un produit)
Soit A et B dans $\mathcal{M}_n(\mathbb{K})$. Alors $\det(AB) = \det(A)\det(B)$.

Et pour terminer, la caractérisation annoncée depuis le début :

Théorème 8.2.46 (Caractérisation des matrices inversibles)
Soit $A \in \mathcal{M}_n(\mathbb{K})$. Alors $A \in \mathrm{GL}_n(\mathbb{K})$ si et seulement si $\det(A) \neq 0$. *Si cette condition est satisfaite, on a alors :*

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Ces règles nous permettent d'établir le fait que le déterminant est un invariant de similitude :

Théorème 8.2.47 (Cohérence relative au choix des bases)
Soit $f \in \mathcal{L}(E)$. Alors pour toute base $\mathcal{B}$ de E,

$$\det(f) = \det(\mathrm{Mat}_{\mathcal{B}}(f)).$$

Pour des matrices décrites par la donnée tabulaire de leurs coefficients, on utilise souvent la notation suivante :

Notation 8.2.48 (Notation tabulaire du déterminant)
Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$. On note

$$\det(A) = \begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix}.$$

Nous transcrivons aussi la description initiale que nous avions donnée pour les familles de vecteurs grâce aux permutations :

Théorème 8.2.49 (Expression du déterminant par les coefficients)
Soit $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$. Alors

$$\det(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} = \sum_{\tau \in \mathfrak{S}_n} \varepsilon(\tau) a_{1,\tau(1)} \cdots a_{n,\tau(n)}.$$

Corollaire 8.2.50 (Expression des déterminant 2×2 et 3×3)

1. *On a :* $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$
2. *(Règle de Sarrus) On a :*

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - gec - hfa - idb,$$

c'est à dire « diagonales descendantes moins diagonales montantes ».

Avertissement 8.2.51
Prenez garde à ne pas généraliser trop vite la règle de Sarrus à des matrices d'ordre supérieur ! Pour $n \geqslant 4$, une règle aussi simpliste est fausse ! Il faut considérer toutes les permutations possibles !

La description explicite par les coefficients permet également d'établir de façon quasi-immédiate :

Théorème 8.2.52 (Invariance du déterminant par transposée)
Soit $A \in \mathcal{M}_n(\mathbb{R})$. On a alors $\det({}^tA) = \det(A)$.

III Calcul des déterminants

Nous abordons maintenant les techniques calculatoires. Nous verrons esentiellement quatre techniques, la première étant une adaptation de la méthode du pivot, la seconde étant une façon de se ramener à des déterminants plus petits lorsque la matrice est triangulaire par blocs, l'importante troisième méthode est le développement suivant une ligne, ramenant le calcul d'un déterminant d'ordre n au calcul de n déterminants d'ordre $n-1$, coefficientiés par les coefficients de la ligne (intéressant surtout lorsque la plupart de ces coefficients sont nuls), et la dernière est l'utilisation du caractère polynomial du déterminant en chacune des coordonnées de la matrice.

III.1 Opérations sur les lignes et colonnes

Proposition 8.3.1 (Déterminant d'une matrice triangulaire)

Le déterminant d'une matrice triangulaire est égal au produit de ses coefficients diagonaux :

$$\begin{vmatrix} \lambda_1 & \bullet & \cdots & \bullet \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bullet \\ 0 & \cdots & 0 & \lambda_n \end{vmatrix} = \lambda_1\lambda_2\cdots\lambda_n.$$

Lemme 8.3.2 (Déterminant des matrices de codage des opérations)

On a :

$$\det(E(i,j)) = -1, \qquad \det(E_i(\lambda)) = \lambda \qquad \det(E_{i,j}(\lambda)) = 1.$$

Corollaire 8.3.3 (Effet des opérations élémentaires sur le déterminant)

(i) Échanger les lignes i et j change le signe du déterminant ;

(ii) Multiplier une ligne par un scalaire λ multiplie le déterminant par ce scalaire

(iii) Faire une combinaison $L_i \leftarrow L_i - \lambda L_j$ ne modifie pas le déterminant (attention à ne pas mettre de coefficient devant L_i).

(iv) De même pour les opérations sur les colonnes.

Méthode 8.3.4 (Calcul du déterminant par la méthode du pivot)

Ainsi, on peut adapter la méthode du pivot pour le calcul du déterminant :

- Comme usuellement, faire des opérations pour échelonner la matrice, en écrivant des égalités entre les déterminants des différentes matrices, mais :
  - Changer le signe lorsqu'on fait un échange de lignes
  - Compenser en divisant à l'extérieur par λ lorsqu'on multiplie une ligne par λ.
- Une fois ramené à une matrice triangulaire, on calcule son déterminant en effectuant le produit des coefficients diagonaux.

Exemple 8.3.5

1. Déterminant de $\begin{pmatrix} 1 & -2 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 2 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$
2. Déterminant de $\begin{pmatrix} a & 1 & \cdots & 1 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \cdots & 1 & a \end{pmatrix}$

III.2 Calcul par blocs

Dans certaines configurations, on peut se ramener à des déterminants plus petits :

Proposition 8.3.6 (Déterminant d'une matrice triangulaire par blocs)

Soit T une matrice triangulaire par blocs, c'est-à-direde la forme

$$T = \begin{pmatrix} A_1 & \bullet & \cdots & \bullet \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bullet \\ 0 & \cdots & 0 & A_k \end{pmatrix}$$

où les A_k sont des matrices carrées. Alors

$$\det(T) = \det(A_1) \cdots \det(A_k).$$

Exemples 8.3.7

1. Déterminant de $\begin{pmatrix} 1 & 2 & 3 & 5 \\ 2 & 2 & 2 & 8 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & -1 & 2 \end{pmatrix}$
2. Déterminant de $\begin{pmatrix} A_1 & \bullet & \cdots & \bullet \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \bullet \\ 0 & \cdots & 0 & A_k \end{pmatrix}$, où $A_i = \begin{pmatrix} \cos(\theta_i) & -\sin(\theta_i) \\ \sin(\theta_i) & \cos(\theta_i) \end{pmatrix}$
3. On peut combiner avec les opérations du pivot, pour éviter la dernière opération (ou les deux dernières), en se ramenant ainsi au calcul d'un déterminant d'ordre 2 ou 3 (par Sarrus).

III.3 Développements suivant une ligne ou une colonne

Sans doute une des techniques les plus importantes, lorsqu'une ligne ou une colonne contient beaucoup de 0. Elle est efficace lorsqu'il y a beaucoup de symétries dans l'expression, pour construire une récurrence, et si ce n'est pas possible, elle peut être combinée à d'autres méthodes pour les calculs des déterminants plus petits obtenus.
Pour exprimer la formule, nous définissons :

Définition 8.3.8 (Mineurs, cofacteurs, comatrice)

Soit $M \in \mathcal{M}_n(\mathbb{K})$.

- Le mineur de position (i,j) de M est le déterminant de la matrice $\tilde{M}_{i,j}$ obtenue en supprimant de M la i-ième ligne et la j-ième colonne. On note $\Delta_{i,j}(M)$ ce mineur
- Le cofacteur de position (i,j) de M est le scalaire $(-1)^{i+j}\Delta_{i,j}(M)$.
- La comatrice de M est la matrice $\mathrm{Com}(M) = ((-1)^{i+j}\Delta_{i,j}(M))_{1 \leqslant i,j \leqslant n}$, c'est à dire la matrice des cofacteurs de M.

Nous obtenons alors la formule de développement :

Théorème 8.3.9 (Développement suivant une ligne)

Soit $M = (m_{i,j}) \in \mathcal{M}_n(\mathbb{K})$, et $i \in [\![1,n]\!]$. Alors

$$\det(M) = \sum_{j=1}^{n} (-1)^{i+j} m_{i,j} \Delta_{i,j}.$$

Théorème 8.3.10 (Développement suivant une colonne)
Soit $M = (m_{i,j}) \in \mathcal{M}_n(\mathbb{K})$, et $j \in [\![1, n]\!]$. Alors

$$\det(M) = \sum_{i=1}^{n} (-1)^{i+j} m_{i,j} \Delta_{i,j}.$$

Corollaire 8.3.11 (Caractérisation de l'inversibilité par la comatrice)
Soit $M \in \mathcal{M}_n(\mathbb{K})$. Alors

$$M\ {}^t\mathrm{Com}(M) = {}^t\mathrm{Com}(M) M = \det(M) I_n.$$

En particulier, M est inversible si et seulement si $\det(M) \neq 0$, et dans ce cas,

$$M^{-1} = \frac{1}{\det(M)} \mathrm{Com}(M).$$

Si la formule elle-même s'avère assez inefficace pour $n > 2$, la caractérisation de l'inversibilité par la non nullité du déterminant est d'une très grande utilité.

Exemples 8.3.12

1. Déterminant de la matrice tridiagonale $\begin{pmatrix} a+b & a & 0 & \cdots & 0 \\ b & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & a \\ 0 & \cdots & 0 & b & a+b \end{pmatrix}$.
2. On peut combiner pivot de Gauss et développement suivant les lignes ou colonnes, en annulant d'abord un grand nombre de termes par les opérations élémentaires avant de développer. Nous illustrons ceci sur le calcul du déterminant de Vandermonde :

$$V(x_1, \ldots, x_n) = \begin{vmatrix} 1 & 1 & \cdots & 1 & 1 \\ x_1 & x_2 & \cdots & x_{n-1} & x_n \\ x_1^2 & x_2^2 & \cdots & x_{n-1}^2 & x_n^2 \\ \vdots & \vdots & & \vdots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_{n-1}^{n-1} & x_n^{n-1} \end{vmatrix}.$$

Cet exemple est un peu plus qu'un exemple, nous le consignons dans la proposition suivante :

Proposition 8.3.13 (Déterminant de Vandermonde)
Le déterminant de Vandermonde, défini dans l'exemple ci-dessus, est donné par :

$$V(x_1, \ldots, x_n) = \prod_{1 \leqslant i < j \leqslant n} (a_j - a_i).$$

Nous voyons dans la section suivante une autre façon de calculer ce déterminant de Vandermonde.

III.4 Caractère polynomial du déterminant

La formule explicite du déterminant nous permet d'affirmer :

Proposition 8.3.14 (« Polynomialitude » du déterminant)

L'application $M \mapsto \det(M)$ est polynomiale en chacune des coordonnées de M. Il s'agit plus précisément d'une fonction polynomiale en les n^2 variables définies par les coordonnées de M, vérifiant :

- *il est homogène de degré total n,*
- *les degrés partiels par rapport à chacune des variables $m_{i,j}$ sont* 1.
- *Les degrés partiels par rapport à l'ensemble des variables constituant une même ligne ou une même colonne sont* 1

Exemple 8.3.15

Calcul de $V_n(x_1, \ldots, x_n)$.

9 Espaces préhilbertiens réels

Un espace préhilbertien réel est un espace vectoriel muni d'un produit scalaire. Cet outil permet de parler d'orthogonalité, et donc d'introduire un certain nombre de concepts permettant de généraliser la géométrie euclidienne du plan à des situations plus abstraites.
Pour commencer nous introduisons donc cette notion abstraite de produit scalaire, ce qui nous oblige à faire quelques rappels sur les formes bilinéaires, déjà rencontrées dans le chapitre précédent.

Tous les espaces vectoriels considérés dans ce chapitre sont des espaces vectoriels sur $\mathbb{R}$.

I Produits scalaires

I.1 Formes bilinéaires

Les résultats de cette sous-section sont valables dans un cadre plus général que les espaces sur $\mathbb{R}$, mais toute la suite du chapitre nécessitant de travailler avec des espaces vectoriels sur $\mathbb{R}$, nous nous limitons également à ce cadre dans les rappels qui suivent.

On rappelle la définition d'une forme bilinéaire :

Définition 9.1.1 (Forme bilinéaire)

Soit E un espace vectoriel sur $\mathbb{R}$. Une forme bilinéaire φ sur E est une application $\varphi : E \times E \to \mathbb{R}$, linéaire par rapport à chaque facteur, l'autre étant fixé, c'est-à-dire :

(i) $\forall (x, y, z) \in E^3, \quad \forall \lambda \in \mathbb{K}, \quad \varphi(\lambda x + y, z) = \lambda \varphi(x, z) + \varphi(y, z)$

(ii) $\forall (x, y, z) \in E^3, \quad \forall \lambda \in \mathbb{K}, \quad \varphi(x, \lambda y + z) = \lambda \varphi(x, y) + \varphi(x, z).$

On rappelle la propriété de bilinéarité généralisée

Lemme 9.1.2 (Bilinéarité généralisée)

Soit φ une forme bilinéaire sur E. Soit $(k, \ell) \in (\mathbb{N}^*)^2$, et soit $(x_1, \dots, x_k, y_1, \dots, y_\ell) \in E^{k+\ell}$, et $(\lambda_1, \dots, \lambda_k, \mu_1, \dots, \mu_\ell) \in \mathbb{K}^{k+\ell}$. Alors

$$\varphi\left(\sum_{i=1}^{k} \lambda_i x_i, \sum_{j=1}^{\ell} \mu_j y_j\right) = \sum_{i=1}^{k} \sum_{j=1}^{\ell} \lambda_i \mu_j \varphi(x_i, y_j).$$

Définition 9.1.3 (Ensemble des formes bilinéaires)
On note $\mathcal{B}(E)$ l'ensemble des formes bilinéaires de E

Nous avons déjà vu le résultat suivant :

Proposition 9.1.4 (Structure de $\mathcal{B}(E)$)
L'ensemble $\mathcal{B}(E)$ est un espace vectoriel sur $\mathbb{R}$.

Voici des exemples particulièrement importants :

Exemples 9.1.5 (Formes bilinéaires)

1. cov : $(X,Y) \mapsto \text{cov}(X,Y)$ sur le $\mathbb{R}$-espace vectoriel $\mathcal{V}$ des variables aléatoires réelles discrètes admettant une variance.
2. $\varphi : (P,Q) \mapsto \displaystyle\int_a^b P(x)Q(x)\,\mathrm{d}x$ sur $\mathbb{R}[X]$, ou sur $\mathcal{C}^0([a,b])$.
3. $\varphi : \left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) \mapsto \displaystyle\sum_{i=1}^n x_i y_i$ sur $\mathbb{R}^n$
4. $p_1 : (x,y) \mapsto x$ sur $\mathbb{R}$
5. $\varphi : (X,Y) \mapsto {}^tXAY$, sur $\mathbb{R}^n$, $A \in \mathcal{M}_n(\mathbb{R})$.

Ce dernier exemple est très important, car si E est de dimension finie, toute forme bilinéaire peut être représenté de cette forme après choix d'une base de E. C'est ce que nous étudions maintenant.
Nous supposons jusqu'à la fin de ce paragraphe que E est de dimension finie n.

Définition 9.1.6 (Matrice associée à une forme bilinaire)
Soit φ une forme bilinéaire sur E, et $\mathcal{B} = (e_1, \ldots, e_n)$ une base de E. Alors on définit la matrice de B relativement à la base $\mathcal{B}$ par :

$$\text{Mat}_{\mathcal{B}}(\varphi) = (\varphi(e_i, e_j))_{1 \leqslant i,j \leqslant n}.$$

Exemples 9.1.7

1. Matrice des variances-covariances si $X_1, \ldots, X_n$ sont linéairement indépendants.
2. Matrice dans la base canonique de $\varphi : (P,Q) \mapsto \displaystyle\int_0^1 P(x)Q(x)\,\mathrm{d}x$ sur $\mathbb{R}_n[X]$.
3. Matrice du produit scalaire canonique de $\mathbb{R}^n$ dans la base canonique ; dans la base $\mathcal{B} = (b_1, \ldots, b_n)$, où b_i est constitué de 1 sur les coordonnées 1 à i, et de 0 ailleurs.

Théorème 9.1.8 (Expression matricielle de $\varphi(x,y)$)
Soit φ une forme bilinéaire sur E, et $\mathcal{B}$ une base de E. Alors, pour tout $(x,y) \in E^2$,

$$\varphi(x,y) = {}^t[x]_{\mathcal{B}}\text{Mat}_{\mathcal{B}}(\varphi)[y]_{\mathcal{B}} = {}^tXMY,$$

où X et Y sont les vecteurs colonnes représentant les coordonnées de x et y dans la base $\mathcal{B}$, et M est la matrice de φ relativement à cette même base $\mathcal{B}$.

Une base $\mathcal{B}$ étant fixée, cette relation caractérise d'ailleurs la matrice $\mathcal{M}_{\mathcal{B}}(\varphi)$:

Théorème 9.1.9 (Caractérisation de $\mathrm{Mat}_{\mathcal{B}}(\varphi)$ par la relation $\varphi(X,Y) = {}^tXMY$)

Soit $M \in \mathcal{M}_n(\mathbb{K})$. La relation $\varphi(x,y) = {}^tXMY$ caractérise la matrice de φ relativement à la base $\mathcal{B}$: si

$$\forall (x,y) \in E^2, \quad \varphi(x,y) = {}^t[x]_{\mathcal{B}} M [y]_{\mathcal{B}},$$

alors $M = \mathrm{Mat}_{\mathcal{B}}(\varphi)$.

Exemple 9.1.10

L'égalité $\mathcal{M}_{bc}(\varphi) = I_n$ se traduit par $\varphi(X,Y) = {}^tXI_nY = {}^tXY$; c'est le produit scalaire usuel.

Corollaire 9.1.11

L'application

$$\begin{array}{rccc} \Phi : & \mathcal{B}(E) & \longrightarrow & \mathcal{M}_n(\mathbb{R}) \\ & \varphi & \longmapsto & \mathrm{Mat}_{\mathcal{B}}(\varphi) \end{array}$$

est un isomorphisme.

Corollaire 9.1.12 (dimension de $\mathcal{B}(E)$)

Si E est de dimension finie n, alors $\mathcal{B}(E)$ est de dimension finie, et

$$\dim(\mathcal{B}(E)) = n^2.$$

Comme pour les applications linéaires, les changements de base s'expriment facilement par des opérations matricielles.

Théorème 9.1.13 (Formule de changement de base pour les formes bilinéaires)

Soit E un espace vectoriel sur $\mathbb{R}$ de dimension finie. Soit φ une forme bilinéaire sur E, et soit $\mathcal{C}$ et $\mathcal{D}$ deux bases de E. Soit P la matrice de passage de $\mathcal{C}$ à $\mathcal{D}$. Alors

$$\mathrm{Mat}_{\mathcal{D}}(\varphi) = {}^tP\mathrm{Mat}_{\mathcal{C}}(\varphi)P.$$

Exemple 9.1.14

Explicitation du changement de la base canonique à la base $\mathcal{B}$ dans l'exemple 9.1.7, pour le produit scalaire canonique.

II Produits scalaire

Pour définir la notion de produit scalaire, on introduit trois propriétés pouvant être vérifiées par une forme bilinéaire. La conjonction de ces trois propriétés définiront un produit scalaire.

II.1 Formes bilinéaires symétriques, définies, positives

Définition 9.2.1 (Symétrie, positivité, caractère défini)

Soit φ une forme bilinéaire sur un espace vectoriel E.

1. On dit que φ est symétrique si : $\forall (x,y) \in E^2, \quad \varphi(x,y) = \varphi(y,x)$.

2. On dit que φ est positive si : $\forall x \in E, \quad \varphi(x,x) \geqslant 0$.
3. On dit que φ est définie si : $\forall x \in E, \quad \varphi(x,x) = 0 \iff x = 0$.

Proposition 9.2.2 (positivité des formes définies)
Soit φ une forme définie. Alors φ est soit positive soit négative.

Pour cette raison, on parle assez rarement de forme définie, cette propriété étant indissociable d'une propriété de positivité ou de négativité. On parlera alors de forme définie positive, ou définie négative.

Proposition 9.2.3 (Caractérisation matricielle de la symétrie)
Soit φ une forme bilinéaire sur un espace vectoriel E de dimension finie. Les propriétés suivantes sont équivalentes :

(i) φ est symétrique ;

(ii) Il existe une base de E telle que $\mathrm{Mat}_{\mathcal{B}}(\varphi)$ est une matrice symétrique

(iii) Pour tout base $\mathcal{B}$ de E, $\mathrm{Mat}_{\mathcal{B}}(\varphi)$ est une matrice symétrique.

Exemples 9.2.4

1. $\varphi : (x,y) \mapsto xy$ sur $\mathbb{R}$ est symétrique, définie, positive.
2. cov est symétrique positive, mais pas définie.
3. Le produit scalaire canonique sur $\mathbb{R}^n$ est symétrique, défini, positif.
4. $\varphi : (f,g) \mapsto \displaystyle\int_a^b f(t)g(t)\,\mathrm{d}t$ sur $\mathcal{C}^0([a,b])$ est symétrique, définie et positive.

Voici enfin un résultat d'une importance capitale, que nous avons déjà rencontré en cours d'année dans différents contextes, qui sont tous des cas particuliers du contexte général exposé ici.

Proposition 9.2.5 (Inégalité de Cauchy-Schwarz)
Soit φ une forme bilinéaire symétrique positive sur un espace E. Alors, pour tout $(x,y) \in \mathrm{E}^2$, on a

$$\varphi(x,y)^2 \leqslant \varphi(x,x) \cdot \varphi(y,y),$$

et on a égalité si et seulement si il existe λ tel que $\varphi(x+\lambda y, x+\lambda y) = 0$.

Exemple 9.2.6
Pour toutes variables aléatoires X et Y admettant une variance, on a :

$$|\mathrm{cov}(X,Y)| \leqslant \sigma(X)\sigma(Y).$$

Nous verrons d'autres exemples plus loin, lorsque nous aurons réexprimé plus simplement le cas d'égalité lorsque φ est un produit scalaire.

II.2 Produits scalaires

Dans cette section, E désigne un espace vectoriel sur $\mathbb{R}$.

Définition 9.2.7 (Produit scalaire)

Un produit scalaire sur E est une forme bilinéaire symétrique, définie et positive.

Remarque 9.2.8

La symétrie et la linéarité par rapport à la première variable entraîne la linéarité par rapport à la seconde variable.

Voici les exemples usuels (les 2 premiers), desquels on peut dériver d'autres exemples.

Exemples 9.2.9 (produits scalaires)

1. $(f,g) \mapsto \displaystyle\int_a^b f(t)g(t)\,\mathrm{d}t$ est un produit scalaire sur $\mathcal{C}^0(\mathbb{R})$ (ou sur $\mathbb{R}[X]$)
2. Le produit scalaire canonique de $\mathbb{R}^n$ est un produit scalaire au sens de cette définition.
3. Un autre produit scalaire sur $\mathbb{R}^n$: $(X,Y) \mapsto {}^tXMY$, où $M = (\min(i,j))_{1\leqslant i,j\leqslant n}$.

Ainsi, on n'a pas unicité d'un produit scalaire sur un espace vectoriel E.

Notation 9.2.10 (Notations fréquentes pour le produit scalaire)

Soit φ un produit scalaire sur E on note souvent :

$$\varphi(x,y) = \langle x,y\rangle \qquad \text{ou} \qquad \varphi(x,y) = (x|y).$$

Soit φ un produit scalaire, noté $\varphi(x,y) = \langle x,y\rangle$. Alors en particulier, φ est une forme bilinéaire, et si $\mathcal{B} = (e_1,\dots,e_n)$ est une base de E :

$$\mathrm{Mat}_{\mathcal{B}}(\varphi) = (\langle e_i,e_j\rangle)_{1\leqslant i,j\leqslant n}$$

Proposition 9.2.11

La matrice d'un produit scalaire dans une base quelconque est symétrique. La réciproque est fausse.

II.3 Normes euclidiennes

Définition 9.2.12 (Norme)

Une norme sur un espace vectoriel E est une application $N : E \to \mathbb{R}$ telle que :

(i) $\forall x \in E,\quad N(x) = 0 \iff x = 0$

(ii) $\forall \lambda \in \mathbb{R},\quad \forall x \in E,\quad N(\lambda x) = |\lambda| N(x)$

(iii) $\forall (x,y) \in E^2,\quad N(x+y) \leqslant N(x) + N(y)$.

Proposition 9.2.13 (Positivité des normes)

Si N est une norme sur E, alors pour tout $x \in E$, $N(x) \geqslant 0$.

Nous allons maintenant définir une norme associée à tout produit scalaire.

Notation 9.2.14 (norme euclidienne associée à un produit scalaire)
Soit $\langle -,- \rangle$ un produit scalaire sur E. Pour tout $x \in E$, on note $\|x\| = \sqrt{\langle x, x \rangle}$.

La fin de ce paragraphe a pour but de justifier que ceci définit bien une norme. Pour commencer, nous obtenons :

Proposition 9.2.15
L'application $x \mapsto \|x\|$ vérifie les points (i) et (ii) de la définition d'une norme.

Il nous reste donc à voir l'inégalité triangulaire, un pei plus délicate.

Lemme 9.2.16 (Formule de polarisation)
Soit $\langle -,- \rangle$ un produit scalaire sur E. Alors, pour tout $(x, y) \in E^2$,

$$\|x+y\|^2 = \|x\|^2 + 2\langle x, y \rangle + \|y\|^2, \qquad \textit{soit:} \qquad \langle x, y \rangle = \frac{1}{2}\left(\|x+y\|^2 - \|x\|^2 - \|y\|^2\right).$$

Cette formule affirme notamment que le produit scalaire est entièrement déterminé par la donnée de la norme euclidienne.
Nous avons cette formule dans le contexte d'une forme non définie (la covariance). Cette formule de polarisation se généralise, comme dans ce contexte particulier, en la formule suivante (à comparer au développement d'un carré dans $\mathbb{R}$) :

Proposition 9.2.17 (carré de la norme d'une somme)
Plus généralement, étant donné $(x_1, \ldots, x_n) \in \mathrm{E}^n$,

$$\|x_1 + \cdots + x_n\|^2 = \sum_{i=1}^{n} \|x_i\|^2 + 2 \sum_{1 \leqslant i < j \leqslant n} \langle x_i, x_j \rangle .$$

Nous pouvons maintenant donner une version améliorée de l'inégalité de Cauchy-Schwarz (en particulier pour l'expression du cas d'égalité)

Théorème 9.2.18 (Inégalité de Cauchy-Schwarz pour les produits scalaires)
Soit $\langle -,- \rangle$ un produit scalaire. Alors, pour tout $(x, y) \in \mathrm{E}^2$, on a

$$\left| \langle x, y \rangle \right| \leqslant \|x\| \cdot \|y\|.$$

L'égalité est obtenue si et seulement si x et y sont colinéaires (de même sens pour un signe positif, de sens opposé pour un signe négatif).

Nous en déduisons l'inégalité triangulaire qui nous manquait encore pour pouvoir affirmer que :

Théorème 9.2.19 (Norme euclidienne associée au produit scalaire)
Soit $\langle -,- \rangle$ un produit scalaire sur E. Alors l'application $\| - \|$ qui à x associe $\|x\|$ est une norme sur E. Cette norme est appelée norme euclidienne associée au produit scalaire $\langle -,- \rangle$.

Comme promis plus haut, nous donnons deux inégalités de Cauchy-Schwarz à connaitre impérativement.

Exemples 9.2.20 (Autres exemples d'inégalités de Cauchy-Schwarz)

1. Inégalité de Cauchy-Schwarz numérique :

$$\forall X = (x_1, \dots, x_n),\ Y = (y_1, \dots, y_n) \in \mathbb{R}^n,\quad |\langle X, Y\rangle| \leqslant \|X\| \cdot \|Y\|,$$

à savoir :

$$\left|\sum_{i=1}^{n} x_i y_i\right| \leqslant \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} y_i^2\right)^{\frac{1}{2}},$$

ou encore :

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 \leqslant \left(\sum_{i=1}^{n} x_i^2\right) \left(\sum_{i=1}^{n} y_i^2\right),$$

avec égalité si et seulement si X et Y sont colinéaires.

2. Inégalité de Cauchy-Schwarz intégrale :

$$\forall (f, g) \in \mathcal{C}^0([a, b]),\ \left|\int_1^b f(t)g(t)\ \mathrm{d}t\right| \leqslant \left(\int_a^b f(t)^2\ \mathrm{d}t\right)^{\frac{1}{2}} \left(\int_a^b g(t)^2\ \mathrm{d}t\right)^{\frac{1}{2}},$$

avec égalité si et seulement si $g = 0$, ou s'il existe λ tel que $f = \lambda g$.

Remarques 9.2.21

1. D'après ce qui précède, si $\|-\|$ est une norme euclidienne associée à un certain produit scalaire, ce produit scalaire est nécessairement défini par

$$\forall (x, y) \in E^2,\quad \langle x, y\rangle = \frac{1}{2}\left(\|x + y\|^2 - \|x\|^2 - \|y\|^2\right)$$

Ainsi, une même norme euclidienne ne peut pas être associée à deux produits scalaires différents

2. Toutes les normes ne sont pas des normes euclidiennes. Pour vérifier si une norme est euclidienne ou non, il suffit de vérifier si le seul candidat à être le produit scalaire associé, à savoir :

$$\forall (x, y) \in E^2,\quad \varphi(x, y) = \frac{1}{2}\left(\|x + y\|^2 - \|x\|^2 - \|y\|^2\right),$$

est effectivement un produit scalaire

Exemple 9.2.22

$N : (x_1, \dots, x_n) \mapsto \max(|x_1|, \dots, |x_n|)$ est une norme sur $\mathbb{R}^n$, mais n'est pas une norme euclidienne.

II.4 Espaces préhilbertiens réels, espaces euclidiens

Définition 9.2.23 (Espace préhilbertien réel)

Un espace préhilbertien réel $(E, \langle \bullet, \bullet\rangle)$ est un espace vectoriel E sur $\mathbb{R}$, muni d'un produit scalaire $\langle \bullet, \bullet\rangle$.

S'il n'y a pas d'ambiguïté sur le produit scalaire, on parlera plus simplement de l'espace préhilbertien E, au lieu de $(E, \langle \bullet, \bullet\rangle)$.

Définition 9.2.24 (Espace euclidien)

Un espace euclidien est un espace préhilbertien réel de dimension finie.

Par exemple $\mathbb{R}^n$ muni du produit scalaire canonique est un espace euclidien. $\mathbb{R}_n[X]$ muni d'un produit scalaire intégral est un espace euclidien. En revanche, $\mathbb{R}[X]$ muni du même produit scalaire n'est qu'un espace préhilbertien réel, ainsi que $\mathcal{C}^0([a,b])$ muni du produit scalaire intégral.

III Orthogonalité

Dans toute cette section, on considère $(E, \langle \bullet, \bullet \rangle)$ un espace préhilbertien. Il sera précisé euclidien dans certains cas.

III.1 Vecteurs orthogonaux

Définition 9.3.1 (Vecteurs orthogonaux)

Soit $(x,y) \in E$ deux vecteurs de E. On dit que x et y sont orthogonaux, et on note $x \perp y$, si $\langle x, y \rangle = 0$.

Exemples 9.3.2

1. $\forall x \in E, \quad x \perp 0$.
2. Soit $(e_1, \dots, e_n)$ la base canonique de $\mathbb{R}^n$. Alors pour tout $i \neq j$, $e_i \perp e_j$.
3. $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \perp \begin{pmatrix} -4 \\ -3 \\ 2 \\ 1 \end{pmatrix}$, pour le produit scalaire usuel.
4. $x \mapsto \sin \pi x$ et $x \mapsto \cos \pi x$ sont orthogonaux pour le produit scalaire $\langle f, g \rangle = \displaystyle\int_{-1}^{1} f(t)g(t)\,\mathrm{d}t$ sur $\mathcal{C}^0([-1,1])$.

Définition 9.3.3 (Famille orthogonale, orthonormale)

1. Soit $\mathcal{F} = (x_1, \dots, x_n)$ une famille (finie) d'éléments de E. On dit que la famille $\mathcal{F}$ est orthogonale si et seulement si les x_i sont deux à deux orthogonaux, c'est-à-dire si et seulement si pour tout $(i,j) \in [\![1,n]\!]$ tel que $i \neq j$, $x_i \perp x_j$.
2. On dit que la famille $\mathcal{F}$ est orthonormale (ou orthonormée) si et seulement si elle est orthogonale, et que pour tout $i \in [\![1,n]\!]$, $\|x_i\| = 1$.

Théorème 9.3.4 (Liberté des familles orthogonales)

Soit $\mathcal{F}$ une famille orthogonale ne contenant pas le vecteur nul. Alors $\mathcal{F}$ est une famille libre.

En particulier, toute famille orthonormale est libre.

Corollaire 9.3.5 (base orthonormale)

On suppose E de dimension finie n (i.e. *E euclidien). Soit $(e_1, \dots, e_n)$ une famille orthonormale de E. Alors $(e_1, \dots, e_n)$ est une base de E. On dit qu'il s'agit d'une base orthonormale, et on abrège en b.o.n..*

Un fait qu'il est important de garder en mémoire est la facilité d'expression vectorielle des objets dans une base orthonormale. Nous pouvons donner dès maintenant l'expression d'un vecteur dans une b.o.n., nous verrons un peu plus loin l'expression de la matrice d'un endomorphisme.

Proposition 9.3.6 (Expression des coordonnées d'un vecteur dans une b.o.n.)

Soit E un espace euclidien. Soit $\mathcal{B} = (b_1, \dots, b_n)$ une base orthonormale de E. Alors :

$$\forall X \in E, \quad X = \sum_{i=1}^{n} \langle X, b_i \rangle \, b_i.$$

Exemple 9.3.7

Dans $\mathbb{R}^2$, il s'agit des projections orthogonales sur chaque axe.

Ce théorème n'est vraiment utile que si on sait déterminer des bases orthonormales. On verra un peu plus loin comment construire une b.o.n. à partir de n'importe quelle base (procédé d'orthonormalisation de Gram-Schmidt). En particulier, ce procédé nous assurera de l'existence d'une b.o.n.
On ne suppose plus ici que E est de dimension finie

Théorème 9.3.8 (Théorème de Pythagore)

Soit E un espace préhilbertien réel.

1. *Soit $(x, y) \in E^2$ tel que $x \perp y$. Alors $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.*
2. *Plus généralement, soit $(x_1, \dots, x_n)$ une famille orthogonale de E. Alors*

$$\left\| \sum_{i=1}^{n} x_i \right\|^2 = \sum_{i=1}^{n} \|x_i\|^2.$$

IV Sous-espaces orthogonaux

Définition 9.4.1 (Sous-espaces orthogonaux)

Soit E un espace préhilbertien réel.

1. Soit $x \in E$, et F un sous-espace vectoriel de E. On dit que x est orthogonal à F si pour tout $y \in F$, $x \perp y$. On note $x \perp F$, ou $x \in F^\perp$ comme on le verra plus tard.
2. Soit F et G deux sous-espaces vectoriels de E. On dit que F et G sont orthogonaux si et seulement si :
$$\forall x \in F, \quad \forall y \in G, \quad x \perp y.$$

Exemples 9.4.2

1. Dans $\mathbb{R}^3$, le plan (Oxy) et la droite (Oz)
2. Dans $\mathbb{R}^3$, le plan d'équation $x + y + z = 0$, et la droite engendrée par $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.
3. Dans $\mathcal{C}^0([-1, 1])$ muni du ps usuel, le sous espace P des fonctions paires, et le sous-espace I des fonctions impaires.

Proposition 9.4.3 (orthogonalité et somme directe)

Soit F et G des sous-espaces vectoriels de E. Si $F \perp G$, alors la somme $F + G$ est directe.

Du fait de la bilinéarité du produit scalaire, il n'est pas nécessaire d'établir l'orthogonalité de toutes les paires de vecteurs de E et F pour obtenir l'orthogonalité des deux espaces. Il suffit en fait de verifier cette orthogonalité sur des vecteurs de familles génératrices. Nous établissonspour cela un lemme, résultant immédiatement de la bilinéarité du produit scalaire.

Proposition 9.4.4 (Stabilité par CL de l'orthogonalité)

1. *Soit $(x, y, z) \in E^3$ tel que $x \perp y$ et $x \perp z$, et soit $\lambda \in \mathbb{R}$. Alors $x \perp \lambda y + z$.*
2. *Soit $x \in E$, et $(y_1, \dots, y_n)$ une famille quelconque telle que pour tout $i \in [\![1, n]\!]$, $x \perp y_i$. Soit $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$. Alors : $x \perp \sum_{i=1}^{n} \lambda_i x_i$.*
3. *Soit $(x_1, \dots, x_m) \in E^m$ et $(y_1, \dots, y_n) \in E^n$ deux familles de E telles que pour tout $(i, j) \in [\![1, m]\!] \times [\![1, n]\!]$, $x_i \perp y_j$. Alors pour tout $(\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$ et pour tout $(\mu_1, \dots, \mu_n) \in \mathbb{R}^n$:*

$$\sum_{i=1}^{m} \lambda_i x_i \perp \sum_{j=1}^{m} \mu_j y_j.$$

Proposition 9.4.5 (Caractérisation de l'orthogonalité par les familles génératrices)

Soit $(x_1, \dots, x_m)$ et $(y_1, \dots y_n)$ deux familles de E. Alors

$$\mathrm{Vect}(x_1, \dots, x_m) \perp \mathrm{Vect}(y_1, \dots, y_n) \text{ si et seulement si } \forall i \in [\![1, m]\!], \forall j \in [\![1, n]\!], \quad x_i \perp y_j.$$

Corollaire 9.4.6 (Caractérisation de l'orthogonalité par des bases)

Soit F et G deux sous-espaces vectoriels de E, et soit $\mathcal{B} = (b_1, \dots, b_n)$ et $\mathcal{C} = (c_1, \dots, c_n)$ des bases de F et G respectivement. Alors $F \perp G$ si et seulement si

$$\forall (i, j) \in [\![1, m]\!] \times [\![1, n]\!], \quad b_i \perp c_j.$$

Définition 9.4.7 (orthogonal de F dans E)

Soit F un sev de E. On note $F^\perp = \{x \in E \mid x \perp F\}$, l'ensemble des vecteurs orthogonaux à F. L'ensemble $F^\perp$ est appelé l'orthogonal de F.

Définition 9.4.8 (orthogonal d'une partie)

Plus généralement, X étant une partie de E, $X^\perp$ désigne l'ensemble des vecteurs orthogonaux à tous les vecteurs de X.

Proposition 9.4.9 (Stabilité de l'orthogonal par Vect)

Soit X une partie de E. Alors $X^\perp = \mathrm{Vect}(X)^\perp$.

Proposition 9.4.10

L'ensemble $F^\perp$ est un sous-espace vectoriel de E, et $F \perp F^\perp$, donc en particulier, la somme $F + F^\perp$ est directe.

En particulier, l'orthogonal de toute partie est un sous-espace vectoriel.

Avertissement 9.4.11

Attention, contrairement à l'idée intuitive qu'on se fait en considérant l'orthogonalité dans $\mathbb{R}^3$, $F^\perp$ n'est pas forcément un supplémentaire de F dans E. On montrera que c'est le cas si E est de dimension finie, mais que cette propriété entre en défaut si E est de dimension infinie.

Exemple 9.4.12

Déterminer l'orthogonal dans $\mathbb{R}^4$ muni de la base canonique de

$$F = \text{Vect}\left(\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\\1\end{pmatrix}\right).$$

Proposition 9.4.13 (caractérisation de l'appartenance à $F^\perp$ par une base)

Soit F un sous-espace vectoriel de dimension fini de E, et $\mathcal{B} = (b_1, \dots, b_n)$ une base de F. Alors $x \in F^\perp$ si et seulement si pour tout $i \in [\![1, n]\!]$, $x \perp b_i$.

IV.1 Projeté orthogonal

Définition 9.4.14 (Projeté orthogonal sur une droite)

Soit x et y deux éléments de E. On dit qu'un vecteur $z \in E$ est le projeté orthogonal de y sur la droite $\text{Vect}(x)$ si et seuement si $y - z \perp x$ et $z \in \text{Vect}(x)$.

Proposition 9.4.15 (Existence et expression du projeté orthogonal sur une droite)

Soit x et y deux éléments de E. Alors, le projeté orthogonal de y sur $\text{Vect}(x)$ existe, est unique, et vaut :

$$z = \langle y, x\rangle \cdot \frac{x}{\|x\|^2} = \left\langle y, \frac{x}{\|x\|}\right\rangle \cdot \frac{x}{\|x\|}.$$

Définition 9.4.16

Soit F un sous-espace vectoriel de E, et $y \in E$. On dit que $z \in E$ est le projeté orthogonal de y sur F si et seulement si :

(i) $z \in F$

(ii) $(y - z) \perp F$.

Théorème 9.4.17 (Existence du projeté orthogonal sur un sous-espace de dimension finie)

*Soit $y \in E$, et F un sous-espace vectoriel **de dimension finie** de E, tel qu'il existe une base orthonormale $(b_1, \dots, b_m)$ de F. Alors le projeté orthogonal de y sur F existe, est unique, et vaut :*

$$z = \sum_{i=1}^{m} \langle y, b_i\rangle\, b_i.$$

Remarques 9.4.18

1. On verra dans la suite du cours qu'un espace vectoriel de dimension finie et muni d'un produit scalaire admet toujours au moins une base orthonormale pour ce produit scalaire : l'hypothèses d'existence de cette base est donc superflue dans le théorème ci-dessus.
2. Si $(b_1, \dots, b_n)$ est une base orthogonale, mais non orthonormale, on obtient la formule suivante pour le projeté orthogonale de y sur F :
$$z = \sum_{i=1}^{m} \left\langle y, \frac{b_i}{\|b_i\|} \right\rangle \cdot \frac{b_i}{\|b_i\|}.$$
3. Si $y \in F$, son projeté est bien entendu lui-même, et on obtient, pour une b.o.n. $(b_1, \dots, b_m)$ de F :
$$y = \sum_{i=1}^{m} \langle y, b_i \rangle \, b_i.$$
Ainsi, on retrouve l'expression des coordonnées d'un vecteur y dans une base orthonormale.

IV.2 Orthonormalisation de Schmidt

Motivation : Étant donné une famille libre $(e_1, \dots, e_n)$ de E, trouver un moyen concret de construire une famille libre orthonormée $(f_1, \dots f_n)$ engendrant le même espace que $(e_1, \dots, e_n)$, c'est-à-dire tel que $(f_1, \dots, f_n)$ est une b.o.n. de $\mathrm{Vect}(e_1, \dots, e_n)$.
En particulier, si $(e_1, \dots, e_n)$ est initialement une base de E, on décrit une façon canonique de construire une b.o.n. de E à partir de cette base.
On fait cette construction étape par étape, de manière à avoir, pour tout $k \in [\![1, n]\!]$,

$$\mathrm{Vect}(e_1, \dots, e_k) = \mathrm{Vect}(f_1, \dots, f_k).$$

Théorème 9.4.19 (Procédé d'orthonormalisation de Schmidt)

Soit E un espace préhilbertien réel. Soit $(e_1, \dots, e_n)$ une famille libre de E. Il existe une unique famille orthonormale $(f_1, \dots, f_n)$ telle que, pour tout $k \in [\![1, n]\!]$, on ait

$$\mathrm{Vect}(f_1, \dots, f_k) = \mathrm{Vect}(e_1, \dots, e_k), \qquad \text{et} \qquad \langle e_k, f_k \rangle \geqslant 0.$$

Cette famille peut être construite explicitement par la description par récurrence suivante :

$$f_1 = \frac{e_1}{\|e_1\|} \qquad \text{et} \qquad \forall k \in [\![2, n]\!], \quad f_k = \frac{e_k - \sum_{i=1}^{k-1} \langle e_k, f_i \rangle \, f_i}{\left\| e_k - \sum_{i=1}^{k-1} \langle e_k, f_i \rangle \, f_i \right\|}.$$

Ainsi, pour tout $k \in [\![1, n]\!]$, $(f_1, \dots, f_k)$ est une b.o.n. de $\mathrm{Vect}(e_1, \dots, e_k)$. En particulier,si initialement, $(e_1, \dots, e_n)$ est une base de E, alors $(f_1, \dots, f_n)$ en est une base orthonormale.

Exemples 9.4.20

1. Base orthonormale de $F = \mathrm{Vect}\left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right)$
2. Orthonormalisée de Schmidt de $(1, X)$ dans $\mathbb{R}_n[X]$ muni de :
$$\langle P, Q \rangle = \int_0^1 P(t)Q(t) \, \mathrm{d}t.$$

V Espaces euclidiens

On suppose ici que E est un espace euclidien (donc de dimension finie). On commence par montrer l'existence de bases orthonormales. On en déduit en particulier la possibilité de projeter orthogonalement sur tout sous-espace vectoriel.

V.1 Bases orthonormales d'un espace euclidien

Théorème 9.5.1 (Existence d'une b.o.n. d'un espace euclidien)

Tout espace euclidien E admet au moins une base orthonormale.

Théorème 9.5.2 (théorème de la base orthonormale incomplète)

Soit E un espace euclidien.

1. *Toute famille libre orthogonale de E peut être complétée en une base orthogonale de E*
2. *Toute famille orthonormale de E peut être complétée en une base orthonormale de E.*

Corollaire 9.5.3 (Existence et unicité d'un supplémentaire orthogonal)

Soit E un espace euclidien. Tout sous-espace F de E admet un unique supplémentaire G tel que $F \perp G$. De plus, on a $G = F^\perp$. On dit que $F^\perp$ est le supplémentaire orthogonal de F.

Proposition 9.5.4 (base orthonormale d'une somme orthogonale)

Soit E un espace euclidien, et F et G deux sous-espaces vectoriels de E tels que $F \perp G$. Soit $(b_1, \dots, b_p)$ une base orthonormale de F, et $(c_1, \dots, c_q)$ une base orthonormale de G. Alors $(b_1, \dots, b_p, c_1, \dots, c_q)$ est une base orthonormale de $F \oplus G$; en particulier, c'est une famille orthonormale de E.

Corollaire 9.5.5

Soit E un espace euclidien, et $F_1, \dots, F_p$ des sous-espaces vectoriels de E, deux à deux orthogonaux. Alors la somme $F_1 + \cdots + F_p$ est directe, et on obtient une base orthonormale de $F_1 \oplus \cdots \oplus F_p$ en juxtaposant des bases orthonormales des espaces $F_1, \dots, F_p$.

Corollaire 9.5.6

Soit E un espace euclidien, F un sous-espace vectoriel de E, et $F^\perp$ son supplémentaire orthogonal. Soit $(b_1, \dots, b_p)$ une base orthonormale de F et $(c_1, \dots, c_q)$ une base orthonormale de $F^\perp$. Alors $(b_1, \dots, b_p, c_1, \dots, c_q)$ est une base orthonormale de E.

Comme on l'a évoqué plus haut, trouver les coordonnées d'un vecteur X dans une b.o.n. est simple (si on sait calculer le produit scalaire). Il en est de la norme.

Théorème 9.5.7 (Coordonnées d'un vecteur dans une b.o.n. et norme)

Soit E un espace euclidien, et $\mathcal{B} = (b_1, \dots, b_n)$ une b.o.n. de E. Soit $x \in E$. Alors :

(i) $x = \sum_{i=1}^{n} \langle x, b_i \rangle \, b_i$, *c'est-à-dire* $[x]_\mathcal{B} = \begin{pmatrix} \langle x, b_1 \rangle \\ \vdots \\ \langle x, b_n \rangle \end{pmatrix}$.

(ii) $\|x\|^2 = \sum_{i=1}^{n} \langle x, b_i \rangle^2$

Ainsi, la norme de x est la racine de la somme des carrés des coordonnées de x dans la b.o.n. considérée. Cela correspond bien à la situation connue dans $\mathbb{R}^n$ muni de son produit scalaire canonique.
La matrice d'un endomrophisme admet également une description très simple :

Théorème 9.5.8 (Matrice d'un endomorphisme relativement à une b.o.n.)
Soit E un espace euclidien muni d'une b.o.n. $\mathcal{B} = (b_1, \dots, b_n)$. Soit $u \in \mathcal{L}(E)$. Alors :

$$\mathcal{M}_{\mathcal{B}}(u) = (\langle b_i, u(b_j)\rangle)_{1\leqslant i,j\leqslant n} = \begin{pmatrix} \langle b_1, u(b_1)\rangle & \cdots & \langle b_1, u(b_n)\rangle \\ \vdots & & \vdots \\ \langle b_n, u(b_1)\rangle & \cdots & \langle b_n, u(b_n)\rangle \end{pmatrix}.$$

Enfin, l'expression de la matrice du produit scalaire dans la base $\mathcal{B}$ permet de caractériser facilement les bases orthonormales.

Théorème 9.5.9 (Matrice du produit scalaire relativement à une b.o.n.)
Soit E un espace euclidien, et $\mathcal{B}$ une base de E. Alors la base $\mathcal{B}$ est orthonormale si et seulement si la matrice du produit scalaire dans $\mathcal{B}$ est I_n.
Dans ce cas, pour tout (x, y) dans E^2, on a :

$$\langle x, y\rangle = {}^tXY,$$

où $X = [x]_{\mathcal{B}}$ et $Y = [y]_{\mathcal{B}}$ sont les vecteurs des coordonnées de x et y dans la base $\mathcal{B}$.

V.2 Changements de base et matrices orthogonales

Les matrices de passage d'une b.o.n. à une autre vérifient une propriété trsè forte :

Théorème 9.5.10 (propriété des matrices de passage d'une b.o.n. à une autre)
Soit $\mathcal{B}$ et $\mathcal{C}$ deux b.o.n. de E. Soit $P = [\mathcal{B} \to \mathcal{C}]$ la matrice de passage de la base $\mathcal{B}$ à la base $\mathcal{C}$. Alors :

$${}^tPP = I_n = P\,{}^tP, \qquad \textit{donc:} \qquad P^{-1} = {}^tP.$$

Cette propriété définit la notion de matrice orthogonale :

Définition 9.5.11 (Matrice orthogonale)
Soit $P \in \mathcal{M}_n(\mathbb{R})$ une matrice carrée d'ordre n. On dit que P est une matrice orthogonale si et seulement si ${}^tPP = I_n$.

De la définition même découle de façon immédiate :

Proposition 9.5.12 (Inverse d'une matrice orthogonale)
Soit P une matrice orthogonale. Alors P est inversible, et $P^{-1} = {}^tP$.

Une matrice orthogonale peut se caractériser également par l'orthonormalité de ses colonnes :

Théorème 9.5.13 (Caractérisation d'une matrice orthogonale par ses colonnes)

Une matrice P est orthogonale si et seulement si ses colonnes forment une base orthonormale de $\mathbb{R}^n = \mathcal{M}_{n,1}(\mathbb{R})$, muni du produit scalaire canonique.

Théorème 9.5.14 (Caractérisation des matrices de passage entre b.o.n. par orthogonalité)

Soit E un espace euclidien.

1. *Toute matrice de passage d'une b.o.n. de E à une autre b.o.n. de E est une matrice orthogonale.*
2. *Réciproquement, soit $\mathcal{B}$ une b.o.n. de E, et P une matrice orthogonale. Alors il existe une unique base $\mathcal{C}$ telle que P soit la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$, et cette base $\mathcal{C}$ est une b.o.n. de E.*

Définition 9.5.15 (Groupe orthogonal)

On note $\mathrm{O}_n(\mathbb{R})$, ou $\mathrm{O}(n)$ l'ensemble des matrices orthogonales de $\mathcal{M}_n(\mathbb{R})$. Cet ensemble $\mathrm{O}_n(\mathbb{R})$ est appelé groupe orthogonal.

Théorème 9.5.16 (Structure de $\mathrm{O}_n(\mathbb{R})$)

L'ensemble $\mathrm{O}_n(\mathbb{R})$ est un groupe multiplicatif, ce qui est cohérent avec la terminologie introduite dans la définition précédente.

Théorème 9.5.17 (Déterminant d'une matrice orthogonale)

Soit $P \in \mathrm{O}_n(\mathbb{R})$. Alors $\det(P) \in \{-1, 1\}$. *Plus précisément,* det *est un morphisme de groupe de* $\mathrm{O}_n(\mathbb{R})$ *dans* $(\{-1, 1\}, \times)$.

Le noyau de ce morphisme est donc un sous-groupe de $\mathrm{O}_n(\mathbb{R})$.

Définition 9.5.18 (Groupe spécial orthogonal)

Le noyau du déterminant défini sur $\mathrm{O}_n(\mathbb{R})$ est appelé groupe spécial orthogonal, et noté $\mathrm{SO}_n(\mathbb{R})$ ou $\mathrm{SO}(n)$. Ainsi, les éléments de $\mathrm{SO}_n(\mathbb{R})$ sont les matrices orthogonales P telles que $\det(P) = 1$.

Le choix d'une matrice orthogonale P telle que $\det(P) = -1$ définit alors une bijection de $\mathrm{SO}(n)$ dans $\mathrm{O}^-(n) = \mathrm{O}(n) \setminus \mathrm{SO}(n)$ par $Q \mapsto PQ$.

V.3 Projecteurs orthogonaux

On rappelle la définition du projeté orthogonal :

Définition 9.5.19

Soit E un espace euclidien, et F un sous-espace vectoriel de E. Soit $x \in E$. On dit qu'un vecteur y de E est un projeté orthogonal de x sur F si et seulement si :

(i) $y \in F$

(ii) $x - y \in F^\perp$.

Comme nous l'avons vu plus haut, étant donnée une base orthonormale d'un sous-espace de dimension finie F de E, on dispose d'une formule simple de projection orthogonale, assurant l'existence et l'unicité de ce projeté. Comme nous avons montré entre temps que tout espace euclidien de dimension finie admet

une b.o.n., on peut réexprimer ce théorème ainsi, en se limitant au cas euclidien (assurant l'hypothèse de la dimension de F) :

Proposition 9.5.20 (Projeté orthogonal dans un espace euclidien)

*Soit E un espace euclidien, et F un sous-espace vectoriel de E. Alors tout vecteur x de E admet un et un seul projeté orthogonal sur F. De plus, étant donné une base **orthonormale** $(b_1, \dots, b_p)$ de F, ce projeté orthogonal $p_F(x)$ s'exprime de la manière suivante :*

$$p_F(x) = \sum_{i=1}^{p} \langle x, b_i \rangle \, b_i.$$

Définition 9.5.21 (projection orthogonale)

Le projeté orthogonal d'un vecteur sur F définit une application $p_F : E \to E$, qui a x associe son projeté orthogonal sur F. Cette application est appelée « projection orthogonale sur F »

Proposition 9.5.22

Évidemment, p_F est un projecteur.

Définition 9.5.23

Rappel : soit p un projecteur quelconque de E. Alors $\mathrm{id}_E - p$ est un projecteur de E, appelé projecteur associé de p.

Proposition 9.5.24 (projecteur associé à un projecteur orthogonal)

Soit E un espace euclidien, et F un sous-espace vectoriel de E. Alors le projecteur associé de la projection orthogonale p_F sur F est la projection orthogonale $p_{F^\perp}$ sur l'orthogonal de F.

Exemple 9.5.25 (Comment déterminer la matrice d'un projecteur othogonal de $\mathbb{R}^n$)

Soit dans $\mathbb{R}^3$ le plan F d'équation $x+2y-z=0$. Déterminer la matrice de p_F, le projecteur orthogonal sur le plan F.

(Réponse : $\frac{1}{6}\begin{pmatrix} 5 & -2 & 1 \\ -2 & 2 & 2 \\ 1 & 2 & 5 \end{pmatrix}$)

- Première méthode : trouver une b.o.n. de F, et utiliser la formule donnant le projeté à l'aide d'une b.o.n.
- Deuxième méthode (à réserver aux hyperplans, mais plutôt plus rapide dans ce cas) : Déterminer une base $F^\perp$ (c'est une droite, si F est un hyperplan). Exprimer le projeté orthogonal à l'aide d'une appartenance à $F^\perp$ (ce qui s'exprime par une colinéarité, si F est un hyperplan) et d'une appartenance à F (ce qui s'exprime par une équation).

V.4 Distance d'un point à un sous-espace vectoriel

Théorème 9.5.26 (Distance d'un point à un sous-espace)

Soit E un espace euclidien, F un sous-espace de E, et $x \in E$. Alors :

$$\forall y \in F, \quad \|x - p_F(x)\| \leqslant \|x - y\|,$$

l'égalité étant réalisée si et seulement si $y = p_F(x)$.
Autrement dit, $p_F(x)$ est l'unique vecteur de F minimisant la distance de x à un point de F.

Définition 9.5.27

On dit que $p_F(x)$ est la meilleure approximation de x dans F. On appelle distance de x à F le réel suivant :

$$d(x, F) = \|x - p_F(x)\| = \min_{y \in F} \|x - y\|.$$

VI Géométrie affine

Comme vous le savez pour l'avoir déjà utilisé en physique, un plan (non vectoriel) peut être défini par la donné d'un de ses points et d'un vecteur normal. Cela nécessite pour commencer une définition rigoureuse du cadre de la définition affine.

VI.1 Sous-espaces affines d'un espace vectoriel

Définition 9.6.1 (Structure d'espace affine)

Soit T un espace vectoriel sur $\mathbb{K}$, et E un ensemble. On dit que E est un espace affine attaché à T s'il est muni d'une loi externe de $T \times E$ dans E, notée $(t, x) \mapsto t + x$ ou $x + t$, telle que

- pour tout $(t, t') \in T^2$, tout $x \in E$, $(t + t') + x = t + (t' + x)$
- pour tout $x \in E$, $0 + x = x$
- pour tout $(x, y) \in E^2$, il existe t tel que $y = t + x$
- $(\forall x \in E,\ t + x = x) \Longrightarrow t = 0$

Proposition 9.6.2

Soit E un espace affine attaché à T. Soit $x \in E$. L'application $t \mapsto t + x$ est une bijection de T sur E.

Proposition/Définition 9.6.3 (Translation)

Soit E un espace affine attaché à T. Soit $t \in T$. L'application $\tau_t : x \mapsto t + x$ est une bijection de E dans lui-même. Cette application est appelée *translation de vecteur t*.

Ainsi, $(T, +)$ est le groupe des translations de E.

Définition 9.6.4 (Structure affine sur un espace vectoriel)

Soit E un espace vectoriel sur $\mathbb{K}$. On peut alors définir une structure d'espace affine sur E, attaché à lui-même, la loi externe correspondant alors à l'addition de E.

Terminologie 9.6.5 (Points et vecteurs)

Soit E un espace vectoriel, muni de sa structure affine usuelle. On distingue les propriétés affines et les propriétés vectorielles de E en considérant, comme dans le cas général d'un espace affine quelconque, que l'esace affine E est distinct de l'espace vectoriel $T = E$ auquel il est rattaché. Ainsi, parlant des éléments de l'espace affine, on parlera de *points* alors que les éléments de l'espace vectoriel E seront appelés *vecteurs*.

La loi externe de l'espace affine E est donc donnée par une relation du type $B = A + \overrightarrow{u}$, où A et B sont deux points de l'espace affine et $\overrightarrow{u}$ un élément de l'espace vectoriel.

Notation 9.6.6

Soit A et B deux points de l'espace affine A. On notera $\overrightarrow{AB}$ l'unique vecteur $\overrightarrow{u}$ de E tel que $B = A + \overrightarrow{u}$.

Ainsi, $\overrightarrow{AB}$ est entièrement déterminé par la relation $B = A + \overrightarrow{AB}$.

Définition 9.6.7 (Translaté d'un sous-ensemble de E)

Soit E un espace vectoriel muni de sa structure affine, et X un sous-ensemble de E, vu comme espace affine. Soit t un vecteur de E. Alors le translaté $\tau_t(X)$ de l'ensemble X est le sous-ensemble de l'espace affine E défini par :

$$\tau_t(X) = \{t + x,\ x \in X\} = \{y \in A \mid \exists x \in X,\ y = t + x\}.$$

Définition 9.6.8 (Sous-espace affine de E)

Soit E un espace vectoriel, muni de sa structure affine. Un sous-espace affine de E est un translaté par un vecteur t d'un sous-espace vectoriel de E.

En d'autres termes, et en dissociant d'avantage la structure vectorielle et la structure affine, un sous-espace affine de E est un ensemble non vide F tel qu'il existe $x \in F$ et V un sous-espace vectoriel de E tels que

$$F = \{x + u \mid u \in V\}.$$

Cette définition est indépendante d'une correspondance stricte entre les éléments de l'espace affine et les éléments de l'espace vectoriel, et reste vraie dans un espace affine général. Dans le cas de la structure affine usuelle d'un espace vecoriel, les points et les vecteurs étant assimilés, la description précédente correspond à l'espace obtenu par translation de V par le vecteur x.

Proposition/Définition 9.6.9 (Direction)

Soit F un sous-espace affine de E. Il existe un unique sous-espace vectoriel V de E tel que F soit un translaté de V. On dit que V est la direction de F, ou encore que F est dirigé par V.

Proposition 9.6.10

Soit F un sous-espace affine de E de direction V, et $A \in F$. Alors pour tout point B de E, $B \in F \iff \overrightarrow{AB} \in V$.

Assez logiquement nous définissons

Définition 9.6.11 (Hyperplan affine)

Un hyperplan affine d'un espace affine est un sous-espace affine dirigé par un hyperplan vectoriel de E.

Théorème 9.6.12 (Intersection de sous-espaces affines)

L'intersection de sous-espaces affines est soit vide, soit égale à un sous-espace affine. Si cette intersection est non vide, sa direction est l'intersection des directions de chacun des sous-espaces affines.

Certains auteurs définissent parfois la notion de sous-variété affine : il s'agit d'un sous-espace affine, ou de l'ensemble vide. Ainsi, l'intersection de sous-variétés affines est toujours une sous-variété affine.

Exemple 9.6.13 (Sous-espaces affines de $\mathbb{R}^2$ et $\mathbb{R}^3$)

Décrivez-les !

Un théorème important fournissant des sous-espaces affines (et on en a déjà vu des cas particuliers) est le suivant :

Théorème 9.6.14 (Fibres d'une application linéaire)

Soit $u \in \mathcal{L}(E, F)$, et $a \in F$. Alors l'image réciproque $u^{-1}(\{a\})$ (appelée fibre en a de u, ou ligne de niveau), est soit l'ensemble vide, soit un sous-espace affine (donc toujours une sous-variété affine).

Les exemples que nous avons déjà rencontrés sont :

Exemple 9.6.15 (Sous-espaces affines obtenus comme fibres)

1. Ensemble des solutions d'un système linéaire
2. Résolution des équations différentielles linéaires non homogènes de degré 1 ou 2.
3. L'ensemble des polynômes interpolateurs en un certain nombre de points.

Définition 9.6.16 (Repère affine)

Un repère affine est la donnée d'un point O de l'espace affine E, et d'une base (vectorielle) $(b_1, \dots, b_n)$ de l'espace vectoriel E.

Définition 9.6.17 (Coordonnées dans un repère affine)

Les coordonnées d'un point A dans un repère $(O, b_1, \dots, b_n)$ sont les coordonnées vectorielles du vecteur $\overrightarrow{OA}$ dans la base $(b_1, \dots, b_n)$.

Proposition 9.6.18

Soit E un espace affine de dimension n. L'application qui à A associe le vecteur de ses coordonnées est une bijection de E dans $\mathbb{R}^n$.

VI.2 Définition d'un hyperplan par vecteur normal

On considère un espace euclidien E, muni de sa structure affine. On commence par préciser les notions d'orthogonalité dans le contexte affine.

Définition 9.6.19 (vecteur orthogonal à un sous-espace affine)

Soit $\overrightarrow{v}$ un vecteur de E, et F un sous-espace affine. On dit que $\overrightarrow{v}$ est orthogonal à F si l'une des trois propriétés équivalentes suivantes est vérifiée :

- $\overrightarrow{v}$ est orthogonal à la direction de F,
- étant donné A fixé, $\overrightarrow{v}$ est orthogonal à tout vecteur $\overrightarrow{AB}$, pour $B \in F$;
- Pour tout $(A, B) \in F^2$, $\overrightarrow{v}$ est orthogonal à $\overrightarrow{AB}$

Proposition 9.6.20 (Hyperplans définis par le vecteur normal $\overrightarrow{n}$)
Soit $\overrightarrow{n}$ un vecteur non nul. Alors les hyperplans affines orthogonaux à $\overrightarrow{n}$ sont exactement les fibres (ou lignes de niveau) de $\overrightarrow{u} \mapsto \langle \overrightarrow{u}, \overrightarrow{n} \rangle$.

Proposition 9.6.21 (Définition d'un hyperplan affine par vecteur normal)
Soit $\overrightarrow{n}$ un vecteur non nul de E, et A un point. Il existe un unique hyperplan affine H de E passant par A et orthogonal à $\overrightarrow{n}$. On dit que $\overrightarrow{n}$ est un vecteur normal à H

La donnée donnée d'un vecteur normal et de A permet de trouver facilement une équation de l'hyperplan dans une base orthonormale, et réciproquement :

Proposition 9.6.22 (Équation d'un hyperplan de vecteur normal $\overrightarrow{n}$)
Soit $\mathcal{B}$ une base orthonormale de l'espace euclidien E, et soit $[\overrightarrow{n}]_{\mathcal{B}} = (a_1, \dots, a_n)$ le vecteur de s coordonnées de $\overrightarrow{n}$. Alors un hyperplan affine H de E admet $\overrightarrow{n}$ comme vecteur normal si et seulement s'il admet une équation du type :

$$B \in H \iff a_1 x_1 + \cdots + a_n x_n = b,$$

où $[\overrightarrow{OB}]_{\mathcal{B}} = (x_1, \dots, x_n)$.

Plus précisément, si on connaît un vecteur A de H, tel que

$$[\overrightarrow{OA}]_{\mathcal{B}} = (y_1, \dots, y_n)$$

la condition $B \in H$ s'écrit $\left\langle \overrightarrow{AB}, \overrightarrow{n} \right\rangle = 0$, ou encore :

$$a_1(x_1 - y_1) + \cdots + a_n(x_n - y_n) = 0, \qquad \text{soit:} \qquad a_1 x_1 + \cdots + a_n x_n = a_1 y_1 + \cdots + a_n y_n.$$

Récirproquement, lire un vecteur normal sur une équation est immédiat !

Exemples 9.6.23
Donner une description par point et vecteur normal dans les cas suivants :

1. D est la droite de $\mathbb{R}^2$ d'équation $y = 3x - 1$
2. P est le plan de $\mathbb{R}^3$ d'équation $x + 2y + z = 2$.

Définition 9.6.24 (Distance à un hyperplan)
Soit H un hyperplan défini par le point A et le vecteur normal **unitaire** $\overrightarrow{n}$. Alors la distance d'un point à H est $|\left\langle \overrightarrow{AM}, \overrightarrow{n} \right\rangle|$.

10 Isométries vectorielles

Nous étudions dans ce chapitre des applications linéaires particulières entre espaces euclidiens, conservant la norme. Dans le cas de la dimension 2, on retrouve les transformations vectorielles du plan (symétries et rotations vectorielles). Les isométries sont intimement liées à la notion de matrice orthogonale.

I Isométries d'un espace euclidien

Définition 10.1.1 (Isométrie vectorielle)

Une isométrie vectorielle d'un espace euclidien E est un endomorphisme $u \in \mathcal{L}(E)$ vérifiant :

$$\forall x \in E, \quad \|f(x)\| = \|x\|.$$

On note $\mathrm{O}(E)$ l'ensemble des isométries de E.

Ainsi, une isométrie est par définition un endomorphisme conservant la norme. La notation $\mathrm{O}(E)$ est très voisine de celle utilisée pour désigner le groupe orthogonal. Cela n'a rien d'anodin, comme on le constatera plus tard.

Définition 10.1.2 (Isométrie affine, HP)

Une isométrie affine d'un espace euclidien E est une application f telle que $x \mapsto f(x) - f(0)$ est une isométrie vectorielle.

Exemples 10.1.3

1. Id_E, $-\mathrm{Id}_E$
2. Dans $\mathbb{R}^2$ muni de la structure eulidienne canonique, les endomorphismes canoniquement associés aux matrices $\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ et $\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$.
3. Les symétries orthogonales par rapport à un sous-espace F, explicitement données en fonction de la projection orthogonale par $s_F = 2p_F - \mathrm{id}_E$.
4. En particulier, les reflexion (symétries par rapport à un hyperplan)
5. Les projecteurs orthogonaux ?

De façon assez immédiate, en considérant la norme d'un élément du noyau, on obtient :

Proposition 10.1.4 (Bijectivité des isométries)
Une isométrie vectorielle u d'une espace euclidien est un isomorphisme, et u^{-1} est encore une isométrie.

De façon tout aussi évidente :

Proposition 10.1.5 (Composée d'isométries)
La composée $v \circ u$ de deux isométries vectorielles de E est encore une isométrie de E.

On en déduit alors que

Corollaire 10.1.6 (Structure de O(E))
L'ensemble $\mathrm{O}(E)$ *des isométries est un sous-groupe de* $\mathrm{GL}(E)$.

Proposition 10.1.7 (Caractérisation des isométries par conservation du produit scalaire)
Soit E un espace euclidien et $u \in \mathcal{L}(E)$. Alors u est une isométrie si et seulement si :

$$\forall (x, y) \in \mathrm{E}^2, \quad \langle f(x), f(y) \rangle = \langle x, y \rangle .$$

Sans avoir défini correctement la notion d'angle entre deux vecteurs, mais par analogie à la situation bien connue de $\mathbb{R}^2$, cette propriété est à voir comme une propriété de conservation de l'angle (non orienté), en plus de la norme (obtenue pour x et y). En particulier, on a la conservation de l'orthogonalité : si $x \perp y$ alors $u(x) \perp u(y)$.
On peut obtenir une caractérisation par l'orthogonalité de la sorte, mais en rajoutant une information permettant de récupérer la conservation des normes, dans toutes les directions

Proposition 10.1.8 (Caractérisation des isométries par conservation des b.o.n.)
Soit $u \in \mathcal{L}(E)$. Les propositions suivantes sont équivalentes :
(i) u est une isométries
(ii) u envoie toute b.o.n. sur une b.o.n.
(iii) il existe une b.o.n. envoyée par u sur une b.o.n.

On en déduit alors la caractérisation matricielle suivante :

Proposition 10.1.9 (Caractérisation matricielle des isométries)
Soit $u \in \mathcal{L}(E)$. Les propositions suivantes sont équivalentes :
(i) u est une isométries
(ii) La matrice de u dans toute b.o.n. $\mathcal{B}$ est orthogonale
(iii) il existe une b.o.n. $\mathcal{B}$ telle que $\mathrm{Mat}_{\mathcal{B}}(u) \in \mathrm{O}(n)$.

Ainsi, le choix d'une base orthonormale $\mathcal{B}$ de E détermine un isomorphisme (de groupes) $\mathrm{O}(E) \xrightarrow{\sim} \mathrm{O}(n)$. Via cet isomorphisme, on peut considérer le sous-groupe de $\mathrm{O}(E)$, image réciproque du sous-groupe $\mathrm{SO}(n)$ des matrices orthogonales positives, c'est-à-dire les isométries u telles que
$det(\mathrm{Mat}_{\mathcal{B}(u)}) = 1$, ou, de façon équivalente, par définition du déterminant d'un endomorphisme, telles que $\det(u) = 1$

Définition 10.1.10 (Isométries vectorielles positives)
Une isométrie vetcorielle u est dite positive si et seulement si $\det(u) = 1$. On note $\mathrm{SO}(E)$ l'ensemble des isométries positives, sous-groupe distingué de $\mathrm{O}(n)$ (en tant que noyau d'un morphisme de groupe).

II Isométries vectorielles en dimension 2

II.1 Description de $\mathrm{O}(2)$

Nous donnons dans cette section une description complète des isométries en dimension 2. Pour cela, nous commençons par déterminer les matrices orthogonales de $\mathcal{M}_2(\mathbb{R})$.

Proposition 10.2.1 (Matrices orthogonales de $\mathcal{M}_2(\mathbb{R})$)
(i) Soit $M \in \mathrm{SO}(2)$, alors il existe $\theta \in \mathbb{R}$ tel que $M = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$;
(ii) Soit $M \in \mathrm{O}(2) \setminus \mathrm{SO}(2)$, alors il existe $\theta \in \mathbb{R}$ tel que $M = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix}$;
Ces descriptions sont uniques modulo 2π.

On a de plus une règle simple pour le produit de matrices de $\mathrm{SO}(2)$. En notant, pour tout $\theta \in \mathbb{R}$,

$$R(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

on obtient :

Proposition 10.2.2 (Inverse et produit dans SO(2))
(i) $R(0) = I_n$
(ii) Pour tout $(\theta, \theta') \in \mathbb{R}$, $R(\theta)R(\theta') = R(\theta + \theta') = R(\theta')R(\theta)$
(iii) Pour tout $\theta \in \mathbb{R}$, $R(\theta)^{-1} = R(-\theta)$.

En particulier, on reconnaît en $\mathrm{SO}(2)$ un groupe qu'on a déjà rencontré.

Théorème 10.2.3 (SO(2) est isomorphe à $\mathbb{U}$)
L'application qui à $R(\theta)$ associe $\mathrm{e}^{\mathrm{i}\,\theta}$ est un isomorphisme de groupe entre $\mathrm{SO}(2)$ et $\mathbb{U}$.

II.2 Isométries positives en dimension 2

Soit E un espace euclidien de dimension 2. On suppose que E est orienté. On obtient la caractérisation matricielle des isométries positives de E :

Théorème 10.2.4 (Caractérisation matricielle des isométries positives en dimension 2)
Soit $u \in \mathcal{L}(E)$. Les propriétés suivantes sont équivalentes :
(i) $u \in \mathrm{SO}(E)$
(ii) Il existe une b.o.n. directe $\mathcal{B}$ et un réel θ tel que $\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$
(iii) Il existe θ tel que pour toute b.o.n. directe $\mathcal{B}$, $\mathrm{Mat}_{\mathcal{B}}(u) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$
De plus, dans ce cas, θ est unique modulo 2π.

Définition 10.2.5 (Rotation)

On appelle rotation (vectorielle) d'angle $\theta \in \mathbb{R}$ (défini modulo 2π) l'application de SO(2), usuellement notée ρ_θ, dont la matrice dans toute base orthonormale directe est $\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$

Dans $\mathbb{R}^2$ muni de son produit scalaire canonique, cette définition correspond bien à la définition géométrique élémentaire et intuitive d'une rotation de centre 0 et d'angle θ. Son effet est de tourner la base canonique d'un angle θ.

Remarque 10.2.6

Si on décide d'inverser l'orientation de E, cela a pour effet sur la rotation de changer son angle θ en $-\theta$.

Les règles de produit matriciel dans SO(2) amènent directement les règles de composition des rotations, dont l'interprétation géométrique est assez intuitive :

Proposition 10.2.7 (Inverse et composée de deux rotations)

(i) $\rho_0 = \mathrm{id}_E$

(ii) *Pour tout* $(\theta, \theta') \in \mathbb{R}$, $\rho_\theta \circ \rho_{\theta'} = \rho_{\theta+\theta'} = \rho_{\theta'} \circ \rho_\theta$

(iii) *Pour tout* $\theta \in \mathbb{R}$, $\rho_\theta^{-1} = \rho_{-\theta}$.

Nous pouvons définir la notion d'angle orienté entre deux vecteurs grâce aux rotations. Pour cela, nous utilisons le lemme suivant :

Lemme 10.2.8

Soit E un espace euclidien orienté, et soit x et y deux vecteurs de norme 1 de E. Il existe une unique rotation ρ telle que $\rho(x) = y$.

Définition 10.2.9 (Angle orienté entre deux vecteurs)

Soit E un espace vectoriel orienté, et x et y deux vecteurs non nuls de E. Alors l'angle orienté $\widehat{(x, y)}$ est l'angle θ, défini modulo 2π, de l'unique rotation ρ telle que

$$\rho\left(\frac{x}{\|x\|}\right) = \frac{y}{\|y\|}.$$

Des règles de composition des rotations découlent immédiatement les trois premières des règles suivantes sur les angles :

Proposition 10.2.10 (Propriétés des angles orientés)

Les égalités ci-dessous sont à lire modulo 2π.

(i) $\forall x \in E \setminus \{0\}$, $\widehat{(x, x)} = 0$

(ii) $\forall x, y, z \in E \setminus \{0\}$, $\widehat{(x, z)} = \widehat{(x, y)} + \widehat{(y, z)}$

(iii) $\forall x, y \in E \setminus \{0\}$, $\widehat{(y, x)} = -\widehat{(x, y)}$.

(iv) $\forall x, y \in E \setminus \{0\}$, *et* $\lambda, \mu \in \mathbb{R}_+^*$, $\widehat{(\lambda x, \mu y)} = \widehat{(x, y)}$.

(v) $\forall x \in E \setminus \{0\}$, $\widehat{(-x, x)} = \pi$

(vi) $\forall x, y \in E \setminus \{0\}$, $\widehat{(-x, y)} = \pi + \widehat{(x, y)}$.

Nous voyons maintenant comment déduire de toutes les définitions ci-dessus l'expression du produit scalaire vue au lycée à l'aide de l'angle entre les vecteurs, ainsi que l'expression du déterminant. Pour cela, nous introduisons une définition générale, valable en dimension n quelconque.
Une matrice de $\mathrm{SO}(n)$ ayant un déterminant égal à 1, la formule de changement de base pour les déterminants permet d'affirmer que le déterminant relativement à une base orthonormée d'une famille de n vecteurs d'un espace euclidien de dimension n ne dépend pas du choix de cette base orthonormale.

Définition 10.2.11 (Produit mixte)

Le produit mixte de n vecteurs d'un espace euclidien orienté de dimension n, noté $[x_1, \dots, x_n]$ ou $\det(x_1, \dots, x_n)$, est la valeur commune des $\det_{\mathcal{B}}(x_1, \dots, x_n)$ dans les b.o.n. directes.

Proposition 10.2.12 (Expression du produit scalaire et du produit mixte par l'angle)

Soit x et y deux vecteurs non nuls de E euclidien orienté de dimension 2. Alors

$$\langle x, y\rangle = \|x\| \cdot \|y\| \cdot \cos\widehat{(x, y)} \qquad \text{et} \qquad [x, y] = \|x\| \cdot \|y\| \cdot \sin\widehat{(x, y)}$$

Ainsi, cette proposition permet une formalisation plus rigoureuse de la conservation de l'angle par les isométries, évoquée plus haut à propos de la conservation du produit scalaire : l'expression ci-dessus affirme en fait la conservation du cosinus de l'angle par une isométrie, donc une conservation de la valeur absolue de l'angle. La conservation ou non du signe du déterminant distingue alors les cas de conservation de l'angle orienté ou de passage à l'opposé ; cela distingue donc les isométries positives et les isométries négatives.
Nous donnons, donc, pour les isométries positives de E de dimension 2 :

Proposition 10.2.13 (Conservation de l'angle par une isométrie positive)

Soit $\rho \in \mathrm{SO}(E)$, E étant un espace euclidien orienté de dimension 2. Alors

$$\forall (x, y) \in E^2, \quad \langle \rho(x), \rho(y)\rangle = \langle x, y\rangle .$$

II.3 Isométries négatives en dimension 2

Nous terminons ce chapitre par la description des isométries négatives du plan. Nous rappelons qu'une reflexion est une symétrie orthogonale par rapport à un hyperplan. En dimension 2, les hyperplans sont des droites. Donc les reflexions sont les symétries orthogonales par rapport à une droite.

Théorème 10.2.14 (Isométries négatives en dimension 2)

Soit E un espace euclidien de dimension 2. Les isométries négatives de E sont les reflexions.

Observez que contrairement au cas des rotations, l'angle θ dépend ici du choix de la base orthonormale choisie.
De façon peut surprenant, les symétries inversent les angles. Pour le démontrer nous utilisons le lemme suivant :

Lemme 10.2.15

Soit σ une reflexion et ρ une rotation. Alors $\sigma \circ \rho \circ \sigma^{-1} = \rho^{-1}$.

On obtient

Proposition 10.2.16 (Inversion des angles par isométries négatives)
Soit σ une isométrie négative de E de dimension 2. Alors :

$$\forall (x,y) \in E^2, \quad \langle \sigma(x), \sigma(y) \rangle = - \langle x, y \rangle .$$