Département MIDO, première année

Algèbre 2

Introduction à l'algèbre linéaire

Alexandre Afgoustidis

(version du 6 janvier 2020)

Alexandre Afgoustidis
CEREMADE, Université Paris-Dauphine, 75016 Paris, France.
E-mail : `afgoustidis@ceremade.dauphine.fr`

Ce polycopié a été créé avec LaTeX; pour la mise en forme, nous avons adapté des fichiers de style fournis par la Société Mathématique de France, notamment la classe `smfbook`.

ALGÈBRE 2

Alexandre Afgoustidis

TABLE DES MATIÈRES

AVANT DE COMMENCER

Ce document a été écrit pour accompagner les cours dispensés à partir de 2017. Il doit beaucoup aux discussions que j'ai eues à son sujet avec Denis Pasquignon, Guillaume Legendre, François-Xavier Vialard, Emeric Bouin, Anne-Marie Boussion, Moulka Tamzali-Lafond, et beaucoup d'autres.

Dans la version actuelle, il reste inévitablement des fautes de frappe, de français, de mathématiques... Je vous serais très reconnaissant de me les signaler.

Ce polycopié est long : il est conçu pour être un *document de travail* plutôt qu'une transcription du cours dispensé en amphi. Il n'a pas vocation à se substituer à ce dernier : en particulier, il comporte des passages qui ne seront pas traités en amphi (et ne seront pas au programme de l'examen). À l'inverse, certains éléments ou exemples développés en cours ne figureront pas dans le polycopié.

Une particularité de ce document, qui explique en partie sa longueur, est la place qu'y occupent les *exemples* : leur étude me semble essentielle pour s'approprier la théorie abstraite exposée dans le cours... et pour comprendre son intérêt.

À la fin de certains chapitres, vous trouverez des *exercices de manipulation*. Leur but est de vous permettre de travailler sur des situations proches des exemples détaillés dans le polycopié, afin de consolider votre compréhension des notions du cours. Ils ne sont donc *pas conçus pour être traités en TD*, mais pour vous accompagner dans votre travail quotidien sur le cours. Il n'existe pas de correction tapée de ces exercices, mais vous pouvez bien sûr solliciter de l'aide à leur sujet (auprès de vos enseignants ou auprès de moi).

CHAPITRE 1

ESPACES VECTORIELS

1. La structure d'espace vectoriel

1.1. Groupes abéliens. —

On souhaite travailler avec des structures dans lesquelles il est possible de parler de la « somme » de deux éléments.

Soit E un ensemble. Une *loi de composition interne sur* E est une application

$$\begin{aligned} E \times E &\to E \\ (x, y) &\mapsto x + y \end{aligned}$$

qui permet de produire un élément de E à partir de deux éléments de E.

On dit que « $(E, +)$ *est un groupe abélien* » lorsque E est un ensemble et $+$ une loi de composition interne sur E vérifiant les quatre propriétés suivantes.

$$\begin{aligned}
&\text{Associativité :} && \textit{pour tout } (x, y, z) \in E^3, \text{ on a } (x + y) + z = x + (y + z)\ ; && (1.1)\\
&\text{Commutativité :} && \textit{pour tout } (x, y) \in E^2, , \text{ on a } x + y = y + x\ ; && (1.2)\\
&\text{Existence d'un neutre :} && \textit{il existe un élément } \mathbf{0}_E \textit{ de } E \textit{ vérifiant : } \forall x \in E,\ x + \mathbf{0}_E = x\ ; && (1.3)\\
&\text{Existence des opposés :} && \textit{pour tout } x \in E, \textit{ il existe un élément } \tilde{x} \in E \textit{ vérifiant : } x + \tilde{x} = \mathbf{0}_E. && (1.4)
\end{aligned}$$

Si x est un élément de E et si $\tilde{x}_1$, $\tilde{x}_2$ sont deux éléments de E vérifiant $x + \tilde{x}_1 = \mathbf{0}_E$ et $x + \tilde{x}_2 = \mathbf{0}_E$, alors

$$\tilde{x}_1 = \tilde{x}_1 + \mathbf{0}_E = \tilde{x}_1 + (x + \tilde{x}_2) = (\tilde{x}_1 + x) + \tilde{x}_2 = \mathbf{0}_E + \tilde{x}_2 = \tilde{x}_2,$$

si bien que l'élément $\tilde{x}$ mentionné dans (1.4) est uniquement déterminé par x. On l'appelle bien sûr *l'opposé* de x et on le note $-x$.

Si x et y sont deux éléments de E, on peut ainsi toujours parler de l'élément $x - y \overset{\text{déf}}{=} x + (-y)$: dans un groupe abélien, toutes les soustractions sont possibles.

1.2. Scalaires ; lois externes. —

Fixons un corps [1] $\mathbb{K}$ (dans ce cours, $\mathbb{K}$ sera toujours l'un des trois corps $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Q}$).

1. Pour cette notion, voir le cours d'analyse du premier semestre, page 21.

Partons d'un ensemble E muni d'une loi de composition interne $+$. On souhaite qu'il soit possible de parler non seulement de la somme de deux éléments de E, mais aussi du produit $\lambda \cdot x$ lorsque λ est un nombre dans $\mathbb{K}$ et x un élément de E : lorsque $\mathbb{K} = \mathbb{R}$, cela permettra par exemple de parler du *double* ou du *triple* d'un élément de E, de *multiplier un élément de E par 0.5, par 17 ou par π*, autrement dit, de faire des *changements d'échelle*. Les *nombres* de $\mathbb{K}$ seront ainsi appelés *scalaires* dans notre contexte. Cela motive la définition suivante.

Une *loi de multiplication des éléments de E par les scalaires de* $\mathbb{K}$ (on dit aussi une « loi externe pour E de corps de base $\mathbb{K}$ ») est une application

$$\begin{aligned} \mathbb{K} \times E &\to E \\ (\lambda, x) &\mapsto \lambda \cdot x \end{aligned}$$

qui, partant d'un scalaire de $\mathbb{K}$ et d'un élément de E, « fabrique » donc un élément de E.

On dit que « la loi $\cdot$ est *compatible avec* $+$ » lorsqu'elle vérifie les quatre propriétés suivantes :

$$\text{Multiplication par 1 : } \textit{pour tout } a \text{ de } E,\ 1 \cdot a = a \text{ ;} \tag{1.5}$$

$$\text{Changements successifs : } \textit{pour tout } (\lambda, \mu) \text{ de } \mathbb{K}^2 \text{ et tout } x \text{ de } E,\ (\lambda\mu) \cdot x = \lambda \cdot (\mu \cdot x) \text{ ;} \tag{1.6}$$

$$\text{Première distributivité : } \textit{pour tout } \text{ de } \mathbb{K} \text{ et tout } (x, y) \text{ de } E^2,\ \lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y \text{ ;} \tag{1.7}$$

$$\text{Seconde distributivité : } \textit{pour tout } (\lambda, \mu) \text{ de } \mathbb{K}^2 \text{ et tout } x \text{ de } E,\ (\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x. \tag{1.8}$$

1.3. Définition des espaces vectoriels. —

Définition 1.1 – Espace vectoriel sur un corps $\mathbb{K}$

Un $\mathbb{K}$-espace vectoriel est un triplet $(E, +, \cdot)$, où

E est un ensemble,

$+$ est une loi de composition interne sur E,

$\cdot$ est une loi de multiplication des éléments de E par les scalaires de $\mathbb{K}$,

qui vérifient :

$(E, +)$ est un groupe abélien.

La loi externe $\cdot$ est compatible avec $+$.

On notera que l'ensemble E contient l'élément neutre $\mathbf{0}_E$ de la loi $+$: un espace vectoriel ne peut jamais être vide.

1.3.1. *Exemples de règles de calcul.* —

Lemme 1.2 – Règles de calcul

Soit $(E, +, \cdot)$ un $\mathbb{K}$-espace vectoriel.

(a) Pour tout x de E, on a : $\quad 0 \cdot x = \mathbf{0}_E$.

(b) Pour tout λ de $\mathbb{K}$, on a : $\quad \lambda \cdot \mathbf{0}_E = \mathbf{0}_E$.

(c) Pour tout (λ, x) de $\mathbb{K} \times E$, $\quad \lambda \cdot x = \mathbf{0}_E \iff (\lambda = 0$ *ou* $x = \mathbf{0}_E)$.

(d) Pour tout x de E, on a : $\quad (-1) \cdot x = (-x)$.

Démonstration. —

(a) Soit x un élément de E. Alors $0 \cdot x = (0+0) \cdot x \overset{(1.8)}{=} (0 \cdot x) + (0 \cdot x)$.

Notons $y = (0 \cdot x)$; alors $y = y + y$, donc $\mathbf{0}_E \overset{(1.4)}{=} y + (-y) = (y+y) + (-y) \overset{(1.1)}{=} y + (y + (-y)) \overset{(1.3)}{=} y + \mathbf{0}_E \overset{(1.4)}{=} y$, cqfd.

(b) Soit λ un élément de $\mathbb{K}$. On a $\lambda \cdot \mathbf{0}_E = \lambda \cdot (\mathbf{0}_E + \mathbf{0}_E) \overset{(1.7)}{=} (\lambda \cdot \mathbf{0}_E) + (\lambda \cdot \mathbf{0}_E)$.

Notant $y = \lambda \cdot \mathbf{0}_E$, on a $y = y + y$, et la deuxième étape de la preuve de (a) montre que $y = \mathbf{0}_E$.

(c) Soit λ un élément de $\mathbb{K}$ et x un élément de E. Nous avons déjà démontré que si $\lambda = 0$ ou $x = \mathbf{0}_E$, alors $\lambda \cdot x = \mathbf{0}_E$. Réciproquement, si $\lambda \cdot x = \mathbf{0}_E$, soit $\lambda = 0$, soit $\lambda \neq 0$ et alors $x \overset{(1.5)}{=} 1 \cdot x = \left(\frac{1}{\lambda} \times \lambda\right) \cdot x \overset{(1.6)}{=} \frac{1}{\lambda} \cdot (\lambda \cdot x) \overset{(\text{hyp})}{=} \frac{1}{\lambda} \cdot \mathbf{0}_E \overset{(b)}{=} \mathbf{0}_E$.

(d) Soit x un élément de E ; on a $x + (-1) \cdot x \overset{(1.5)}{=} 1 \cdot x + (-1) \cdot x \overset{(1.4)}{=} (1 + (-1)) \cdot x = 0 \cdot x \overset{(a)}{=} \mathbf{0}$.

□

1.3.2. *Remarques sur la terminologie.* —

- Lorsque $(E, +, \cdot)$ est un $\mathbb{K}$-espace vectoriel, on parle souvent de *vecteurs* pour les éléments de E.
- Il faut noter que les *vecteurs* de E peuvent être des objets très différents de ce que la géométrie classique appelait « vecteurs » en son temps : l'espace E peut être un ensemble de nombres, de fonctions, de matrices, d'images, de sons, de polynômes, etc... Ce changement de vocabulaire est tout sauf anodin : en travaillant avec un espace vectoriel de fonctions (resp. de matrices, de sons, d'images, etc), le simple fait de *voir pour l'occasion les fonctions (resp. matrices, sons, images, etc) comme des objets géométriques* fournit de nouvelles images mentales et de nouveaux outils pour penser.
- On ne met généralement pas de flèche sur les éléments de E, même si on les appelle des vecteurs. On prendra bien garde à ne pas confondre scalaires et vecteurs dans une expression du type

$$\lambda \cdot x + \mu \cdot y$$

où, par exemple, λ et μ sont des scalaires et x et y des vecteurs.

1.3.3. *Remarques sur les notations.* —

- On s'autorise souvent à dire « Soit E un $\mathbb{K}$-espace vectoriel » plutôt que « Soit $(E, +, \cdot)$ un $\mathbb{K}$- espace vectoriel » lorsqu'il n'y a pas d'ambiguité sur le fait que les symboles $+$ et $\cdot$ vont désigner dans le raisonnement les deux lois qui interviennent dans la structure d'espace vectoriel.
- On omettra souvent le point pour la loi externe : si $(E, +, \cdot)$ est un $\mathbb{R}$-espace vectoriel et x est un élément de E, on notera généralement $3x$ ce qui, formellement, devrait être noté $3 \cdot x$. On conserve cependant toujours l'ordre dans ce cas, et on écrit le scalaire « à gauche » : en général, une expression du type $a \times 3$ n'a pas de sens.
- On s'autorisera souvent à écrire simplement 0 pour l'élément $\mathbf{0}_E$; il faut dans ce cas faire attention à ne pas confondre le *scalaire zéro* et le *vecteur nul*.

Ces entorses au formalisme permettent d'éviter des détours de langage, car il est rare par exemple que le texte

« *Soit E un $\mathbb{R}$-espace vectoriel. Considérons deux vecteurs x, y de E,*
notons $u = \sqrt{3}x + 2y$ et $v = -\frac{3}{2}x - \sqrt{3}y$;
alors $\sqrt{3}u + 2v = 0$. »

comporte une ambiguité.

1.3.4. *Remarques d'ordre psychologique.* —

- Il est *normal* de trouver que la structure d'espace vectoriel est (très) abstraite. C'est précisément parce qu'elle est abstraite qu'elle permet d'aborder des situations aussi diverses que les exemples de l'introduction. Pour comprendre les avantages *pratiques* (et pas seulement « intellectuels » ou

« esthétiques ») qu'il y a à développer la théorie des espaces vectoriels de façon abstraite, on pourra garder à l'esprit les deux remarques suivantes.

— Naturellement, un théorème démontré de façon abstraite est vrai *une fois pour toutes* : si on observe un phénomène commun à des espaces vectoriels très divers et si la démonstration est à peu près la même dans chaque cas particulier, le démontrer abstraitement une fois pour toutes représente une grande économie de temps et de pensée.

— Plus important : un problème pratique qui semble difficile et porte sur des objets compliqués peut, lorsqu'on le reformule dans des termes abstraits, apparaître comme *structurellement identique* à un problème qui porte sur des objets beaucoup plus simples, face auquel l'intuition est bien plus à l'aise. L'habitude de l'abstraction est alors un *outil pour penser* à ce problème pratique. Nous en verrons un exemple spectaculaire avec le *problème de la compression d'images.*

- Il est *normal* de trouver difficile de retenir les propriétés (1.1) à (1.8)... et les retenir n'est pas utile.

  En effet, il est *très rare* de devoir vérifier complètement ces propriétés pour reconnaître un espace vectoriel.

  Le plus souvent, lorsqu'on rencontre un ensemble E sur lequel on dispose d'une loi de composition interne $+$ et d'une loi externe $\cdot$ de corps de base $\mathbb{K}$, l'ensemble E apparaît comme sous-ensemble d'un $\mathbb{K}$-espace vectoriel déjà connu. On verra dans le §4 qu'il y a une manière très naturelle de savoir si $(E, +, \cdot)$ est un espace vectoriel, et qu'elle ne nécessite pas de revérifier (1.1) à (1.8).
- Une mise en garde : dans un espace vectoriel, on peut multiplier un vecteur par un nombre, mais...

  La structure d'espace vectoriel ne donne pas de notion de « multiplication des vecteurs ».

2. Exemples fondamentaux

2.1. Espaces numériques $\mathbb{K}^n$ $(n \in \mathbb{N}^\star)$. —

Soit n un élément de $\mathbb{N}^\star$. Les éléments de $\mathbb{K}^n$ sont les *n-uplets d'éléments de* $\mathbb{K}$.

Si x est un élément de $\mathbb{K}^n$, on écrira souvent $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ où les x_i, $i \in \{1, .., n\}$, sont des éléments de $\mathbb{K}$.

Lorsque cette notation en colonnes est trop encombrante, on notera $x = (x_1, ...x_n)$ pour le même vecteur.

Rappelons que l'ordre compte : dans $\mathbb{R}^3$, les vecteurs $\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}$ et $\begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}$ ne sont pas les mêmes.

Si $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ sont deux vecteurs de $\mathbb{K}^n$ et si $\lambda \in \mathbb{K}$, on note

$$x + y = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} \quad \text{et } \lambda \cdot x \ = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

On définit ainsi une loi interne sur $\mathbb{K}^n$ et une loi $\cdot$ de multiplication des éléments de $\mathbb{K}^n$ par les scalaires de $\mathbb{K}$. On constate alors que $(\mathbb{K}^n, +, \cdot)$ est un $\mathbb{K}$-espace vectoriel, dont le vecteur nul est $\mathbf{0}_{\mathbb{K}^n} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$.

Les propriétés (1.1) à (1.8) découlent directement des propriétés de l'addition et de la multiplication dans $\mathbb{K}$.

2.2. Espaces de matrices $\mathcal{M}_{n,p}(\mathbb{K})$ $(n \in \mathbb{N}^\star,\ p \in \mathbb{N}^\star)$. —

Soient $A = (a_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le p}}$ et $B = (b_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le p}}$ deux matrices à coefficients dans $\mathbb{K}$ comportant n lignes et p colonnes. Soit λ un élément de $\mathbb{K}$. Rappelons que $A + B$ et λA sont les matrices

$$A + B = \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1p} + b_{1p} \\ \vdots & & \vdots \\ a_{n1} + b_{n1} & \dots & a_{np} + b_{np} \end{pmatrix} \quad \text{et} \quad \lambda A = \begin{pmatrix} \lambda a_{11} & \dots & \lambda a_{1p} \\ \vdots & & \vdots \\ \lambda a_{n1} & \dots & \lambda a_{np} \end{pmatrix}.$$

On définit ainsi une loi interne sur $\mathcal{M}_{n,p}(\mathbb{K})$, une loi $\cdot$ de multiplication des éléments de $\mathcal{M}_{n,p}(\mathbb{K})$ par les scalaires de $\mathbb{K}$, et on constate que $(\mathcal{M}_{n,p}(\mathbb{K}), +, \cdot)$ est un $\mathbb{K}$-espace vectoriel, dont le vecteur nul est la matrice nulle.

Quelques rappels de notions vues au premier semestre :

- Il est très important d'être au point sur la résolution des systèmes linéaires et sur la manière classique de mettre une matrice ou un système linéaire sous forme échelonnée (« *méthode du pivot* »). Cela servira constamment dans ce cours.
- Il faut être au point sur la notion de *transposée* M^T d'une matrice M.
  Pour tous A, B de $\mathcal{M}_n(\mathbb{R})$ et pour tout λ de $\mathbb{K}$, on a :
  ▷ $(A + B)^T = A^T + B^T$,
  ▷ $(\lambda A)^T = \lambda A^T$,

▷ $(A^T)^T = A$,
▷ $(AB)^T = B^T A^T$: attention à l'inversion de l'ordre.

- La *trace* $\mathrm{Tr}(M)$ d'une matrice carrée M est la somme de ses coefficients diagonaux.
  Pour tous A, B de $\mathcal{M}_n(\mathbb{R})$ et pour tout λ de $\mathbb{K}$, on a :
  ▷ $\mathrm{Tr}(A+B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)$,
  ▷ $\mathrm{Tr}(\lambda A) = \lambda \mathrm{Tr}(A)$,
  ▷ $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

2.3. Espaces de polynômes $\mathbb{K}[X]$. —

- Si p est un entier naturel, et si $a_0, ...a_p$ sont des éléments de $\mathbb{K}$, on peut considérer le *le polynôme à coefficients dans* $\mathbb{K}$
$$P(X) = a_p X^p + a_{p-1} X^{p-1} + ... + a_1 X + a_0.$$
  L'ensemble des polynômes à coefficients dans $\mathbb{K}$ se note $\mathbb{K}[X]$.
- Dans ce cours, les notations P et $P(X)$ ont la même signification (et l'objet qu'elles désignent est un polynôme, alors que $P(0)$ ou $P(1)$ sont des éléments de $\mathbb{K}$, des « nombres »).
- Rappelons que si $P(X) = a_p X^p + a_{p-1} X^{p-1} + ... + a_1 X + a_0$ avec les $a_i, 0 \leq i \leq p$ dans $\mathbb{K}$, le *degré* de P, noté $\deg(P)$, est le plus grand entier k tel que $a_k \neq 0$. Rappelons aussi qu'il faut faire attention aux notations : si P s'écrit comme dans la phrase précédente, alors $\deg(P) \leq p$, mais si on ne demande pas que a_p soit non nul, on ne sait rien de plus.
- Rappelons par ailleurs que si $P(X) = a_p X^p + a_{p-1} X^{p-1} + ... + a_1 X + a_0$ et $Q(X) = b_q X^q + b_{q-1} X^{q-1} + ... + b_1 X + b_0$ sont deux polynômes de $\mathbb{K}[X]$, on peut définir le polynôme *somme* $P+Q$, comme suit : on note $a_k = 0$ pour tout entier $k > p$, on note $b_k = 0$ pour tout entier $k > q$, et on pose
$$(P+Q)(X) = \sum_{k \in \mathbb{N}} (a_k + b_k) X^k;$$
  la somme ci-dessus est en fait une somme finie, puisqu'on a $a_k + b_k = 0$ pour tout $k > \max(p, q)$.
- Si λ est un élément de $\mathbb{K}$ et $P(X) = a_p X^p + a_{p-1} X^{p-1} + ... + a_1 X + a_0$ est un élément de $\mathbb{K}[X]$ on définit le polynôme $\lambda \cdot P$ comme $(\lambda \cdot P)(X) = (\lambda a_p) X^p + ... + (\lambda a_1) X + (\lambda a_0)$.
- On définit ainsi une loi interne sur $\mathbb{K}[X]$, une loi $\cdot$ de multiplication des éléments de $\mathbb{K}[X]$ par les scalaires de $\mathbb{K}$, et on constate que $(\mathbb{K}[X], +, \cdot)$ est un $\mathbb{K}$-espace vectoriel, dont le vecteur nul est le *polynôme nul.*

Quelques rappels supplémentaires :

- Dire que deux polynômes sont égaux, c'est dire qu'ils ont les mêmes coefficients.
- Si P est un polynôme, on note $\deg(P)$ son degré. C'est un entier naturel, sauf si P est le polynôme nul.
  Le polynôme nul est, par convention, de degré $-\infty$.
- Si P et Q sont deux polynômes et λ un élément *non nul* de $\mathbb{K}$, alors on a $\deg(\lambda P) = \deg(P)$
  et $\deg(PQ) = \deg(P) + \deg(Q)$ (avec l'exception évidente si l'un des polynômes est nul).
  On a $\deg(P+Q) \leq \max(\deg(P), \deg(Q))$, mais pas égalité en général : songez (par exemple) à $X^3 + 1$ et $-X^3 + 2X$.
- Un polynôme de $\mathbb{K}[X]$ de degré $n \in \mathbb{N}$ admet au plus n racines dans $\mathbb{K}$.
  Le seul polynôme de degré n ayant strictement plus de n racines est le polynôme nul.
  Tout polynôme à coefficients dans $\mathbb{C}$ admet au moins une racine *complexe.*

2.4. Espaces de fonctions $\mathcal{F}(A, \mathbb{K})$ où A est un ensemble non vide quelconque. —

Soit A un ensemble non vide. Lorsque f et g sont deux applications de A dans $\mathbb{K}$ et lorsque λ est un élément de $\mathbb{K}$, on peut définir

$$f+g \quad : \quad \begin{aligned} A &\to \mathbb{K} \\ x &\mapsto f(x)+g(x) \end{aligned} \qquad \text{et} \qquad \lambda\cdot f \quad : \quad \begin{aligned} A &\to \mathbb{K} \\ x &\mapsto \lambda f(x) \end{aligned}$$

(pour chaque x de A, l'addition $f(x)+g(x)$ ou la multiplication $\lambda f(x)$ se font *dans* $\mathbb{K}$, donc on les connaît déjà, alors qu'on est *en train de définir* $f+g$ et $\lambda \cdot f$).

On définit ainsi une loi interne $+$ sur $\mathcal{F}(A,\mathbb{K})$, une loi $\cdot$ de multiplication des éléments de $\mathcal{F}(A,\mathbb{K})$ par les scalaires de $\mathbb{K}$.
Il est alors un peu long, mais pas difficile, de vérifier que le triplet $(\mathcal{F}(A,\mathbb{K}),+,\cdot)$ est un $\mathbb{K}$-espace vectoriel.

Remarque 1.3. — les *vecteurs de* $\mathcal{F}(A,\mathbb{K})$ sont donc des *fonctions*. Le *vecteur nul* de $\mathcal{F}(A,\mathbb{K})$ est la *fonction nulle*

$$\mathbf{0}_{\mathcal{F}(A,\mathbb{K})} : \begin{aligned} A &\to \mathbb{K} \\ x &\mapsto 0. \end{aligned}$$

Exemple 1.4. — Nous rencontrerons souvent *l'espace* $\mathcal{F}(\mathbb{R},\mathbb{R})$ *des fonctions de* $\mathbb{R}$ *dans* $\mathbb{R}$, ainsi que l'espace $\mathcal{F}(\mathbb{N},\mathbb{R})$ des applications de $\mathbb{N}$ dans $\mathbb{R}$, autrement dit, *l'espace* $\mathcal{F}(\mathbb{N},\mathbb{R})$ *des suites réelles.*

2.5. Espaces de fonctions $\mathcal{F}(A,E)$ où A est un ensemble et E un $\mathbb{K}$-espace vectoriel. —

Soit A un ensemble non vide et $(E,\underset{E}{+},\underset{E}{\cdot})$ un $\mathbb{K}$-espace vectoriel. Lorsque f et g sont deux applications de A dans E et lorsque λ est un élément de $\mathbb{K}$ on peut définir

$$f+g : \begin{aligned} A &\to E \\ x &\mapsto f(x)\underset{E}{+}g(x) \end{aligned} \qquad \text{et} \qquad \lambda\cdot f : \begin{aligned} A &\to E \\ x &\mapsto \lambda \underset{E}{\cdot} [f(x)]. \end{aligned}$$

On définit ainsi une loi interne $+$ sur $\mathcal{F}(A,E)$, une loi $\cdot$ de multiplication des éléments de $\mathcal{F}(A,E)$ par les scalaires de $\mathbb{K}$. Il est alors un peu long, mais pas difficile, de vérifier que le triplet $(\mathcal{F}(A,E),+,\cdot)$ est un $\mathbb{K}$-espace vectoriel.

Exemple 1.5. — L'espace $\mathcal{F}(A,\mathbb{R}^3)$ des fonctions de A dans $\mathbb{R}^3$ est un $\mathbb{R}$-espace vectoriel.

2.6. Espaces vectoriels produits. —

Soient n un entier naturel non nul et E_1, ... E_n des $\mathbb{K}$-espaces vectoriels. Sur le *produit cartésien* $\prod_{i=1}^{n} E_i = E_1 \times E_2 \times ... \times E_n$,

- On peut définir une loi de composition interne :
  si $(x_1,x_2,...,x_n)$ et $(y_1,y_2,...,y_n)$ et sont deux éléments de $E_1 \times E_2 \times ... \times E_n$, on pose

$$(x_1,x_2,...,x_n)+(y_1,y_2,...,y_n) \overset{\text{def}}{=} (x_1+y_1,x_2+y_2,...,x_n+y_n)$$

(où x_1+y_1 est défini à l'aide de la structure d'espace vectoriel de E_1, x_2+y_2 à partir de celle de E_2, etc).

- On peut aussi définir une loi externe de corps de base $\mathbb{K}$:
  si $(x_1, x_2, ..., x_n)$ est un élément de $E_1 \times E_2 \times ... \times E_n$ et si λ est un élément de $\mathbb{K}$, on pose
$$\lambda \cdot (x_1, x_2, ..., x_n) \overset{\text{def}}{=} (\lambda \cdot x_1,\ \lambda \cdot x_2, ...,\ \lambda \cdot x_n)$$
  (où $\lambda \cdot x_1$ est défini à l'aide de la structure d'espace vectoriel de E_1, $\lambda \cdot x_2$ à partir de celle de E_2, etc).

Il est alors long, mais pas difficile, de vérifier que le triplet $\left(\prod_{i=1}^{n} E_i\ ,\ +\ ,\ \cdot\right)$ est un $\mathbb{K}$-espace vectoriel.

Remarque 1.6. — La structure déjà décrite sur $\mathbb{K}^n$ est bien la structure d'espace vectoriel produit correspondant à l'égalité
$$\mathbb{K}^n = \underbrace{\mathbb{K} \times \mathbb{K} \times ... \times \mathbb{K}}_{n \text{ fois}}.$$

3. Combinaisons linéaires

3.1. Familles d'éléments d'un ensemble. —

Rappelons que si E est un ensemble non vide et I un ensemble, une *famille d'éléments de E indexée par I* est la donnée, pour chaque i de I, d'un élément x_i de E. On note $(x_i)_{i\in I}$ ou $(x_i \mid i \in I)$ une telle famille [2].

L'ordre compte : dans $\mathbb{R}^2$, les familles $\left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\2\end{pmatrix}\right)$ et $\left(\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right)$ sont différentes.

La famille vide n'est pas très intéressante, mais elle existe.

On rappelle aussi que si A est une partie de E, on peut la décrire comme la famille $(a \mid a \in A)$ d'éléments de E.

On s'autorisera souvent à dire « Soit $(x_i)_{i\in I}$ une famille d'éléments de E » plutôt que « Soit I un ensemble et $(x_i)_{i\in I}$ une famille d'éléments de E indexée par I », en sous-entendant qu'on fixe l'ensemble I dans la suite de la discussion.

3.2. Combinaisons linéaires : cas d'une famille finie. —

3.2.1. *Définition et remarques.* —

Définition 1.7 – Combinaison linéaire d'une famille finie de vecteurs

Soient E un $\mathbb{K}$-espace vectoriel, p un entier naturel non nul et $(x_1, ...x_p)$ une famille de p éléments de E.
On dit qu'un élément de E est *combinaison linéaire de la famille* $(x_1, ...x_p)$ lorsqu'il peut s'écrire sous la forme

$$\lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_p x_p$$

où $\lambda_1, ...\lambda_p$ sont des éléments de $\mathbb{K}$.

Remarque 1.8. — On n'oubliera pas que le scalaire zéro est « autorisé » dans les combinaisons linéaires : par exemple, même si $p > 1$, x_1 est toujours combinaison linéaire de $(x_1, ...x_p)$, puisqu'il s'écrit $1x_1 + 0x_2 + ... + 0x_p$.

Pour la même raison,

Le vecteur nul est combinaison linéaire de toutes les familles.

Remarque 1.9. — Pour la famille vide, on introduit la convention suivante : la seule combinaison linéaire possible est le vecteur nul.

Remarque 1.10. — Si $\mathbb{A}$ est une partie de $\mathbb{K}$, on parlera de combinaisons linéaires de $(x_1, ...x_p)$ *à coefficients dans* $\mathbb{A}$ pour désigner les vecteurs de la forme $\lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_p x_p$ avec $\lambda_1, ...\lambda_p$ $\underline{\text{dans } \mathbb{A}}$: par exemple, si E est un $\mathbb{R}$-espace vectoriel et (x_1, x_2, x_3) une famille de trois vecteurs de E, $2x_1 + 7x_2 - 18x_3$ est une combinaison linéaire à coefficients dans $\mathbb{Z}$ (on dit aussi : à coefficients entiers) de (x_1, x_2, x_3), tandis que $\frac{2}{7}x_1 - x_2 + \frac{9}{5}x_3$ en est une combinaison linéaire à coefficients rationnels.

2. Formellement, une famille indexée par I d'éléments de E est donc *exactement la même chose* qu'une application de I dans E... mais dans de nombreux contextes, le mot « famille » est agréable pour avoir une juste représentation des objets qu'on manipule.

3.2.2. *Premiers exemples concrets.* —

Exemple 1.11. — Dans $\mathbb{R}^2$, le vecteur $\begin{pmatrix}1\\5\end{pmatrix}$ est combinaison linéaire de la famille $\left(\begin{pmatrix}2\\-1\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right)$, car $\begin{pmatrix}1\\5\end{pmatrix} = (-5)\begin{pmatrix}2\\-1\end{pmatrix} + 11\begin{pmatrix}1\\0\end{pmatrix}$.

Exemple 1.12. — Dans $\mathbb{C}^3$, le vecteur $\begin{pmatrix}1-i\\3+i\\2-5i\end{pmatrix}$ *n'est pas* combinaison linéaire de $\begin{pmatrix}2\\-1\\0\end{pmatrix}$ et $\begin{pmatrix}5+i\\i\\0\end{pmatrix}$, car toute combinaison linéaire de ces vecteurs est de la forme $\begin{pmatrix}2\alpha+(7+i)\beta\\-\alpha+i\beta\\0\end{pmatrix}$, où α et β sont des nombres complexes, et ne peut donc pas avoir pour troisième coordonnée $2-5i$.

Exemple 1.13. — Dans $\mathcal{M}_2(\mathbb{R})$, la matrice $\begin{pmatrix}3&-4\\4&3\end{pmatrix}$ est combinaison linéaire de $\begin{pmatrix}1&0\\0&1\end{pmatrix}$ et $\begin{pmatrix}0&1\\-1&0\end{pmatrix}$.

En revanche, $\begin{pmatrix}17&-4\\4&3\end{pmatrix}$ n'est pas combinaison linéaire de ces deux matrices. Voyez-vous pourquoi ?

Exemple 1.14. — Dans $\mathbb{C}^2$, le vecteur $\begin{pmatrix}1+i\\-1\end{pmatrix}$ est combinaison linéaire *à coefficients complexes* de $\begin{pmatrix}1\\0\end{pmatrix}$ et $\begin{pmatrix}1\\i\end{pmatrix}$,

mais il n'est pas combinaison linéaire *à coefficients réels* des mêmes vecteurs.

Exemple 1.15. — Dans $\mathbb{R}^3$, le vecteur $u = \begin{pmatrix}3\\6\\9\end{pmatrix}$ est combinaison linéaire de $\left(\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\-1\\0\end{pmatrix}\right)$. On peut bien sûr le voir en écrivant

$$\begin{pmatrix}3\\6\\9\end{pmatrix} = 3\begin{pmatrix}1\\2\\3\end{pmatrix} + 0\begin{pmatrix}1\\0\\1\end{pmatrix} + 0\begin{pmatrix}1\\-1\\0\end{pmatrix};$$

cela dit, on a aussi

$$\begin{pmatrix}3\\6\\9\end{pmatrix} = 2\begin{pmatrix}1\\2\\3\end{pmatrix} + 3\begin{pmatrix}1\\0\\1\end{pmatrix} - 2\begin{pmatrix}1\\-1\\0\end{pmatrix}.$$

Il y a donc plusieurs « jeux de coefficients » qui permettent d'exprimer le vecteur u comme combinaison linéaire de la famille donnée. On peut bien sûr relier cette absence d'unicité de l'écriture à l'égalité (visible en prenant la différence)

$$\begin{pmatrix}0\\0\\0\end{pmatrix} = \begin{pmatrix}1\\2\\3\end{pmatrix} - 3\begin{pmatrix}1\\0\\1\end{pmatrix} + 2\begin{pmatrix}1\\-1\\0\end{pmatrix};$$

nous y reviendrons au §7

3.2.3. *Exemples plus abstraits.* —

Exemple 1.16. — Dans $\mathcal{F}(\mathbb{R}, \mathbb{R})$, dire qu'une fonction est *affine*,
c'est dire qu'elle est combinaison linéaire des deux fonctions $\mathrm{id}_{\mathbb{R}} : x \mapsto x$ et $\mathrm{cste}_1 : x \mapsto 1$.

Exemple 1.17. — Dans $\mathcal{M}_n(\mathbb{R})$, dire qu'une matrice est *diagonale*, c'est dire que c'est une combinaison linéaire des matrices

$$\begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & 0 \end{pmatrix}, ..., \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Exemple 1.18. — Dans $\mathbb{R}[X]$, si n appartient à $\mathbb{N}$, les combinaisons linéaires de la famille $(1, X, X^2, ...X^n)$ sont les polynômes de la forme

$$a_0 + a_1 X + ... + a_n X^n$$

où $(a_0, a_1, ...a_n)$ appartient à $\mathbb{R}^{n+1}$. Ce sont donc exactement les polynômes de degré *inférieur*[3] *ou égal à* n.

Exemple 1.19. — Soit A une matrice à n lignes et p colonnes. Notons $C_1, ...C_n$ les colonnes de A : Si $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$ est un vecteur de $\mathbb{R}^p$, alors AX est un vecteur de $\mathbb{R}^n$; la définition du produit matriciel donne

$$AX = x_1 C_1 + x_2 C_2 + ... + x_p C_p.$$

Un produit matrice-vecteur, c'est une combinaison linéaire des colonnes de la matrice,
avec pour coefficients les coordonnées du vecteur.

3.3. Combinaisons linéaires : cas d'une famille infinie. —

Définition 1.20 – Combinaison linéaire d'une famille infinie de vecteurs

Soient E un $\mathbb{K}$-espace vectoriel, I un ensemble *quelconque* et $(x_i)_{i \in I}$ une famille d'éléments de E indexée par I.
On dit qu'un vecteur v de E est *combinaison linéaire de la famille* $(x_i)_{i\in I}$ s'il existe une *sous-famille finie* de $(x_i)_{i\in I}$ dont v est combinaison linéaire, autrement dit s'il existe p dans $\mathbb{N}^\star$, une sous-famille $(x_{i_1}, x_{i_2}, ...x_{i_p})$ de $(x_i)_{i\in I}$ et des scalaires $\lambda_1, ...\lambda_p$ de $\mathbb{K}$ qui vérifient :

$$v = \lambda_1 x_{i_1} + \lambda_2 x_{i_2} + ... + \lambda_p x_{i_p}.$$

Remarque 1.21. — Les combinaisons linéaires *d'un nombre infini de vecteurs* restent donc des sommes *finies*. Il n'est pas possible de parler d'une combinaison « infinie » sans qu'une notion de limite intervienne, ce qui dépasse le cadre de ce cours de première année.

3. Noter que rien n'impose à a_n d'être non-nul, donc que le polynôme ci-dessus n'est pas forcément de degré n.

Dans les combinaisons linéaires d'une famille, même infinie, il n'y a qu'un nombre fini de termes.

Attention, en revanche : la manipulation des familles infinies et de leurs combinaisons linéaires nécessite des précautions, notamment pour ce qui concerne les notations. Les exemples qui suivent devraient le montrer clairement.

Exemple 1.22. — Plaçons-nous dans l'espace $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$.

Considérons la famille $(c_\lambda)_{\lambda\in\mathbb{R}}$ d'éléments de E, où pour chaque $\lambda \in \mathbb{R}$ la fonction c_λ est donnée par : $\forall x \in \mathbb{R}, c_\lambda(x) = \cos(\lambda x)$.

Notons f la fonction définie par : $\forall x \in \mathbb{R},\ f(x) = 2 - 7\cos(\sqrt{3}x) + \sqrt{2}\cos(150000x)$.

Alors le vecteur f de de $\mathcal{F}(\mathbb{R}, \mathbb{R})$ est combinaison linéaire de la famille $(c_\lambda)_{\lambda\in\mathbb{R}}$.

En effet, c'est une combinaison linéaire de la sous-famille de $(c_\lambda)_{\lambda\in\mathbb{R}}$ formée des trois fonctions $c_0 : x \mapsto \cos(0x)$, $c_{\sqrt{3}} : x \mapsto \cos(\sqrt{3}x)$ et $c_{150000} : x \mapsto \cos(150000x)$.

Exemple 1.23. — Dans $\mathbb{K}[X]$, *tout* polynôme est combinaison linéaire de la famille $(X^i)_{i\in\mathbb{N}}$.

Par exemple, le polynôme $3 + 9X^2 - 8X^{17} + 2.3 \cdot X^{987654}$ peut s'écrire sous la forme $\lambda_1 X^{i_1} + \lambda_2 X^{i_2} + \lambda_3 X^{i_3} + \lambda_4 X^{i_4}$

avec $(i_1, i_2, i_3, i_4) = (0, 2, 17, 987654)$ et $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (3, 9, -8, 2.3)$.

On notera que :

- Le *degré d'un polynôme donné* peut être *aussi grand qu'on veut*, mais il est toujours *fini*.
- Il ne faut pas confondre le *degré d'un polynôme* et le *nombre de coefficients non nuls* dans son écriture comme combinaison des puissances de X.

Exemple 1.24. — Dans l'espace $\mathcal{F}(\mathbb{N}, \mathbb{R})$ des suites réelles, définissons pour chaque p de $\mathbb{N}$ une suite $u^{[p]}$ par

$$u_n^{[p]} = \delta_{pn} = \begin{cases} 1 \text{ si } p = n, \\ 0 \text{ si } p \neq n. \end{cases}$$

(attention à bien comprendre les notations : pour p fixé, $u^{[p]}$ est une suite. Par exemple, $u^{[1]}$ est la suite dont les termes sont $0, 1, 0, 0, ...$, tandis que $u^{[5]}$ est la suite dont les termes sont $0, 0, 0, 0, 0, 1, 0, 0, 0....$).

Alors les combinaisons linéaires de la famille $(u^{[p]}, p \in \mathbb{N})$ sont les suites *nulles à partir d'un certain rang.*

Notion de famille à support fini de scalaires et écriture des combinaisons linéaires

- Soient I un ensemble quelconque et $(\lambda_i)_{i\in I}$ une famille indexée par I d'éléments de $\mathbb{K}$. On dit que $(\lambda_i)_{i\in I}$ est *à support fini* (on dit aussi : *presque nulle*) lorsque tous les λ_i sont nuls sauf un nombre fini d'entre eux, autrement dit, lorsqu'il existe un sous-ensemble *fini* J de I vérifiant : $\forall i \in (I \setminus J),\ \lambda_i = 0$.
- Soit E un $\mathbb{K}$-espace vectoriel et $(x_i)_{i\in I}$ une famille de vecteurs de E indexée par I. Alors la somme

$$\sum_{i\in I} \lambda_i x_i \tag{3.1}$$

  a un sens, puisque c'est en fait la somme finie $\sum_{i\in J} \lambda_i x_i$. Cette somme est une combinaison linéaire de la famille $(x_i)_{i\in J}$, donc aussi de la famille $(x_i)_{i\in I}$.

- Réciproquement, partons d'une combinaison linéaire de $(x_i)_{i\in I}$. Puisque c'est une combinaison linéaire d'une sous-famille finie $(x_{i_1},\dots x_{i_p})_{i=1,..p}$ pour un certain p de $\mathbb{N}^\star$ et pour un certain sous-ensemble $\{i_1,\dots i_p\}$ de I, elle peut s'écrire $\sum\limits_{i\in I}\lambda_i x_i$ avec $\lambda_i = 0$ sauf éventuellement si i est égal à i_1, à i_2, ... ou à i_p.
- Ainsi, les combinaisons linéaires de la famille $(x_i)_{i\in I}$ sont exactement les vecteurs de E qui peuvent s'écrire sous la forme (3.1) pour une certaine famille *à support fini* $(\lambda_i)_{i\in I}$ de scalaires de $\mathbb{K}$.

Ce point de vue simplifie considérablement les notations, et nous nous en servirons régulièrement dans les démonstrations théoriques. Je n'ai pas voulu l'adopter pour définition, cependant : lorsqu'une belle notation cache une partie de la complexité du sujet, gare aux erreurs d'interprétation...

4. Sous-espaces vectoriels

4.1. Définition et premier exemple. —

Définition 1.25 – Sous-espace vectoriel

Soient E un $\mathbb{K}$-espace vectoriel et F une partie de E. On dit que F est un *sous-espace vectoriel de* E lorsque

(a) La partie F est non vide

(b) Si x et y sont deux vecteurs de F, alors $x+y$ appartient à F

(c) Si x est un vecteur de F, alors pour tout λ de $\mathbb{K}$, $\lambda\cdot x$ appartient à F

Exemple 1.26. —

- La partie $F_1 = \left\{\begin{pmatrix}x\\y\end{pmatrix}\in\mathbb{R}^2\ ,\ y=0\right\}$ est un sous-espace vectoriel de $\mathbb{R}^2$,
- La partie $F_2 = \left\{\begin{pmatrix}x\\y\end{pmatrix}\in\mathbb{R}^2\ ,\ y=1\right\}$ n'en est pas un (en effet, $\begin{pmatrix}1\\1\end{pmatrix}$ et $\begin{pmatrix}3\\1\end{pmatrix}$appartiennent à F_2, mais pas leur somme).
- La partie $F_2' = \left\{\begin{pmatrix}x\\y\end{pmatrix}\in\mathbb{R}^2\ ,\ x\in\mathbb{Z} \text{ et } y\in\mathbb{Z}\right\}$ n'en est pas un non plus (en effet$\begin{pmatrix}1\\0\end{pmatrix}$ appartient à F_2', mais pas $\sqrt{2}\begin{pmatrix}1\\0\end{pmatrix}$).

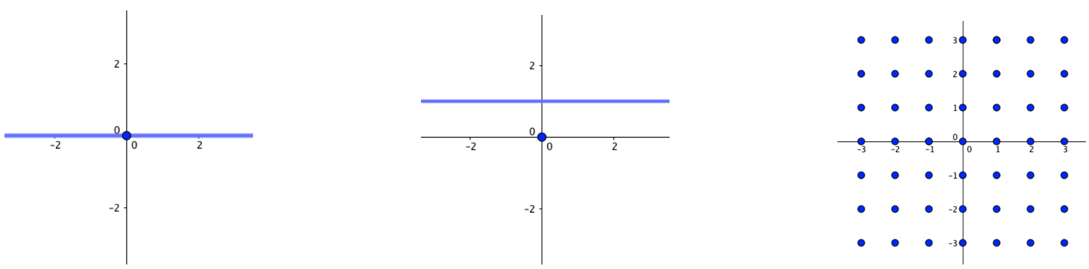


FIGURE 1. Dessin des parties F_1, F_2 et F_2' de l'exemple 1.26

4.2. Conséquences et reformulations de la définition. —
Soient $(E, +, \cdot)$ un $\mathbb{K}$-espace vectoriel et F un sous-espace vectoriel de E.

- La condition (a) dit qu'il existe un élément u dans F. Mais alors, grâce à (c), le vecteur $0 \cdot u$, appartient à F. Or, $0 \cdot u$ n'est autre que le vecteur nul $\mathbf{0}_E$ (voir les règles de calcul page **??**) , et ainsi :

> Un sous-espace vectoriel contient automatiquement le vecteur nul.

- La condition (b) dit que la restriction $+_{F^2}$ de $+$ à $F \times F$ est à valeurs dans F, si bien que $+_{F^2}$ est une loi de composition interne sur F ; elle vérifie alors *automatiquement* les propriétés (1.1) à (**??**), et ainsi $(F, +_{F^2})$ est un groupe abélien.
- La condition (c) dit, quant à elle, que la restriction $\cdot_{\mathbb{K}\times F}$ de $\cdot$ à $\mathbb{K} \times F$ est à valeurs dans F ; ainsi $\cdot_{\mathbb{K}\times F}$ définit une loi de multiplication des éléments de F par les scalaires de $\mathbb{K}$. Elle vérifie *automatiquement* les propriétés (1.5) à (1.8).

> Par conséquent, si F est un sous-espace vectoriel de E, il hérite une structure d'espace vectoriel de celle de E.

- Soit $(x_i)_{i\in I}$ une famille de vecteurs de F et v une combinaison linéaire de la famille $(x_i)_{i\in I}$: rappelons que cela signifie qu'il existe un entier p, des scalaires $\lambda_1, ...\lambda_p$ et des indices $i_1, ...i_p$ tels que
$$v = \lambda_1 x_{i_1} + \lambda_2 x_{i_2} + ... + \lambda_p x_{i_p}.$$
D'après (c), chacun des $\lambda_k x_{i_k}$ $(k \in \{1, ..p\})$ appartient à F, et d'après (b), une somme d'éléments de F est dans F. Ainsi,

> Dire que F est un sous-espace vectoriel de E, c'est dire que :
> toute combinaison linéaire d'éléments de F est encore un élément de F.

Pour parler de cette propriété, on dit aussi que F est *stable par combinaisons linéaires.*

Il est très utile de retenir la reformulation suivante de la définition.

Proposition 1.27 – Sous-espace vectoriel : en pratique

Soit E un $\mathbb{K}$-espace vectoriel. Une partie F de E est un sous-espace vectoriel de E si et seulement si elle vérifie les deux conditions suivantes :

(i) Le vecteur nul $\mathbf{0}_E$ appartient à F.

(ii) Pour tout (λ, μ) de $\mathbb{K}^2$ et pour tout (x, y) de F^2, $\lambda x + \mu y$ appartient à F.

Démonstration. — Nous avons en effet déjà vu que (a), (b) et (c) impliquent (i) et (ii) ; dans l'autre sens, (i) implique bien sûr (a), et si F vérifie (ii), alors :
- pour tout (x, y) de F^2, $1x + 1y$ appartient à F, donc F vérifie (b).
- pour tout (λ, x) de $\mathbb{K} \times F$, $\lambda x + 0y$ appartient à F, donc F vérifie (c).

□

Exercices de manipulation. — *I.1, I.2.*

4.3. Exemples fondamentaux de sous-espaces vectoriels. —

4.3.1. *Sous-espaces* $\{\mathbf{0}_E\}$ *et* E. — Il y a d'abord deux exemples « triviaux » à ne pas oublier : si E est un espace vectoriel,

> Les parties E et $\{\mathbf{0}_E\}$ sont toujours des sous-espaces vectoriels de E.

4.3.2. *Droites et plans passant par l'origine.* —

- Plaçons-nous d'abord dans l'espace vectoriel réel $\mathbb{R}^2$.

  Si $\vec{u}$ est un vecteur non nul de $\mathbb{R}^2$, rappelons que la *droite de* $\mathbb{R}^2$ *passant par l'origine et dirigée par* $\vec{u}$ est l'ensemble

  $$\mathbb{R}\vec{u} = \{\vec{x} \in \mathbb{R}^2, \exists t \in \mathbb{R}, \vec{x} = t\vec{u}\}.$$

  C'est un sous-espace vectoriel de $\mathbb{R}^2$: d'abord, $\mathbf{0}_{\mathbb{R}^2} = 0\vec{u}$ appartient à $\mathbb{R}\vec{u}$. De plus, si λ et μ sont deux nombres réels et si $\vec{x}$ et $\vec{y}$ appartiennent à $\mathbb{R}\vec{u}$, alors il existe deux réels t_x et t_y tels que $\vec{x} = t_x\vec{u}$ et $\vec{y} = t_y\vec{u}$; mais alors $\lambda\vec{x} + \mu\vec{y} = (\lambda t_x + \mu t_y)\,\vec{u}$ appartient aussi à $\mathbb{R}\vec{u}$.

- Plaçons-nous maintenant dans l'espace vectoriel réel $\mathbb{R}^3$.
  - — Une *droite de* $\mathbb{R}^3$ *passant par l'origine* est un sous-ensemble de la forme $\mathbb{R}\vec{u} = \{\vec{r} \in \mathbb{R}^3, \exists t \in \mathbb{R}, \vec{r} = t\vec{u}\}$, osûr $\vec{u}$ est un vecteur non nul de $\mathbb{R}^3$; pour la même raison que précédemment, c'est un sous-espace vectoriel de $\mathbb{R}^3$.
  - — Un *plan de* $\mathbb{R}^3$ *passant par l'origine* est un sous-ensemble de $\mathbb{R}^3$ de la forme

    $$\mathcal{P} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3,\ ax + by + cz = 0 \right\},$$

    où (a, b, c) est un triplet de nombres réels.

    Un tel plan est un sous-espace vectoriel de $\mathbb{R}^3$: bien sûr il contient toujours $\mathbf{0}_{\mathbb{R}^3}$; si de plus $\vec{r}_1 = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$ et $\vec{r}_2 = \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}$ sont deux vecteurs de $\mathcal{P}$ et si λ et μ sont deux réels, on a

    $$a(\lambda x_1 + \mu x_2) + b(\lambda y_1 + \mu y_2) + c(\lambda z_1 + \mu z_2) = \lambda(ax_1 + by_1 + cz_1) + \mu(ax_2 + by_2 + cz_2) = 0,$$

    et donc $\lambda\vec{r}_1 + \mu\vec{r}_2 = \begin{pmatrix} \lambda x_1 + \mu x_2 \\ \lambda y_1 + \mu y_2 \\ \lambda z_1 + \mu z_2 \end{pmatrix}$ vérifie l'équation de $\mathcal{P}$, donc il appartient à $\mathcal{P}$.

> Dans $\mathbb{R}^2$ et dans $\mathbb{R}^3$,
> les droites passant par l'origine et les plans passant par l'origine sont des sous-espaces vectoriels.

FIGURE 2. Droites et plans passant par l'origine sont des sous-espaces vectoriels

4.3.3. *Espace des solutions d'un système linéaire homogène.* —
Soient n et m deux entiers naturels non nuls. Rappelons qu'un *système de m équations linéaires homogènes à n inconnues à coefficients dans* $\mathbb{K}$ est un système d'équations de la forme

$$(S) : \begin{cases} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = 0 \\ \quad\vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = 0 \end{cases}$$

où les *coefficients* $a_{ij}, i \in \{1,..m\}, j \in \{1,..n\}$ (des éléments de $\mathbb{K}$) sont des données du problème et où $x_1, x_2, ...x_n$ sont les *inconnues* ; rappelons également qu'en notant A la matrice à n lignes et p colonnes dont les coefficients sont les a_{ij}, résoudre le système (S), c'est résoudre l'équation

$$(S') : AX = \mathbf{0}_{\mathbb{K}^n}$$

d'inconnue $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$. Bien sûr le vecteur $\mathbf{0}_{\mathbb{K}^n} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ est solution de (S'), car les seconds membres de (S') sont nuls.

De plus, λ et μ sont deux éléments de $\mathbb{K}$ et si X et X' sont deux solutions de (S'), alors AX_1 et AX_2 sont nuls ; les propriétés du produit matriciel donnent alors

$$A\,(\lambda X_1 + \mu X_2) = \lambda(AX_1) + \mu(AX_2) = \lambda\mathbf{0}_{\mathbb{K}^n} + \mu\mathbf{0}_{\mathbb{K}^n} = \mathbf{0}_{\mathbb{K}^n}$$

si bien que $\lambda X_1 + \mu X_2$ est une solution de (S'). Nous avons obtenu le résultat suivant :

Proposition 1.28 – Ensemble des solutions d'un système linéaire homogène

Soit (S) un système de m équations linéaires homogènes à n inconnues à coefficients dans $\mathbb{K}$. L'ensemble des solutions de (S) est un sous-espace vectoriel de $\mathbb{K}^n$.

C'est un fait important, car nous verrons plus loin (voir page 64) que *tout* sous-espace vectoriel de $\mathbb{K}^n$ peut être décrit comme l'ensemble des solutions d'un système linéaire homogène.

Exemple 1.29. — Si $\vec{u} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ est un vecteur non nul de $\mathbb{R}^3$, la droite $\mathbb{R}\vec{u}$ peut être vue comme $\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3, \begin{array}{r} bx - ay = 0 \\ cy - bz = 0 \end{array} \right\}$.

4.3.4. *L'espace* $\mathbb{K}_n[X]$. —

Proposition 1.30 – L'espace $\mathbb{K}_n[X]$ est un sous-espace vectoriel de $\mathbb{K}[X]$

Si n est un entier naturel, le sous-ensemble $\mathbb{K}_n[X]$ de $\mathbb{K}[X]$ formé des polynômes de degré inférieur ou égal à *n est un sous-espace vectoriel de $\mathbb{K}[X]$.*

Démonstration. — • Le polynôme nul est bien de degré inférieur ou égal à n.

- Soient P et Q deux polynômes de degré inférieur ou égal à n et λ et μ deux scalaires de $\mathbb{K}$. On peut écrire

$$P(X) = a_0 + a_1X + ... + a_nX^n \text{ et } Q(X) = b_0 + b_1X + ... + b_nX^n$$

avec $(a_0, a_1, ...a_n)$ et $(b_0, b_1, ...b_n)$ dans $\mathbb{K}^{n+1}$ (*attention,* a_n *et* b_n *peuvent très bien être nuls ici*).

Mais alors

$$\lambda P(X) + \mu Q(X) = (\lambda a_0 + \mu b_0) + (\lambda a_0 + \mu b_0)X + ... + (\lambda a_n + \mu b_n)X^n$$

où les nombres $(\lambda a_0+\mu b_0), (\lambda a_1+\mu b_1), ...(\lambda a_n+\mu b_n)$ sont des éléments de $\mathbb{K}$. Cela prouve que $\lambda P+\mu Q$ est de degré inférieur ou égal à n.

□

Attention. — L'espace des polynômes de degré *égal* à n n'est pas un sous-espace vectoriel de $\mathbb{K}[X]$, ne serait-ce que parce qu'il ne contient pas le polynôme nul [4].

4.3.5. *Fonctions continues, dérivables, suites convergentes.* —

Dans $\mathcal{F}(\mathbb{R}, \mathbb{R})$,

- L'ensemble des fonctions continues est un sous-espace vectoriel.
  En effet, vous avez vu dans le cours d'analyse du premier semestre que la fonction nulle est continue et que si λ et μ sont deux réels et f et g deux fonctions continues, alors la fonction $\lambda f + \mu g$ est encore continue.
- L'ensemble des fonctions dérivables est un sous-espace vectoriel (voir le cours d'analyse).

Dans l'espace $\mathcal{F}(\mathbb{N}, \mathbb{R})$ des suites réelles,

- L'ensemble des suites convergentes est un sous-espace vectoriel (voir le cours d'analyse).

4.3.6. *Solutions d'une équation différentielle linéaire homogène.* —

Soit n un entier naturel ; soient φ_0, φ_1, ... φ_n des fonctions de $\mathbb{R}$ dans $\mathbb{R}$. L'ensemble des fonctions $y \in \mathcal{F}(\mathbb{R}, \mathbb{R})$ qui sont n fois dérivables sur $\mathbb{R}$ et vérifient

$$\forall x \in \mathbb{R}, \qquad \varphi_n(x) \cdot y^{(n)}(x) + ... + \varphi_2(x) \cdot y^{''}(x) + \varphi_1(x) \cdot y^{'}(x) + \varphi_0(x) \cdot y(x) = 0 \tag{4.1}$$

est un sous-espace vectoriel de $\mathcal{F}(\mathbb{R}, \mathbb{R})$.

Si φ_n n'est pas la fonction nulle, on dit que l'équation (4.1) est une *équation différentielle linéaire homogène d'ordre* n.

Par exemple, l'équation différentielle $y' = y$ est d'ordre 1, l'équation différentielle $y'' + y = 0$ est d'ordre 2. Vous en connaissez déjà des solutions : exp pour la première, sin et cos pour la seconde. Nous y reviendrons.

Exercice de manipulation. — I.3.

4. On peut aussi remarquer que deux polynômes de degré n, comme $X^n + 3X + 1$ et $-X^n + 5X^2$ (pour cet exemple on suppose $n > 2$), peuvent avoir pour somme un polynôme de degré *strictement inférieur* à n (ici $5X^2 + 3X + 1$).

4.3.7. *Suites vérifiant une relation de récurrence linéaire.* —

Soit k un entier naturel ; soient a_1, ... a_{k-1} des nombres réels.
L'ensemble des suites $(u_n)_{n\in\mathbb{N}}$ qui vérifient

$$\forall n \in \mathbb{N}, \qquad u_{n+k} = a_0u_n + a_1u_{n+1} + \ldots + a_{k-1}u_{n+(k-1)} \tag{4.2}$$

est un sous-espace vectoriel de l'espace $\mathcal{F}(\mathbb{N},\mathbb{R})$ des suites réelles.

On dit que la relation (4.2) est une *relation de récurrence linéaire d'ordre* k.
Les suites vérifiant une relation de récurrence linéaire d'ordre 1 sont les *suites géométriques*.
La *suite de Fibonacci*, définie par $u_0 = u_1 = 0$ et $u_{n+2} = u_{n+1} + u_n$ pour tout n de $\mathbb{N}$, est une suite récurrente linéaire d'ordre 2.

Exercice de manipulation. — I.4.

4.4. Autres exemples et non-exemples. —

4.4.1. *Sous-ensembles « non-plats », « asymétriques » ou « à trous ».* —

- Dans $\mathbb{R}^2$, les sous-ensembles suivants ne sont pas des sous-espaces vectoriels :

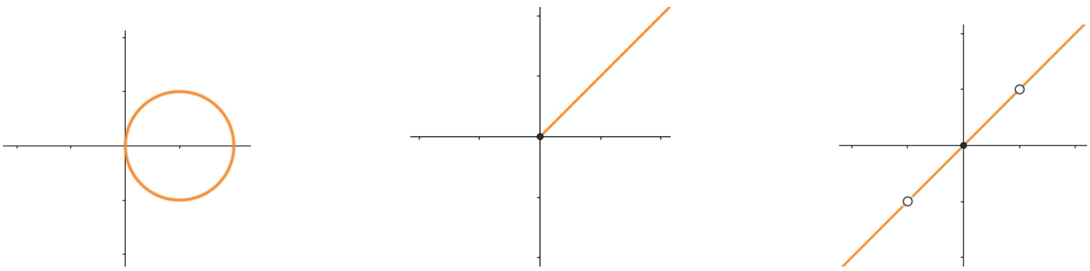


Figure 3. Un cercle, une demi-droite, une droite à deux trous

le *cercle* $\left\{\begin{pmatrix}x\\y\end{pmatrix} \in \mathbb{R}^2, (x-1)^2+y^2=1\right\}$, la *demi-droite* $\left\{\begin{pmatrix}x\\y\end{pmatrix} \in \mathbb{R}^2, y=x \text{ et } x\geq 0\right\}$, la *droite à deux trous* $\left\{\begin{pmatrix}x\\y\end{pmatrix} \in \mathbb{R}^2, y=x \text{ et } x^2\neq 1\right\}$,

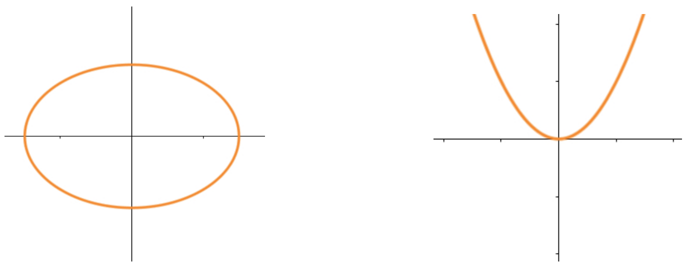


Figure 4. Une ellipse et une parabole

l'*ellipse* $\left\{\begin{pmatrix}x\\y\end{pmatrix} \in \mathbb{R}^2, \frac{x^2}{a^2}+\frac{y^2}{b^2}=1\right\}$, la *parabole* $\left\{\begin{pmatrix}x\\y\end{pmatrix} \in \mathbb{R}^2, y=x^2\right\}$,

- Dans $\mathbb{R}^3$, le *tore* $\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3,\ (4-\sqrt{x^2+y^2})^2+z^2=1 \right\}$, le *cône* $\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3,\ xy-z^2=0 \right\}$ non plus.

FIGURE 5. Un cône et un tore

Les seuls sous-espaces vectoriels de $\mathbb{R}^2$ sont, outre $\mathbb{R}^2$ et $\{0\}$, les droites qui passent par l'origine

. De même,

Les seuls sous-espaces vectoriels de $\mathbb{R}^3$ sont, outre $\mathbb{R}^3$ et $\{0\}$, les droites et plans passant par $\mathbf{0}_{\mathbb{R}^3}$.

4.4.2. *Systèmes linéaires* NON-*homogènes.* —
Si $a_{ij}, i \in \{1,..m\}, j \in \{1,..n\}$ et $b_i, i \in \{1,..m\}$ sont des éléments de $\mathbb{K}$, alors l'ensemble des éléments $(x_1, x_2, ..x_n)$ de $\mathbb{K}^n$ qui vérifient

$$(S) : \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n & = & b_2 \\ \quad\vdots \qquad\quad \vdots \qquad\qquad\qquad \vdots & \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n & = & b_m \end{cases}$$

est un sous-espace vectoriel de $\mathbb{R}^n$ si *et seulement si* tous les b_i sont nuls.

4.4.3. *Ensembles de fonctions ou de suites.* —

- Dans $\mathcal{F}(\mathbb{R}, \mathbb{R})$, l'espace des *fonctions qui valent zéro en zéro* est un sous-espace vectoriel,
  mais l'espace des *fonctions positives* n'est pas un sous-espace vectoriel.

- Dans $\mathcal{F}(\mathbb{N}, \mathbb{R})$, l'espace des suites qui *convergent vers zéro* est un sous-espace vectoriel,
  mais l'espace des suites qui *convergent vers* $\sqrt{2}$ n'est pas un sous-espace vectoriel,
  et l'espace des suites qui *convergent vers un nombre entier* n'est pas un sous-espace vectoriel.

Exercices de manipulation. — I.3, I.4

4.5. Intersection de sous-espaces vectoriels. —

Proposition 1.31 – Une intersection de sous-espaces vectoriels en est un

Si E est un $\mathbb{K}$-espace vectoriel et si $(F_i)_{i\in I}$ est une famille de sous-espaces vectoriels de E, alors l'intersection $\bigcap_{i\in I} F_i$ est un sous-espace vectoriel de E.

Démonstration. —
- Le vecteur nul appartient à chacun des F_i, donc il appartient à $\bigcap_{i\in I} F_i$.
- Soient λ, μ deux éléments de $\mathbb{K}^2$ et x, y deux éléments de $\bigcap_{i\in I} F_i$.

  Alors pour tout i de I, x et y appartiennent à F_i, donc $\lambda x + \mu y$ appartient à F_i.
  Ainsi, le vecteur $\lambda x + \mu y$ appartient à tous les F_i, donc il appartient à $\bigcap_{i\in I} F_i$. □

□

Attention. — La *réunion* de deux sous-espaces vectoriels n'est « presque jamais » un sous-espace vectoriel.

Figure 6. L'intersection de deux plans distincts passant par l'origine est une droite qui passe par l'origine ; c'est bien un sous-espace vectoriel de $\mathbb{R}^3$. La réunion de deux droites distinctes, par contre, n'est pas un sous-espace vectoriel de $\mathbb{R}^3$.

5. Sous-espace vectoriel engendré

5.1. Définition et exemples. —

Définition 1.32 – Sous-espace engendré par une partie

Soit E un $\mathbb{K}$-espace vectoriel et $(x_i)_{i\in I}$ une famille de vecteurs de E. Le *sous-espace vectoriel de E engendré par* $(x_i)_{i\in I}$ est l'ensemble des éléments de E qui peuvent s'écrire comme combinaison linéaire de la famille $(x_i)_{i\in I}$. On le notera $\mathbf{Vect}[(x_i)_{i\in I}]$.

Remarque 1.33. — La façon dont les vecteurs sont indexés dans $(x_i)_{i\in I}$ n'a pas d'influence sur $\mathbf{Vect}[(x_i)_{i\in I}]$, à cause de la commutativité de l'addition. Par exemple, si n appartient à $\mathbb{N}^\star$ et si $(x_1, x_2, ...x_n)$ est une famille de n vecteurs de E,

$$\mathbf{Vect}[x_1, x_2, ..., x_n] = \{v \in E \quad / \quad \exists(\lambda_1, \lambda_2, ..., \lambda_n) \in \mathbb{K}^n, \quad v = \lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_n x_n \quad \}$$

et on a toujours $\mathbf{Vect}[x_1, x_2, ..., x_n] = \mathbf{Vect}[x_2, x_1, ..., x_n] = \mathbf{Vect}[x_n, x_2, ..., x_1]$.

Remarque 1.34. — Si A est une partie de E, on rappelle qu'on peut voir A comme une famille de vecteurs de E (voir le §3.1). On notera $\mathbf{Vect}[A]$ le sous-espace vectoriel engendré par A.

Exemple 1.35. — Dans $\mathbb{R}^3$, $\mathbf{Vect}\left[\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right]$ est l'ensemble des vecteurs qui peuvent s'écrire sous la forme $\alpha\begin{pmatrix}1\\0\\0\end{pmatrix} + \beta\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}\alpha\\0\\\beta\end{pmatrix}$ avec α, β dans $\mathbb{R}$, c'est donc le plan $\left\{\begin{pmatrix}x\\y\\z\end{pmatrix} \in \mathbb{R}^3 \quad / \quad y = 0\right\}$.

***Exemple 1.36* (Notion de droite vectorielle).** — Dans $\mathbb{R}^2$, $\mathbf{Vect}\left[\begin{pmatrix}2\\1\end{pmatrix}\right]$ est l'ensemble des vecteurs *proportionnels à* $\begin{pmatrix}2\\1\end{pmatrix}$, c'est-à-dire la droite passant par l'origine dont un vecteur directeur est $\begin{pmatrix}2\\1\end{pmatrix}$.

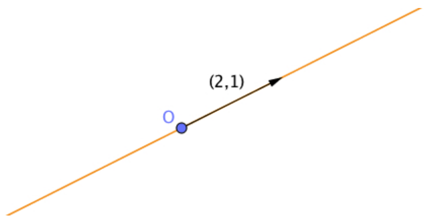


FIGURE 7. Une droite vectorielle de $\mathbb{R}^2$.

Plus généralement, si u est un vecteur *non nul* dans un espace vectoriel E, alors $\mathbf{Vect}[u]$ est l'ensemble des vecteurs de E qui sont proportionnels à u : par analogie avec l'exemple (b), on l'appellera la *droite vectorielle engendré par u*.

On appelle *droite vectorielle de E* tout sous-espace vectoriel de E engendré par un seul vecteur non nul.

***Exemple 1.37* (Notion de plan vectoriel).** — On peut également généraliser la notion de plan passant par l'origine.

Soient E un $\mathbb{K}$-espace vectoriel et u et v deux vecteurs de E. On dit que u et v sont *colinéaires* s'il existe soit un élément α de $\mathbb{K}$ tel qu'on ait $u = \alpha v$, soit un élément β de $\mathbb{K}$ tel qu'on ait $v = \beta u$. On notera que *le vecteur nul est colinéaire à tous les vecteurs.*

On dit alors qu'une partie P de E est un *plan vectoriel* s'il existe deux vecteurs *non colinéaires* u et v de E qui vérifient : $P = \mathbf{Vect}[u, v]$. Ainsi :

On appelle *plan vectoriel de E* tout sous-espace vectoriel de E engendré par deux vecteurs non colinéaires.

***Exemple 1.38* (Espace des fonctions affines).** — Dans $\mathcal{F}(\mathbb{R}, \mathbb{R})$, en reprenant les notations de l'exemple (a) page 11, le sous-espace $\mathbf{Vect}[\mathrm{cste}_1, \mathrm{id}_{\mathbb{R}}]$ est l'ensemble des fonctions affines.

***Exemple 1.39* (Exemples avec des combinaisons à coefficients dans $\mathbb{Q}$)**

Dans le $\mathbb{Q}$-espace vectoriel $\mathbb{R}$, le sous-espace $\mathbf{Vect}\left[1, \sqrt{2}\right]$ est l'ensemble des nombres réels qui peuvent s'écrire sous la forme $a + b\sqrt{2}$ avec a et b rationnels.

Dans le $\mathbb{Q}$-espace vectoriel $\mathbb{R}$, on a par ailleurs $\mathbf{Vect}\left[7, \sqrt{2}\right] = \mathbf{Vect}\left[1, \sqrt{2}\right]$ et $\mathbf{Vect}[-4] = \mathbb{Q}$.

***Exemple 1.40* (Matrices).** — Dans $\mathcal{M}_2(\mathbb{R})$, $\mathbf{Vect}\left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right]$ est l'ensemble des matrices triangulaires supérieures.

Exercices de manipulation : — I.5, I.6, I.7.

5.2. Propriétés théoriques. — Il faut d'abord remarquer que la terminologie est cohérente : si A est une partie de E, alors $\mathbf{Vect}[A]$ est bien un sous-espace vectoriel de E. En effet :

- Le vecteur nul en est bien toujours combinaison linéaire des vecteurs de A.
- Soient λ et μ deux éléments de $\mathbb{K}$ et u et v des éléments de $\mathbf{Vect}[A]$.

  Dire que u et v sont dans $\mathbf{Vect}[A]$, c'est dire qu'il existe deux familles à support fini $(\alpha_i)_{i\in I}$ et $(\beta_i)_{i\in I}$ de scalaires telles que $u = \sum \alpha_i x_i$ et $v = \sum \beta_i x_i$. En introduisant les ensembles *finis* $J_u = \{i \in I,\ \alpha_i \neq 0\}$ et $J_v = \{i \in I,\ \beta_i \neq 0\}$, on peut alors écrire

  $$\lambda u + \mu v = \sum_{i\in I} (\lambda\alpha_i + \mu\beta_i) x_i = \sum_{i\in J_u \cup J_v} (\lambda\alpha_i + \mu\beta_i) x_i$$

  puisque si i n'est ni dans J_u ni dans J_v, alors $\alpha_i = \beta_i = 0$. L'ensemble $J_u \cup J_v$ étant fini, cela montre que $\lambda u + \mu v$ est combinaison linéaire de la famille (x_i), donc qu'il appartient bien à $\mathbf{Vect}[A]$.

Nous isolons maintenant ce qui est peut-être la propriété la plus utile du sous-espace engendré.

Proposition 1.41 – Propriété fondamentale du sous-espace engendré

Si F est un sous-espace vectoriel de E et si tous les vecteurs de A appartiennent à F, alors $\mathbf{Vect}[A] \subset F$.

On peut ainsi dire que si A est une partie de E, $\mathbf{Vect}[A]$ est le *plus petit sous-espace vectoriel de E qui contienne A*, au sens de la relation d'inclusion : il est inclus dans tous les sous-espaces vectoriels de E qui contiennent A.

Démonstration. — Si F est un sous-espace vectoriel de E et si tous les vecteurs de A appartiennent à F, alors toutes les combinaisons linéaires des vecteurs de A appartiennent à F. Comme $\mathbf{Vect}[A]$ est exactement l'ensemble des combinaisons linéaires de vecteurs de A, il est bien contenu dans F. □

***Exemple 1.42* (Premier exemple d'utilisation).** — Soit $F = \{(x, y, z, t) \in \mathbb{R}^4,\ x + y + z + t = 0\}$; c'est un sous-espace vectoriel de $\mathbb{R}^4$, puisque c'est l'ensemble des solutions d'un système linéaire homogène.

Du fait que les vecteurs $\begin{pmatrix}1\\-1\\0\\0\end{pmatrix}$, $\begin{pmatrix}1\\0\\-1\\0\end{pmatrix}$ et $\begin{pmatrix}1\\0\\0\\-1\end{pmatrix}$ appartiennent à F, on peut donc déduire :

$$\mathbf{Vect}\left[\begin{pmatrix}1\\-1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\-1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\\-1\end{pmatrix}\right] \subset F.$$

***Exemple 1.43* (Montrer une égalité de $\mathbf{Vect}[\cdot]$).** — Dans $\mathbb{R}^3$, montrons l'égalité $\mathbf{Vect}\left[\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}1\\0\\-1\end{pmatrix}\right] = \mathbf{Vect}\left[\begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}0\\2\\4\end{pmatrix}\right]$. Notons $\vec{u} = \begin{pmatrix}1\\2\\3\end{pmatrix}$, $\vec{v} = \begin{pmatrix}1\\0\\-1\end{pmatrix}$, $\vec{a} = \begin{pmatrix}1\\1\\1\end{pmatrix}$ et $\vec{b} = \begin{pmatrix}0\\2\\4\end{pmatrix}$.

- On a $\vec{a} = \frac{1}{2}(\vec{u} + \vec{v})$, donc $\vec{a}$ appartient à $\mathbf{Vect}[\vec{u}, \vec{v}]$. De plus, $\vec{b} = \vec{u} - \vec{v}$, donc $\vec{b}$ appartient aussi à $\mathbf{Vect}[\vec{u}, \vec{v}]$.

  La proposition donne ainsi : $\mathbf{Vect}\left[\vec{a}, \vec{b}\right] \subset \mathbf{Vect}[\vec{u}, \vec{v}]$.
- Pour l'autre inclusion, on constate que $\vec{u} = \vec{a} + \frac{1}{2}\vec{b}$ et $\vec{v} = \vec{a} - \frac{1}{2}\vec{b}$.

  Ainsi, $\vec{u}$ et $\vec{v}$ appartiennent tous les deux à $\mathbf{Vect}\left[\vec{a}, \vec{b}\right]$, et $\mathbf{Vect}[\vec{u}, \vec{v}] \subset \mathbf{Vect}\left[\vec{a}, \vec{b}\right]$.

Exercices de manipulation : — I.8, I.9.

Voici deux corollaires importants de la propriété fondamentale du sous-espace engendré.

Proposition 1.44 – Comportement de $\mathbf{Vect}[A]$ si on ajoute ou retire des vecteurs à A

(a) ***(Compatibilité avec les inclusions)***

Si A et B sont deux parties de E et si $A \subset B$, alors $\mathbf{Vect}[A] \subset \mathbf{Vect}[B]$.

(b) ***(Propriété d'élimination des redondances)***

Si A est une partie de E et si x est une combinaison linéaire des vecteurs de A, alors $\mathbf{Vect}[A \cup \{x\}] = \mathbf{Vect}[A]$.

Démonstration. —

- Pour (a), il suffit de constater que si A et B sont deux parties de E et si les vecteurs de A appartiennent tous à B, alors toute combinaison linéaire des vecteurs de A est aussi une combinaison linéaire des vecteurs de B.

- Pour (b), si A est une partie de E, dire que x est combinaison linéaire des vecteurs de A, c'est dire que x appartient à $\mathbf{Vect}[A]$. Notons $F = \mathbf{Vect}[A]$; tous les vecteurs de $A \cup \{x\}$ appartiennent alors à F, donc d'après la proposition précédente, $\mathbf{Vect}[A \cup \{x\}] \subset F$, c'est-à-dire $\mathbf{Vect}[A \cup \{x\}] \subset \mathbf{Vect}[A]$. Pour l'autre inclusion, il suffit de constater que A est contenue dans $A \cup \{x\}$ et d'utiliser la partie (a) que nous venons de démontrer : cela donne $\mathbf{Vect}[A] \subset \mathbf{Vect}[A \cup \{x\}]$.

□

Exemple 1.45 **(Utilisation de la propriété d'élimination des redondances.)**

Dans $\mathbb{R}^4$, notons $F = \mathbf{Vect}\left[\begin{pmatrix}1\\2\\1\\1\end{pmatrix}, \begin{pmatrix}-2\\0\\3\\4\end{pmatrix}, \begin{pmatrix}5\\6\\0\\-1\end{pmatrix}\right]$. Alors on a l'égalité $F = \mathbf{Vect}\left[\begin{pmatrix}1\\2\\1\\1\end{pmatrix}, \begin{pmatrix}-2\\0\\3\\4\end{pmatrix}\right]$.

En effet, $\begin{pmatrix}5\\6\\0\\-1\end{pmatrix} = 3\begin{pmatrix}1\\2\\1\\1\end{pmatrix} - \begin{pmatrix}-2\\0\\3\\4\end{pmatrix}$.

Le sous-espace F, bien qu'engendré par trois vecteurs, est donc un plan vectoriel.

Exercices de manipulation : — I.10

6. Familles génératrices

6.1. Définition et exemples fondamentaux. —

Définition 1.46 – Famille génératrice d'un espace vectoriel

Soit E un $\mathbb{K}$-espace vectoriel et $(x_i)_{i\in I}$ une famille de vecteurs de E. On dit que $(x_i)_{i\in I}$ *engendre* E (ou *est génératrice de* E) lorsque $E = \mathbf{Vect}[(x_i)_{i\in I}]$.

Nous donnons de premiers exemples importants.

Exemple 1.47. —

La famille $\left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right)$ est une famille génératrice de $\mathbb{K}^2$: pour tout $\begin{pmatrix}x\\y\end{pmatrix}$ de $\mathbb{K}^2$, on a $\begin{pmatrix}x\\y\end{pmatrix} = x\begin{pmatrix}1\\0\end{pmatrix} + y\begin{pmatrix}0\\1\end{pmatrix}$.

La famille $\left(\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right)$ engendre $\mathbb{K}^3$: pour tout $\begin{pmatrix}x\\y\\z\end{pmatrix}$ de $\mathbb{K}^3$, on a $\begin{pmatrix}x\\y\\z\end{pmatrix} = x\begin{pmatrix}1\\0\\0\end{pmatrix} + y\begin{pmatrix}0\\1\\0\end{pmatrix} + z\begin{pmatrix}0\\0\\1\end{pmatrix}$. Plus généralement, pour tout n de $\mathbb{N}^\star$, on peut définir une famille de n vecteurs de $\mathbb{K}^n$ en posant

$$e_1 = \begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \; e_2 = \begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}, \ldots, \; e_n = \begin{pmatrix}0\\\vdots\\0\\1\end{pmatrix}.$$

La famille $(e_1, e_2, \ldots e_n)$ est alors génératrice de $\mathbb{K}^n$.

Exemple 1.48. — Pour $n \in \mathbb{N}^\star$, la famille $(1, X, X^2, ...X^n)$ engendre $\mathbb{K}_n[X]$ (voir page 11), tandis que la famille infinie $(X^i)_{i\in\mathbb{N}}$ engendre $\mathbb{K}[X]$ (voir l'exemple (b) page 12).

Exemple 1.49. — La famille $\left(\begin{pmatrix}1&0\\0&0\end{pmatrix}, \begin{pmatrix}0&1\\0&0\end{pmatrix}, \begin{pmatrix}0&0\\1&0\end{pmatrix}, \begin{pmatrix}0&0\\0&1\end{pmatrix}\right)$ engendre $\mathcal{M}_2(\mathbb{K})$. En effet, tout élément de $\mathcal{M}_2(\mathbb{K})$ peut s'écrire $\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix}$ avec $a_{11}, a_{12}, a_{21}, a_{22}$ dans $\mathbb{K}$, et $\begin{pmatrix}a_{11}&a_{12}\\a_{21}&a_{22}\end{pmatrix} = a_{11}\begin{pmatrix}1&0\\0&0\end{pmatrix} + a_{12}\begin{pmatrix}0&1\\0&0\end{pmatrix} + a_{21}\begin{pmatrix}0&0\\1&0\end{pmatrix} + a_{22}\begin{pmatrix}0&0\\0&1\end{pmatrix}$.

Plus généralement, pour tout (n,p) de $(\mathbb{N}^\star)^2$ et pour tout (i,j) de $\{1, ..n\}\times\{1, ...p\}$, notons E_{ij} la matrice de taille $n \times p$ dont tous les coefficients valent zéro sauf celui situé en position (i,j), qui vaut 1. La famille $(E_{ij})_{\substack{1\le i\le n\\1\le j\le p}}$ comporte np matrices. Comme toute matrice $A = (a_{ij})_{1\le i\le n 1\le j\le p}$ à coefficients dans $\mathbb{K}$ peut s'écrire $A = \sum\limits_{\substack{1\le i\le n\\1\le j\le p}} a_{ij}E_{ij}$, la famille $(E_{ij})_{\substack{1\le i\le n\\1\le j\le p}}$ engendre $\mathcal{M}_{n,p}(\mathbb{K})$.

Exemple 1.50. — Dans le $\mathbb{R}$-espace vectoriel $\mathbb{C}$, la famille $(1, i)$ est génératrice : tout élément de $\mathbb{C}$ peut s'écrire comme combinaison linéaire *à coefficients réels* de 1 et de i. En revanche, la famille (i) n'engendre que l'ensemble des nombres imaginaires purs. Dans le $\mathbb{C}$-espace vectoriel $\mathbb{C}$, la famille (i) (qui comporte un seul vecteur) est une famille génératrice. Il faut bien comprendre la différence avec l'exemple précédent : si on voit $\mathbb{C}$ comme un $\mathbb{C}$-espace vectoriel, les combinaisons *complexes* sont autorisées.

6.2. Un exemple concret détaillé. —

Vérifions, uniquement à partir de la définition, que la famille $\left(\begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\0\\-1\end{pmatrix}, \begin{pmatrix}0\\2\\2\end{pmatrix}, \begin{pmatrix}0\\0\\1\end{pmatrix}\right)$ engendre $\mathbb{R}^3$.

Ce qu'il faut montrer, c'est que tout vecteur de $\mathbb{R}^3$ peut s'écrire comme combinaison linéaire des vecteurs de la famille.

Soit $\begin{pmatrix}x\\y\\z\end{pmatrix}$ un vecteur quelconque de $\mathbb{R}^3$. Nous voulons vérifier qu'il existe des réels $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ tels qu'on ait

$$\begin{pmatrix}x\\y\\z\end{pmatrix} = \lambda_1\begin{pmatrix}1\\1\\1\end{pmatrix} + \lambda_2\begin{pmatrix}1\\0\\-1\end{pmatrix} + \lambda_3\begin{pmatrix}0\\2\\2\end{pmatrix} + \lambda_4\begin{pmatrix}0\\0\\1\end{pmatrix}$$

autrement dit

$$\begin{cases}\lambda_1 + \lambda_2 & = x,\\ \lambda_1 \qquad + 2\lambda_3 & = y,\\ \lambda_1 - \lambda_2 + \lambda_3 + \lambda_4 & = z.\end{cases}$$

Nous devons donc étudier *l'existence* d'une solution à ce *système linéaire* dont les *inconnues* sont $\lambda_1, \lambda_2, \lambda_3, \lambda_4$.

Si nous appliquons la procédure habituelle (« méthode du pivot »), nous avons

$$\begin{cases} \lambda_1 + \lambda_2 & = x \\ \lambda_1 \quad + 2\lambda_3 & = y \\ \lambda_1 - \lambda_2 + \lambda_3 + \lambda_4 & = z \end{cases} \iff \begin{cases} \lambda_1 + \lambda_2 & = x \\ \quad -\lambda_2 + 2\lambda_3 & = y - x \\ \quad -2\lambda_2 + \lambda_3 + \lambda_4 & = z - x \end{cases} \iff \begin{cases} \lambda_1 + \lambda_2 & = x \\ \quad -\lambda_2 + 2\lambda_3 & = y - x \\ \quad\quad -3\lambda_3 + \lambda_4 & = x - 2y + z \end{cases}$$

On constate alors qu'une solution possible est donnée par $\begin{cases} \lambda_3 = 0 \\ \lambda_4 = -x + 2y - z \\ \lambda_2 = x - y \\ \lambda_1 = y \end{cases}$ (attention, ce n'est pas la seule solution).

Ainsi, il y a *toujours au moins une solution* au système envisagé : effectivement, on a bien

$$y \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + (x - y) \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + 0 \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} + (x - 2y + z) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Conclusion : tous les vecteurs de $\mathbb{R}^3$ peuvent s'écrire comme combinaison linéaire de notre famille, et notre famille est génératrice de $\mathbb{R}^3$.

Exercices de manipulation. — I.11, I.12, I.13

6.3. Ajout ou retrait de vecteurs à une famille génératrice. —

Bien sûr, la notion de famille génératrice est très directement liée à celle de sous-espace vectoriel engendré par une famille. Voici par exemple une reformulation immédiate de la dernière proposition que nous avons rencontrée au numéro 1.6.

Proposition 1.51 – Ajouter ou retirer des vecteurs à une famille génératrice

a. ***(Ajouter des vecteurs à une famille génératrice)***
Une famille qui contient une partie génératrice est elle-même génératrice.

b. ***(Ôter un vecteur* redondant *à une famille génératrice).*** *Soit A une partie de E et x un élément de E.*
Si $A \cup \{x\}$ engendre E et si x est combinaison linéaire des vecteurs de A, alors A elle-même engendre E.

Exemple 1.52. — Nous avons vu que la famille $\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$ engendre $\mathbb{R}^3$.

Mais le vecteur $\begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}$ est combinaison linéaire des autres vecteurs de la famille : on a $\begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - 2 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} - 2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Par conséquent, $\left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right)$ engendre $\mathbb{R}^3$.

L'exemple précédent nous mène à la notion suivante, qui prendra une grande importance dans le chapitre 2 :

Définition 1.53 – Famille génératrice minimale

Une famille génératrice est dite *minimale* lorsqu'il est impossible de lui ôter un vecteur sans détruire son caractère générateur.

Exercice de manipulation. — I.14.

7. Familles libres, indépendance linéaire

7.1. Définition, premiers exemples, interprétation de la liberté. —

Définition 1.54 – Famille libre dans un espace vectoriel

Soit E un $\mathbb{K}$-espace vectoriel.
Cas d'une famille finie. Soit n un entier naturel non nul et $(x_1, x_2, ..., x_n)$ une famille de n vecteurs de E.
On dit que $(x_i)_{i\in I}$ est *libre* si pour tout $(\lambda_1, \lambda_2, ...\lambda_n)$ de $\mathbb{K}^n$, on a

$$(\lambda_1 x_1 + \lambda_2 x_2 + ... + \lambda_n x_n = \mathbf{0}_E) \implies (\lambda_1 = \lambda_2 = ... = \lambda_n = 0)\,.$$

Cas général. Soit $(x_i)_{i\in I}$ une famille de vecteurs de E.
On dit que $(x_i)_{i\in I}$ est *libre* si toute sous-famille finie de $(x_i)_{i\in I}$ est libre.

Remarque 1.55. — Par convention, la famille vide est libre.

***Exemple 1.56* (Un exemple typique dans $\mathbb{R}^3$).** — Dans $\mathbb{R}^3$, la famille $\left(\begin{pmatrix}1\\2\\4\end{pmatrix}, \begin{pmatrix}3\\7\\2\end{pmatrix}, \begin{pmatrix}1\\-5\\1\end{pmatrix}\right)$ est libre.

En effet, soient $\lambda_1, \lambda_2, \lambda_3$ trois réels. Supposons l'égalité

$$\lambda_1 \begin{pmatrix}1\\2\\4\end{pmatrix} + \lambda_2 \begin{pmatrix}3\\7\\2\end{pmatrix} + \lambda_3 \begin{pmatrix}1\\-5\\1\end{pmatrix} = \mathbf{0}_{\mathbb{R}^3}$$

et montrons que $\lambda_1 = \lambda_2 = \lambda_3 = 0$. L'égalité que nous avons supposée signifie

$$\begin{cases} \lambda_1 + 3\lambda_2 + \lambda_3 & = 0 \\ 2\lambda_1 + 7\lambda_2 - 5\lambda_3 & = 0 \\ 4\lambda_1 + 2\lambda_2 + \lambda_3 & = 0. \end{cases}$$

Ce système équivaut (méthode de Gauss) à

$$\begin{cases} \lambda_1 + 3\lambda_2 + \lambda_3 & = 0 \\ \lambda_2 - 7\lambda_3 & = 0 \\ -10\lambda_2 - 3\lambda_3 & = 0 \end{cases} \quad \text{ou encore} \quad \begin{cases} \lambda_1 + 3\lambda_2 + \lambda_3 & = 0 \\ \lambda_2 - 7\lambda_3 & = 0 \\ -73\lambda_3 & = 0. \end{cases}$$

La seule possibilité est d'avoir $\lambda_1 = \lambda_2 = \lambda_3 = 0$, ce qu'il fallait démontrer.

***Exemple 1.57* (Une famille de matrices).** — Dans $\mathcal{M}_2(\mathbb{R})$, la famille $\left(\begin{pmatrix}1 & 2\\0 & 1\end{pmatrix}, \begin{pmatrix}2 & 5\\-3 & 0\end{pmatrix}, \begin{pmatrix}-1 & 4\\0 & 0\end{pmatrix}\right)$ est libre. En effet, soient λ, μ, ν trois réels. Supposons l'égalité

$$\lambda \begin{pmatrix}1 & 2\\0 & 1\end{pmatrix} + \mu \begin{pmatrix}2 & 5\\-3 & 0\end{pmatrix} + \nu \begin{pmatrix}-1 & 4\\0 & 0\end{pmatrix} = \mathbf{0}_{\mathcal{M}_2(\mathbb{R})}.$$

On a alors $\begin{pmatrix} \lambda+2\mu-\nu & 2\lambda+2\mu+4\nu \\ -3\mu & \lambda \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$, donc $\begin{cases} \lambda+2\mu-\nu=0 \\ -3\mu=0 \\ 2\lambda+2\mu+4\nu=0 \\ \lambda=0 \end{cases}$

et cela implique $\lambda=\mu=\nu=0$.

***Exemple 1.58* (Une famille de fonctions).** — Dans $\mathcal{F}(\mathbb{R},\mathbb{R})$, la famille $(\sin,\cos)$ est libre. En effet, soient λ,μ deux réels. Supposons l'égalité

$$\lambda\sin+\mu\cos=\mathbf{0}_{\mathcal{F}(\mathbb{R},\mathbb{R})}.$$

Attention, nous travaillons dans $\mathcal{F}(\mathbb{R},\mathbb{R})$, donc ceci est une égalité de *fonctions*, qui signifie :
$\forall x\in\mathbb{R}, \lambda\sin(x)+\mu\cos(x)=0_{\mathbb{R}}$.

En l'évaluant en zéro, on obtient : $\lambda\sin(0)+\mu\cos(0)=0$, et comme $\sin(0)=0$ et $\cos(0)=1$, cela signifie $\mu=0$.

En l'évaluant en $\frac{\pi}{2}$, on obtient $\lambda\sin(\frac{\pi}{2})+\mu\cos(\frac{\pi}{2})=0$, et comme $\sin(\frac{\pi}{2})=1$ et $\cos(\frac{\pi}{2})=1$, cela signifie $\lambda=0$.

On en déduit que λ et μ sont tous deux nuls ; cela montre bien que $(\sin,\cos)$ est libre.

Exercices de manipulation. — I.15, I.16

Lorsqu'une famille n'est pas libre, on dit aussi qu'elle est *liée.*

Proposition 1.59 – Comment voir « immédiatement » qu'une famille est liée

a. *Toute famille contenant le vecteur nul est liée.*

b. *Si l'un des vecteurs de $(x_i)_{i\in I}$ est combinaison linéaire des autres vecteurs de $(x_i)_{i\in I}$, alors $(x_i)_{i\in I}$ est liée.*

Exemple 1.60. — Dans $\mathbb{R}^3$, la famille $\left(\begin{pmatrix}1\\2\\3\end{pmatrix},\begin{pmatrix}4\\5\\6\end{pmatrix},\begin{pmatrix}6\\9\\12\end{pmatrix}\right)$ est liée, car $\begin{pmatrix}7\\8\\9\end{pmatrix}=2\begin{pmatrix}1\\2\\3\end{pmatrix}+\begin{pmatrix}4\\5\\6\end{pmatrix}$.

Exercices de manipulation. — I.17

Proposition 1.61 – Trois interprétations de la liberté

Les trois assertions suivantes sont équivalentes.

(a) La famille $(x_i)_{i\in I}$ est libre.

(b) ***Aucun vecteur de la famille n'est combinaison linéaire des autres :***

pour tout i_0 de I, $x_{i_0} \notin \mathbf{Vect}[(x_i)_{i\in I, i\neq i_0}]$.

(c) ***Unicité de l'écriture des combinaisons linéaires :***

si $(\alpha_i)_{i\in I}$ et $(\beta_i)_{i\in I}$ sont deux familles à support fini *de scalaires, alors*

$$\left(\sum_{i\in I}\alpha_i x_i = \sum_{i\in I}\beta_i x_i\right) \implies (\forall i\in I, \alpha_i = \beta_i)$$

Remarque 1.62. — • À cause de la propriété (b), dire que des vecteurs forment une famille libre, c'est dire qu'il n'existe pas de *relation de dépendance linéaire* entre eux. On dit donc aussi qu'ils sont *linéairement indépendants.*

- La propriété (c) sous-tend beaucoup des démarches « par identification » que vous avez déjà rencontrées.

Démonstration de la proposition. — Montrons que (a) implique (b), que (b) implique (c), que (c) implique (a).

- Vérifions que non(b) implique non(a) : s'il existe i_0 de I tel qu'on ait $x_{i_0} \in \mathbf{Vect}[x_i, i\neq i_0]$, alors il existe une sous-famille finie $(i_1, ... i_n)$ de $(x_i, i\neq i_0)$ et des scalaires $\lambda_{i_1}, ... \lambda_{i_n}$ tels qu'on ait
$$x_{i_0} = \sum_{k=1}^{n} \lambda_{i_k} x_{i_k}$$
autrement dit
$$x_{i_0} - \lambda_{i_1}x_{i_1} - ... - \lambda_{i_n}x_{i_n} = 0$$
ce qui fournit une combinaison linéaire de $(x_i)_{i\in I}$ dont les coefficients ne sont pas tous nuls et qui donne le vecteur nul, donc prouve que $(x_i)_{i\in I}$ est liée.
- Montrons que non(c) implique non(b). Supposons non(c), ce qui signifie qu'il existe deux familles à support fini (α_i) et (β_i) de scalaires telles qu'on ait $\sum_{i\in I}\alpha_i x_i = \sum_{i\in I}\beta_i x_i$, mais qu'il y ait un i_0 de I pour lequel $\alpha_{i_0} \neq \beta_{i_0}$. Dans ce cas on a $\sum_{i\in I}(\alpha_i - \beta_i)x_i = 0$, mais $\sum_{i\in I}(\alpha_i-\beta_i)x_i$ peut s'écrire $(\alpha_{i_0} - \beta_{i_0})x_{i_0} + \sum_{i\neq i_0}(\alpha_i - \beta_i)x_i$, si bien que :
$$x_{i_0} = \frac{-1}{\alpha_{i_0} - \beta_{i_0}} \sum_{i\neq i_0}(\alpha_i - \beta_i)x_i.$$
La famille $(\alpha_i - \beta_i)_{i\in I}$ est une famille à support fini de scalaires, le membre de droite est donc une combinaison linéaire de la famille $(x_i, i\neq i_0)$. L'égalité ci-dessus montre donc que x_{i_0} appartient à $\mathbf{Vect}[(x_i)_{i\in I, i\neq i_0}]$. Ainsi non(b) est vérifiée.
- Enfin, montrons que non(a) implique non(c) : si $(x_i)_{i\in I}$ est liée, alors il existe une famille à support fini $(\alpha_i)_{i\in I}$ telle que les α_i ne soient pas tous nuls, mais qu'on ait $\sum_{i\in I}\alpha_i x_i = 0$. Si $(\beta_i)_{i\in I}$ est la famille de scalaires où il n'y a que des zéros, on a $\left(\sum_{i\in I}\alpha_i x_i = \sum_{i\in I}\beta_i x_i\right)$ sans que les familles $(\alpha_i)_{i\in I}$ et $(\beta_i)_{i\in I}$ ne coïncident.

□

7.2. Exemples généraux de familles libres. —

7.2.1. *Familles comportant un seul vecteur non nul.* —

Si x est un vecteur *non nul* de E, alors la famille $\{x\}$ est libre.

En effet, nous avons vu (propriété (c) page **??**) que si λ est un scalaire et $\lambda x = \mathbf{0}_E$, alors $\lambda = 0$.

7.2.2. *Familles de deux vecteurs non colinéaires.* —

Soient x et y deux vecteurs de E.
Dire que la famille (x, y) est libre, c'est dire que les vecteurs x et y ne sont pas colinéaires.

Nous l'avons déjà vu page 22.

7.2.3. *Familles de p vecteurs de $\mathbb{K}^n$ et méthode « du pivot ».* — Soit $(x_1, ...x_p)$ une famille de p vecteurs de $\mathbb{K}^n$. Écrivons-les sous la forme

$$x_1 = \begin{pmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{n1} \end{pmatrix}, ..., x_p = \begin{pmatrix} x_{1p} \\ x_{2p} \\ \vdots \\ x_{np} \end{pmatrix},$$

où les x_{ij}, $(i, j) \in \{1, ..n\} \times \{1, ..p\}$ sont des éléments de $\mathbb{K}$.

Savoir si la famille $(x_1, ...x_p)$ est libre dans $\mathbb{K}^n$, c'est savoir s'il existe des scalaires $\lambda_1, ..., \lambda_p$ qui vérifient $\lambda_1 x_1 + ... + \lambda_p x_p = 0$ sans être tous nuls. C'est donc savoir si le système linéaire

$$\begin{cases} x_{11}\lambda_1 + ... + x_{1p}\lambda_p & = 0 \\ & \vdots \\ x_{n1}\lambda_1 + ... + x_{np}\lambda_p & = 0 \end{cases}$$

d'inconnues $\lambda_1, ..., \lambda_p$, admet une autre solution que $\lambda_1 = \lambda_2 = ... = \lambda_p = 0$.

Cela ramène la question à la résolution d'un système linéaire.

Proposition 1.63 – Liberté d'une famille de polynômes à degrés distincts

Soit $(P_i)_{i \in I}$ une famille de polynômes non nuls à coefficients dans $\mathbb{K}$.
Si les degrés de ces polynômes sont deux à deux distincts, c'est-à-dire si pour $i \neq j$, on a $\deg(P_i) \neq \deg(P_j)$,
alors la famille $(P_i)_{i \in I}$ est libre dans $\mathbb{K}[X]$.

Exemple 1.64. —

Attention. — La réciproque est fausse, une famille de polynômes peut être libre sans être à degrés deux à deux distincts : voir l'exercice 28(3) de la première feuille de TD et les célèbres *polynômes de Lagrange* au chapitre 2

Démonstration de la proposition.. — Nous devons montrer que toute sous-famille finie de $(P_i)_{i \in I}$ est libre.

- Pour chaque k de $\mathbb{N}$, notons $\mathcal{P}(k)$ l'énoncé

« Si $(P_i)_{i\in J}$ est une sous-famille de $(P_i)_{i\in I}$ telle que pour tout i de J, on ait $\deg P_i \leq k$, alors $(P_i)_{i\in J}$ est libre ».

Si nous parvenons à prouver que cet énoncé est vrai pour tout k de $\mathbb{N}$, alors en partant d'une sous-famille finie quelconque $(P_i)_{i\in J}$ de $(P_i)_{i\in I}$, on pourra poser $k = \max\limits_{i\in J}(\deg P_i)$, et appliquer $\mathcal{P}(k)$ montrera sa liberté.

• L'énoncé $\mathcal{P}(0)$ est vrai : si $(P_i)_{i\in J}$ est une sous-famille de $(P_i)_{i\in I}$ telle que pour tout i de J, on ait $\deg P_i \leq 0$, on doit avoir $\deg P_i = 0$ pour tout i de J (rappel : tous les P_i sont non-nuls). À cause de l'hypothèse sur les degrés distincts, ce n'est possible que si $(P_i)_{i\in J}$ est soit vide, soit formée d'un seul polynôme non-nul. Dans les deux cas, elle est libre.

• Soit k un entier naturel. Supposons $\mathcal{P}(k)$ vrai.
Soit $(P_i)_{i\in J}$ une sous-famille de $(P_i)_{i\in I}$ telle que pour tout i de J, on ait $\deg(P_i) \leq (k+1)$. Montrons que $(P_i)_{i\in J}$ est libre.

- Si on a en fait $\deg(P_i) \leq k$ pour tout i de J, alors $\mathcal{P}(k)$ permet déjà de conclure que $(P_i)_{i\in J}$ est libre.
- Sinon, il existe un i_0 de I tel que P_{i_0} soit de degré $k+1$, et alors tous les P_i, $i \in J - \{i_0\}$, sont de degré inférieur à k.

  Soit $(\lambda_i)_{i\in J}$ une famille de scalaires vérifiant $\sum\limits_{i\in J} \lambda_i P_i = 0$. Nous devons montrer que tous les λ_i, $i \in J$, sont nuls.

  Le polynôme $\sum\limits_{i\in J} \lambda_i P_i$ est une somme de polynômes de degrés au plus $k+1$, il est donc de degré au plus $k+1$;

  de plus, son coefficient dominant est λ_{i_0}, puisque la seule contribution au coefficient de X^{k+1} y est celle de P_{i_0}.

  Ainsi $\lambda_{i_0} = 0$, et de plus $\sum\limits_{i\in J-\{i_0\}} \lambda_i P_i = 0$. D'après $\mathcal{P}(k)$, la famille $(P_i)_{i\in J-\{i_0\}}$ est libre, donc tous les λ_i sont bien nuls. Nous avons prouvé $\mathcal{P}(k+1)$.

$\square$

7.2.4. *Familles de matrices avec des coefficients en positions complémentaires.* —

Proposition 1.65 – Famille de matrices avec coefficients en positions complémentaires

Soit $(A^{[\alpha]})_{\alpha\in A}$ *une famille d'éléments non nuls de* $\mathcal{M}_{np}(\mathbb{K})$. *Pour chaque* α *de* A, *notons* $A^{[\alpha]} = (c_{ij}^{[\alpha]})_{\substack{1\leq i\leq n \\ 1\leq j\leq p}}$.
Si pour chaque (i,j) *de* $\{1,..,n\} \times \{1,...,p\}$, *il y a* ***au plus*** *un* α *de* A *tel que* $c_{ij}^{[\alpha]}$ *soit non nul, alors la famille* $(A^{[\alpha]})_{\alpha\in A}$ *est libre dans* $\mathcal{M}_{np}(\mathbb{K})$.

Exemple 1.66. —

- Dans $\mathcal{M}_2(\mathbb{R})$, la famille $\left(\begin{pmatrix} 0 & 8 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -3 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & \pi \end{pmatrix}\right)$ est libre.

- Dans $\mathcal{M}_{23}(\mathbb{R})$, la famille $\left(\begin{pmatrix} 1 & 0 & 5 \\ 0 & 0 & -7 \end{pmatrix}, \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 99 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & -0.76 & 0 \end{pmatrix}\right)$ est libre.

Attention. — La réciproque est bien sûr fausse : dans l'exemple 1.57, nous avons rencontré une famille libre de matrices 2×2 dont les coefficients ne sont pas en positions complémentaires.

Démonstration de la proposition. — Soit $(\lambda_\alpha)_{\alpha\in A}$ une famille à support fini de scalaires. Supposons $\sum_{\alpha\in A}\lambda_\alpha A_\alpha = 0$. Dans la matrice $M = \sum_{\alpha\in A}\lambda_\alpha A_\alpha$, chaque coefficient est soit un zéro, soit de la forme $c_{ij}^\alpha\lambda_\alpha$ pour un (et un seul) des λ_α. Chacun des λ_α apparaît dans au moins un des coefficients de M, puisqu'aucune des matrices A_α n'est nulle. Ainsi tous les λ_α sont nuls. □

7.2.5. *Famille de fonctions à supports deux à deux disjoints.* —

Proposition 1.67 – Liberté d'une famille de fonctions à supports disjoints

Soit $(f_i)_{i\in I}$ une famille de fonctions non nulles de $\mathbb{R}$ dans $\mathbb{R}$. Si pour tout x de $\mathbb{R}$, il existe au plus un i de I tel que $f_i(x)$ soit non-nul, alors la famille $(f_i)_{i\in I}$ est libre dans $\mathcal{F}(\mathbb{R},\mathbb{R})$.

Exemple 1.68. — Pour chaque a de $\mathbb{R}$, définissons $\delta_a : \mathbb{R}\to\mathbb{R}$ en posant, pour tout x de $\mathbb{R}$, $\delta_a(x) = \begin{cases} 1 \text{ si } x = a, \\ 0 \text{ sinon.}\end{cases}$

La famille $(\delta_a)_{a\in\mathbb{R}}$ est libre dans $\mathcal{F}(\mathbb{R},\mathbb{R})$.

Exemple 1.69. — Pour chaque k de $\mathbb{Z}$, définissons $\varphi_k : \mathbb{R}\to\mathbb{R}$ en posant, pour tout x de $\mathbb{R}$, $\varphi_k(x) = \begin{cases} \exp\left(\frac{1}{(x-k)(k+1-x)}\right) & \text{si } x\in]k,k+1[, \\ 0 \text{ sinon.}\end{cases}$ Chacune des fonctions φ_k, $k\in\mathbb{Z}$, est continue (et même $\mathcal{C}^\infty$) sur $\mathbb{R}$. La famille $(\varphi_k)_{k\in\mathbb{Z}}$ est libre dans $\mathcal{F}(\mathbb{R},\mathbb{R})$.

Attention. — La réciproque est bien sûr fausse : dans l'exemple (c) page 29, nous avons rencontré une famille libre de fonctions dont les supports ne sont pas deux à deux disjoints.

Démonstration de la proposition. — Soit $(\lambda_i)_{i\in I}$ une famille à support fini de scalaires. Supposons $\sum_{i\in I}\lambda_i f_i = 0$. Les hypothèses sur $(f_i)_{i\in I}$ disent que pour chaque i de I, il existe exactement [5] un x_i de $\mathbb{R}$ tel que $f_i(x_i)\neq 0$.

Si on note g la fonction $\sum_{i\in I}\lambda_i f_i$, on a alors $g(x_i) = f_i(x_i)\lambda_i$.

L'hypothèse $\sum_{i\in I}\lambda_i f_i = 0$ signifie que g est la fonction nulle sur $\mathbb{R}$; ainsi pour tout i de I on a $g_i(x_i) = 0$, mais comme $g(x_i) = f_i(x_i)\lambda_i$ avec $f_i(x_i)\neq 0$, tous les λ_i doivent être nuls. □

7.3. Ajout ou retrait de vecteurs à une famille libre. —

5. « Au plus un » vient du fait qu'en un point x de $\mathbb{R}$, on ne peut pas avoir deux des f_i qui soient simultanément non-nulles, et « au moins un » vient du fait que f_i n'est pas la fonction nulle.

Proposition 1.70 – Effet sur la liberté de l'ajout ou du retrait d'un vecteur

a. ***Ôter des vecteurs à une famille libre :***

Une sous-famille d'une famille libre est elle-même libre.

b. ***Ajouter un vecteur* indépendant *à une famille libre :*** *Soit A une partie de E et x un élément de E.*

*On suppose que A est libre. Alors $A \cup \{x\}$ est libre si et seulement si x n'appartient pas à **Vect**$[A]$.*

La seconde propriété signifie que pour conserver la liberté en ajoutant un vecteur, il faut et il suffit que le vecteur ajouté ne soit pas combinaison linéaire des vecteurs initiaux.

***Exemple 1.71* (Exemple d'ajout d'un vecteur indépendant).** — Dans $\mathbb{R}^3$, la famille $\left(\begin{pmatrix}1\\2\\1\end{pmatrix}, \begin{pmatrix}3\\-1\\3\end{pmatrix}\right)$ est libre, car c'est une famille de deux vecteurs et que ces deux vecteurs ne sont pas colinéaires.

De plus, pour que $\begin{pmatrix}x\\y\\z\end{pmatrix}$ puisse appartenir à **Vect** $\left[\begin{pmatrix}1\\2\\1\end{pmatrix}, \begin{pmatrix}3\\-1\\3\end{pmatrix}\right]$, il faut avoir $x = z$.

Ainsi, $\begin{pmatrix}1\\2\\3\end{pmatrix}$ n'appartient pas à **Vect** $\left[\begin{pmatrix}1\\2\\1\end{pmatrix}, \begin{pmatrix}3\\-1\\3\end{pmatrix}\right]$; par conséquent, la famille $\left(\begin{pmatrix}1\\2\\1\end{pmatrix}, \begin{pmatrix}3\\-1\\3\end{pmatrix}, \begin{pmatrix}1\\2\\3\end{pmatrix}\right)$ est libre.

Exercices de manipulation. — I.18, I.19

Démonstration de la proposition.. — a. Soit $(x_i)_{i\in I}$ une famille libre. Soit $(x_i)_{i\in I^\flat}$, où $I^\flat \subset I$, une sous-famille de $(x_i)_{i\in I}$. Soit $(\lambda_i)_{i\in I^\flat}$ une famille *presque nulle* de scalaires telle que $\sum\limits_{i\in I^\flat} \lambda_i x_i = 0$. Nous devons vérifier que tous les λ_i sont nuls.

Pour tout i de I, posons $\mu_i = \lambda_i$ si i appartient à $I^\flat$, $\mu_i = 0$ sinon. Alors la famille $(\mu_i)_{i\in I}$ est une famille à support fini de scalaires, et on a

$$\sum_{i\in I} \mu_i x_i = \sum_{i\in I^\flat} \lambda_i x_i = 0.$$

Puisque la famille $(x_i)_{i\in I}$ est libre, on en déduit que tous les μ_i sont nuls, donc tous les λ_i aussi.

b. Soit A une partie libre de E et x un élément de E.

• Vérifions que si x appartient à **Vect**$[A]$, alors $A\cup\{x\}$ n'est pas libre. Dire que x appartient à **Vect**$[A]$, c'est dire qu'il existe des vecteurs $y_1, ...y_k$ de A et des scalaires $\lambda_1, ..., \lambda_k$ tels que $x = \lambda_1 y_1 + ... + \lambda_k y_k$. Si c'est vrai, alors on a

$$-x + \lambda_1 y_1 + ... + \lambda_k y_k = 0$$

et ceci exprime le vecteur nul comme combinaison linéaire de vecteurs de $A \cup \{x\}$; les coefficients de la combinaison ne sont pas tous nuls, puisque le premier vaut -1. Ainsi la partie $A \cup \{x\}$ n'est pas libre.

• Réciproquement, vérifions que si $A \cup \{x\}$ n'est pas libre, alors x appartient à **Vect**$[A]$. Dire que la partie $A\cup\{x\}$ n'est pas libre, c'est dire qu'il existe des vecteurs $y_1, ...y_k$ de A et des scalaires $\lambda_0, \lambda_1, ..., \lambda_k$ *non tous nuls* tels qu'on ait

$$\lambda_0 x + \lambda_1 y_1 + ... + \lambda_k y_k = 0.$$

Si $\lambda_0 = 0$, alors $\lambda_1 y_1 + ... + \lambda_k y_k = 0$, ce qui n'est pas possible puisque A est libre, que tous les y_i appartiennent à A et que les λ_i ne peuvent pas être tous nuls. Ainsi $\lambda_0 \neq 0$, mais alors

$$x = -\frac{\lambda_1}{\lambda_0} y_1 - ... - \frac{\lambda_k}{\lambda_k} y_k$$

et x appartient à $\mathbf{Vect}[A]$.

□

Définition 1.72 – Famille libre maximale

Une famille libre est dite *maximale* lorsqu'il est impossible de lui ajouter un vecteur sans détruire sa liberté.

7.4. Critère de liberté pour les familles infinies indexées par $\mathbb{N}$. — Dans le cas où une famille $(x_i)_{i\in I}$ de E comporte une infinité de vecteurs, dire qu'elle est libre, c'est dire que pour tout p de $\mathbb{N}^\star$, pour toute sous-famille $(x_{i_1}, ... x_{i_p})$ à p éléments de $(x_i)_{i\in I}$ et pour tout $(\lambda_1, \lambda_2, ... \lambda_p)$ de $\mathbb{K}^p$, on a

$$\left(\lambda_1 x_{i_1} + \lambda_2 x_{i_2} + ... + \lambda_p x_{i_p} = \mathbf{0}_E\right) \implies (\lambda_1 = \lambda_2 = ... = \lambda_p = 0)\,.$$

Cela fait beaucoup de notations. Dans un cas important, celui où la famille est en fait donnée par une *suite* d'éléments de E, il est possible d'alléger les discussions :

Proposition 1.73 – Caractérisation de la liberté pour les familles indexées par $\mathbb{N}$

Soit $(x_i)_{i\in\mathbb{N}}$ une famille de vecteurs de E indexée par $\mathbb{N}$.
La famille $(x_i)_{i\in\mathbb{N}}$ est libre si et seulement si pour tout n de $\mathbb{N}^\star$, la famille $(x_0, ... x_n)$ est libre.

Démonstration. — Si nous supposons que $(x_i)_{i\in\mathbb{N}}$ est libre, alors pour tout n de $\mathbb{N}^\star$, la famille $(x_0, ... x_n)$ est une sous-famille *finie* de $(x_i)_{i\in\mathbb{N}}$, donc elle est libre (cela fait partie de la définition de « famille libre »). Réciproquement, supposons que pour tout n de $\mathbb{N}^\star$, la famille $(x_0, ... x_n)$ est libre. Pour montrer que $(x_i)_{i\in\mathbb{N}}$ est libre, considérons un élément p de $\mathbb{N}^\star$, une sous-famille $(x_{i_1}, ... x_{i_p})$ à p éléments de $(x_i)_{i\in\mathbb{N}}$ et des éléments $\lambda_1, \lambda_2, ... \lambda_p$ de $\mathbb{K}$. On peut supposer $i_1 \leq i_2 \leq ... \leq i_p$. Si on suppose

$$\lambda_1 x_{i_1} + \lambda_2 x_{i_2} + ... + \lambda_p x_{i_p} = \mathbf{0}_E$$

et si on pose $n = i_p$, alors le membre de gauche est une combinaison linéaire de la famille $(x_1, ... x_n)$ (à compléter par beaucoup de zéros pour les entiers entre 1 et n qui ne sont pas parmi les $i_k, k = 1..p$). Comme cette famille est libre, les coefficients de la combinaison sont tous nuls, et $\lambda_{i_1} = ... = \lambda_{i_p} = 0$. □

***Exemple 1.74* (Premier exemple d'utilisation).** — Plaçons-nous dans l'espace $\mathcal{F}(\mathbb{N}, \mathbb{R})$ des suites réelles, et reprenons la famille $(u^{[p]})_{p\in\mathbb{N}}$ définie à la page 12 :

$u^{[0]}$ est la suite dont les termes sont $1, 0, 0, 0, ...,$

$u^{[1]}$ est la suite dont les termes sont $0, 1, 0, 0, ...,$

$u^{[2]}$ est la suite dont les termes sont $0, 0, 1, 0, ...,$

etc.

Vérifions que cette famille est libre.

D'après la proposition, il suffit de vérifier que pour tout n de $\mathbb{N}^\star$, la famille $(u^{[1]}, u^{[2]}, ..., u^{[n]})$ est libre.

Soit n un élément de $\mathbb{N}^\star$ et $\lambda_0, \lambda_1, ...\lambda_n$ une famille de n nombres réels. Supposons l'égalité

$$\lambda_0 u^{[0]} + \lambda_1 u^{[1]} + ... + \lambda_n u^{[n]} = 0.$$

C'est une égalité de *suites* : elle affirme que la suite $v = \lambda_0 u^{[0]} + \lambda_1 u^{[1]} + ... + \lambda_n u^{[n]}$ est nulle.

La suite v vérifie : $v_0 = \lambda_0 \cdot 1 + \lambda_0 \cdot 0 + ... + \lambda_n \cdot 0 = \lambda_0$.

De même, $v_1 = \lambda_1$, $v_2 = \lambda_2$, ..., $v_n = \lambda_n$.

De plus, pour tout $k > n$, on a $v_k = 0$.

Dire que la suite v est nulle, c'est dire que tous ses termes sont nuls ; ainsi, $\lambda_0 = \lambda_1 = ... = \lambda_n = 0$, ce qu'il fallait démontrer.

Exercice de manipulation : — I.20

Exemple 1.75 **(Plus difficile : indépendance des sinus de fréquences différentes)**
La proposition ci-dessus permet, entre autres, de procéder par récurrence pour démontrer qu'une famille indexée par $\mathbb{N}$ est libre. Voici un exemple.

Pour tout k de $\mathbb{N}^\star$, notons s_k la fonction $x \mapsto \sin(kx)$ de $\mathbb{R}$ dans $\mathbb{R}$.
Alors la famille $(s_k)_{k\in\mathbb{N}}$ est libre dans $\mathcal{F}(\mathbb{R},\mathbb{R})$.

Pour chaque n de $\mathbb{N}^\star$, notons $\mathcal{P}(n)$ l'assertion « la famille $(s_0, s_1, s_2, ...s_n)$ est libre dans $\mathcal{F}(\mathbb{R},\mathbb{R})$ ».

L'assertion $\mathcal{P}(1)$ est bien sûr vraie, car la famille (s_0) comporte un seul vecteur de $\mathcal{F}(\mathbb{R},\mathbb{R})$ et que ce vecteur est non-nul.

Soit maintenant n un élément de $\mathbb{N}^\star$; supposons $\mathcal{P}(n)$ vraie. Soient $\lambda_1, ...\lambda_{n+1}$ des nombres réels. Supposons l'égalité

$$\sum_{k=1}^{n+1} \lambda_k s_k = \mathbf{0}_{\mathcal{F}(\mathbb{R},\mathbb{R})} \tag{7.1}$$

et rappelons que c'est une égalité de *fonctions*. Si l'on dérive deux fois (c'est possible !), et si l'on se rappelle que la dérivée de $s_k : x \mapsto \sin(kx)$ est $x \mapsto k\cos(kx)$, qui se dérive à son tour en $x \mapsto -k^2 \sin(kx)$, on obtient

$$\sum_{k=1}^{n+1} \lambda_k \cdot (k^2 \cdot s_k) = \mathbf{0}_{\mathcal{F}(\mathbb{R},\mathbb{R})}. \tag{7.2}$$

En soustrayant $(n+1)^2$ fois (7.1) à (7.2), le terme $\lambda_{n+1} \cdot (n+1)^2 \cdot s_{n+1}$ disparaît et il vient

$$\sum_{k=1}^{n} ((n+1)^2 - k^2) \cdot \lambda_k \cdot s_k) = \mathbf{0}_{\mathcal{F}(\mathbb{R},\mathbb{R})} \tag{7.3}$$

(on notera que la somme s'arrête au rang n, et que $((n+1)^2 - k^2))$ n'est jamais nul si $k \leq n$). Comme nous avons supposé $\mathcal{P}(n)$ vraie, cela implique $\lambda_1 = \lambda_2 = ... = \lambda_n = 0$. Si on observe alors ce que (7.1) donne en $x = \frac{\pi}{2(n+1)}$, on obtient dans la foulée $\lambda_{n+1} = 0$.

Nous avons bien vérifié $\mathcal{P}(n+1)$. On en déduit (par récurrence) que $\mathcal{P}(n)$ est vraie pour tout n de $\mathbb{N}^\star$; c'est ce qu'il fallait montrer.

Exercices de manipulation du cours du chapitre 1

Exercice de manipulation 1.1. —

$$\text{Vérifier que } \left\{ \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} \in \mathbb{R}^4, \quad \begin{array}{r} x + 2y + z = 0 \\ -x + 3t = 0 \end{array} \right\} \text{ est un sous-espace vectoriel de } \mathbb{R}^4.$$

Exercice de manipulation 1.2. —
Vérifier que $\{P \in \mathbb{R}[X],\ P(-7) = 0\}$ est un sous-espace vectoriel du $\mathbb{R}$-espace vectoriel $\mathbb{R}[X]$,
mais que $\{P \in \mathbb{R}[X],\ P(-7) = 1\}$ n'en est pas un.

Exercice de manipulation 1.3. —

(a) Dans $\mathbb{R}^2$, vérifier que les sous-ensembles suivants ne sont pas des sous-espaces vectoriels :

i. les *cercles* $\big\{(x, y) \in \mathbb{R}^2, x^2 + y^2 = 1\big\}$ et $\big\{(x, y) \in \mathbb{R}^2, (x-1)^2 + y^2 = 1\big\}$.

ii. la *demi-droite* $\big\{(x, y) \in \mathbb{R}^2, y = x \text{ et } x \geq 0\big\}$, la *droite épointée* $\big\{(x, y) \in \mathbb{R}^2, y = x \text{ et } x \neq 0\big\}$,

iii. l'*ellipse* $\Big\{(x, y) \in \mathbb{R}^2, \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\Big\}$, la *parabole* $\big\{(x, y) \in \mathbb{R}^2, y = x^2\big\}$,

(b) Dans $\mathbb{R}^3$, vérifier que les sous-ensembles suivants ne sont pas des sous-espaces vectoriels :

i. le *tore* $\Big\{(x, y, z) \in \mathbb{R}^3,\ (4 - \sqrt{x^2 + y^2})^2 + z^2 = 1\Big\}$,

ii. le *cône* $\big\{(x, y, z) \in \mathbb{R}^3,\ xy - z^2 = 0\big\}$.

Exercice de manipulation 1.4. —
Vérifier les affirmations suivantes (voir page 19) :

- Dans $\mathcal{F}(\mathbb{R}, \mathbb{R})$, l'espace des *fonctions qui valent zéro en zéro* est un sous-espace vectoriel,
  mais l'espace des *fonctions positives* n'est pas un sous-espace vectoriel.
- Dans $\mathcal{F}(\mathbb{N}, \mathbb{R})$, l'espace des suites qui *convergent vers zéro* est un sous-espace vectoriel,
  mais l'espace des suites qui *convergent vers* $\sqrt{2}$ n'est pas un sous-espace vectoriel,
  et l'espace des suites qui *convergent vers un nombre entier* n'est pas un sous-espace vectoriel.

Exercice de manipulation 1.5. —
On se place dans $\mathbb{R}^3$.

(a) Montrer que si $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ appartient à $\mathbf{Vect}\left[\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix}\right]$, alors $x + 2y - 3z = 0$.

(b) Montrer réciproquement que si $x + 2y - 3z = 0$, alors $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ appartient à $\mathbf{Vect}\left[\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix}\right]$.

Exercice de manipulation 1.6. —
Montrer que dans $\mathbb{R}^3$, $\begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix}$ appartient à $\mathbf{Vect}\left[\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}\right]$ mais pas à $\mathbf{Vect}\left[\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ -\pi \\ -\pi \end{pmatrix}, \begin{pmatrix} 6 \\ 1 \\ 1 \end{pmatrix}\right]$.

Exercice de manipulation 1.7. —
Dans $\mathcal{M}_2(\mathbb{R})$, quel est le sous-espace $\mathbf{Vect}\left[\begin{pmatrix}1&0\\0&0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix},\begin{pmatrix}0&1\\1&0\end{pmatrix}\right]$?

Exercice de manipulation 1.8. —
On note $F=\left\{\begin{pmatrix}x\\y\\z\\t\end{pmatrix}\in\mathbb{R}^4,\quad \begin{array}{l}x+2y+3z-t=0\\x+y-z=0\end{array}\right\}$.

(a) Pourquoi F est-il un sous-espace vectoriel de $\mathbb{R}^4$?

(b) Montrer l'inclusion $\mathbf{Vect}\left[\begin{pmatrix}0\\1\\1\\5\end{pmatrix},\begin{pmatrix}-1\\1\\0\\1\end{pmatrix},\begin{pmatrix}1\\-2\\-1\\-6\end{pmatrix}\right]\subset F$.

Exercice de manipulation 1.9. —
On se place dans $\mathbb{R}^4$. Montrer l'égalité : $\mathbf{Vect}\left[\begin{pmatrix}1\\1\\-3\\0\end{pmatrix},\begin{pmatrix}1\\0\\-3\\-2\end{pmatrix}\right]=\mathbf{Vect}\left[\begin{pmatrix}-1\\1\\3\\4\end{pmatrix},\begin{pmatrix}1\\2\\-3\\2\end{pmatrix},\begin{pmatrix}0\\1\\0\\2\end{pmatrix}\right]$.

Exercice de manipulation 1.10. —
Dans $\mathbb{R}^3$, montrer l'égalité $\mathbf{Vect}\left[\begin{pmatrix}1\\-2\\1\end{pmatrix},\begin{pmatrix}3\\-1\\0\end{pmatrix},\begin{pmatrix}1\\3\\-2\end{pmatrix},\begin{pmatrix}-2\\-1\\1\end{pmatrix}\right]=\mathbf{Vect}\left[\begin{pmatrix}1\\-2\\1\end{pmatrix},\begin{pmatrix}3\\-1\\0\end{pmatrix}\right]$.

Exercice de manipulation 1.11. —
Vérifier, uniquement à partir de la définition, que la famille $\left(\begin{pmatrix}1\\1\\2\end{pmatrix},\begin{pmatrix}2\\4\\2\end{pmatrix},\begin{pmatrix}-1\\0\\1\end{pmatrix}\right)$ engendre $\mathbb{R}^3$.

Exercice de manipulation 1.12. —

(a) Vérifier que la famille $\left(\begin{pmatrix}2\\5\\0\end{pmatrix},\begin{pmatrix}3\\2\\0\end{pmatrix},\begin{pmatrix}4\\-1\\0\end{pmatrix}\right)$ n'est pas une famille génératrice de $\mathbb{R}^3$.

(b) Même question avec $\left(\begin{pmatrix}1\\2\\1\end{pmatrix},\begin{pmatrix}-1\\3\\-1\end{pmatrix},\begin{pmatrix}2\\7\\2\end{pmatrix},\begin{pmatrix}0\\-5\\0\end{pmatrix}\right)$.

Exercice de manipulation 1.13. —
On se place dans $\mathbb{R}^4$ et on note $F=\{(x,y,z,t)\in\mathbb{R}^4, x+y+z+t=0\}$.

(a) Pourquoi est-ce un sous-espace vectoriel de $\mathbb{R}^4$?

(b) On considère les vecteurs $u=(1,-1,-1,1)$, $v=(1,0,-1,0)$ et $w=(0,1,-1,0)$.

Vérifier que la famille (u,v,w) est une famille génératrice de F, mais que ce n'est pas une famille génératrice de $\mathbb{R}^4$.

Exercice de manipulation 1.14. —

1. Vérifier sans calcul que la famille $\left(\begin{pmatrix}2\\1\\6\end{pmatrix},\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}1\\-3\\3\end{pmatrix},\begin{pmatrix}0\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\end{pmatrix}\right)$ engendre $\mathbb{R}^3$.
2. En déduire (sans repasser par la définition) que la famille $\left(\begin{pmatrix}2\\1\\6\end{pmatrix},\begin{pmatrix}1\\0\\0\end{pmatrix},\begin{pmatrix}1\\-3\\3\end{pmatrix}\right)$ engendre $\mathbb{R}^3$.

Exercice de manipulation 1.15. —

(a) Démontrer que dans $\mathbb{R}^3$, la famille $\left(\begin{pmatrix}-1\\0\\2\end{pmatrix},\begin{pmatrix}1\\3\\1\end{pmatrix},\begin{pmatrix}0\\1\\-1\end{pmatrix}\right)$ est libre. Même question avec $\left(\begin{pmatrix}2\\1\\1\end{pmatrix},\begin{pmatrix}1\\3\\1\end{pmatrix},\begin{pmatrix}2\\3\\0\end{pmatrix}\right)$.

(b) Dans $\mathbb{R}^3$, la famille $\left(\begin{pmatrix}1\\-2\\-3\end{pmatrix},\begin{pmatrix}4\\1\\-3\end{pmatrix},\begin{pmatrix}-2\\-7\\3\end{pmatrix}\right)$ est-elle libre ? Et la famille $\left(\begin{pmatrix}1\\-2\\-3\end{pmatrix},\begin{pmatrix}4\\1\\-3\end{pmatrix},\begin{pmatrix}-2\\-7\\-5\end{pmatrix}\right)$?

Exercice de manipulation 1.16. —
Définissons trois suites u, v et w par $u_n = n^2 - 1$, $v_n = (-1)^n$ et $w_n = 2^n$ pour $n \in \mathbb{N}$.
Démontrer que dans l'espace $\mathcal{F}(\mathbb{N}, \mathbb{R})$ des suites réelles, la famille (u, v, w) est libre.

Exercice de manipulation 1.17. —
Montrer que les familles suivantes de $\mathbb{R}^3$ sont liées :

(a) $\left(\begin{pmatrix}5\\-3\\2\end{pmatrix},\begin{pmatrix}-15\\9\\-6\end{pmatrix}\right)$

(b) $\left(\begin{pmatrix}5\\-3\\2\end{pmatrix},\begin{pmatrix}-15\\9\\-6\end{pmatrix},\begin{pmatrix}10\\-11\\12\end{pmatrix}\right)$

(c) $\left(\begin{pmatrix}1\\2\\3\end{pmatrix},\begin{pmatrix}0\\0\\0\end{pmatrix},\begin{pmatrix}4\\5\\6\end{pmatrix}\right)$

(d) $\left(\begin{pmatrix}1\\2\\-1\end{pmatrix},\begin{pmatrix}1\\0\\1\end{pmatrix},\begin{pmatrix}-1\\2\\-3\end{pmatrix}\right)$

(e) $\left(\begin{pmatrix}1\\0\\1\end{pmatrix},\begin{pmatrix}1\\1\\0\end{pmatrix},\begin{pmatrix}0\\1\\1\end{pmatrix},\begin{pmatrix}1\\0\\-1\end{pmatrix}\right)$

(f) $\left(\begin{pmatrix}2\\0\\3\end{pmatrix},\begin{pmatrix}-1\\2\\-1\end{pmatrix},\begin{pmatrix}0\\-4\\-1\end{pmatrix}\right)$

Exercice de manipulation 1.18. —

Déduire du résultat de l'exercice **I.5** que les familles $\left(\begin{pmatrix}1\\1\\1\end{pmatrix},\begin{pmatrix}1\\4\\3\end{pmatrix},\begin{pmatrix}1\\1\\0\end{pmatrix}\right)$ et $\left(\begin{pmatrix}1\\1\\1\end{pmatrix},\begin{pmatrix}1\\4\\3\end{pmatrix},\begin{pmatrix}2\\2\\-1\end{pmatrix}\right)$ sont libres dans $\mathbb{R}^3$.

Attention, en revanche la famille $\left(\begin{pmatrix}1\\1\\1\end{pmatrix},\begin{pmatrix}1\\4\\3\end{pmatrix},\begin{pmatrix}1\\1\\0\end{pmatrix},\begin{pmatrix}2\\2\\-1\end{pmatrix}\right)$ est liée : pourquoi ?

Exercice de manipulation 1.19. —

Dire pourquoi la famille $\left(\begin{pmatrix}1 & 2\\0 & 3\end{pmatrix}, \begin{pmatrix}1 & -2\\0 & 3\end{pmatrix}\right)$ de $\mathcal{M}_2(\mathbb{R})$ est libre.

En déduire que la famille $\left(\begin{pmatrix}1 & 2\\0 & 3\end{pmatrix}, \begin{pmatrix}1 & -2\\0 & 3\end{pmatrix}, \begin{pmatrix}0 & 0\\1 & 0\end{pmatrix}\right)$ est aussi libre.

Exercice de manipulation 1.20. —

(a) On se place dans $\mathbb{R}[X]$ et pour chaque i de $\mathbb{N}$, on note $P_i(X) = X^{i+2} - 3$. Vérifier de deux façons différentes que la famille $(P_i)_{i\in\mathbb{N}}$ est libre dans $\mathbb{R}[X]$.

(b) On se place dans $\mathcal{F}(\mathbb{N}, \mathbb{R})$ et pour chaque p de $\mathbb{N}$, on note $u^{[p]}$ la suite définie par :

$$\forall n \in \mathbb{N},\ u_n^{[p]} = \begin{cases} 1 & \text{si } n = p \\ -1 & \text{si } n = p+1 \\ 0 & \text{sinon.} \end{cases}$$

Montrer que la famille $(u^{[p]})_{p\in\mathbb{N}}$ est libre dans $\mathcal{F}(\mathbb{N}, \mathbb{R})$.

CHAPITRE 2

BASES ET DIMENSION

1. Bases d'un espace vectoriel, coordonnées dans une base

1.1. Définition et notion de coordonnées. —

Définition 2.1 – Base d'un espace vectoriel

Soient E un espace vectoriel et $(x_i)_{i\in I}$ une famille de vecteurs de E.
On dit que la famille $(x_i)_{i\in I}$ est une *base* de E si elle est à la fois *libre* et *génératrice de* E.

Si $\mathcal{B} = (x_i)_{i\in I}$ est une base de E, alors le fait qu'elle soit génératrice de E indique que tout vecteur de E peut s'écrire comme combinaison linéaire des vecteurs de $\mathcal{B}$; le fait qu'elle soit libre indique (interprétation c. de la liberté, page 30) qu'il n'est pas possible de l'écrire de deux façons différentes comme combinaison linéaire des vecteurs de $\mathcal{B}$. Ainsi,

La famille $\mathcal{B}$ est une base de E

$\Updownarrow$

Tout vecteur v de E peut s'écrire *de façon unique* comme une combinaison linéaire des vecteurs de $\mathcal{B}$.

Les coefficients de cette combinaison linéaire sont appelés les *coordonnées de v dans la base $\mathcal{B}$*.

Par exemple, si l'on considère une famille finie $(b_1, b_2, ..., b_n)$ de vecteurs d'un $\mathbb{K}$-espace vectoriel E, dire que c'est une base de E, c'est dire que pour chaque v de E, *il existe un unique* n-uplet $(\lambda_1, \lambda_2, ...\lambda_n)$ de scalaires de $\mathbb{K}$ tel que $v = \lambda_1 b_1 + \lambda_2 v_2 + ... + \lambda_n b_n$. Les coordonnées de v dans la base $\mathcal{B}$ sont alors données par le n-uplet $(\lambda_1, \lambda_2, ...\lambda_n)$.

***Exemple 2.2* (Une base de $\mathbb{R}^2$).** — La famille $\mathcal{B} = \left(\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}-1\\1\end{pmatrix}\right)$ est une base de $\mathbb{R}^2$. En effet, pour tout $v = \begin{pmatrix}x\\y\end{pmatrix}$ de $\mathbb{R}^2$, il existe un unique couple (λ_1, λ_2) de réels qui vérifie $v = \lambda_1 \begin{pmatrix}1\\2\end{pmatrix} + \lambda_2 \begin{pmatrix}-1\\1\end{pmatrix}$: le système linéaire $\begin{cases} x = \lambda_1 - \lambda_2 \\ y = 2\lambda_1 + \lambda_2 \end{cases}$ admet une unique solution, à savoir $\lambda_1 = \dfrac{x+y}{3}$, $\lambda_2 = \dfrac{-2x+y}{3}$.

On constate ici que si x et y sont deux réels, les coordonnées du vecteur $v = \begin{pmatrix}x\\y\end{pmatrix}$ dans $\mathcal{B}$ sont données par le couple $\left(\dfrac{x+y}{3}, \dfrac{-2x+y}{3}\right)$.

Une fois qu'on sait que ces deux vecteurs forment une base, pour trouver les coordonnées du vecteur $w = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$, il suffit de constater que $\begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} - 2\begin{pmatrix} -1 \\ 1 \end{pmatrix}$: on sait alors d'avance qu'il s'agit de la seule manière d'exprimer $\begin{pmatrix} 3 \\ 0 \end{pmatrix}$ comme combinaison linéaire des vecteurs de $\mathcal{B}$, donc que les coordonnées du vecteur w sont données par le couple $(1, -2)$.

FIGURE 1. Illustration de l'exemple ci-dessus. *Gauche :* on constate que $(3,0) = (1,2) - 2(-1,1)$. *Droite :* la famille $\mathcal{B}$ est une base de $\mathbb{R}^2$ et définit un repère.

***Exemple 2.3* (Un exemple dans $\mathbb{R}^3$).** — La famille $\mathcal{B} = \left(\begin{pmatrix} 97 \\ 98 \\ 99 \end{pmatrix}, \begin{pmatrix} 97 \\ 98 \\ 97 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right)$ est une base de $\mathbb{R}^3$. En effet, pour tout $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ de $\mathbb{R}^3$, il existe un unique triplet $(\lambda_1, \lambda_2, \lambda_3)$ de réels qui vérifie $v = \lambda_1 \begin{pmatrix} 97 \\ 98 \\ 99 \end{pmatrix} + \lambda_2 \begin{pmatrix} 97 \\ 98 \\ 97 \end{pmatrix} + \lambda_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$: le système linéaire

$$\begin{cases} 97\lambda_1 + 97\lambda_2 + \lambda_3 & = x \\ 98\lambda_1 + 98\lambda_2 + \lambda_3 & = y \\ 99\lambda_1 + 97\lambda_2 + \lambda_3 & = z \end{cases} \iff \begin{cases} 97\lambda_1 + 97\lambda_2 + \lambda_3 & = x \\ \lambda_1 + \lambda_2 & = y - x \\ 2\lambda_1 & = z - x \end{cases}$$

admet une unique solution, à savoir $\lambda_1 = \dfrac{z-x}{2}$, $\lambda_2 = \dfrac{-x+2y-z}{2}$, $\lambda_3 = 98x - 97y$. Les coordonnées du vecteur $\begin{pmatrix} 194 \\ 194 \\ 196 \end{pmatrix}$ dans la base $\mathcal{B}$ sont données par le triplet $(1, 1, 0)$.

Exercices de manipulation. — II.1, II.2.

Les deux exemples précédents devraient montrer clairement que :

- La notion de base est une (vaste) généralisation de celle de *repère du plan ou de l'espace à trois dimensions.*
- Un vecteur peut avoir des coordonnées *très compliquées* dans une base et *très simples* dans une **autre** base.

La notion de base permet donc notamment d'envisager des *repères de l'espace des fonctions*, des *repères de l'espace des sons*, des *repères de l'espace des images*, etc.

De plus, dans certains de ces « repères », il est tout à fait possible que des objets intéressants (pour la théorie... ou pour la pratique) aient des coordonnées particulièrement simples.

1.2. Caractérisation des bases parmi les familles génératrices et parmi les familles libres. —

Proposition 2.4 – Base = famille libre maximale = famille génératrice minimale

Soit E un espace vectoriel et $(x_i)_{i\in I}$ une famille de vecteurs de E. Les trois assertions suivantes sont équivalentes.

(a) La famille $(x_i)_{i\in I}$ est une base de E.

(b) La famille $(x_i)_{i\in I}$ est une famille génératrice minimale de E.

(c) La famille $(x_i)_{i\in I}$ est une famille libre maximale de E.

Démonstration. —

1. Supposons (a) et montrons (b). Si $(x_i)_{i\in I}$ est une base de E et si i_0 appartient à I, alors x_{i_0} n'est pas combinaison linéaire de $(x_i, i \neq i_0)$, à cause de la liberté de la famille : il n'appartient donc pas à $\mathbf{Vect}[x_i, i \neq i_0]$. Ainsi, la famille $(x_i, i \neq i_0)$ ne peut pas être génératrice de E : dès qu'on retire un vecteur à $(x_i)_{i\in I}$, elle perd son caractère générateur.
2. Supposons (b) et montrons (a). Si $(x_i)_{i\in I}$ est génératrice minimale, alors pour tout i_0 de I, x_{i_0} ne peut pas appartenir à $\mathbf{Vect}[x_i, i \neq i_0]$, sinon $(x_i, i \neq i_0)$ serait encore génératrice (voir b. page 26). D'après le critère (b) page 30, la famille $(x_i)_{i\in I}$ est libre : elle est donc à la fois libre et génératrice, et c'est une base.
3. Supposons (a) et montrons (c). Si $(x_i)_{i\in I}$ est une base de E, alors pour tout y de E, y appartient à $\mathbf{Vect}[(x_i)_{i\in I}]$, donc la famille qu'on obtient en ajoutant y à $(x_i)_{i\in I}$ ne peut pas être libre, et $(x_i)_{i\in I}$ est donc une famille libre maximale.
4. Supposons (c) et montrons (a). Si $(x_i)_{i\in I}$ est une famille libre maximale de E et si y est un vecteur quelconque de E, alors la famille obtenue en ajoutant y à $(x_i)_{i\in I}$ est liée : d'après la propriété (b) page 30, y appartient à $\mathbf{Vect}[(x_i)_{i\in I}]$. La famille est donc génératrice en plus d'être libre, et c'est une base.

□

1.3. Des exemples classiques de bases, et quelques remarques. —

1.3.1. *Base canonique de $\mathbb{K}^n$.* —
Nous avons déjà vu que si $n \geq 1$ et si on pose

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},\ e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots,\ e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

alors la famille $(e_1, e_2, \dots e_n)$ est génératrice de $\mathbb{K}^n$. De plus, elle est libre : si $\lambda_1, \dots, \lambda_n$ sont des scalaires vérifiant $\lambda_1 e_1 + \dots + \lambda_n e_n = 0$, alors il suffit de noter que $\lambda_1 e_1 + \dots + \lambda_n e_n = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}$ pour voir que tous les λ_i doivent être nuls.

La famille $(e_1, \dots e_n)$ est donc une base de $\mathbb{K}^n$, célèbre sous le nom [1] de *base canonique de* $\mathbb{K}^n$.

1.3.2. *Il existe cependant bien d'autres bases de $\mathbb{K}^n$.* —

Par exemple, dans $\mathbb{R}^3$, nous avons déjà vu que $\mathcal{B} = \left(\begin{pmatrix} 97 \\ 98 \\ 99 \end{pmatrix}, \begin{pmatrix} 97 \\ 98 \\ 97 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right)$ est une base de $\mathbb{R}^3$.

On peut aussi obtenir beaucoup de bases à partir d'une base donnée :

1. Il est courant, en première année, d'être amusé par l'origine commune à ce mot et à des termes religieux (« droit canonique ») ou non (« les canons de la beauté »). Ici, cela signifie simplement « la base classique », « la base habituelle »...

$\left(\begin{pmatrix}-97\\-98\\-99\end{pmatrix}, \begin{pmatrix}97\\98\\97\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}\right)$ est aussi une base de $\mathbb{R}^3$, $\left(\begin{pmatrix}0.97\\0.98\\0.99\end{pmatrix}, \begin{pmatrix}97\\98\\-97\end{pmatrix}, \begin{pmatrix}\sqrt{3}\\\sqrt{3}\\\sqrt{3}\end{pmatrix}\right)$ également...

Insérons d'ailleurs ici une remarque générale (2) :

Si $E \neq \{0\}$, alors dès qu'il admet une base, il en admet beaucoup d'autres.
Une expression comme « la base », sans préciser laquelle, n'a donc *jamais* de sens.

1.3.3. *Base canonique de $\mathcal{M}_{n,p}(\mathbb{K})$.* — Soit (n,p) un couple d'entiers naturels non nuls. Pour chaque (i,j) de $\{1,...,n\} \times \{1,...,p\}$, notons E_{ij} la matrice dont le coefficient en position (i,j) (au croisement de la $i^{\text{ème}}$ ligne et de la $j^{\text{ème}}$ colonne) vaut 1, et dont tous les autres coefficients valent zéro.

Par exemple si $n=2$, $E_{11} = \begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}$, $E_{12} = \begin{pmatrix}0 & 1\\0 & 0\end{pmatrix}$, $E_{21} = \begin{pmatrix}0 & 0\\1 & 0\end{pmatrix}$, $E_{22} = \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}$.

La famille $(E_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq p}}$ est libre, car elle vérifie le critère de la page 32.

De plus, elle engendre $\mathcal{M}_{np}(\mathbb{K})$, car si $A = (a_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq p}}$ est une matrice $n \times p$ quelconque à coefficients dans $\mathbb{K}$, on peut écrire

$$A = \sum_{\substack{1\leq i\leq n\\1\leq j\leq p}} a_{ij}E_{ij}.$$

et donc exprimer A comme combinaison linéaire de la famille $(E_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq p}}$.

La famille $(E_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq p}}$ est donc une base de $\mathcal{M}_{n,p}(\mathbb{K})$; on l'appelle souvent la *base canonique* de $\mathcal{M}_{n,p}(\mathbb{K})$.

1.3.4. *Il existe cependant beaucoup d'autres bases de $\mathcal{M}_{n,p}(\mathbb{K})$.* — Voici un exemple de base de $\mathcal{M}_2(\mathbb{R})$. Considérons

$$M_1 = \begin{pmatrix}1 & 1\\1 & 1\end{pmatrix}, M_2 = \begin{pmatrix}1 & 1\\-1 & -1\end{pmatrix}, M_3 = \begin{pmatrix}1 & -1\\-1 & 1\end{pmatrix}, M_4 = \begin{pmatrix}1 & -1\\-1 & -1\end{pmatrix}.$$

Soit $A = \begin{pmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{pmatrix}$ un élement de $\mathcal{M}_2(\mathbb{R})$.

Peut-on l'écrire comme combinaison linéaire à coefficients réels de la famille (M_1, M_2, M_3, M_4) ?

Soient $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ quatre réels. Alors

$$A = \sum_{i=1}^{4} \lambda_i M_i \iff \begin{pmatrix}a_{11} & a_{12}\\a_{21} & a_{22}\end{pmatrix} = \begin{pmatrix}\lambda_1+\lambda_2+\lambda_3+\lambda_4 & \lambda_1+\lambda_2-\lambda_3-\lambda_4\\\lambda_1-\lambda_2-\lambda_3-\lambda_4 & \lambda_1-\lambda_2+\lambda_3-\lambda_4\end{pmatrix} \iff \begin{cases}\lambda_1+\lambda_2+\lambda_3+\lambda_4 = a_{11}\\\lambda_1+\lambda_2-\lambda_3-\lambda_4 = a_{12}\\\lambda_1-\lambda_2-\lambda_3-\lambda_4 = a_{21}\\\lambda_1-\lambda_2+\lambda_3-\lambda_4 = a_{22}\end{cases}$$

ce qui équivaut à $\lambda_1 = \dfrac{a_{11}+a_{21}}{2}$, $\lambda_2 = \dfrac{a_{12}-a_{21}}{2}$, $\lambda_3 = \dfrac{a_{22}-a_{21}}{2}$ et $\lambda_4 = \dfrac{a_{11}-a_{22}+a_{21}-a_{12}}{2}$.

Chaque matrice A peut donc s'écrire de façon unique comme combinaison linéaire de (M_1, M_2, M_3, M_4) ; notre famille est bien une base de $\mathcal{M}_2(\mathbb{R})$.

Si l'on en croit le calcul précédent, les coordonnées de la matrice $\begin{pmatrix}1 & 5\\7 & 3\end{pmatrix}$ dans cette base sont données par le quadruplet $\left(\dfrac{1+7}{2}, \dfrac{5-7}{2}, \dfrac{3-7}{2}, \dfrac{1-3+7-5}{2}\right) = (4, -1, -2, 0)$.

Et effectivement, on a $4\begin{pmatrix}1 & 1\\1 & 1\end{pmatrix} - 1\begin{pmatrix}1 & 1\\-1 & -1\end{pmatrix} - 2\begin{pmatrix}1 & -1\\-1 & 1\end{pmatrix} + 0\begin{pmatrix}1 & -1\\-1 & -1\end{pmatrix} = \begin{pmatrix}1 & 5\\7 & -3\end{pmatrix}$.

2. Rappelons que dans ce cours $\mathbb{K}$ est $\mathbb{R}, \mathbb{C}$ ou $\mathbb{Q}$: on peut préciser la remarque en ajoutant que dès que $E \neq \{0\}$ admet une base, alors il en admet une infinité d'autres.

1.3.5. *Base canonique de $\mathbb{K}[X]$, de $\mathbb{K}_n[X]$.* —

• Nous avons déjà vu que la famille $(X^k)_{k\in\mathbb{N}} = (1, X, X^2, ...)$ est une famille génératrice de $\mathbb{K}[X]$, et que c'est une famille libre (car c'est une famille de polynômes à degrés deux à deux distincts). Ainsi, c'est une base de $\mathbb{K}[X]$, souvent appelée *base canonique de* $\mathbb{K}[X]$.

• De même, pour chaque n de $\mathbb{N}$, nous avons vu que la famille $(1, X, X^2, ...X^n)$ engendre $\mathbb{K}_n[X]$, et elle est libre pour la même raison ; c'est donc une base de $\mathbb{K}_n[X]$, souvent appelée *base canonique de* $\mathbb{K}_n[X]$.

• Par exemple, les coordonnées du polynôme $1299X^7 - \pi X^3 + 3X^2$ dans la base canonique de $\mathbb{K}_{10}[X]$ sont données par le 11-uplet $(0, 0, 3, -\pi, 0, 0, 0, 1299, 0, 0, 0)$.

1.3.6. *Il existe cependant beaucoup d'autres bases de $\mathbb{K}[X]$ et de $\mathbb{K}_n[X]$.* — Voici une manière simple d'en fabriquer.

Définition 2.5 – Famille échelonnée en degré de $\mathbb{K}_n[X]$

Une famille *échelonnée en degré de* $\mathbb{K}_n[X]$ est une famille $(P_0, P_1, ...P_n)$ de polynômes à coefficients dans $\mathbb{K}$ vérifiant : pour tout i de $\{0, ..., n\}$, $\deg(P_i) = i$.

Proposition 2.6 – Une famille échelonnée en degré donne une base de $\mathbb{K}_n[X]$

Pour tout n de $\mathbb{N}$, si $(P_0, P_1, ...P_n)$ est une famille échelonnée en degré de $\mathbb{K}_n[X]$, alors $(P_0, P_1, ...P_n)$ est une base de $\mathbb{K}_n[X]$.

Attention. — La réciproque est fausse : il existe des bases de $\mathbb{K}_n[X]$ qui ne sont pas échelonnées en degré. Par exemple, la famille $(X^2, X^2 - X - 1, X^2 - 1)$ est une base de $\mathbb{K}_2[X]$, alors que les trois polynômes de la famille ont le même degré.

Démonstration. — Nous allons raisonner par récurrence sur n.

- Si $n = 0$, $\mathbb{K}_0[X]$ est l'ensemble des polynômes constants, et si (P_0) est une famille comportant un seul polynôme de degré zéro, c'est bien une base de $\mathbb{K}[X]$ (le polynôme est non nul car son degré n'est pas $-\infty$, donc cette famille d'un seul vecteur est libre, et bien sûr elle est génératrice).

- Soit $n \in \mathbb{N}$. Supposons que toute famille échelonnée en degré de $\mathbb{K}_n[X]$ est une base de $\mathbb{K}_n[X]$.
  Montrons que toute famille échelonnée en degré de $\mathbb{K}_{n+1}[X]$ est une base de $\mathbb{K}_{n+1}[X]$.

  Soit $(P_0, P_1, ...P_{n+1})$ une famille échelonnée en degré de $\mathbb{K}_{n+1}[X]$.
  — D'abord, c'est une famille de polynômes à degrés deux à deux distincts, donc elle est libre.
  — Montrons que c'est une famille génératrice de $\mathbb{K}_{n+1}[X]$.
  Effectuons la division euclidienne de X^{n+1} par P_{n+1} : comme ces deux polynômes sont de même degré $n+1$, le quotient est un polynôme constant et le reste est un polynôme de degré $\leq n$. On peut donc écrire
  $$X^{n+1} = \lambda P_{n+1}(X) + R(X)$$
  où R appartient à $\mathbb{K}_n[X]$, donc par l'hypothèse de récurrence, à $\mathbf{Vect}[P_0, P_1, ...P_n]$.
  Ainsi X^{n+1} appartient à $\mathbf{Vect}[P_0, P_1, ...P_n, P_{n+1}]$.
  Mais pour tout k entre zéro et n, X^k appartient à $\mathbb{K}_n[X]$, donc à $\mathbf{Vect}[P_0, P_1, ...P_n]$, donc aussi à $\mathbf{Vect}[P_0, P_1, ...P_n, P_{n+1}]$.
  Conclusion : $1, X, X^2, ... X^n$ et X^{n+1} appartiennent à $\mathbf{Vect}[P_0, P_1, ...P_n, P_{n+1}]$, donc $\mathbf{Vect}\left[1, X, X^2, ...X^n, X^{n+1}\right]$ est contenu dans $\mathbf{Vect}[P_0, P_1, ...P_n, P_{n+1}]$. Comme $\mathbf{Vect}\left[1, X, X^2, ...X^n, X^{n+1}\right] = \mathbb{K}_{n+1}[X]$ et comme on a bien sûr $\mathbf{Vect}[P_0, P_1, ...P_n, P_{n+1}] \subset \mathbb{K}_{n+1}[X]$, on obtient $\mathbb{K}_{n+1}[X] = \mathbf{Vect}[P_0, P_1, ...P_n, P_{n+1}]$:
  on a bien démontré que $(P_0, P_1, ...P_{n+1})$ est une famille génératrice de $\mathbb{K}_{n+1}[X]$.

  Ainsi, $(P_0, P_1, ...P_{n+1})$ est une base de $\mathbb{K}_{n+1}[X]$, ce qu'il fallait démontrer. □

***Exemple 2.7* (Formule de Taylor pour les polynômes).** — Soit a un élément de $\mathbb{K}$.

La famille $\mathcal{B} = \big((X-a)^k\big)_{k=0,\dots n}$ est une famille échelonnée en degré de $\mathbb{K}_n[X]$. C'est donc une base de $\mathbb{K}_n[X]$.
Vous connaissez la *formule de Taylor pour les polynômes* :
si P est un élément de $\mathbb{K}_n[X]$, alors $P(X) = \sum\limits_{k=0}^{n} \frac{P^{(k)}(a)}{k!}(X-a)^k$.
On peut la réinterpréter comme indiquant les coordonnées de P dans la base $\mathcal{B}$:
elles sont données par le $(n+1)$-uplet $\left(P(0), P'(0), \dfrac{P''(0)}{2}, ..., \dfrac{P^{(n)}(0)}{n!}\right)$.

Exercice de manipulation. — II.3.

1.3.7. *Bases de $\mathbb{C}$ en tant que $\mathbb{R}$-espace vectoriel.* —
• Si l'on voit $\mathbb{C}$ comme un $\mathbb{R}$-espace vectoriel, alors dire que tout élément de $\mathbb{C}$ peut s'écrire de manière unique sous la forme $a+ib$ avec a et b réels, c'est dire que la famille $(1, i)$ est une base du $\mathbb{R}$-espace vectoriel $\mathbb{C}$. Les coordonnées d'un élément z de $\mathbb{C}$ dans la base $(1, i)$ sont ainsi données par $(\mathfrak{Re}(z), \mathfrak{Im}(z))$.
• Il y a bien sûr beaucoup d'autres bases possibles : par exemple, la famille $(1-i, 1+i)$ est une base du $\mathbb{R}$-espace vectoriel $\mathbb{C}$, car pour tout z de $\mathbb{C}$, il existe un unique couple (λ, μ) de réels tel qu'on ait $z = \lambda(1-i) + \mu(1+i)$: il est donné par $\lambda = \frac{\mathfrak{Re}(z)-\mathfrak{Im}(z)}{2}$ et $\mu = \frac{\mathfrak{Re}(z)+\mathfrak{Im}(z)}{2}$.

1.3.8. *Hic sunt leones.* — Vous remarquerez que je ne donne ici aucune base de l'espace $\mathcal{F}(\mathbb{R}, \mathbb{R})$, aucune base de l'espace $\mathcal{F}(\mathbb{N}, \mathbb{R})$, aucune base du $\mathbb{Q}$-espace vectoriel $\mathbb{R}$. Ce n'est pas un hasard : savoir s'il existe des bases de ces espaces, et surtout comprendre si on peut en trouver de concrètes ou d'utiles, nous emmènerait trop loin.

2. Espaces vectoriels de dimension finie

2.1. La définition. —

Définition 2.8 – Espace de dimension finie

Soit E un espace vectoriel.
On dit que E est *de dimension finie* s'il admet une famille génératrice de cardinal fini,
et on dit qu'il est *de dimension infinie* s'il n'admet pas de famille génératrice de cardinal fini.

La terminologie peut sembler mystérieuse à ce stade, puisqu'on parle de *dimension finie* sans avoir encore défini la dimension. Ce mystère sera levé dans trois pages. Relevons en attendant que nous avons déjà vu des exemples :

- Les espaces $\mathbb{K}^n$, $\mathcal{M}_{n,p}(\mathbb{R})$, $K_n[X]$ sont de dimension finie, puisque nous avons déjà rencontré des parties génératrices de cardinal fini de chacun d'entre eux.
- L'espace $\mathbb{K}[X]$ est de dimension infinie : en effet, s'il existait une famille génératrice finie $(P_1, ...P_k)$ de $\mathbb{K}[X]$, on pourrait noter $d = \max(\deg P_1, \deg P_2, ..., \deg P_k)$, et alors *tout* polynôme de $\mathbb{K}[X]$ serait une combinaison linéaire de polynômes de degré inférieur ou égal à d, donc serait lui-même de degré inférieur ou égal à d. Mais ce n'est pas le cas de X^{d+1}.

Les résultats de ce paragraphe sont faciles à retenir et à utiliser, mais difficiles à démontrer.
Il ne faut pas en sous-estimer l'importance : ils sont, à bien des égards, la clef de voûte du cours.

2.2. Théorème de la base incomplète. —

Théorème 2.9 – Théorème-clé

Soit E un espace vectoriel de dimension finie. Soient $\mathcal{L}$ une famille libre de E et $\mathcal{G}$ une famille génératrice de E. Il est possible de compléter $\mathcal{L}$ en une base de E en lui ajoutant des vecteurs de $\mathcal{G}$.

Remarque 2.10. — Ce théorème est vrai aussi lorsque l'espace n'est pas de dimension finie, mais la démonstration est (pour de bonnes raisons) beaucoup plus abstraite [3], et nous emmènerait trop loin de l'esprit de ce cours.

Afin d'exploiter l'hypothèse « E de dimension finie » pour pouvoir démontrer le théorème, il est souhaitable de se ramener au cas où $\mathcal{G}$ est de cardinal fini. Cela peut être fait, grâce au lemme suivant :

Lemme 2.11 – Lemme de réduction à une famille génératrice finie

Soit E un espace vectoriel de dimension finie.
De toute famille génératrice de E, on peut extraire une famille génératrice de cardinal fini.

Démonstration du lemme. — Soit $\mathcal{G} = (g_i)_{i \in I}$ une famille génératrice de E.

Comme E est de dimension finie, il existe une autre famille génératrice $\mathcal{S} = (s_1, s_2, ...s_n)$ de E qui, elle, est finie.

• Les vecteurs de $\mathcal{S}$ n'ont aucune raison d'être dans $\mathcal{G}$; cependant, ils sont tous dans $E = \mathbf{Vect}[\mathcal{G}]$, donc chacun est combinaison linéaire *d'un nombre fini* de vecteurs de $\mathcal{G}$: pour tout k de $\{1, ..n\}$, il existe un sous-ensemble FINI J_k de I qui vérifie : $s_k \in \mathbf{Vect}[(g_j,\ j \in J_k)]$.

3. À cause du statut particulier de *l'axiome du choix* qui y intervient – voir par exemple *The Princeton Companion to Mathematics*, chapitre III.1 (disponible sur le site de la bibliothèque).

• Considérons alors la famille $\mathcal{G}' = (g_j,\ j \in J_1 \cup J_2 \cup ... \cup J_n)$. C'est une sous-famille FINIE de $\mathcal{G}$, puisque l'ensemble $J_1 \cup J_2 \cup ... \cup J_n$ est fini.

• De plus, tous les vecteurs de $\mathcal{S}$ appartiennent à $\mathbf{Vect}[\mathcal{G}']$, puisque chaque s_k, $k \in \{1,..n\}$, est combinaison linéaire de certains des vecteurs de $\mathcal{G}'$.

Ainsi, on a $\mathbf{Vect}[\mathcal{S}] \subset \mathbf{Vect}[\mathcal{G}']$, et comme $\mathcal{S}$ engendre E, on en déduit $\mathbf{Vect}[\mathcal{G}'] = E$. Donc $\mathcal{G}'$ engendre E. □

Démonstration du théorème-clé.. — Soient $\mathcal{L}$ une famille libre de E et $\mathcal{G}$ une famille génératrice de E. Extrayons d'abord de $\mathcal{G}$, grâce au lemme, une famille génératrice finie $(g_1, ...g_n)$ de E. On souhaite fabriquer une base de E en ajoutant à $\mathcal{L}$ certains des vecteurs $g_1, ...g_n$ de $\mathcal{G}$.

- Il faut bien sûr que la famille obtenue soit libre : par exemple, si g_1 appartient à $\mathbf{Vect}[\mathcal{L}]$, on ne peut pas l'ajouter à $\mathcal{L}$ sans détruire la liberté de la famille.
- Si on ajoute g_1 à $\mathcal{L}$ et si on tente ensuite d'ajouter d'autres vecteurs parmi $g_2, ...g_n$ pour obtenir une base de E, il n'est raisonnable d'ajouter g_2 que si g_2 n'appartient pas à $\mathbf{Vect}[\mathcal{L} \cup \{g_1\}]$, sinon la famille obtenue n'est pas libre.

Adoptons donc l'algorithme suivant.

Commencer avec $\mathcal{B}_0 \stackrel{\text{def}}{=} \mathcal{L}$.

Pour i de 1 à n :
- ⇝ Si $g_i \notin \mathbf{Vect}[\mathcal{B}_{i-1}]$, l'ajouter à $\mathcal{B}_{i-1}$: poser $\mathcal{B}_i = \mathcal{B}_{i-1} \cup \{g_i\}$.
- ⇝ Sinon, ne rien changer : poser $\mathcal{B}_i = \mathcal{B}_{i-1}$, et passer à l'étape suivante.

À la sortie de la $i^{\text{ème}}$ étape de l'algorithme,
- la propriété b. page 34 garantit que la nouvelle famille $\mathcal{B}_i$ est encore libre si $\mathcal{B}_{i-1}$ l'était après l'étape précédente ;
- la nouvelle famille $\mathcal{B}_i$ vérifie toujours $g_i \in \mathbf{Vect}[\mathcal{B}_i]$.

À la fin de l'algorithme, on obtient donc une famille $\mathcal{B}_{\text{finale}} \stackrel{\text{def}}{=} \mathcal{B}_n$ qui a les deux propriétés suivantes :
- La famille $\mathcal{B}_{\text{finale}}$ est libre
- Le sous-espace $\mathbf{Vect}[\mathcal{B}_{\text{finale}}]$ contient tous les g_i, $i = 1..n$, donc il contient $\mathbf{Vect}[g_1, ...g_n]$, donc $\mathbf{Vect}[\mathcal{B}_{\text{finale}}] = E$.

Cette famille est donc libre et engendre E : c'est une base de E. □

Théorème 2.12 – Conséquences du théorème-clé

1. ***Existence de bases.***

   Si E est un espace vectoriel de dimension finie, alors E admet une base.
2. ***Théorème de la base incomplète.***

   Toute famille libre de E peut être complétée en une base de E.
3. ***Théorème de la base extraite.***

   De toute famille génératrice de E, on peut extraire une base.

Démonstration. —

1. La famille vide est libre et il existe au moins une famille génératrice de E ; en appliquant à ces deux familles le théorème-clé, on obtient l'existence d'une base de E.
2. Si $\mathcal{L}$ est une famille libre de E, il suffit d'appliquer le théorème-clé à $\mathcal{L}$ et à une famille génératrice arbitraire de E afin de compléter $\mathcal{L}$ en une base de E.
3. Si $\mathcal{G}$ est une famille génératrice de E, on peut appliquer le théorème-clé à la famille vide (qui est libre) et à $\mathcal{G}$: on obtient l'existence d'une base obtenue en ajoutant à l'ensemble vide des vecteurs de $\mathcal{G}$, qui est donc formée uniquement de vecteurs de $\mathcal{G}$.

□

2.3. Théorème de Steinitz et notion de dimension. —

Théorème 2.13 – Théorème de Steinitz

Si $\mathcal{L}$ est une famille libre de E de cardinal fini et si $\mathcal{G}$ est une famille génératrice de E de cardinal fini, alors

$$\mathrm{Card}(\mathcal{L}) \leq \mathrm{Card}(\mathcal{G}).$$

Pour démontrer le théorème de Steinitz, nous allons donner un énoncé plus précis mais moins utile à retenir :

Lemme 2.14 – Lemme d'échange de Steinitz

Si $\mathcal{L} = (v_1, ..., v_m)$ est une famille libre de E de cardinal m et si $\mathcal{G}$ est une partie génératrice de E de cardinal n, alors
- *on a nécessairement $m \leq n$;*
- *il existe une numérotation $(w_1, ...w_n)$ des éléments de $\mathcal{G}$ qui fait de $(v_1, ...v_m, w_{m+1}, ...w_n)$ une base de E.*

Comme le second point est une conséquence immédiate du théorème-clé ci-dessus si le premier point est vérifié, le seul intérêt de cette reformulation est de permettre de comprendre la stratégie de la démonstration. C'est d'elle que vient le nom du lemme : il dit qu'on peut obtenir une base de E en « remplaçant » m des vecteurs de $\mathcal{G}$ par les m vecteurs de $\mathcal{L}$.

Démonstration du lemme d'échange (attention c'est difficile). — Nous allons procéder progressivement : on va d'abord remplacer un vecteur de $\mathcal{G}$ par un vecteur de $\mathcal{L}$, puis un autre, et continuer jusqu'à greffer tous les vecteurs de $\mathcal{L}$ dans une famille génératrice dont le reste est formé de vecteurs de $\mathcal{G}$.
- Pour tout entier k entre 1 et m, notons $\mathcal{P}(k)$ l'énoncé

« D'une part, on a nécessairement $k \leq n$;
d'autre part, il existe une manière de numéroter $(w_1, ...w_n)$ les éléments de $\mathcal{G}$ qui fasse de
$\tilde{\mathcal{G}}_{\text{k greffes}} \stackrel{\text{def}}{=} (v_1, ...v_k, w_{k+1}, ...w_n)$ une famille génératrice de E. »

- Montrons que $\mathcal{P}(1)$ est vrai [4]. La première partie est bien sûr vraie : on a $1 \leq n$.
  — Partons d'une numérotation arbitraire $(w_1, ...w_n)$ des vecteurs de $\mathcal{G}$ et observons le premier vecteur de $\mathcal{L}$. La famille $\mathcal{G}$ est génératrice, donc il existe des scalaires $\lambda_1, ...\lambda_n$ tels que
  $$v_1 = \sum_{i=1}^{n} \lambda_i w_i.$$
  — Comme v_1 n'est pas nul (sinon $\mathcal{L}$ serait liée), l'un au moins des λ_i est non-nul. Quitte à re-numéroter les éléments de $\mathcal{G}$, on peut supposer que c'est λ_1. De $v_1 = \lambda_1 w_1 + \lambda_2 w_2 + ... + \lambda_n w_n$, on peut alors déduire
  $$w_1 = \frac{1}{\lambda_1} v_1 + \left(\frac{-\lambda_2}{\lambda_1} w_2 + ... + \frac{-\lambda_n}{\lambda_1} w_n \right)$$
  si bien que w_1 appartient à $\mathbf{Vect}[v_1, w_2, ...w_n]$.

4. On pourrait définir $\mathcal{P}(k)$ à partir de $k = 0$ et l'initialisation serait alors triviale, ce qui raccourcirait de moitié notre démonstration, mais la rendrait plus difficile à comprendre...

— Mais alors les vecteurs $w_1, w_2, ...w_n$ de $\tilde{\mathcal{G}}_{0\ \mathrm{greffe}}$ appartiennent tous à $\mathbf{Vect}[v_1, w_2, ...w_n] = \mathbf{Vect}\left[\tilde{\mathcal{G}}_{1\ \mathrm{greffe}}\right]$. Ainsi $\mathbf{Vect}\left[\tilde{\mathcal{G}}_{0\ \mathrm{greffe}}\right] \subset \mathbf{Vect}\left[\tilde{\mathcal{G}}_{1\ \mathrm{greffe}}\right]$, et comme $\tilde{\mathcal{G}}_{0\ \mathrm{greffe}}$ est génératrice, on en déduit que $\tilde{\mathcal{G}}_{1\ \mathrm{greffe}}$ l'est aussi.

- Soit k un entier entre 1 et $m-1$. Supposons $\mathcal{P}(k)$ vrai et montrons que $\mathcal{P}(k+1)$ est alors vrai.
  — Observons le vecteur v_{k+1}. Comme la famille $\tilde{\mathcal{G}}_{\mathrm{k\ greffes}} = (v_1, ...v_k, w_{k+1}, ...w_n)$ est génératrice, il existe des scalaires $\lambda_1, ...\lambda_k, \mu_{k+1}, ...\mu_n$ tels que

$$v_{k+1} = \sum_{i=1}^{k} \lambda_i v_i + \sum_{j=k+1}^{n} \mu_j w_j. \tag{2.1}$$

  — Comme la famille $\mathcal{L}$ est libre, il n'est pas possible que cette égalité permette d'exprimer v_{k+1} comme combinaison linéaire de $(v_1, ...v_k)$, et donc il est impossible que la deuxième somme ne comporte aucun terme et soit nulle. Ainsi on a nécessairement $k+1 \leq n$, c'est la première partie de $\mathcal{P}(k+1)$.
  — De plus, l'un des μ_j doit être non-nul ; quitte à renuméroter les vecteurs de $\mathcal{G}$, on peut supposer que c'est μ_{k+1}. De l'égalité (2.1), on peut alors déduire

$$w_{k+1} = \frac{1}{\mu_{k+1}} v_{k+1} - \frac{1}{\mu_{k+1}} \left(\sum_{i=1}^{k} \lambda_i v_i + \sum_{j=k+2}^{n} \mu_j w_j \right)$$

  ce qui exprime w_{k+1} comme combinaison linéaire de $\tilde{\mathcal{G}}_{(\mathrm{k+1})\ \mathrm{greffes}} = (v_1, v_2, ...v_{k+1}, w_{k+2}, ...w_n)$.
  Mais alors les vecteurs $v_1, ...v_k, w_{k+1}, ...w_n$ de $\tilde{\mathcal{G}}_{\mathrm{k\ greffes}}$ sont tous contenus dans $\mathbf{Vect}\left[\tilde{\mathcal{G}}_{(\mathrm{k+1})\ \mathrm{greffes}}\right]$, donc $E = \mathbf{Vect}\left[\tilde{\mathcal{G}}_{\mathrm{k\ greffes}}\right]$ est aussi contenu dans $\mathbf{Vect}\left[\tilde{\mathcal{G}}_{(\mathrm{k+1})\ \mathrm{greffes}}\right]$, si bien que $\mathbf{Vect}\left[\tilde{\mathcal{G}}_{(\mathrm{k+1})\ \mathrm{greffes}}\right]$ est bien génératrice.
  Cela prouve la seconde partie de l'énoncé $\mathcal{P}(k+1)$.
- Par récurrence, $\mathcal{P}(k)$ est vrai pour tout $k \leq m$. L'énoncé $\mathcal{P}(m)$ est le lemme d'échange qu'on voulait montrer. □

La conséquence la plus importante du théorème de Steinitz est le résultat suivant, qui est au fondement de toute l'algèbre linéaire.

Théorème 2.15 – Théorème fondamental sur le cardinal des bases

Si E est un espace vectoriel de dimension finie, alors toutes les bases de E ont le même cardinal.

Démonstration. — Commençons par remarquer que toutes les bases de E sont finies : si $\mathcal{B}$ est une base de E, alors c'est une partie génératrice de E, donc on peut en extraire une famille génératrice finie $(b_1, ...b_n)$ (lemme 2.11). Mais comme $\mathcal{B}$ est une base de E, c'est une partie génératrice minimale de E (proposition 2.1.2), donc la famille $(b_1, ...b_n)$ ne peut rester génératrice que si on a $\mathcal{B} = (b_1, ..b_n)$, donc $\mathcal{B}$ est finie.
Si maintenant $\mathcal{B}_1$ et $\mathcal{B}_2$ sont deux bases de E, on sait alors que
- $\mathcal{B}_1$ est libre et $\mathcal{B}_2$ est génératrice de E, donc par le théorème de Steinitz, $\mathrm{Card}(\mathcal{B}_1) \leq \mathrm{Card}(\mathcal{B}_2)$.
- $\mathcal{B}_2$ est libre et $\mathcal{B}_1$ est génératrice de E, donc par le théorème de Steinitz, $\mathrm{Card}(\mathcal{B}_2) \leq \mathrm{Card}(\mathcal{B}_1)$.

□

Définition 2.16 – Notion de dimension

Lorsque E est un espace vectoriel de dimension finie, on appelle *dimension de E*, et on note $\dim(E)$, le cardinal commun à toutes les bases de E. C'est un entier.

Attention. — À ce stade de vos études, la notion de dimension n'a de sens que pour les espaces vectoriels.

Remarque 2.17. — Si E est un $\mathbb{C}$-espace vectoriel, on peut aussi le voir comme un $\mathbb{R}$-espace vectoriel ; la dimension dépend du corps de base choisi (voir la proposition 2.31).

3. Exemples d'espaces de dimension finie

3.1. Les exemples fondamentaux : colonnes, matrices, polynômes. —

Dimension de $\mathbb{K}^n$ ($n \in \mathbb{N}^\star$). —

Nous avons déjà rencontré une base de $\mathbb{K}^n$, la *base canonique* $\left(\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix}, \dots, \begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix}\right)$.

Elle comporte n vecteurs. Par conséquent,

Pour l'espace $\mathbb{K}^n$, on a $\dim(\mathbb{K}^n) = n$.

Dimension de $\mathcal{M}_{n,p}(\mathbb{K})$ ($n, p \in \mathbb{N}^\star$). —

Nous avons déjà rencontré une base de $\mathcal{M}_{n,p}(\mathbb{K})$, la base canonique $(E_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq p}}$.
Cette base comporte np vecteurs. Par conséquent,

Pour l'espace $\mathcal{M}_{n,p}(\mathbb{K})$ des matrices $n \times p$, on a $\dim(\mathcal{M}_{n,p}(\mathbb{K}) = np$.

En particulier,

L'espace $\mathcal{M}_n(\mathbb{K})$ des matrices carrées à n lignes et n colonnes est de dimension n^2.

Dimension de $\mathbb{K}_{\mathbf{n}}[\mathbf{X}]$ ($n \in \mathbb{N}$). —

Nous avons déjà rencontré au moins une base de $\mathbb{K}_n[X]$, la base canonique $(1, X, X^2, \dots, X^n)$.
Elle comporte $n + 1$ vecteurs. Par conséquent,

Pour l'espace $\mathbb{K}_n[X]$ des polynômes de degré $\leq n$, on a $\dim(\mathbb{K}_n[X]) = n + 1$.

Attention aux notations. — pour préciser un élément de $\mathbb{K}_2[X]$, il faut *trois* nombres (par exemple ses coordonnées dans la base canonique : le coefficient constant, le coefficient de X et le coefficient de X^2).

3.2. L'espace des matrices symétriques. —

• Soit $\mathcal{S}_n(\mathbb{R}) = \left\{A \in \mathcal{M}_n(\mathbb{R}), A^T = A\right\}$ l'ensemble des matrices *symétriques* de taille n à coefficients réels. Par exemple,

$$\mathcal{S}_3(\mathbb{R}) = \left\{\begin{pmatrix} a & u & w\\ u & b & v\\ w & v & c\end{pmatrix}, (a, b, c, u, v, w) \in \mathbb{R}^6\right\}.$$

L'ensemble $\mathcal{S}_n(\mathbb{R})$ est un sous-espace vectoriel de $\mathcal{M}_n(\mathbb{R})$.
• En observant la forme des éléments de $\mathcal{S}_3(\mathbb{R})$, on constate qu'il faut exactement six nombres pour décrire une matrice symétrique de taille 3, et cela laisse penser que $\mathcal{S}_3(\mathbb{R})$ est de dimension six.

• Pour le démontrer, il faut exhiber une base de $\mathcal{S}_3(\mathbb{R})$ qui comporte six matrices. Notons :

$$S_{11}=\begin{pmatrix}1&0&0\\0&0&0\\0&0&0\end{pmatrix},S_{12}=\begin{pmatrix}0&1&0\\1&0&0\\0&0&0\end{pmatrix},S_{13}=\begin{pmatrix}0&0&1\\0&0&0\\1&0&0\end{pmatrix},S_{22}=\begin{pmatrix}0&0&0\\0&1&0\\0&0&0\end{pmatrix},S_{23}=\begin{pmatrix}0&0&0\\0&0&1\\0&1&0\end{pmatrix},S_{33}=\begin{pmatrix}0&0&0\\0&0&0\\0&0&1\end{pmatrix}.$$

On constate que la famille $(S_{11}, S_{12}, S_{13}, S_{22}, S_{23}, S_{33})$ est une famille de matrices à coefficients en positions complémentaires (voir l'exemple 1.8.2.E page 32). C'est donc une famille libre de $\mathcal{M}_3(\mathbb{R})$ et de $\mathcal{S}_3(\mathbb{R})$.

Par ailleurs, si $A=\begin{pmatrix}a&u&w\\u&b&v\\w&v&c\end{pmatrix}$ est une matrice symétrique, alors

$$A = aS_{11} + bS_{22} + cS_{33} + uS_{12} + vS_{23} + wS_{13},$$

donc tout élément de $\mathcal{S}_3(\mathbb{R})$ est combinaison linéaire de $(S_{11}, S_{12}, S_{13}, S_{22}, S_{23}, S_{33})$, qui est donc une famille génératrice de $S_n(\mathbb{R})$.

Ainsi, notre famille de six matrices est une base de $\mathcal{S}_3(\mathbb{R})$, cet espace est donc bien de dimension six.

• Plus généralement, revenons à n quelconque : on conjecture que connaître complètement une matrice symétrique, c'est connaître précisément les coefficients situés soit sur la diagonale, soit strictement au-dessus.

Il y a $\dfrac{n(n+1)}{2}$ coefficients à ces positions, on conjecture donc que la dimension de $\mathcal{S}_n(\mathbb{R})$ est $\dfrac{n(n+1)}{2}$.

Démonstration de cette conjecture. — Pour déterminer la dimension de $\mathcal{S}_n(\mathbb{R})$, il faut exhiber une base de cet espace.

- Pour tout (i,j) de $\{1,..,n\}^2$ vérifiant $i<j$, notons S_{ij} la matrice $E_{ij}+E_{ji}$, et pour tout i de $\{1,..n\}$, notons $S_{ii}=E_{ii}$.

  On obtient toujours une matrice symétrique : si $i<j$ on a $S_{ij}^T=(E_{ij}+E_{ij}^T)^T=E_{ij}^T+(E_{ij}^T)^T=E_{ij}^T+E_{ij}=S_{ij}$, et si $i=j$ on a $S_{ij}^T=E_{ii}^T=E_{ii}=S_{ij}$.

  La famille $(S_{ij})_{1\leq i\leq j\leq n}$ comporte $\dfrac{n(n+1)}{2}$ matrices.
- C'est une famille de matrices à coefficients en positions complémentaires, donc c'est une famille libre de $\mathcal{S}_n(\mathbb{R})$.
- De plus, si $A=(a_{ij})_{\substack{1\leq i\leq n\\1\leq j\leq n}}$ est une matrice symétrique quelconque, on a $A=\sum\limits_{1\leq i\leq j\leq n} a_{ij}S_{ij}$, donc la famille $(S_{ij})_{1\leq i\leq j\leq n}$ est une famille génératrice de $\mathcal{S}_n(\mathbb{R})$; c'est bien une base de $\mathcal{S}_n(\mathbb{R})$.

□

Cas des matrices antisymétriques. —

Soit $\mathcal{A}_n(\mathbb{R})=\{A\in\mathcal{M}_n(\mathbb{R}), A^T=-A\}$ l'ensemble des matrices *antisymétriques* de taille n. Par exemple,

$$\mathcal{A}_3(\mathbb{R})=\left\{\begin{pmatrix}0&u&w\\-u&0&v\\-w&-v&0\end{pmatrix},(u,v,w)\in\mathbb{R}^3\right\}.$$

Cette fois, on constate qu'il suffit de connaître les coefficients situés *strictement* au-dessus de la diagonale pour connaître complètement une matrice antisymétrique, car la diagonale est automatiquement nulle. On peut donc penser que la dimension de $\mathcal{A}_n(\mathbb{R})$ est $\dfrac{n(n-1)}{2}$. L'exercice II.5 vous permettra de le vérifier vous-même.

Résumons les résultats obtenus ci-dessus :

- L'espace des matrices symétriques $n\times n$ est de dimension $\frac{n(n+1)}{2}$.
- L'espace des matrices antisymétriques $n\times n$ est de dimension $\frac{n(n-1)}{2}$.

Exercices de manipulation. — II.4, II.5.

3.3. ($\star$) Complément : étude des suites récurrentes linéaires d'ordre 2. —

- Soit (a, b) un couple de nombres réels. Rappelons que l'ensemble

$$E_{a,b} = \{u \in \mathcal{F}(\mathbb{N}, \mathbb{R}), \forall n \in \mathbb{N}, u_{n+2} = au_{n+1} + bu_n\}$$

est un sous-espace vectoriel de $\mathcal{F}(\mathbb{N}, \mathbb{R})$ (voir page 18).

- On peut penser que pour connaître un élément u de $E_{a,b}$, il suffit de connaître u_0 et u_1 : le reste de la suite u sera entièrement déterminé par la relation de récurrence. On peut donc penser que la dimension de $E_{a,b}$ est 2.
- Pour aller plus loin que cette idée intuitive et *démontrer* que $E_{a,b}$ est de dimension 2, il faut trouver une base de $E_{a,b}$.
  Nous allons voir que cela nécessite du travail.
- Cherchons s'il peut y avoir dans $E_{a,b}$ des *suites géométriques de premier terme 1* : y a-t-il des éléments u de $E_{a,b}$ pour lesquels il existe un réel non nul r vérifiant : pour tout n de $\mathbb{N}$, $u_n = r^n$?
  Si c'est le cas, on doit avoir $r^{n+2} = ar^{n+1} + br^n$ pour tout n, donc r doit être solution de *l'équation caractéristique*

$$r^2 = ar + b.$$

Cas 1 : le discriminant $a^2 + 4b$ de cette équation du second degré est strictement positif.

Il y a alors deux solutions réelles ρ_1 et ρ_2 à l'équation caractéristique, et on obtient deux éléments distincts de $E_{a,b}$: les suites v et w définies par $v_n = \rho_1^n$ et $w_n = \rho_2^n$ pour tout n de $\mathbb{N}$.

Soit u un élément quelconque de $E_{a,b}$. Peut-on l'écrire comme combinaison linéaire de v et w ? Si c'est le cas, il doit exister des réels α et β tels qu'on ait $u = \alpha v + \beta w$. Ceci est une égalité de suites ; elle implique en particulier $u_0 = \alpha + \beta$ et $u_1 = \alpha\rho_1 + \beta\rho_2$, et on constate que la seule possibilité est $\alpha = \dfrac{\rho_2 u_0 - u_1}{\rho_2 - \rho_1}$ et $\beta = \dfrac{\rho_1 u_0 - u_1}{\rho_1 - \rho_2}$.

Si on note $\tilde{u}$ la suite $\alpha v + \beta w$, on constate alors que $z = u - \tilde{u}$ est encore un élément de $E_{a,b}$, mais que $z_0 = 0$ et $z_1 = 0$. Par conséquent $z_2 = az_1 + bz_0 = 0$, de même $z_3 = 0$, etc : par une récurrence immédiate, on en déduit que z est la suite nulle.

Conclusion : on a l'égalité $u = \alpha v + \beta w$, et comme nous avons vu qu'il y a une seule valeur possible pour les coefficients α et β, il s'agit de la seule manière d'exprimer la suite u comme combinaison linéaire des suites v et w.

Ainsi, la famille (v, w) est une base de $E_{a,b}$. Donc $E_{a,b}$ est de dimension 2.

Cas 2 : le discriminant $a^2 + 4b$ de l'équation caractéristique est strictement négatif.

Alors aucune suite géométrique *à valeurs réelles* ne peut vérifier la relation de récurrence définissant $E_{a,b}$,
donc il ne peut y avoir aucune suite géométrique dans $E_{a,b}$ (sous-espace de $\mathcal{F}(\mathbb{N}, \mathbb{R})$).

En revanche, il existe des suites géométriques *à valeurs complexes* vérifiant la même relation de récurrence : notons $\zeta_1 = \rho e^{i\theta}$ et $\zeta_2 = \rho e^{-i\theta}$ (avec $\theta \notin \pi\mathbb{Z}$) les racines complexes conjuguées de l'équation caractéristique, alors le raisonnement précédent (où le fait que les racines étaient réelles n'est jamais intervenu dans les calculs algébriques) reste valable et montre que si u est une suite *à valeurs complexes* vérifiant :

$$\forall n \in \mathbb{N}, u_{n+2} = au_{n+1} + bu_n, \tag{3.1}$$

on a, pour chaque n de $\mathbb{N}$,

$$\begin{aligned}u_n &= \left[\frac{\zeta_2 u_0 - u_1}{\zeta_2-\zeta_1}\zeta_1^n + \frac{\zeta_1 u_0 - u_1}{\zeta_1-\zeta_2}\zeta_2^n\right]\\ &= \left[u_0\left(\frac{\zeta_2\zeta_1^n - \zeta_1\zeta_2^n}{\zeta_2-\zeta_1}\right) + u_1\frac{\zeta_2^n-\zeta_1^n}{\zeta_1-\zeta_2}\right]\\ &= \left[u_0\cdot\rho^n\cdot\frac{e^{i(n-1)\theta}-e^{i(1-n)\theta}}{e^{-i\theta}-e^{i\theta}} + u_1\cdot\rho^{n-1}\cdot\frac{e^{-in\theta}-e^{in\theta}}{e^{-i\theta}-e^{i\theta}}\right]\end{aligned}$$

Nous avons obtenu le renseignement suivant : dès que (u_n) est une suite à valeurs complexes vérifiant (2.1), on a

$$\forall n\in\mathbb{N}, u_n = \left(\frac{u_0}{\sin(\theta)}\right)\cdot\rho^n\sin\left[(1-n)\theta\right] + \left(\frac{u_1}{\rho\sin(\theta)}\right)\cdot\rho^n\sin(n\theta). \tag{3.2}$$

Dans le cas où u_0 et u_1 sont des nombres réels, cette expression ne fait plus intervenir que des nombres réels.
Par conséquent, si (u_n) est un élément de $E_{a,b}$, l'égalité (3.2) est vérifiée et montre que u est combinaison linéaire des deux suites v et w définies par [5] $v_n = \rho^n\sin((1-n)\theta)$ et $w_n = \rho^n\sin(n\theta)$ pour $n\in\mathbb{N}$.

En écrivant $v_n = \mathfrak{Im}(\zeta_1\zeta_2^n)$ et $w_n = \mathfrak{Im}(\zeta_1^n)$ et l'équation caractéristique, on vérifie que v et w appartiennent à $E_{a,b}$.

Ainsi $E_{a,b} = \mathbf{Vect}[v,w]$, donc la famille (v,w) engendre $E_{a,b}$.

De plus, les suites v et w ne sont pas colinéaires dans l'espace $\mathcal{F}(\mathbb{N},\mathbb{R})$ (ce qui signifierait qu'elles ne diffèrent l'une de l'autre que par une constante multiplicative) ; la famille (v,w) de deux vecteurs de $E_{a,b}$ est donc libre, c'est une base de $E_{a,b}$. L'espace $E_{a,b}$ est donc bien de dimension 2.

On peut trouver une base plus commode à retenir en rappelant : $\forall n\in\mathbb{N},\ \sin((1-n)\theta) = \sin(\theta)\cos(n\theta) - \cos(\theta)\sin(n\theta)$, ce qui montre que la suite v est combinaison linéaire de w et de la suite $\tilde{v}$ définie par $\tilde{v}_n = \rho^n\cos(\theta)$ pour $n\in\mathbb{N}$. Ainsi $\mathbf{Vect}[v,w]\subset\mathbf{Vect}[\tilde{v},w]$; la même formule donne aussi $\mathbf{Vect}[v,w]\subset\mathbf{Vect}[\tilde{v},w]$. Ainsi $E_{a,b} = \mathbf{Vect}[\tilde{v},w]$ et $(\tilde{v},w)$ est une base de $E_{a,b}$.

Cas 3 : le discriminant a^2+4b de l'équation caractéristique est nul et [6] $a\neq 0$.

Dans ce cas il y a exactement une racine réelle $\rho = \frac{a}{2}$ à l'équation caractéristique, et il y a exactement une suite géométrique de premier terme 1 dans $E_{a,b}$, la suite v définie par $v_n = \rho^n$ pour $n\in\mathbb{N}$.

Cela force-t-il pour autant E à être égal à $\mathbf{Vect}[v]$? Autrement dit, si u appartient à E, la suite $\frac{u}{v}$ est-elle constante ?

Soit u un élément de E et n dans $\mathbb{N}^\star$. Calculons l'écart entre deux termes consécutifs de $\frac{u}{v}$:

$$\frac{u_{n+1}}{v_{n+1}} - \frac{u_n}{v_n} = \left[\frac{au_n+bu_{n-1}}{(a/2)v_n} - \frac{u_n}{v_n}\right] = \left[\frac{2u_n}{v_n} + \frac{bu_{n-1}}{(a/2)^2v_{n-1}} - \frac{u_n}{v_n}\right] = \left[\frac{u_n}{v_n} + \underbrace{\frac{4b}{a^2}}_{\substack{=-1\\ \text{car } a^2+4b=0}}\frac{u_{n-1}}{v_{n-1}}\right] = \frac{u_n}{v_n} - \frac{u_{n-1}}{v_{n-1}}.$$

Ainsi l'écart entre deux termes consécutifs de $\frac{u}{v}$ est toujours égal à $\lambda = \frac{u_1}{\rho} - \frac{u_0}{1}$. On en déduit que $\frac{u}{v}$ n'est pas nécessairement constante mais qu'elle est très particulière : $\frac{u_n}{v_n}$ est égal, pour chaque n, à $\frac{u_0}{v_0} + \lambda n$, si bien que

$$\forall n\in\mathbb{N},\qquad u_n = v_n\cdot\left[u_0 + n\left(\frac{u_1}{\rho} - \frac{u_0}{1}\right)\right]\qquad\text{et donc}\qquad u_n = u_0\cdot v_n + \lambda\cdot w_n$$

---

5. Rappelons qu'ici ρ et θ sont déterminés par a et b, puisqu'ils donnent le module et un argument de l'une des racines de l'équation caractéristique.

6. Si $a^2+4b=0$ et $a=0$, alors $u_n = 0$ pour $n\geq 2$.
Dans ce cas « trivial », les suites v et w dont les termes sont $(1,0,0,...)$ et $(0,1,0,0,...)$ forment une base de $E_{a,b}$.

où w_n est la suite définie par $w_n = nv_n = n \cdot \rho^n$ pour tout n de $\mathbb{N}$.

Tout élément u de $E_{a,b}$ peut donc s'écrire comme combinaison linéaire des deux suites v et w. Cette fois encore nous avons identifié deux suites v et w, qui ne sont pas colinéaires dans $\mathcal{F}(\mathbb{N}, \mathbb{R})$ et qui vérifient $E_{a,b} = \mathbf{Vect}[v, w]$.

Résumé : Dans tous les cas, l'espace $E_{a,b}$ est de dimension 2.

Cas 1 : l'équation caractéristique $r^2 = ar + b$ admet deux racines réelles distinctes ρ_1 et ρ_2.

Dans ce cas, une base de $E_{a,b}$ est donnée par la famille (v, w) où v et w sont les suites définies par

$v_n = \rho_1^n$ et $w_n = \rho_2^n$ pour chaque n de $\mathbb{N}$.

Cas 2 : l'équation caractéristique $r^2 = ar + b$ admet deux racines complexes conjuguées $\rho e^{i\theta}$ et $\rho e^{-i\theta}$.

Dans ce cas, une base de $E_{a,b}$ est donnée par la famille (v, w) où v et w sont les suites définies par

$v_n = \rho^n \cos(n\theta)$ et $w_n = \rho^n \sin(n\theta)$ pour chaque n de $\mathbb{N}$.

Cas 3 : l'équation caractéristique $r^2 = ar + b$ admet une racine double ρ.

Dans ce cas, une base de $E_{a,b}$ est donnée par la famille (v, w) où v et w sont les suites définies par

$v_n = \rho^n$ et $w_n = n\rho^n$ pour chaque n de $\mathbb{N}$.

4. Cardinal des familles libres et génératrices en dimension finie

Avant de poursuivre les développements théoriques, prenons le temps de souligner trois conséquences immédiates du long développement théorique de la section 2. Elles ont un intérêt pratique considérable.

Proposition 2.18 – Familles libres/génératrices et nombre de vecteurs.

Soit E un espace vectoriel de dimension n.

(a) Aucune famille comportant strictement plus de n vecteurs n'est libre.

(b) Aucune famille comportant strictement moins de n vecteurs n'est génératrice de E.

Démonstration. — Si $\mathcal{L}$ est une famille libre de E ; nous avons vu qu'il existe une base $\mathcal{B}$ de E qui contient $\mathcal{L}$ (théorème **??** de la base incomplète) ; comme $\mathcal{B}$ comporte n vecteurs, $\mathcal{L}$ comporte au plus n vecteurs.

De même, si $\mathcal{G}$ est une famille génératrice de E, nous avons vu qu'il existe une base $\mathcal{B}$ de E contenue dans $\mathcal{G}$ (théorème théorème **??** de la base extraite) ; comme $\mathcal{B}$ comporte n vecteurs, $\mathcal{G}$ comporte au moins n vecteurs.

□

***Exemple 2.19* (Exemples d'utilisation dans $\mathbb{R}^3$).** —

- La famille $\left(\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}4\\5\\6\end{pmatrix}\right)$ comporte deux vecteurs ; comme $\mathbb{R}^3$ est de dimension 3, elle ne peut pas être génératrice de $\mathbb{R}^3$.
- La famille $\left(\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}4\\5\\6\end{pmatrix}, \begin{pmatrix}7\\8\\9\end{pmatrix}, \begin{pmatrix}10\\11\\12\end{pmatrix}\right)$ comporte quatre vecteurs ; comme $\mathbb{R}^3$ est de dimension 3, elle ne peut pas être libre.

Exercices de manipulation. — II.6, II.7.

Proposition 2.20 – Comment vérifier qu'un espace est de dimension infinie

Soit E un espace vectoriel. Il y a équivalence entre

(i) il existe une famille libre de E de cardinal infini,

(ii) E est de dimension infinie.

En particulier, s'il existe une famille libre $(x_n)_{n\in\mathbb{N}}$ indexée par $\mathbb{N}$, alors E est de dimension infinie.

Démonstration. — L'implication (i) $\implies$ (ii) est claire, car sa contraposée « si E est de dimension finie, alors toute famille libre de E est de cardinal fini » est la partie (a) ci-dessus.

Pour montrer (ii) $\implies$ (i), supposons E de dimension infinie et construisons par récurrence une famille libre $(x_n)_{n\in\mathbb{N}}$ indexée par n.

- Partons d'un vecteur x_0 non nul de E ; on n'a pas $E = \mathbf{Vect}[x_0]$, sinon E serait de dimension 1. Ainsi il existe x_1 dans E tel que (x_1, x_0) soit libre.
- Soit $n \in \mathbb{N}$. Supposons construite une famille libre $(x_0, ..., x_n)$ de E comportant n vecteurs. On n'a pas $\mathbf{Vect}[x_0, ..., x_n] = E$, sinon E serait de dimension finie. Il existe donc un vecteur x_{n+1} de E

n'appartenant pas à $\mathbf{Vect}[x_0, ..., x_n]$. D'après le critère b. page 34, la famille $(x_0, ..., x_{n+1})$ est donc libre.

- On peut donc construire une famille $(x_i)_{i\in\mathbb{N}}$ telle que pour tout i de $\mathbb{N}$, la famille $(x_0, ...x_n)$ soit libre : d'après la caractérisation de la liberté pour les familles indexées par $\mathbb{N}$, c'est une famille libre de E, et elle est de cardinal infini.

□

***Exemple 2.21* (L'espace de tous les polynômes).** — Dans $\mathbb{K}[X]$, la famille $(X^n)_{n\in\mathbb{N}} = (1, X, X^2, ...)$ est libre. Comme elle comporte une infinité de vecteurs, $\mathbb{K}[X]$ est de dimension infinie (nous avions déjà vérifié page 47 que $\mathbb{K}[X]$ est de dimension infinie, mais ceci est plus « direct »).

***Exemple 2.22* (L'espace $\mathcal{F}(\mathbb{R}, \mathbb{R})$ est de dimension infinie).** — Dans $\mathcal{F}(\mathbb{R}, \mathbb{R})$, nous avons vérifié page 36 que $(x \mapsto \sin(kx))_{k\in\mathbb{N}}$ est libre. Comme elle comporte une infinité de vecteurs, $\mathcal{F}(\mathbb{R}, \mathbb{R})$ est de dimension infinie.

***Exemple 2.23* (L'espace des suites est de dimension infinie).** — Dans l'espace $\mathcal{F}(\mathbb{N}, \mathbb{R})$, nous avons défini page 12 une famille $(u^{[p]})_{p\in\mathbb{N}}$ qui comporte une infinité de suites, et nous avons vérifié qu'elle est libre à la page 35. On en déduit que $\mathcal{F}(\mathbb{N}, \mathbb{R})$ est de dimension infinie.

Exercices de manipulation. — II.8, II.9.

Proposition 2.24 – Pour les familles de n vecteurs en dimension n, « libre = génératrice »

Soit E un espace vectoriel de dimension n.
Si $(v_1, ...v_n)$ est une famille comportant exactement n vecteurs *de E, alors les assertions suivantes sont équivalentes.*

(i) La famille $(v_1, ...v_n)$ est une base de E,

(ii) La famille $(v_1, ...v_n)$ est libre,

(iii) La famille $(v_1, ...v_n)$ est génératrice de E.

Démonstration. — Utilisons la caractérisation des bases parmi les familles libres ou génératrices (proposition 2.4). Si $(v_1, ...v_n)$ est une famille libre de n vecteurs de E, alors c'est une famille libre maximale : la partie (a) de la proposition précédente dit qu'aucune famille ne peut contenir strictement $(v_1, ...v_n)$ et rester libre (elle aurait trop de vecteurs). Ainsi (ii) implique (i). De même, si $(v_1, ...v_n)$ est une famille génératrice, alors c'est une famille génératrice minimale grâce à la partie (b) de la proposition ci-dessus, et (iii) implique (i). Bien sûr (i) implique (ii) et (iii). □

***Exemple 2.25* (Une application dans $\mathbb{R}^3$).** — La famille $\left(\begin{pmatrix}1\\2\\0\end{pmatrix}, \begin{pmatrix}0\\1\\2\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}\right)$ comporte trois vecteurs de $\mathbb{R}^3$. Pour vérifier que c'est une base, il suffit de vérifier qu'elle est libre.

Si $\lambda_1, \lambda_2, \lambda_3$ sont trois réels qui vérifient $\lambda_1 \begin{pmatrix}1\\2\\0\end{pmatrix} + \lambda_2 \begin{pmatrix}0\\1\\2\end{pmatrix} + \lambda_3 \begin{pmatrix}1\\0\\1\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$, alors $\begin{cases}\lambda_1 + \lambda_3 = 0\\ 2\lambda_1 + \lambda_2 = 0\\ 2\lambda_2 + \lambda_3 = 0\end{cases}$, d'où on déduit facilement $\lambda_1 = \lambda_2 = \lambda_3 = 0$. Donc la famille est libre, et c'est *automatiquement* une base de E (pas besoin de vérifier qu'elle est génératrice).

***Exemple 2.26* (Famille de polynômes échelonnée en degré).** — Si $(P_0, P_1, ...P_n)$ est une famille échelonnée en degré de $\mathbb{K}_n[X]$, alors c'est une famille de polynômes à degrés deux à deux distincts, donc elle est libre ; nous avons vu que la dimension de $\mathbb{K}_n[X]$ est $n+1$ et cette famille libre comporte $n+1$ polynômes, c'est donc automatiquement une base de $\mathbb{K}_n[X]$. Cela fournit une alternative à la démonstration page 45 (qui est très simple en apparence, mais dont la simplicité s'appuie sur le gros travail théorique de ce chapitre).

***Remarque 2.27* (Utilisation pratique de ce résultat).** — Étant donnés un espace vectoriel de dimension finie E et une famille $(x_1, .., x_n)$ de vecteurs de E, lrosqu'on se demande si $(x_1, .., x_n)$ est une base de E, il ne faut pas confondre les deux situations suivantes :

(a) Si on *connaît déjà la dimension de* E et si cette dimension est égale au nombre n de vecteurs de la famille, il suffit de vérifier soit que la famille est libre, soit qu'elle est génératrice, mais il est inutile de vérifier les deux séparément.
(si au contraire la dimension de E n'est pas n, la famille $(x_1, .., x_n)$ n'a aucune chance d'être une base de E).

(b) Si on *ne connaît pas déjà la dimension de* E, on ne peut pas utiliser ce résultat : impossible d'échapper à vérifier séparément que la famille est libre puis qu'elle est génératrice, ou alors à vérifier que tout vecteur peut s'écrire de façon unique comme combinaison linéaire de la famille. Si on arrive à vérifier que $(x_1, .., x_n)$ est une base de E, on pourra en *déduire* que la dimension de E est n.

Exercices de manipulation. — II.10, II.11, II.12.

5. Dimension d'un sous-espace, dimension d'un produit, dimension complexe

Proposition 2.28 – Dimension d'un sous-espace et cas d'égalité

Soit E un espace vectoriel de dimension finie.

(i) Si F est un sous-espace vectoriel de E, alors F est de dimension finie et $\dim(F) \leq \dim(E)$ *;*

(ii) De plus, si F est un sous-espace vectoriel de E et si $\dim(F) = \dim(E)$*, alors $F = E$.*

La propriété (ii) est très souvent utilisée pour montrer que deux sous-espaces sont égaux en ne vérifiant que l'une des deux inclusions.

***Exemple 2.29* (Utilisation de (ii)).** — Reprenons l'exemple de la page 23 : dans $\mathbb{R}^3$, comparons $\mathbf{Vect}\left[\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}1\\0\\-1\end{pmatrix}\right]$ et $\mathbf{Vect}\left[\begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}0\\2\\4\end{pmatrix}\right]$.

Chacun de ces deux sous-espaces est de dimension 2 (car engendré par deux vecteurs non colinéaires).

Pour montrer que les deux sous-espaces sont égaux, il suffit donc de montrer l'une des inclusions, l'autre se déduira automatiquement de l'égalité des dimensions.

Écrire $\begin{pmatrix}1\\2\\3\end{pmatrix} + \begin{pmatrix}1\\0\\-1\end{pmatrix} = 2\begin{pmatrix}1\\1\\1\end{pmatrix}$ et $\begin{pmatrix}1\\2\\3\end{pmatrix} - \begin{pmatrix}1\\0\\-1\end{pmatrix} = \begin{pmatrix}0\\2\\4\end{pmatrix}$ prouve $\mathbf{Vect}\left[\begin{pmatrix}1\\1\\1\end{pmatrix}, \begin{pmatrix}0\\2\\4\end{pmatrix}\right] \subset \mathbf{Vect}\left[\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}1\\0\\-1\end{pmatrix}\right]$, donc l'égalité.

Démonstration. —

- Si F est un sous-espace vectoriel de E, toute famille libre de F est une famille libre de E.

Par conséquent, aucune famille libre de F ne peut être de cardinal infini, et F est de dimension finie ; de plus, toute famille libre de F est de cardinal inférieur ou égal à $\dim(E)$, donc le cardinal commun à toutes les bases de F est inférieur ou égal à $\dim(E)$, ce qui prouve (i).

- Si $\dim(F) = \dim(E)$ et si $(u_1, ...u_n)$ est une base de F, alors la famille $(u_1, ...u_n)$ est une famille libre de n vecteurs dans E qui est de dimension n : c'est donc aussi une base de E, donc une famille génératrice de E. La famille $(u_1, ...u_n)$ est donc génératrice de F et de E, ainsi $F = \mathbf{Vect}[u_1, ...u_n] = E$.

□

Exercices de manipulation. — II.13, II.14.

Proposition 2.30 – Dimension d'un espace produit

Si E et F sont deux espaces vectoriels de dimension finie, alors

$$\dim(E \times F) = \dim(E) + \dim(F).$$

Démonstration. — Nous devons montrer qu'il existe une base de $E \times F$ dont le cardinal est $\dim(E)+\dim(F)$.

Partons d'une base $\mathcal{E} = (e_1, e_2, ...e_n)$ de E et d'une base $\mathcal{F} = (f_1, f_2, ...f_p)$ de F.

Les vecteurs $(e_1, \mathbf{0}_F)$, $(e_2, \mathbf{0}_F)$, $...(e_n, \mathbf{0}_F)$, $(\mathbf{0}_E, f_1)$, $(\mathbf{0}_E, f_1)$, ... $(\mathbf{0}_E, f_p)$ sont des vecteurs de $E \times F$. Il y en a $n + p$.

Pour montrer qu'ils forment une base de $E \times F$, il suffit de montrer que si (u, v) est un élément de $E \times F$, il existe un unique $(n + p)$-uplet $(\lambda_1, ...\lambda_n, \mu_1, ...\mu_p)$ de $n + p$ scalaires tel que

$$(u, v) = \lambda_1(e_1, \mathbf{0}_F) + ... + \lambda_n(e_n, \mathbf{0}_F) + \mu_1(\mathbf{0}_E, f_1) + ... + \mu_p(\mathbf{0}_E, f_p). \tag{5.1}$$

Comme le membre de droite est le vecteur $\left(\sum_{i=1}^{n} \lambda_i e_i, \sum_{j=1}^{p} \mu_j f_j\right)$, nous sommes en train de chercher les n-uplets $(\lambda_1, ...\lambda_n)$ de scalaires qui vérifient

$$u = \sum_{i=1}^{n} \lambda_i e_i \tag{5.2}$$

et les p-uplets $(\mu_1, ...\mu_p)$ de scalaires qui vérifient

$$v = \sum_{j=1}^{p} \mu_j f_j. \tag{5.3}$$

Puisque $(e_1, ...e_n)$ est une base de E, il y a bien une unique solution $(\lambda_1, ...\lambda_n)$ pour (5.2) : elle est donnée par les coordonnées de u dans la base $\mathcal{E}$. Puisque Comme $(f_1, ...f_p)$ est une base de F, il y a bien une unique solution $(\mu_1, ...\mu_p)$ pour (5.3) : elle est donnée par les coordonnées de v dans la base $\mathcal{F}$. □

Proposition 2.31 – Dimension sur $\mathbb{C}$ et dimension sur $\mathbb{R}$

Soit E un $\mathbb{C}$-espace vectoriel de dimension finie. Si l'on note $\dim_{\mathbb{C}}(E)$ la dimension de E lorsqu'on voit E comme $\mathbb{C}$-espace vectoriel, et $\dim_{\mathbb{R}}(E)$ la dimension de E lorsqu'on voit E comme $\mathbb{R}$-espace vectoriel, alors

$$\dim_{\mathbb{R}}(E) = 2\dim_{\mathbb{C}}(E).$$

Démonstration. —

- Soit $(x_1, ..., x_n)$ une base du $\mathbb{C}$-espace vectoriel E. Comme E est un $\mathbb{C}$-espace vectoriel, les vecteurs $i \cdot x_1, i \cdot x_2, ..., i \cdot x_n$ ont un sens. Considérons alors la famille $(x_1, i \cdot x_1,\ ...,\ x_n,\ i \cdot x_n)$: c'est une famille de $2n$ vecteurs de E. Montrons que c'est une base du $\mathbb{R}$-espace vectoriel E.
- Soit x un élément de E. Comme $(x_1, ..., x_n)$ est une base du $\mathbb{C}$-espace vectoriel E, il existe un unique n-uplet $(\lambda_1, ..., \lambda_n)$ de *nombres complexes* vérifiant $x = \sum\limits_{i=1}^{n} \lambda_i x_i$. Mais alors, on a $x = \mathfrak{Re}(\lambda_1)x_1 + \mathfrak{Im}(\lambda_1)(i \cdot x_1) + ... + \mathfrak{Re}(\lambda_n)x_n + \mathfrak{Im}(\lambda_n)(i \cdot x_1)$, donc x est combinaison linéaire *à coefficients réels* de $(x_1, i \cdot x_1 ..., x_n, i \cdot x_n)$.
- Y a-t-il une autre combinaison possible ? Si $x = \alpha_1 x_1 + \beta_1(i \cdot x_1) + .. + \alpha_n x_n + \beta_n(i \cdot x_n)$ avec les α_i et β_i réels, on a $x = \sum\limits_{i=1}^{n}(\alpha_i + i\beta_i)x_i$, et comme la seule écriture possible de x comme combinaison complexe des x_i est $x = \sum\limits_{i=1}^{n} \lambda_i x_i$, on a $\alpha_i = \mathfrak{Re}(\lambda_i)$ et $\beta_i = \mathfrak{Im}(\lambda_i)$.
- Ainsi, il y a une unique façon d'exprimer x comme combinaison linéaire à coefficients réels de la famille $(x_1, i \cdot x_1,\ ...,\ x_n,\ i \cdot x_n)$, c'est ce qu'il fallait démontrer.

$\square$

6. Rang d'une famille de vecteurs ou d'une matrice

6.1. Définition et premières propriétés. —

Définition 2.32 – Rang d'une famille de vecteurs

Soient E un espace vectoriel et $(x_i)_{i\in I}$ une famille de vecteurs de E.
Lorsque $\mathbf{Vect}[(x_i)_{i\in I}]$ est de dimension finie, on appelle *rang* de $(x_i)_{i\in I}$ sa dimension. On note :

$$\mathbf{rg}[(x_i)_{i\in I}] \stackrel{\text{def}}{=} \dim \mathbf{Vect}[(x_i)_{i\in I}]\,.$$

Remarque 2.33. —
- Lorsque E est de dimension finie, le rang d'une famille de vecteurs de E est toujours inférieur ou égal à $\dim E$.
- Lorsque $\mathbf{Vect}[(x_i)_{i\in I}]$ est de dimension infinie, on dit que la famille $(x_i)_{i\in I}$ est *de rang infini.*

Exemple 2.34. —
- Dans $\mathbb{R}^3$, la famille $\left(\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}\right)$ est de rang 2.
- Dans $\mathbb{R}^3$, on a : $\mathbf{rg}\left[\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}8\\5\\0\end{pmatrix}\right] = 2.$

Exemple 2.35. — Dans $\mathbb{K}[X]$, on a : $\mathbf{rg}\big[1, X, 3X, X^2, 5X^2+X+8\big] = 3$ et $\mathbf{rg}\big[1, X, X^4, X^7, X^10\big] = 5$.

Proposition 2.36 – Rang d'une famille VS cardinal de la famille et cas d'égalité

Si k est un entier naturel non nul et si $(x_1, ..., x_k)$ est une famille de k vecteurs de E, alors $\mathbf{rg}[x_1, ...x_k] \le k$.
Pour qu'il y ait égalité, il faut et il suffit que la famille $(x_1, ..., x_k)$ soit libre.

Démonstration. — Notons $\mathcal{G}$ la famille $(x_1, ...x_k)$. Elle engendre $\mathbf{Vect}[x_1, ...x_k]$, donc $k = \mathrm{Card}(\mathcal{G})$ est supérieur ou égal à la dimension de $\mathbf{Vect}[x_1, ...x_k]$ (voir le point (b) de la proposition 2.18). Puisque $\dim(\mathbf{Vect}[x_1, ...x_k])$ est par définition le rang de la famille $(x_1, .., x_k)$, on obtient l'inégalité $\mathbf{rg}[x_1, ...x_k] \le k$. Compte tenu de la proposition 2.24, dire qu'il y a de plus égalité, c'est dire que $\mathcal{G}$ est en fait une base de $\mathbf{Vect}[x_1, ...x_k]$, autrement dit, que $(x_1, ..., x_k)$ est libre.

$\square$

Définition 2.37 – Rang d'une matrice

Soit A une matrice à n lignes et p colonnes à coefficients dans $\mathbb{K}$.
Le *rang de* A, noté $\mathbf{rg}[A]$, est la dimension du sous-espace vectoriel de $\mathbb{K}^n$ engendré par les colonnes de A.

Remarque 2.38. — Vous avez vu au premier semestre une autre notion de rang d'une matrice (le nombre de lignes non nulles de la réduite échelonnée de A). Il n'est pas difficile de voir que si A est échelonnée réduite, les deux notions coïncident. Nous montrerons dans la suite de ce chapitre que les deux notions coïncident en fait pour toutes les matrices.

Remarque 2.39. — La notion de rang pour une matrice apparaît, grâce à la définition et la remarque ci-dessus, comme un cas particulier de la notion de rang pour une famille de vecteurs.

Pour les familles finies de vecteurs de $\mathbb{K}^n$, la réciproque est vraie : le rang d'une famille peut être étudié matriciellement.

En effet, supposons donnés p vecteurs $x_1, ... x_p$ de $\mathbb{K}^n$.

Écrivons $x_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix}$, $x_2 = \begin{pmatrix} a_{12} \\ \vdots \\ a_{n2} \end{pmatrix}$, ... $x_p = \begin{pmatrix} a_{1p} \\ \vdots \\ a_{np} \end{pmatrix}$, si bien que la matrice $A = \begin{pmatrix} a_{11} & \dots & a_{1p} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{np} \end{pmatrix}$ a pour colonnes les coordonnées des x_i dans la base canonique. À l'évidence, le rang de A et le rang de la famille $(x_1, ..., x_n)$ sont identiques.

Pour chercher le rang de la famille $(x_1, ..., x_n)$, on peut donc chercher le rang de A.

Exercices de manipulation. — II.15, II.16.

6.2. Opérations élémentaires et détermination pratique du rang. — Dans ce paragraphe, on explique comment calculer concrètement le rang d'une famille de vecteurs de $\mathbb{K}^n$ ou d'une matrice à coefficients dans $\mathbb{K}$.

- Rappelons qu'on dit qu'une matrice est *échelonnée en colonnes* lorsque
  - chaque colonne commence par strictement plus de zéros que la précédente,
  - si une colonne est nulle, alors toutes les suivantes sont nulles.

Par exemple, les matrices $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & 0 \\ 3 & 1 & 0 \\ 4 & 2 & 0 \\ 5 & 3 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 1 & 2 & 0 \\ 8 & 7 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 8 & 5 & 0 & 0 \\ 4 & 3 & 0 & 0 \end{pmatrix}$ sont échelonnées en colonnes,

tandis que les matrices $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 4 & 3 & 2 & 1 \end{pmatrix}$, $\begin{pmatrix} 0 & 5 \\ 3 & 0 \\ 0 & 1 \\ 4 & 2 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 & 6 \\ 2 & 4 & 0 \\ 3 & 0 & 0 \\ 1 & 5 & 1 \end{pmatrix}$ ne sont pas échelonées en colonnes.

- Nous partons de la remarque suivante :

Si la matrice A est échelonnée en colonnes, le rang de A est égal au nombre de colonnes non nulles.

Démonstration (je recommande de se convaincre du résultat sur les exemples ci-dessus avant de la lire...)

Si la matrice A est échelonnée en colonnes et si les colonnes non-nulles de A sont les r premières, notées $C_1, ..., C_r$, alors elles forment une famille libre : si $\lambda_1, ..., \lambda_r$ sont des scalaires vérifiant $\lambda_1 C_1 + ... + \lambda_r C_r = \mathbf{0}_{\mathbb{K}^n}$, et si l'on considère le plus petit entier i tel que a_{i1} soit non-nul, on constate que le coefficient en $i$ème position dans tous les C_j, $j > 1$, est nul, si bien que le coefficient en $i$ème position dans $\lambda_1 C_1 + ... + \lambda_r C_r$ est $\lambda_1 a_{i1}$: comme le vecteur $\lambda_1 C_1 + ... + \lambda_r C_r$ est nul, ce coefficient doit être nul, donc λ_1 doit être nul. En poursuivant ainsi, on constate que tous les λ_i doivent être nuls. □

• Expliquons maintenant comment ramener la recherche du rang d'une famille ou matrice quelconque au cas échelonné.

Rappelons que les *opérations élémentaires sur les colonnes* sont les suivantes :
- Effectuer une *transvection*, c'est-à-dire remplacer l'une des colonnes C_j par $C_j + \alpha C_i$ où $i \neq j$ et $\alpha \in \mathbb{K}$;
- Effectuer une *dilatation*, c'est-à-dire remplacer l'une des colonnes C_j par λC_j où $\lambda \in \mathbb{K}^\star$;
- Effectuer une *permutation*, c'est-à-dire échanger deux colonnes C_i et C_j, $i \neq j$.

La remarque suivante est essentielle :
- on a $\mathbf{Vect}[C_1, ..., C_p] = \mathbf{Vect}[C_1, ..., C_{j-1}, C_j + \alpha C_i, C_{j+1} ..., C_p]$
  (toute combinaison linéaire de la deuxième famille est évidemment combinaison linéaire de la première, et la réciproque est vraie : tous les $C_k, k \neq j$ ainsi que $C_j = \alpha C_i + (C_j - \alpha C_i)$ appartiennent à $\mathbf{Vect}[C_1, ..., C_{j-1}, C_j + \alpha C_i, C_{j+1} ..., C_p]$).
- De même, $\mathbf{Vect}[C_1, ..., C_p] = \mathbf{Vect}[C_1, ..., C_{j-1}, \lambda C_j, C_{j+1} ..., C_p]$.
- Enfin, une permutation n'a visiblement pas d'effet sur le sous-espace engendré par les colonnes.

Par conséquent, le sous-espace vectoriel engendré par les colonnes est inchangé lorsqu'on effectue une des opérations élémentaires. Sa dimension est donc aussi inchangée :

Les opérations élémentaires sur les colonnes préservent le rang.

• Or, on sait que partant d'une matrice quelconque, il existe une suite finie d'opérations élémentaires (« méthode du pivot », voir le premier semestre) qui mène à une matrice échelonnée en colonnes. Cette « réduite échelonnée » a donc le même rang que la matrice de départ, et aucun travail n'est nécessaire pour déterminer son rang.

La « méthode du pivot », effectuée sur les colonnes, permet de trouver le rang d'une matrice.

• **Colonnes ou lignes ?**
Nous verrons plus loin (chapitre 6) que le rang d'une matrice est le même que le rang de sa transposée. Or, pour effectuer une opération élémentaire sur les *lignes* d'une matrice A, on peut prendre le chemin détourné suivant : transposer A, effectuer l'opération élémentaire voulue sur les colonnes de A^T, puis prendre la transposée du résultat. Du fait qu'aucune de ces opérations ne change le rang, on déduit que les opérations élémentaires sur les lignes préservent également le rang. On peut alors utiliser la « méthode du pivot » pour se ramener à une matrice échelonnée en lignes, dont le rang (toujours grâce au théorème à démontrer plus tard sur le rang de la transposée) est le nombre de lignes non nulles.

Pour trouver le rang d'une matrice, on peut aussi effectuer des opérations élémentaires sur les lignes, comme au premier semestre.

Exemple 2.40. — Déterminons le rang de la matrice $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 2 & 3 \\ 1 & 1 & 1 & 2 \\ 0 & 1 & 1 & 1 \end{pmatrix}$.

D'après les résultats que nous venons de montrer, cette matrice a le même rang que les matrices

$$\begin{pmatrix}1&2&3&4\\1&1&2&3\\1&1&1&2\\0&1&1&1\end{pmatrix},\begin{pmatrix}1&0&0&0\\1&-1&-1&-1\\1&-1&-2&-2\\0&1&1&1\end{pmatrix},\begin{pmatrix}1&0&0&0\\1&-1&0&0\\1&-1&-1&-1\\0&1&0&0\end{pmatrix},\begin{pmatrix}1&0&0&0\\1&-1&0&0\\1&-1&-1&0\\0&1&0&0\end{pmatrix}.$$

et la dernière matrice, échelonnée en colonnes, est visiblement de rang 3, donc le rang de A est 3.

On aurait aussi pu passer par des opérations élémentaires sur les lignes : la matrice A a le même rang que les matrices

$$\begin{pmatrix}1&2&3&4\\1&1&2&3\\1&1&1&2\\0&1&1&1\end{pmatrix},\begin{pmatrix}1&2&3&4\\0&-1&-1&-1\\0&-1&-2&-2\\0&1&1&1\end{pmatrix},\begin{pmatrix}1&2&3&4\\0&-1&-1&-1\\0&0&-1&-1\\0&0&0&0\end{pmatrix}$$

et la dernière est échelonnée en lignes et visiblement de rang 3.

Exercices de manipulation. — II.17, II.18.

7. Description des sous-espaces de $\mathbb{K}^n$: repères ou équations cartésiennes

Dans ce paragraphe, on démontre que tout sous-espace vectoriel de $\mathbb{K}^n$ est l'ensemble des solutions d'un système linéaire homogène à coefficients dans $\mathbb{K}$; on explique comment passer concrètement de la description d'un sous-espace à l'aide d'une famille génératrice à la description du même sous-espace comme ensemble des solutions d'un système linéaire.

7.1. La question. — Soient F un sous-espace vectoriel de $\mathbb{K}^n$ et $(u_1, ...u_p)$ une famille génératrice finie de F.

Dire qu'un vecteur $v = \begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}$ de $\mathbb{K}^n$ appartient à F, c'est donc dire que v appartient à $\mathbf{Vect}[u_1, ...u_p]$, c'est-à-dire qu'il existe des éléments $\lambda_1, ..., \lambda_p$ de $\mathbb{K}$ vérifiant : $v = \lambda_1 u_1 + ... + \lambda_p u_p$.

Notons comme précédemment $u_1 = \begin{pmatrix}a_{11}\\ \vdots\\ a_{n1}\end{pmatrix}$, $u_2 = \begin{pmatrix}a_{12}\\ \vdots\\ a_{n2}\end{pmatrix}$, ... $u_p = \begin{pmatrix}a_{1p}\\ \vdots\\ a_{np}\end{pmatrix}$. La question à étudier est donc celle de

l'existence d'au moins une solution au système linéaire $(S) : \begin{cases} a_{11}\lambda_1 + ... + a_{1p}\lambda_p = & x_1 \\ \quad\vdots & \vdots \\ a_{n1}\lambda_1 + ... + a_{np}\lambda_p = & x_n \end{cases}$, d'inconnues $\lambda_1, ..., \lambda_p$.

Il s'agit donc de chercher les conditions sur le membre de droite pour que ce système admette au moins une solution.

7.2. Un exemple concret. — .

Plaçons-nous dans $\mathbb{R}^4$, considérons $u_1 = \begin{pmatrix}1\\1\\-1\\-1\end{pmatrix}, u_2 = \begin{pmatrix}1\\-1\\1\\-1\end{pmatrix}, u_3 = \begin{pmatrix}1\\-1\\-1\\1\end{pmatrix}, u_4 = \begin{pmatrix}1\\-1\\0\\0\end{pmatrix}$ et le sous-espace $F = \mathbf{Vect}[u_1, u_2, u_3, u_4]$.

Un vecteur $\begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix}$ de $\mathbb{R}^4$ appartient à F si et seulement s'il existe des réels $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ vérifiant :

$$\begin{cases} \lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 & = x, \\ \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4 & = y, \\ -\lambda_1 + \lambda_2 - \lambda_3 & = z, \\ -\lambda_1 - \lambda_2 + \lambda_3 & = t. \end{cases}$$

Ce système équivaut à

$$\begin{cases} \lambda_1 + \lambda_2 \quad + \lambda_3 + \lambda_4 & = x \\ \quad -2\lambda_2 - 2\lambda_3 - 2\lambda_4 & = y - x \\ \quad +2\lambda_2 \qquad + \lambda_4 & = z + x \\ \qquad +2\lambda_3 \ + \lambda_4 & = t + x \end{cases}$$

ou encore, en allant au bout de la méthode du pivot,

$$\begin{cases} \lambda_1 + \lambda_2 \quad + \lambda_3 + \lambda_4 & = x \\ \quad -2\lambda_2 - 2\lambda_3 - 2\lambda_4 & = y - x \\ \qquad -2\lambda_3 \ - \lambda_4 & = z + y \\ \qquad 0 & = t + x + (z + y) \end{cases}$$

Le dernier système est échelonné ; visiblement, pour qu'il ait une solution, il est *nécessaire* qu'on ait $x + y + z + t = 0$, et si cette condition est vérifiée, *il y a bien au moins une solution*, par exemple donnée par $\lambda_4 = 0$, $\lambda_3 = -\frac{z+y}{2}$, $\lambda_2 = \frac{z+x}{2}$, $\lambda_1 = \frac{y+x}{2}$.
On en déduit que v appartient à F si et seulement si ses coordonnées vérifient l'équation $x + y + z + t = 0$.

7.3. Une stratégie générale. — .

- Pour étudier l'existence d'au moins une solution au système linéaire $(S) : \begin{cases} a_{11}\lambda_1 + ... + a_{1p}\lambda_p = & x_1 \\ \vdots & \vdots \\ a_{n1}\lambda_1 + ... + a_{np}\lambda_p = & x_n \end{cases}$
d'inconnues $\lambda_1, ..., \lambda_p$, on peut naturellement utiliser la « méthode du pivot ».
- Vous savez que des opérations élémentaires sur les lignes permettent de voir que (S) est équivalent à un système *échelonné en lignes* dont le nombre de lignes non nulles est, vu le numéro 2.6, le rang r de la famille $(u_1, ..., u_p)$.
- Ainsi (S) est équivalent à un système de la forme

$$(S') : \begin{cases} \tilde{a}_{11}\lambda_1 + \dots\dots\dots\dots\dots\dots + \tilde{a}_{1p}\lambda_p = & y_1 \\ 0 + \tilde{a}_{2j_2}\lambda_{j_2} + \dots\dots\dots\dots + \tilde{a}_{2p}\lambda_p = & y_2 \\ \quad \vdots & \vdots \\ 0 + \dots + 0 + \tilde{a}_{rj_r}\lambda_{j_r} + \dots + \tilde{a}_{rp}\lambda_p = & y_r \\ 0 + \dots\dots\dots\dots\dots\dots\dots\dots + 0 = & y_{r+1} \\ \quad \vdots & \vdots \\ 0 + \dots\dots\dots\dots\dots\dots\dots\dots + 0 = & y_n \end{cases}$$

ayant les deux propriétés suivantes :

(a) sur chaque ligne i, l'entier $j_i = \min(j \in \{1, ..p\} \mid a_{ij} \neq 0)$ qui donne la position du premier coefficient non nul est strictement supérieur à celui j_{i-1} de la ligne précédente ;

(b) les seconds membres y_i sont obtenus à partir des coordonnées x_i en faisant des combinaisons linéaires :
pour chaque i, il existe des scalaires $\beta_{i1}, ...\beta_{ip}$ tels que y_i soit égal à $\beta_{i1}x_1 + ... + \beta_{ip}x_p$.

La propriété (a) traduit le fait que (S') est échelonné en lignes, la propriété (b) traduit le fait que les opérations élémentaires sur les lignes ne font intervenir que des combinaisons linéaires des lignes.

- On constate alors que pour qu'il y ait une solution au système (S'), il est *nécessaire* que les seconds membres des $(n-r)$ dernières lignes soient nuls, c'est-à-dire que les x_i vérifient les $n-r$ équations

$$(C)\begin{cases} \beta_{(r+1)1}\ x_1 + \ldots + \beta_{(r+1)p}\ x_p & = 0 \\ \vdots & \vdots \\ \beta_{n1}\quad x_1 + \ldots + \quad \beta_{np}\quad x_p & = 0 \end{cases}$$

- C'est également *suffisant* : si ces équations sont vérifiées, alors on obtient une solution de (S'), donc de (S),
en choisissant $\lambda_{j_r} = \dfrac{y_r}{\tilde{a}_{rj_r}}$ et tous les λ_k nuls pour $k > j_r$, ce qui garantit que la $r^{\text{ème}}$ équation de (S') est vérifiée,
puis en choisissant $\lambda_{j_{r-1}} = y_{r-1} - \tilde{a}_{(r-1)j_r}\lambda_{j_r}$ et tous les λ_k nuls pour k strictement compris entre j_{r-1} et j_r,
ce qui garantit que la $(r-1)^{\text{ème}}$ équation de (S') est vérifiée, et en poursuivant ainsi jusqu'à la première ligne.

- Ainsi le système (S) admet une solution, autrement dit v est dans F, si et seulement si les **conditions de compatibilité** (C) sont vérifiées. Ces conditions sont des équations linéaires homogènes portant sur les coordonnées de v.
- Il y a $(n-r)$ équations dans le système (C). Rappelons le lien avec la dimension de F : l'entier r était le rang de $(u_1, ...u_p)$, c'est donc la dimension de $\mathbf{Vect}[u_1, ..., u_p]$, mais cette famille est génératrice de F, donc $r = \dim F$.

Nous avons démontré le résultat suivant :

Théorème 2.41 – Description des S.E.V. de $\mathbb{K}^n$ par des équations linéaires homogènes

Tout sous-espace vectoriel de $\mathbb{K}^n$ est l'ensemble des solutions d'un système linéaire homogène à coefficients dans $\mathbb{K}$.
Si F est un sous-espace vectoriel de dimension r de $\mathbb{K}^n$,
alors il existe un système linéaire homogène de $(n-r)$ équations dont F est l'ensemble des solutions.

Exercices de manipulation. — II.19, II.20.

Exercices de manipulation du cours du chapitre 2

Exercice de manipulation 2.1. —
Montrer que la famille $\left(\begin{pmatrix}-1\\1\\1\end{pmatrix},\begin{pmatrix}1\\-1\\1\end{pmatrix},\begin{pmatrix}1\\1\\-1\end{pmatrix}\right)$ est une base de $\mathbb{R}^3$. Quelles sont les coordonnées du vecteur $\begin{pmatrix}8\\4\\2\end{pmatrix}$ dans cette base ?

Exercice de manipulation 2.2. —
Montrer que la famille $\left[\begin{pmatrix}1&0\\0&0\end{pmatrix},\begin{pmatrix}1&1\\0&0\end{pmatrix},\begin{pmatrix}1&1\\1&0\end{pmatrix},\begin{pmatrix}1&1\\1&1\end{pmatrix}\right]$ est une base de $\mathcal{M}_2(\mathbb{R})$.

Quelles sont les coordonnées de $\begin{pmatrix}10&9\\7&4\end{pmatrix}$ dans cette base ? Celles de $\begin{pmatrix}7&7\\7&7\end{pmatrix}$?

Exercice de manipulation 2.3. —
Dans $\mathbb{R}_3[X]$, on considère les quatre polynômes $\mathbf{P}(X)=1$, $\mathbf{Q}(X)=1+X$, $\mathbf{U}(X)=1+X+X^2$, $\mathbf{V}(X)=1+X+X^2+X^3$.

1. Montrer que la famille $(\mathbf{P},\mathbf{Q},\mathbf{U},\mathbf{V})$ est une base de $\mathbb{R}_3[X]$.
2. Quelles sont les coordonnées du polynôme X^3 dans cette base ? Celles du polynôme X^3+2X^2+X ?

Exercice de manipulation 2.4. —
Quelle est la dimension de l'espace des matrices de $\mathcal{M}_2(\mathbb{R})$ qui n'ont que des zéros sur la diagonale ?
On demande la réponse et une démonstration.

Exercice de manipulation 2.5. —
Dimension de l'espace des matrices antisymétriques (voir page 52).

Pour tout (i,j) de $\{1,..,n\}^2$ vérifiant $i<j$ (remarquer l'inégalité stricte), notons A_{ij} la matrice $E_{ij}-E_{ji}$.

(i) Vérifier que c'est automatiquement une matrice antisymétrique.
(ii) Vérifier que la famille $(A_{ij})_{1\leq i\leq j\leq n}$ est une base de $\mathcal{A}_n(\mathbb{R})$.
(iii) En déduire que la dimension de $\mathcal{A}_n(\mathbb{R})$ est $\dfrac{n(n-1)}{2}$.

Exercice de manipulation 2.6. —
Dire d'abord quelles sont les familles dont on peut affirmer *simplement en regardant le nombre de vecteurs* qu'elles ne sont pas libres ou pas génératrices. Essayer ensuite de voir sans prendre de stylo lesquelles sont libres, génératrices.

- Dans $\mathbb{R}^4$,
  — $((1,2,3,4),(5,6,7,8))$,
  — $((0,-1,0,2),(1,1,1,1),(1,2,3,4),(9,8,9,11))$,
  — $((1,2,3,4),(1,1,2,3),(1,1,1,2),(1,1,1,1),(0,0,0,1))$,
  — $((1,0,0,0),(0,1,0,0),(0,0,4,5))$,
  — $((2,2,2,2),(3,3,3,3))$,
  — $((0,-1,0,2),(1,1,1,1),(0,0,0,0),(9,9,9,10))$,
  — $((0,-1,0,2),(1,1,1,1),(0,0,0,0),(0,-4,2,7),(9,9,9,10))$.
- Dans $\mathbb{R}_2[X]$,
  — $(1,\ X,\ X^2,\ 2X^2+3X)$,
  — $(3X,\ 7X^2+3)$,
  — $(1,\ 6X+4,\ -X^2+5X+9)$,

— $(3,\ 8X+2,\ 9X^2,\ X^2-3)$
— $(3X,\ 8X+2,\ 9X^2,\ X^2-3X)$,
— $(X^2-3,\ X^2+X,\ X^2)$,
— $(X^2-3,\ X^2+X,\ X^2+X+7)$.

- Dans $\mathcal{M}_2(\mathbb{R})$,

— $\left(\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}1&5\\-3&2\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix}\right)$,
— $\left(\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}0&5\\0&2\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix},\begin{pmatrix}1&1\\0&1\end{pmatrix},\begin{pmatrix}1&2\\0&4\end{pmatrix}\right)$,
— $\left(\begin{pmatrix}1&1\\0&1\end{pmatrix},\begin{pmatrix}2&3\\0&4\end{pmatrix},\begin{pmatrix}9&-9\\0&12\end{pmatrix},\begin{pmatrix}8&8\\0&0\end{pmatrix}\right)$
— $\left(\begin{pmatrix}1&2\\0&1\end{pmatrix},\begin{pmatrix}0&-3\\0&7\end{pmatrix},\begin{pmatrix}0&1\\\pi&0\end{pmatrix},\begin{pmatrix}0&31\\0&9\end{pmatrix}\right)$

Exercice de manipulation 2.7. —
Soit A une matrice de $\mathcal{M}_3(\mathbb{R})$ vérifiant : $A^7=0$.
Décrire la famille $(A^n)_{n\in\mathbb{N}}$. Peut-elle être génératrice de $\mathcal{M}_3(\mathbb{R})$?

Exercice de manipulation 2.8. —

1. Pour chaque a de $\mathbb{R}^+$, on note f_a la fonction de $\mathbb{R}$ dans $\mathbb{R}$ définie par : $f_\alpha(x)=\begin{cases}1 & \text{si } x=-\alpha \text{ ou } x=\alpha,\\ 0 & \text{sinon.}\end{cases}$
   Dire pourquoi la famille $(f_\alpha)_{\alpha\in\mathbb{R}}$ est libre dans $\mathcal{F}(\mathbb{R},\mathbb{R})$ (on pourra saisir l'occasion pour relire le §7.2).
2. En déduire que l'espace des *fonctions paires* de $\mathbb{R}$ dans $\mathbb{R}$ est de dimension infinie.

Exercice de manipulation 2.9. —

1. Pour chaque k de $\mathbb{N}^\star$, on note ℓ_k la fonction $x\mapsto \ln(x)^k$ de $\mathbb{R}_\star^+$ dans $\mathbb{R}$. Montrer que la famille $(\ell_k)_{k\in\mathbb{N}^\star}$ est libre dans $\mathcal{F}(\mathbb{R}_\star^+,\mathbb{R})$ (raisonner par récurrence et dériver, comme dans l'exemple de la page 36).
2. En déduire que l'espace $\left\{f\in\mathcal{F}(\mathbb{R}_\star^+,\mathbb{R}), f(1)=0\right\}$ est de dimension infinie.

Exercice de manipulation 2.10. —
Démontrer sans aucun calcul que la famille $\left(\begin{pmatrix}2\\3\end{pmatrix},\begin{pmatrix}-\sqrt{2}\\1\end{pmatrix}\right)$ est une base de $\mathbb{R}^2$.

Exercice de manipulation 2.11. —
Pour chacune des familles suivantes de $\mathbb{R}^3$, dire si elle forme une base (si possible, essayer de faire les vérifications de tête).

- $((1,2,3),(4,5,6),(7,8,9))$,
- $((1,2,1),(3,5,3),(-2,9,-2))$,
- $((1,2,1),(3,5,0),(1,0,0))$,
- $((1,1,-2),(-3,0,3),(-8,3,5))$
- $((1,1,-2),(4,4,4),(3,3,-6))$
- $\big((0,3,0),(-2,0,\sqrt{3}),(5,0,1)\big)$.

Exercice de manipulation 2.12. —
Pour chacune des familles suivantes de $\mathbb{R}_2[X]$, dire si elle forme une base de l'espace considéré (si possible, essayer de faire les vérifications de tête).

- $(8, 3X-5, 11X^2+4X+1)$,
- $(8, X^2+3X-5, X^2+3X+3)$,
- $(8, 3X+5, -11X+7)$,
- $(1, X, X^2)$,
- $(X^2-3,\ X^2+X,\ X^2)$,
- $(X^2-3,\ X^2+X,\ X^2+X+7)$.

Exercice de manipulation 2.13. —
Démontrer avec le moins possible de calculs que dans $\mathbb{R}^3$, on a $\mathbf{Vect}\left[\begin{pmatrix}1\\-1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\-1\end{pmatrix}\right] = \mathbf{Vect}\left[\begin{pmatrix}1\\-2\\1\end{pmatrix}, \begin{pmatrix}-3\\1\\2\end{pmatrix}\right]$.

Exercice de manipulation 2.14. —

Soit E un espace vectoriel de dimension $n \geq 1$, $(x_1, ..., x_p)$ une famille de vecteurs de E et y un élément de E.
On suppose que $H = \mathbf{Vect}[x_1, ..., x_p]$ est de dimension $n-1$.
Montrer qu'on a soit $\mathbf{Vect}[x_1, ..., x_p, y] = H$, soit $\mathbf{Vect}[x_1, ..., x_p, y] = E$.

Exercice de manipulation 2.15. —
Reprendre les familles des exercices 6, 11 et 12 et donner leur rang dans chaque cas.
Lorsque c'est possible, essayer de se passer de stylo.

Exercice de manipulation 2.16. —
On fixe deux réels a et b et on note A la matrice $\begin{pmatrix}1 & 1 & -1 & 2\\ 3 & 0 & 1 & -4\\ a & 2 & -2 & b\end{pmatrix}$.

(a) Dire sans aucun calcul pourquoi le rang de A est soit 2, soit 3.

(b) Déterminer une condition nécessaire et suffisante sur a et b pour que le rang de A soit égal à 2.

Exercice de manipulation 2.17. —
Déterminer le rang de la matrice $\begin{pmatrix}1 & -1 & 2 & -1\\ 0 & 0 & 1 & -1\\ 1 & -1 & 1 & 0\\ 1 & -1 & 1 & 0\end{pmatrix}$.

Exercice de manipulation 2.18. —
Dans $\mathbb{R}^4$, on considère les vecteurs $u = \begin{pmatrix}1\\2\\1\\-1\end{pmatrix}$, $v = \begin{pmatrix}2\\-1\\-3\\2\end{pmatrix}$, $w = \begin{pmatrix}5\\5\\0\\-1\end{pmatrix}$.

Déterminer le rang de la famille (u, v, w).

Exercice de manipulation 2.19. —
On reprend la famille (u, v, w) de l'exercice II.18.

Trouver un système linéaire homogène dont l'ensemble des solutions est $\mathbf{Vect}[u, v, w]$.

Exercice de manipulation 2.20. —
On se place dans $\mathbb{R}^4$. On note $u_1 = \begin{pmatrix} 1 \\ 2 \\ 1 \\ 2 \end{pmatrix}$, $u_2 = \begin{pmatrix} -2 \\ 0 \\ 3 \\ 4 \end{pmatrix}$, $u_3 = \begin{pmatrix} 4 \\ 3 \\ -6 \\ -7 \end{pmatrix}$ et $F = \mathbf{Vect}[u_1, u_2, u_3]$.

(a) Décrire F comme ensemble des solutions d'un système linéaire homogène.

(b) Soit G le sous-espace $\left\{ (x, y, z, t) \in \mathbb{R}^4, \begin{array}{r} x - 2y + z = 0 \\ 2x + z = 0 \end{array} \right\}$. Déduire de (a) une base de $F \cap G$ et sa dimension.

CHAPITRE 3

SOMMES ET SUPPLÉMENTAIRES

1. Somme de deux sous-espaces vectoriels

Définition 3.1 – Sous-espace somme de deux sous-espaces vectoriels

Soit E un espace vectoriel ; soient F_1, F_2 deux sous-espaces vectoriels de E.
Le *sous-espace somme de* F_1 *et* F_2, noté $F_1 + F_2$, est l'ensemble des vecteurs qui peuvent s'écrire comme somme d'un vecteur de F_1 et d'un vecteur de F_2 :

$$F_1 + F_2 = \left\{ x_1 + x_2 \;\middle|\; x_1 \in F_1,\ x_2 \in F_2 \right\}$$

Il s'agit toujours d'un sous-espace vectoriel. — En effet, reprenant les notations précédentes, si $x = x_1 + x_2$ et $y = y_1 + y_2$ sont deux éléments de $F_1 + F_2$ (avec, donc, x_1 et y_1 dans F_1 et x_2, y_2 dans F_2), et si λ et μ sont deux scalaires, on a $\lambda x + \mu y = (\lambda x_1 + \mu y_1) + (\lambda x_2 + \mu y_2)$; puisque F_1 et F_2 sont des sous-espaces vectoriels, $\lambda x_1 + \mu y_1$ appartient à F_1 et $\lambda x_2 + \mu y_2$ à F_2. Ainsi $\lambda x + \mu y$ appartient à $F_1 + F_2$, ce qu'il fallait démontrer.

Proposition 3.2 – Reformulation abstraite de la définition

On a l'égalité $F_1 + F_2 = \mathbf{Vect}[F_1 \cup F_2]$.
Ainsi, $F_1 + F_2$ *est le plus petit sous-espace vectoriel de* E *qui contienne à la fois* F_1 *et* F_2.

Démonstration. — Le sous-espace $F_1 + F_2$ est un sous-espace vectoriel de E qui contient $F_1 \cup F_2$; d'après la propriété fondamentale du sous-espace engendré, on a $\mathbf{Vect}[F_1 \cup F_2] \subset F_1 + F_2$. Réciproquement, tout élément de $F_1 + F_2$ est la somme d'un élément de F_1 et d'un élément de F_2, donc combinaison linéaire de deux éléments de $F_1 \cup F_2$: ainsi $F_1 + F_2 \subset \mathbf{Vect}[F_1 \cup F_2]$. □

***Exemple 3.3* (Avec deux droites).** — Dans $\mathbb{R}^3$, les droites vectorielles $\mathcal{D}_1 = \mathbf{Vect}\left[\begin{pmatrix}1\\0\\1\end{pmatrix}\right]$ et $\mathcal{D}_2 = \mathbf{Vect}\left[\begin{pmatrix}0\\1\\0\end{pmatrix}\right]$ ont pour somme le plan $\mathcal{P} = \left\{ \begin{pmatrix}x\\y\\z\end{pmatrix} \in \mathbb{R}^3, x = z \right\}$.

En effet, $\mathcal{P} = \left\{ \begin{pmatrix} x \\ y \\ x \end{pmatrix},\ x, y \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} x \\ 0 \\ x \end{pmatrix} + \begin{pmatrix} 0 \\ y \\ 0 \end{pmatrix},\ x, y \in \mathbb{R} \right\} = \{u + v,\ u \in \mathcal{D}_1, v \in \mathcal{D}_2\} = \mathcal{D}_1 + \mathcal{D}_2$.

***Exemple 3.4* (Avec deux plans).** —
Toujours dans $\mathbb{R}^3$, considérons les plans $\mathcal{H} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3,\ z = 0 \right\}$ et $\mathcal{V} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3,\ y = 0 \right\}$. On a $\mathcal{H} + \mathcal{V} = \mathbb{R}^3$. En effet, on a bien sûr $\mathcal{H} + \mathcal{V} \subset \mathbb{R}^3$, et pour tout vecteur $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ de $\mathbb{R}^3$, on peut écrire $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \underbrace{\begin{pmatrix} x/2 \\ y \\ 0 \end{pmatrix}}_{\in \mathcal{H}} + \underbrace{\begin{pmatrix} x/2 \\ 0 \\ z \end{pmatrix}}_{\in \mathcal{V}}$, ainsi $\mathbb{R}^3 \subset \mathcal{H} + \mathcal{V}$.

***Exemple 3.5* (Polynômes pairs et impairs).** — Dans $\mathbb{R}[X]$, considérons les sous-espaces

$$\mathcal{P} = \{P \in \mathbb{R}_6[X], P(-X) = P(X)\} = \left\{a + cX^2 + eX^4 + gX^6, \quad (a, c, e, g) \in \mathbb{R}^4\right\}$$
$$\mathcal{I} = \{P \in \mathbb{R}_6[X], P(-X) = -P(X)\} = \left\{bX + dX^3 + fX^5, \quad (b, d, f) \in \mathbb{R}^3\right\}$$

des polynômes *pairs* et *impairs* de degré inférieur ou égal à 6. On a $\mathcal{P} + \mathcal{I} = \mathbb{R}_6[X]$.

***Exemple 3.6* (Un cas en dimension infinie).** —
Toujours dans $\mathbb{R}[X]$, considérons $\mathcal{V}_0 = \{P \in \mathbb{R}[X], P(0) = 0\}$ et $\mathcal{V}_{-5} = \{P \in \mathbb{R}[X], P(-5) = 0\}$. Ce sont des sous-espaces vectoriels, et on a $\mathbb{R}[X] = \mathcal{V}_0 + \mathcal{V}_{-5}$. En effet, pour tout polynôme P de $\mathbb{R}[X]$, on peut écrire $P(X) = \underbrace{\left(P(X) - \frac{P(0)}{5}(X+5)\right)}_{\in \mathcal{V}_0} + \underbrace{\left(\frac{P(0)}{5}(X+5)\right)}_{\in \mathcal{V}_{-5}}$.

Proposition 3.7 – Sous-espace engendré par une réunion de parties

Si A et B sont deux parties de E (pas nécessairement des sous-espaces vectoriels), on a

$$\mathbf{Vect}[A \cup B] = \mathbf{Vect}[A] + \mathbf{Vect}[B].$$

Démonstration. — Le sous-espace $\mathbf{Vect}[A \cup B]$ contient $\mathbf{Vect}[A]$ et $\mathbf{Vect}[B]$, donc (par la reformulation abstraite de la définition) il contient $\mathbf{Vect}[A] + \mathbf{Vect}[B]$; par ailleurs, tout élément de $\mathbf{Vect}[A] + \mathbf{Vect}[B]$ est somme d'une combinaison linéaire d'éléments de A et d'une combinaison linéaire d'éléments de B, c'est donc une combinaison linéaire d'éléments de $A \cup B$, ainsi tout élément de $\mathbf{Vect}[A] + \mathbf{Vect}[B]$ appartient à $\mathbf{Vect}[A \cup B]$. □

Exemple 3.8. — Dans $\mathbb{R}^3$, on a l'égalité $\mathbf{Vect}\left[\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}\right] + \mathbf{Vect}\left[\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right] = \mathbf{Vect}\left[\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right]$.

Comme $3\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$, on a donc en fait $\mathbf{Vect}\left[\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}\right] + \mathbf{Vect}\left[\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right] = \mathbf{Vect}\left[\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}\right]$.

Exercices de manipulation. — III.1, III.2, III.3, III.4, III.5.

2. Formule de Grassmann pour la dimension du sous-espace somme

Théorème 3.9 – Formule de Grassmann

Soit E un espace vectoriel de dimension finie. Si F_1 et F_2 sont deux sous-espaces vectoriels de E, alors

$$\dim(F_1 + F_2) = \dim(F_1) + \dim(F_2) - \dim(F_1 \cap F_2).$$

Démonstration. —

- Soit $\mathcal{B} = (b_1, ..., b_n)$ une base de $F_1 \cap F_2$. On peut la voir comme une famille de vecteurs de F_1, et c'est une famille libre : d'après le théorème de la base incomplète, il existe donc une base $\tilde{\mathcal{B}_1} = (b_1, ..., b_n, x_1, ..., x_p)$ de F_1 qui contient $\mathcal{B}$. De même, il existe une base $\tilde{\mathcal{B}_2} = (b_1, ..., b_n, y_1, ..., y_q)$ de F_2 qui contient $\mathcal{B}$.
- La famille $\tilde{\mathcal{B}_1} \cup \tilde{\mathcal{B}_2} = (b_1, ..., b_n, x_1, ..., x_q, y_1, ..., x_q)$ est génératrice de $F_1 + F_2$, car $\mathbf{Vect}[\mathcal{B}_1 \cup \mathcal{B}_2] = \mathbf{Vect}[\mathcal{B}_1] + \mathbf{Vect}[\mathcal{B}_2] = F_1 + F_2$.
- J'affirme qu'elle est de plus libre. En effet, soient $\lambda_1, ..., \lambda_n, \alpha_1, .., \alpha_p, \beta_1, ..., \beta_q$ des scalaires ; supposons

  $$\sum_{i=1}^{n} \lambda_i b_i + \sum_{i=1}^{p} \alpha_i x_i + \sum_{i=1}^{q} \beta_i y_i = 0 \qquad ; \qquad \text{on a alors } \sum_{i=1}^{q} \beta_i y_i = -\left(\sum_{i=1}^{n} \lambda_i b_i + \sum_{i=1}^{p} \alpha_i x_i\right).$$

  Dans cette égalité, le membre de droite est une combinaison linéaire de vecteurs de F_1 ; ainsi $\sum_{i=1}^{q} \beta_i y_i$ appartient à F_1, et comme les y_i sont dans F_2, $\sum_{i=1}^{q} \beta_i y_i$ appartient à $F_1 \cap F_2$. Comme $\mathcal{B}$ est une base de $F_1 \cap F_2$, on en déduit l'existence de scalaires $\mu_1, ..., \mu_n$ vérifiant

  $$\sum_{i=1}^{q} \beta_i y_i = \sum_{i=1}^{n} \mu_i b_i$$

  Mais alors la combinaison linéaire $(\beta_1 y_1 + \beta_2 y_2 + ... + \beta_n y_n) + (-\mu_1 b_1 - ... - \mu_n b_n)$ de la famille $\tilde{\mathcal{B}_1}$ est nulle, et $\tilde{\mathcal{B}_1}$ est libre, donc tous les coefficients β_i, $i = 1..q$ sont nuls (ainsi que tous les μ_i, $i = 1..n$). L'égalité de départ devient maintenant

  $$\sum_{i=1}^{n} \lambda_i b_i + \sum_{i=1}^{p} \alpha_i x_i = 0$$

  mais elle ne fait intervenir que les vecteurs de $\tilde{\mathcal{B}_2}$, qui est libre, donc tous les λ_i, $i = 1..n$ et tous les α_i, $i = 1..p$ sont nuls, ce qu'il fallait démontrer.
- Ainsi $\tilde{\mathcal{B}_1} \cup \tilde{\mathcal{B}_2}$ est une base de $F_1 + F_2$, tandis que $\tilde{\mathcal{B}_1}$ est une base de F_1, $\tilde{\mathcal{B}_2}$ est une base de F_2 et $\mathcal{B} = \tilde{\mathcal{B}_1} \cap \tilde{\mathcal{B}_2}$ est une base de $F_1 \cap F_2$. De la formule combinatoire

  $$\text{Card}\left(\tilde{\mathcal{B}_1} \cup \tilde{\mathcal{B}_2}\right) = \text{Card}(\tilde{\mathcal{B}_1}) + \text{Card}(\tilde{\mathcal{B}_2}) - \text{Card}\left(\tilde{\mathcal{B}_1} \cup \tilde{\mathcal{B}_2}\right)$$

  et de la définition de la dimension, on déduit alors la formule de Grassmann.

□

Exercices de manipulation. — 3.6, 3.7.

3. Somme directe (cas de deux sous-espaces)

3.1. Définition et premiers exemples. —

Définition 3.10 – Situation de somme directe pour deux sous-espaces

Soit E un espace vectoriel, soient F_1 et F_2 des sous-espaces vectoriels de E.
On dit que F_1 et F_2 sont *en somme directe* lorsque la propriété suivante est vérifiée.

$$\begin{aligned} &\textit{si} \quad x_1 + x_2 = \mathbf{0}_E \quad \textit{où} \quad x_1 \in F_1 \textit{ et } x_2 \in F_2, \\ &\textit{alors} \quad x_1 = \mathbf{0}_E \text{ et } x_2 = \mathbf{0}_E. \end{aligned}$$

Examinons cette notion sur les exemples 3.3 à 3.6.

Exemple 3.11. — Dans $\mathbb{R}^3$, les droites vectorielles $\mathcal{D}_1 = \mathbf{Vect}\left[\begin{pmatrix}1\\0\\1\end{pmatrix}\right]$ et $\mathcal{D}_2 = \mathbf{Vect}\left[\begin{pmatrix}0\\1\\0\end{pmatrix}\right]$ sont en somme directe.

En effet, si $x_1 = \begin{pmatrix}\alpha\\0\\\alpha\end{pmatrix}$ est un vecteur de $\mathcal{D}_1$ et $x_2 = \begin{pmatrix}0\\\beta\\0\end{pmatrix}$ est un vecteur de $\mathcal{D}_2$, si $\begin{pmatrix}\alpha\\0\\\alpha\end{pmatrix} + \begin{pmatrix}0\\\beta\\0\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$, alors bien sûr $\alpha = \beta = 0$ et $x_1 = x_2 = 0$.

Exemple 3.12. — Dans $\mathbb{R}^3$, les plans $\mathcal{H} = \left\{\begin{pmatrix}x\\y\\z\end{pmatrix} \in \mathbb{R}^3,\ z = 0\right\}$ et $\mathcal{V} = \left\{\begin{pmatrix}x\\y\\z\end{pmatrix} \in \mathbb{R}^3,\ y = 0\right\}$ ne sont pas en somme directe. En effet, $\begin{pmatrix}1\\0\\0\end{pmatrix}$ appartient à $\mathcal{H}$, $\begin{pmatrix}-1\\0\\0\end{pmatrix}$ appartient à $\mathcal{V}$, et on a $\begin{pmatrix}1\\0\\0\end{pmatrix} + \begin{pmatrix}-1\\0\\0\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}$ alors que ces vecteurs ne sont pas nuls.

Exemple 3.13. — Dans $\mathbb{R}[X]$, les sous-espaces

$$\mathcal{P} = \{P \in \mathbb{R}_6[X], P(-X) = P(X)\} = \left\{a + cX^2 + eX^4 + gX^6, (a,c,e,g) \in \mathbb{R}^4\right\}$$

et

$$\mathcal{I} = \{P \in \mathbb{R}_6[X], P(-X) = -P(X)\} = \left\{bX + dX^3 + fX^5, (b,d,f) \in \mathbb{R}^3\right\}$$

sont en somme directe. En effet, si $U(X) = a + cX^2 + eX^4 + gX^6$ est un vecteur de $\mathcal{P}$ et $V(X) = bX + dX^3 + fX^5$ est un vecteur de $\mathcal{I}$ et si leur somme est nulle, alors tous les coefficients a à g sont nuls, donc les polynômes U et V sont nuls.

Exemple 3.14. — Dans $\mathbb{R}[X]$, les sous-espaces $\mathcal{V}_0 = \{P \in \mathbb{R}[X], P(0) = 0\}$ et $\mathcal{V}_{-5} = \{P \in \mathbb{R}[X], P(-5) = 0\}$ ne sont pas en somme directe. En effet, on a $\underbrace{X(X+5)}_{\in \mathcal{V}_0} + \underbrace{(-X)(X+5)}_{\in \mathcal{V}_{-5}} = 0$ alors que les deux termes de cette somme sont non-nuls.

***Remarque 3.15* (Notation $\oplus$).** — L'usage, lorsque F_1 et F_2 sont en somme directe, est de noter $F_1 \oplus F_2$ la somme qui d'habitude se note $F_1 + F_2$.

Ainsi, écrire « $F_1 + F_2 = F_1 \oplus F_2$ », c'est affirmer que F_1 et F_2 sont en somme directe.

Attention à utiliser cette notation à bon escient : dans ce cours,

- on ne peut parler de $F_1 \oplus F_2$ que si on a déjà vérifié que F_1 et F_2 sont en somme directe [1] ;
- si on veut « montrer que $A = B \oplus C$ », il faut vérifier que $A = B + C$ *et* que B et C sont en somme directe.

3.2. Caractérisations de la situation de somme directe. —

La notion de somme directe est extrêmement utile, car elle admet l'interprétation suivante.

Proposition 3.16 – Somme directe et unicité des décompositions des vecteurs de $F + G$

Soit E un espace vectoriel, soient F_1 et F_2 des sous-espaces vectoriels de E. Il y a équivalence entre :

(i) F_1 et F_2 sont en somme directe.

(ii) Si $x_1 + x_2 = x'_1 + x'_2$ avec $(x_1, x'_1) \in F_1^2$ et $(x_2, x'_2) \in F_2^2$, alors $x_1 = x'_1$ et $x_2 = x'_2$.

(iii) Pour tout vecteur z de $F_1 + F_2$, il existe un unique *x_1 de F_1 et un* unique *x_2 de F_2 tels que $z = x_1 + x_2$.*

Démonstration. —

- Supposons (i) et montrons (ii) : si $x_1 + x_2 = x'_1 + x'_2$ avec $(x_1, x'_1) \in F_1^2$ et $(x_2, x'_2) \in F_2^2$, alors $(x_1 - x'_1) + (x_2 - x'_2) = 0$; comme $(x_1 - x'_1)$ appartient à F_1 et $(x_2 - x'_2)$ à F_2, la définition donne $x_1 - x'_1 = 0$ et $x_2 - x'_2 = 0$, comme voulu.
- Supposons (ii) et montrons (iii) : pour tout vecteur z de $F + G$, il existe $x_1 \in F_1$ et $x_2 \in F_2$ vérifiant $z = x_1 + x_2$ (c'est la définition de $F + G$) ; par ailleurs, (ii) dit qu'il s'agit de la seule écriture possible de z comme somme d'un élément de F_1 et d'un élément de F_2.
- Supposons (iii) et montrons (i). Si $x_1 + x_2 = 0$ avec $x_1 \in F_1$ et $x_2 \in F_2$, alors nous avons une écriture du vecteur nul de $F_1 + F_2$ comme somme d'un vecteur de F_1 et d'un autre de F_2 ; écrire $0 = \underset{\in F_1}{0} + \underset{\in F_2}{0}$ donne une autre écriture. Par unicité, on obtient $x_1 = 0$ et $x_2 = 0$.

$\square$

Exercices de manipulation. — III.8, III.9.

Le résultat qui vient simplifie considérablement l'utilisation de la notion de somme directe.

Proposition 3.17 – Somme directe $\iff$ intersection réduite à $\{0\}$

Soit E un espace vectoriel, soient F_1 et F_2 des sous-espaces vectoriels de E. Il y a équivalence entre :

(i) F_1 et F_2 sont en somme directe.

(ii) $F_1 \cap F_2 = \{\mathbf{0}_E\}$.

Démonstration. —

- Supposons que F_1 et F_2 soient en somme directe. Soit x un élément de $F_1 \cap F_2$. On a alors $\underset{\in F_1}{x} + \underset{\in F_2}{0} = \underset{\in F_1}{0} + \underset{\in F_2}{x}$; par la propriété (ii) de la proposition 3.16, on en déduit $x = \mathbf{0}_E$.

1. Dans d'autres contextes mathématiques, cette notation peut avoir un autre sens.

- Supposons $F_1 \cap F_2 = \{\mathbf{0}_E\}$. Si x_1 appartient à F_1, x_2 à F_2 et si $x_1 + x_2 = 0$, alors $x_1 = -x_2$. Dans ce cas x_1 est l'opposé d'un élément de F_2 qui est un sous-espace vectoriel, donc x_1 appartient à F_2, et bien sûr x_1 appartient aussi à F_1, donc d'après notre hypothèse sur $F_1 \cap F_2$ on doit avoir $x_1 = \mathbf{0}_E$. De même x_2 doit être nul, et nous avons montré que F_1 et F_2 sont en somme directe.

□

***Exemple 3.18* (Somme de deux droites vectorielles).** — Si $\mathcal{D}$ et $\mathcal{D}'$ sont deux droites vectorielles *distinctes* de $\mathbb{R}^3$, alors $\mathcal{D}$ et $\mathcal{D}'$ sont en somme directe. En effet, $\mathcal{D} \cap \mathcal{D}'$ est contenu dans $\mathcal{D}$ et il est aussi contenu dans $\mathcal{D}'$; il est donc de dimension 0 ou 1. S'il est de dimension 1, il doit être égal à $\mathcal{D}$ (voir b. page **??**) et il doit être aussi égal à $\mathcal{D}'$ pour la même raison ; donc si $\mathcal{D} \neq \mathcal{D}'$ on doit avoir $\mathcal{D} \cap \mathcal{D}' = \{0\}$.

***Exemple 3.19* (Somme d'un plan et d'une droite de $\mathbb{R}^3$).** — Si $\mathcal{P}$ est un plan de $\mathbb{R}^3$, si $\mathcal{D}$ est une droite vectorielle de $\mathbb{R}^3$ et si $\mathcal{D}$ n'est pas contenue dans $\mathcal{P}$, alors $\mathcal{D}$ et $\mathcal{P}$ sont en somme directe. En effet, $\mathcal{D} \cap \mathcal{P}$ est contenu dans $\mathcal{D}$ qui est de dimension 1, donc c'est soit $\{0\}$ soit $\mathcal{D}$. Dire que $\mathcal{D} \cap \mathcal{P} = \mathcal{D}$, c'est dire que $\mathcal{D}$ est contenue dans $\mathcal{P}$. Ainsi, si $\mathcal{D}$ n'est pas contenue dans $\mathcal{P}$ alors $\mathcal{D} \cap \mathcal{P} = \{0\}$.

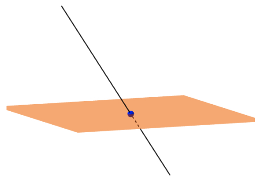


FIGURE 1. Un plan et une droite en somme directe dans $\mathbb{R}^3$.

Les deux exemples ci-dessus peuvent se généraliser de la manière suivante.

Proposition 3.20 – Somme directe lorsque l'un des sous-espaces est une droite

si F est un sous-espace vectoriel d'un espace vectoriel E et si x est un vecteur qui n'appartient pas à F, alors F et la droite $\mathbf{Vect}[x]$ sont en somme directe.

Démonstration. — Le sous-espace $\mathbf{Vect}[x] \cap F$ est contenu dans $\mathbf{Vect}[x]$, donc il est de dimension 0 ou 1. S'il était de dimension 1, ce serait un sous-espace vectoriel contenu dans $\mathbf{Vect}[x]$ et de même dimension que $\mathbf{Vect}[x]$, donc on aurait $\mathbf{Vect}[x] \cap F = \mathbf{Vect}[x]$; mais alors $\mathbf{Vect}[x]$ serait contenu dans F, ce qui n'est visiblement pas le cas puisque $x \notin F$. Ainsi, $\mathbf{Vect}[x] \cap F$ est de dimension zéro, donc c'est $\{0\}$. □

Continuons à illustrer le lien entre somme directe et intersection.

***Exemple 3.21* (Somme de deux plans de $\mathbb{R}^3$).** — Dans l'espace vectoriel réel $\mathbb{R}^3$, deux plans ne sont jamais en somme directe : si P_1 et P_2 sont deux plans de $\mathbb{R}^3$, la dimension de $P_1 + P_2$ est soit 2 soit 3 ; de la formule de Grassmann on déduit que $P_1 \cap P_2$ est de dimension 1 ou 2, donc qu'on n'a jamais $P_1 \cap P_2 = \{\mathbf{0}_{\mathbb{R}^3}\}$.

***Exemple 3.22* (Un exemple abstrait important).** — Plus généralement, soit E un espace vectoriel de dimension finie. Si F_1 et F_2 sont deux sous-espaces vectoriels de E et si $\dim(F_1) + \dim(F_2) > \dim(E)$, alors F_1 et F_2 ne peuvent pas être en somme directe : $F_1 + F_2$ étant un sous-espace vectoriel de E, sa dimension ne dépasse pas $\dim(E)$, et la formule de Grassmann $\dim(F \cap G) = \dim(F) + \dim(G) - \dim(F + G)$ montre qu'on ne peut avoir $\dim(F \cap G) = 0$ dans le cas ici envisagé.

FIGURE 2. *Gauche :* Deux plans de $\mathbb{R}^3$ ne sont jamais en somme directe. *Droite :* deux droites en somme directe déterminent un plan.

***Exemple 3.23* (Retour sur la page 71).** — Reprenons les exemples 3.13 et 3.14 pour voir si le lien entre somme directe et intersection peut simplifier les discussions.

- Dans $\mathbb{R}[X]$, pour vérifier que les sous-espaces

$$\mathcal{P} = \{P \in \mathbb{R}_6[X], P(-X) = P(X)\} \quad \text{et} \quad \mathcal{I} = \{P \in \mathbb{R}_6[X], P(-X) = -P(X)\}$$

  sont en somme directe, il suffit d'étudier l'intersection.

  Or, si P appartient à $\mathcal{P} \cap \mathcal{I}$, alors on a $P(X) = P(-X) = -P(X)$, si bien que $2P(X) = 0$ et P est le polynôme nul. Ainsi $\mathcal{P} \cap \mathcal{I} = \{\mathbf{0}_{\mathbb{R}_6[X]}\}$.
- Dans $\mathbb{R}[X]$, les sous-espaces $\mathcal{V}_0 = \{P \in \mathbb{R}[X], P(0) = 0\}$ et $\mathcal{V}_{-5} = \{P \in \mathbb{R}[X], P(-5) = 0\}$ ne sont pas en somme directe. En effet, $\mathcal{V}_0 \cap \mathcal{V}_{-5}$ contient le polynôme $X(X+5)$, donc il n'est pas réduit à $\{\mathbf{0}_{\mathbb{R}[X]}\}$.

Exercices de manipulation : — III.10, III.11, III.12, III.13.

3.3. Cas de la dimension finie. —

Proposition 3.24 – Sommes directes : dimensions et bases

Soit E un espace vectoriel de dimension finie , soient F_1 et F_2 des sous-espaces vectoriels de E. Il y a équivalence entre :

(i) *F_1 et F_2 sont en somme directe.*

(ii) $\dim(F_1 + F_2) = \dim(F_1) + \dim(F_2)$.

(iii) *Si $\mathcal{B}_1$ est une base de F_1 et $\mathcal{B}_2$ est une base de F_2, alors $\mathcal{B}_1 \cap \mathcal{B}_2 = \emptyset$ et $\mathcal{B}_1 \cup \mathcal{B}_2$ est une base de $F_1 + F_2$.*

Démonstration. — L'équivalence entre (i) et (ii) résulte immédiatement de la formule de Grassmann et du lien entre somme directe et intersection que nous venons de signaler. Le fait que (iii) implique (ii) est immédiat.

Montrons que (i) implique (iii). Si $\mathcal{B}_1$ est une base de F_1, $\mathcal{B}_2$ une base de F_2 et s'il y a un vecteur dans $\mathcal{B}_1 \cap \mathcal{B}_2$, il doit appartenir à $F_1 \cap F_2$ et il ne peut être nul puisqu'une famille contenant le vecteur nul est liée : c'est impossible, et $\mathcal{B}_1 \cap \mathcal{B}_2$ est vide. Par ailleurs, dans la démonstration de la formule de Grassmann, nous avons vu que si $\mathcal{B}$ est une base de $F_1 \cap F_2$, $\tilde{\mathcal{B}}_1$ une base de F_1 qui contient $\mathcal{B}$ et $\tilde{\mathcal{B}}_2$ une base de F_2 qui contient $\mathcal{B}$, alors $\tilde{\mathcal{B}}_1 \cup \tilde{\mathcal{B}}_2$ est une base de $F_1 + F_2$: en appliquant ceci à $\mathcal{B} = \emptyset$, $\tilde{\mathcal{B}}_1 = \mathcal{B}_1$ et $\tilde{\mathcal{B}}_2 = \mathcal{B}_2$, on obtient (iii). □

Exercices de manipulation : — III.14.

Décomposition d'un espace vectoriel en somme directe de deux sous-espaces vectoriels. — Étant donné un espace vectoriel E, il n'est pas rare que l'existence d'une décomposition $E = F_1 \oplus F_2$ en somme directe de deux sous-espaces « intéressants » soit un renseignement précieux. Nous étudierons en détail cette situation dans le prochain paragraphe.

Je résume ici ce que disent de ce cas quelques-uns des résultats ci-dessus.

Proposition 3.25 – Cas d'une somme directe donnant E tout entier

Il y a équivalence entre

(a) $E = F_1 \oplus F_2$,

(b) $F_1 \cap F_2 = \{0\}$ *et* $\dim(E) = \dim(F_1) + \dim(F_2)$

(c) il existe une base de E dont les k premiers vecteurs sont dans F_1 et les $n - k$ derniers dans F_2.

4. Supplémentaires d'un sous-espace vectoriel

4.1. Définition et remarques. —

Définition 3.26 – Supplémentaire d'un sous-espace vectoriel

Soient E un espace vectoriel et F un sous-espace vectoriel de E.
Un *supplémentaire de F dans E* est un sous-espace vectoriel S de E qui vérifie : $E = F \oplus S$.

Interprétation. — dire que S est un supplémentaire de F dans E, c'est dire que tout vecteur de E peut s'écrire *de façon unique* comme somme d'un vecteur de F et d'un vecteur de S.

Attention. — Ne pas confondre *les éventuels supplémentaires de F dans E* (une notion d'algèbre linéaire, qui donne toujours des sous-espace vectoriels) et *le complémentaire de F dans E* (une notion de théorie des ensembles, qui ne donne jamais un sous-espace vectoriel).

Remarque 3.27. — D'après la partie (b) de la proposition 3.25, tous les supplémentaires éventuels de F dans E ont la même dimension, à savoir $\dim(E) - \dim(F)$.

Vocabulaire : — *deux sous-espaces F_1 et F_2 de E* sont supplémentaires *quand $E = F_1 \oplus F_2$. C'est la situation rencontrée dans la proposition 3.25.*

4.2. Exemples classiques. —

Exemple 3.28 **(Droite et plan de $\mathbb{R}^3$).** — **Droites et plans de $\mathbb{R}^3$**

Soient a, b et c sont trois réels non tous nuls ; considérons le plan $\mathcal{P} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3,\ ax + by + cz = 0 \right\}$ de $\mathbb{R}^3$. Le vecteur $\vec{u} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ n'appartient pas à $\mathcal{P}$. Montrons que la droite $\mathcal{D} = \mathbb{R}\vec{u} = \mathbf{Vect}[\vec{u}]$ est un supplémentaire de $\mathcal{P}$ dans $\mathbb{R}^3$.

La droite $\mathcal{D}$ n'est pas contenue dans le plan $\mathcal{P}$, donc $\mathcal{D}$ et $\mathcal{P}$ sont en somme directe (voir l'exemple 3.18).

De plus, on a l'égalité $\dim(\mathcal{P})+\dim(\mathcal{D}) = 2+1 = \dim(\mathbb{R}^3)$. D'après le critère (b) de la proposition 3.25, on a $\mathbb{R}^3 = \mathcal{P} \oplus \mathcal{D}$, ce qui signifie bien que $\mathcal{D}$ est un supplémentaire de $\mathcal{P}$ dans $\mathbb{R}^3$.

FIGURE 3. *Gauche :* une droite et un plan supplémentaire dans $\mathbb{R}^3$. *Droite :* le plan orange admet une infinité de supplémentaires disincts

Plus généralement, si $\mathcal{P}$ est un plan de $\mathbb{R}^3$, *toute* droite non contenue dans $\mathcal{P}$ est un supplémentaire de $\mathcal{P}$ dans $\mathbb{R}^3$ (le raisonnement est exactement le même qu'à partir de la troisième ligne de l'exemple 3.28).

Si $\mathcal{P}$ est un plan de $\mathbb{R}^3$, il admet donc une infinité de supplémentaires différents !

***Exemple 3.29* (Matrices symétriques et antisymétriques). —**
Plaçons-nous dans $\mathcal{M}_n(\mathbb{R})$, $n \geq 1$, considérons l'espace $\mathcal{S} = \left\{M \in \mathcal{M}_n(\mathbb{R}),\ M^T = M\right\}$ des matrices symétriques et l'espace $\mathcal{A} = \left\{M \in \mathcal{M}_n(\mathbb{R}),\ M^T = -M\right\}$ des matrices antisymétriques.

- Si une matrice M appartient à $\mathcal{S}$ et à $\mathcal{A}$, on a $M = M^T = -M$, donc M doit être la matrice nulle : $\mathcal{S} \cap \mathcal{A} = \{0\}$.
- De plus, nous avons vu page 52 que $\dim(\mathcal{S}) = \frac{n(n+1)}{2}$ et $\dim(\mathcal{A}) = \frac{n(n-1)}{2}$; on a donc $\dim(\mathcal{S}) + \dim(\mathcal{A}) = n^2$.
- Comme $\mathcal{M}_n(\mathbb{R})$ est de dimension n^2, le critère (b) de la proposition 3.25 garantit alors que $\mathcal{S}$ et $\mathcal{A}$ sont supplémentaires.

Pour chaque matrice M, on peut préciser concrètement l'unique décomposition de M comme somme d'un élément de $\mathcal{S}$ et d'un élément de $\mathcal{A}$: on constate que

$$M = \underbrace{\frac{M+M^T}{2}}_{\text{symétrique}} + \underbrace{\frac{M-M^T}{2}}_{\text{antisymétrique}},$$

et on *sait d'avance* que c'est la seule manière d'écrire M sous la forme $M_{\mathcal{S}} + M_{\mathcal{A}}$ avec $M_{\mathcal{S}}$ symétrique et $M_{\mathcal{A}}$ antisymétrique.

***Exemple 3.30* (Fonctions paires et impaires). —**
Plaçons-nous dans $\mathcal{F}(\mathbb{R},\mathbb{R})$, et considérons l'espace $\mathcal{P}$ des fonctions *paires* et l'espace $\mathcal{I}$ des fonctions *impaires*. Si f est une fonction à la fois paire et impaire et si x est un nombre réel, on a $f(x) = f(-x) = -f(x)$, donc $f(x) = 0$, si bien que f est la fonction nulle. Ainsi $\mathcal{P} \cap \mathcal{I} = \{0\}$ et $\mathcal{P}$ et $\mathcal{I}$ sont *en somme directe*.

Montrons que $\mathcal{P}$ et $\mathcal{I}$ sont supplémentaires. Comme $\mathcal{F}(\mathbb{R},\mathbb{R})$ est de dimension infinie, il n'est pas possible d'utiliser un argument de dimension. Montrons que toute fonction $f \in \mathcal{F}(\mathbb{R},\mathbb{R})$ peut s'écrire *de façon unique* comme
$f = f_{\mathcal{P}} + f_{\mathcal{I}}$ avec $f_{\mathcal{P}} \in \mathcal{P}$ et $f_{\mathcal{I}} \in \mathcal{I}$.

Si une telle décomposition est disponible, alors pour tout x de $\mathbb{R}$, on a

$$f(x) = f_{\mathcal{P}}(x) + f_{\mathcal{I}}(x)$$

et comme $f_{\mathcal{P}}(-x) = f_{\mathcal{P}}(x)$ et $f_{\mathcal{I}}(-x) = -f_{\mathcal{I}}(x)$,

$$f(-x) = f_{\mathcal{P}}(x) - f_{\mathcal{I}}(x)$$

ce qui donne *nécessairement* $f_{\mathcal{P}}(x) = \dfrac{f(x)+f(-x)}{2}$ et $f_{\mathcal{I}}(x) = \dfrac{f(x)-f(-x)}{2}$.
Ainsi les *seuls candidats possibles* pour $f_{\mathcal{P}}$ et $f_{\mathcal{I}}$ sont les fonctions $x \mapsto \dfrac{f(x)+f(-x)}{2}$ et $x \mapsto \dfrac{f(x)-f(-x)}{2}$.
Comme on a bien

$$\forall x \in \mathbb{R},\ f(x) = \frac{f(x)+f(-x)}{2} + \frac{f(x)-f(-x)}{2},$$

il existe une *unique* décomposition de f en somme d'une fonction paire et d'une fonction impaire. Cela montre bien que $\mathcal{P}$ et $\mathcal{I}$ sont supplémentaires.

On appelle parfois *partie paire* de f la fonction $x \mapsto \dfrac{f(x)+f(-x)}{2}$ et *partie impaire* de f la fonction $x \mapsto \dfrac{f(x)-f(-x)}{2}$:
par exemple, la partie paire de exp est la fonction cosh et sa partie paire est la fonction sinh.

***Exemple 3.31* (Constantes et fonctions qui valent zéro en zéro).** —
Restons dans $\mathcal{F}(\mathbb{R},\mathbb{R})$ et considérons l'espace $\mathcal{E}_0$ des fonctions *qui valent zéro en zéro* et l'espace $\mathcal{C}$ des fonctions *constantes*. Montrons que $\mathcal{F}(\mathbb{R},\mathbb{R}) = \mathcal{E}_0 \oplus \mathcal{C}$. Il s'agit de savoir si toute fonction de $\mathbb{R}$ dans $\mathbb{R}$ peut s'écrire de façon unique comme somme d'une fonction valant zéro en zéro et d'une fonction constante.

Si f est une fonction de $\mathbb{R}$ dans $\mathbb{R}$ et s'il existe des éléments $g \in \mathcal{E}_0$ et $h \in \mathcal{C}$ vérifiant $f = f_0 + c_f$, on a $f(0) = g(0) + h(0) = 0 + h(0)$. Comme h doit être une fonction constante, c'est nécessairement la fonction constante $x \mapsto f(0)$, et g est nécessairement la fonction $x \mapsto f(x) - f(0)$. Il y a donc une seule possibilité pour g et une seule possibilité pour h : constatant qu'on a bien

$$\forall x \in \mathbb{R},\ f(x) = (f(x) - f(0)) + f(0)$$

on obtient l'existence d'une *unique* décomposition de f en somme d'un élément de $\mathcal{E}_0$ et d'un élément de $\mathcal{C}$.

On a donc bien $\mathcal{F}(\mathbb{R},\mathbb{R}) = \mathcal{E}_0 \oplus \mathcal{C}$, autrement dit, $\mathcal{E}_0$ et $\mathcal{C}$ sont supplémentaires dans $\mathcal{F}(\mathbb{R},\mathbb{R})$.

Exercices de manipulation. — III.15, III.16, III.17

4.3. Existence et non-unicité. —

Proposition 3.32 – Supplémentaires : existence toujours, mais unicité presque jamais

Soit E un espace vectoriel, soit F un sous-espace vectoriel de E.

- *Il existe au moins un supplémentaire de F dans E*
- *Si $\dim(E) \geq 2$ et si F n'est ni $\{0\}$ ni E, alors F admet une infinité de supplémentaires différents.*

Démonstration. — On ne montrera ce résultat qu'en dimension finie (il est vrai aussi en dimension infinie).

- *Existence.* Soit $\mathcal{B}$ une base de F. D'après le théorème de la base incomplète, on peut la compléter en une base $\tilde{\mathcal{B}}$ de E. Posons $S = \mathbf{Vect}\left[\tilde{\mathcal{B}} \backslash \mathcal{B}\right]$. D'après la proposition 3.25, on a $E = F \oplus S$, et S est un supplémentaire de F.

- *Absence d'unicité.* Supposons que $\dim(E) \geq 2$ et F différent de $\{0\}$ et de E. Soit S un supplémentaire de F ; on a donc $S \neq E$ et $S \neq \{0\}$. Fixons un vecteur non nul x dans F (il en existe).
  Pour chaque scalaire λ non nul, notons $S_\lambda = \mathbf{Vect}[\{s + \lambda x, s \in S\}]$. On définit ainsi un sous-espace vectoriel de E.
  Si $(s_1, ..., s_k)$ est une base de S, alors $(\tilde{s}_1, \tilde{s}_2, ..., \tilde{s}_k) = (s_1 + \lambda x, s_2 + \lambda x, ..., s_k + \lambda x)$ est une base de S_λ, donc $\dim(S_\lambda) = \dim(S)$.
  Montrons que $S_\lambda \cap F = \{0\}$: si $y = \alpha_1 \tilde{s}_1 + ... + \alpha_k \tilde{s}_k$ appartient à $S_\lambda \cap F$, alors $z = y - (\alpha_1 + ... + \alpha_k)\lambda x$ appartient aussi à $S_\lambda \cap F$, puisque $x = \frac{1}{\lambda}(0 + \lambda x)$ appartient à $S_\lambda \cap F$ qui est un sous-espace vectoriel. Le vecteur z appartient donc à F, mais $z = \alpha_1 s_1 + ... + \alpha_k s_k$ appartient aussi à S ; comme $S \cap F = \{0\}$ on en déduit que $z = 0$, puis, comme $(s_1, ..., s_k)$ est une base de S, que tous les α_i sont nuls, si bien que $y = 0$. On a donc $S_\lambda \cap F = \{0\}$ et $\dim(S_\lambda) + \dim(F) = \dim(S) + \dim(F) = \dim(E)$: d'après le critère (b) de la proposition 3.25, S_λ est un supplémentaire de F dans E.

$\square$

Remarque 3.33 **(Comment trouver un supplémentaire en pratique)**

La partie « existence » de la démonstration indique que pour obtenir concrètement un supplémentaire de F dans E, on peut par exemple chercher une base de F puis la compléter en une base de E.

Exercice de manipulation. — III. 18.

5. Somme et somme directe (cas de plusieurs sous-espaces)

Nous généralisons ici les résultats vus dans ce chapitre au cas où on somme strictement plus de deux sous-espaces.

L'intérêt des résultats de ce paragraphe n'est pas évident en première lecture ; ils nous seront cependant indispensables au dernier chapitre (à partir de la proposition **??**).

5.1. Définitions. —

Définition 3.34 – Sous-espace somme pour k sous-espace vectoriels

Soit E un espace vectoriel, soient k un élément de $\mathbb{N}^\star$ et F_1, F_2, ..., F_k des sous-espaces vectoriels de E. Le *sous-espace somme* $F_1 + F_2 + ... + F_k$ est l'ensemble

$$\left\{x_1 + x_2 + ... + x_k \;\middle|\; x_1 \in F_1,\ x_2 \in F_2, ..., x_k \in F_k\right\}.$$

On a l'égalité $F_1 + F_2 + ... + F_k = \mathbf{Vect}[F_1 \cup F_2 \cup ... \cup F_k]$.

Définition 3.35 – Situation de somme directe pour k sous-espaces vectoriels

On dit que F_1, F_2, ..., F_k sont *en somme directe* lorsque

$$\text{si } x_1 + x_2 + ... + x_k = 0 \quad \text{avec} \quad x_1 \in F_1, x_2 \in F_2, ..., x_k \in F_k,$$
$$\text{alors } x_1 = x_2 = ... = x_k = 0.$$

Dans ce cas, on note $F_1 \oplus ... \oplus F_k$ à la place de $F_1 + ... + F_k$.

Attention. — Il ne suffit pas, pour que F_1, F_2, ..., F_k soient en somme directe, que les F_i soient deux à deux en somme directe : voir l'exercice 21 de la feuille de TD 3. De même,

L'étude de $F_1 \cap F_2 \cap ... \cap F_k$ ne permet pas de savoir si F_1, F_2, ..., F_k sont ou pas en somme directe.

5.2. Caractérisations de la situation de somme directe. —

Proposition 3.36 – Caractérisations de la situation de somme directe

Soit E un espace vectoriel, soit k un élément de $\mathbb{N}^\star$ et F_1, F_2, ..., F_k des sous-espaces vectoriels de E. Il y a équivalence entre

(i) La somme $F_1 + ... + F_k$ est directe

(ii) ***(Unicité des décompositions)*** *Si* $\underbrace{x_1}_{\in F_1} + ... + \underbrace{x_k}_{\in F_k} = \underbrace{x'_1}_{\in F_1} + ... + \underbrace{x'_k}_{\in F_k}$, *alors* $x_1 = x'_1$, ..., $x_k = x'_k$.

(iii) ***(Pas de perte de dimension)*** $\dim(F_1 + ... + F_k) = \dim(F_1) + ... + \dim(F_k)$

(iv) ***(Concaténation de bases)*** *Si $\mathcal{B}_1$, $\mathcal{B}_2$, ..., $\mathcal{B}_k$ sont des bases de F_1, F_2,..., F_k, alors $\mathcal{B}_i \cap \mathcal{B}_j = \emptyset$ pour chaque $i \neq j$ et $\mathcal{B}_1 \cup ... \cup \mathcal{B}_k$ est une base de $F_1 + ... + F_k$.*

Ces résultats peuvent être démontrés par récurrence sur le nombre k de sous-espaces vectoriels qui intervient dans la somme, grâce au lemme suivant :

Lemme 3.37 – Construire progressivement une somme directe de plusieurs sous-espaces

Soit E un espace vectoriel, soient k un élément de $\mathbb{N}^\star$ et F_1, F_2, ..., F_{k+1} des sous-espaces vectoriels de E. Si

- *F_1, F_2, ..., F_k sont en somme directe*
- *F_{k+1} et $(F_1 \oplus ... \oplus F_k)$ sont en somme directe,*

alors F_1, F_2, ..., F_k, F_{k+1} sont en somme directe.

Démonstration. — Si x_1, ..., x_k, x_{k+1} sont des vecteurs de F_1, ..., F_k, F_{k+1} vérifiant $x_1 + ... + x_k + x_{k+1} = 0$, alors $0 = \underbrace{(x_1 + ... + x_k)}_{\in F_1 \oplus ... \oplus F_k} + \underbrace{x_{k+1}}_{\in F_{k+1}}$.
Comme $F_{k+1} \cap (F_1 \oplus ... \oplus F_k) = \{0\}$, F_{k+1} et $(F_1 \oplus ... \oplus F_k)$ sont en somme directe, donc cela implique $x_{k+1} = 0$ et $x_1 + ... + x_k = 0$; puisque $F_1, ..., F_k$ sont en somme directe, cela implique que tous les x_i sont nuls. $\square$

Exercices de manipulation. — III. 19, III.20.

Exercices de manipulation du cours du chapitre 3

Exercice de manipulation 3.1. —
On se place dans $\mathbb{R}^4$.

1. Soit $F = \left\{ (x,y,z,t) \in \mathbb{R}^4, \begin{array}{r} x+y+z+t=0 \\ t=0 \end{array} \right\}$ et $G = \mathbf{Vect}\left[\begin{pmatrix}1\\1\\0\\0\end{pmatrix}\right]$. Montrer que $F+G = \big\{(x,y,z,t) \in \mathbb{R}^4, t=0\big\}$.
2. Même question lorsque $G = \mathbf{Vect}\left[\begin{pmatrix}1\\2\\3\\0\end{pmatrix}, \begin{pmatrix}3\\2\\1\\0\end{pmatrix}\right]$.

Exercice de manipulation 3.2. —
On se place dans l'espace vectoriel $\mathbb{R}[X]$.
Soit $E = \{P \in \mathbb{R}[X], P(0) = 0 \text{ et } P \text{ est pair }\}$ et $F = \mathbf{Vect}\big[1,\ (X+3),\ (X^2+X),\ (X^3-5X),\ X^4\big]$.
Quel est le sous-espace $E+F$?

Exercice de manipulation 3.3. —
On se place dans $\mathcal{M}_3(\mathbb{R})$. Soit T l'ensemble des matrices triangulaires supérieures et U l'ensemble des matrices de diagonale nulle.
Montrer l'égalité $T+U = \mathcal{M}_3(\mathbb{R})$.

Exercice de manipulation 3.4. —
Soit E un espace vectoriel de dimension finie $n \geq 1$, soient F et G deux sous-espaces vectoriels de E vérifiant :

(i) F est de dimension $n-1$

(ii) $G \not\subset F$.

Montrer l'égalité $F+G = E$.

Exercice de manipulation 3.5. —
On se place dans $\mathbb{R}^4$. Soit $F = \mathbf{Vect}\left[\begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\\-1\\1\end{pmatrix}\right]$ et $G = \mathbf{Vect}\left[\begin{pmatrix}0\\0\\0\\1\end{pmatrix}, \begin{pmatrix}3\\2\\2\\0\end{pmatrix}\right]$. Montrer l'égalité $F+G = \big\{(x,y,z,t) \in \mathbb{R}^4, y=z\big\}$.

Exercice de manipulation 3.6. —
On se place dans $\mathbb{R}^4$. Notons $F = \left\{ (x,y,z,t) \in \mathbb{R}^4, \begin{array}{r} x+y+z+t=0 \\ x+2y+z-2t=0 \end{array} \right\}$ et $G = \mathbf{Vect}\left[\begin{pmatrix}1\\1\\-1\\-1\end{pmatrix}, \begin{pmatrix}0\\2\\-4\\2\end{pmatrix}\right]$.
Déterminer la dimension de $F \cap G$ (utiliser le fait que c'est un sous-espace de G). En déduire la dimension de $F+G$.

Exercice de manipulation 3.7. —
On se place dans $\mathbb{R}[X]$. Soit $E = \mathbf{Vect}\big[1, X, (X^2+3X-7)\big]$ et $F = \mathbf{Vect}\big[4,\ (X^2+X),\ (X^2+3X+8),\ (X^5-2X^3)\big]$.
Décrire $E+F$. Quelle est la dimension de $E \cap F$?

Exercice de manipulation 3.8. —
Soient E un espace vectoriel réel et u, v, s, t quatre vecteurs de E.
Montrer (en utilisant uniquement la définition) que si la famille (u, v, s, t) est libre, alors les sous-espaces $\mathbf{Vect}[u, v]$ et $\mathbf{Vect}[t, s]$ sont en somme directe.

Exercice de manipulation 3.9. —
On se place dans $\mathcal{M}_n(\mathbb{R})$. On note
- E l'ensemble des matrices triangulaires supérieures,
- F l'ensemble des matrices diagonales de trace nulle,
- G l'ensemble des matrices triangulaires supérieures dont tous les coefficients diagonaux sont identiques.

Montrer l'égalité : $E = F \oplus G$.

Exercice de manipulation 3.10. —

1. Considérons $F = \left\{(x,y,z,t) \in \mathbb{R}^4,\ \begin{array}{r} x+y+z+t=0 \\ -2x-y+z-2t=0 \end{array}\right\}$ et $G = \left\{(x,y,z,t) \in \mathbb{R}^4,\ \begin{array}{r} 3x+4y+4z+3t=0 \\ x-y+z-t=0 \end{array}\right\}$.
   Montrer qu'ils sont en somme directe.
2. Considérons $F = \left\{(x,y,z,t) \in \mathbb{R}^4,\ \begin{array}{r} x+y+z+t=0 \\ -2x-y+z-2t=0 \end{array}\right\}$ et $G = \mathbf{Vect}[u,v]$ avec $u = (1,2,3,0)$ et $v = (0,1,0,2)$.
   Montrer qu'ils sont en somme directe.

Exercice de manipulation 3.11. —
Soient $E = \{P \in \mathbb{R}[X], P(1) = 0\}$ et $F = \{P \in \mathbb{R}[X], P(X+1) = P(X)\}$. Montrer qu'ils sont en somme directe.

Exercice de manipulation 3.12. —
Soient F et G deux sous-espaces vectoriels de $\mathbb{R}^9$. On suppose qu'ils sont tous les deux de dimension 5.
Peuvent-ils être en somme directe ?

Exercice de manipulation 3.13. —
Démontrer l'affirmation de l'exemple 1(c) page 76.

Exercice de manipulation 3.14. —
On reprend les sous-espaces de l'exercice 3.10 :

1. Considérons $F = \left\{(x,y,z,t) \in \mathbb{R}^4,\ \begin{array}{r} x+y+z+t=0 \\ -2x-y+z-2t=0 \end{array}\right\}$ et $G = \left\{(x,y,z,t) \in \mathbb{R}^4,\ \begin{array}{r} 3x+4y+4z+3t=0 \\ x-y+z-t=0 \end{array}\right\}$.
   On a montré à l'exercice III.10 qu'ils sont en somme directe. En déduire une base de $F+G$. Quel est ce sous-espace ?

2. Même question avec $F = \left\{ (x,y,z,t) \in \mathbb{R}^4, \begin{array}{r} x+y+z+t=0 \\ -2x-y+z-2t=0 \end{array} \right\}$ et $G = \mathbf{Vect}[u,v]$ avec $u = (1,2,3,0)$ et $v = (0,1,0,2)$.

Exercice de manipulation 3.15. —
On se place dans $\mathbb{R}^3$ et on considère le plan $\mathcal{P} = \{(x,y,z) \in \mathbb{R}^3, x+2y+3z=0\}$.

Montrer de deux façons différentes que pour tout $\alpha \neq -9$, la droite $\mathcal{D}_\alpha = \mathbf{Vect}[(\alpha,2,3)]$ est un supplémentaire de $\mathcal{P}$ dans $\mathbb{R}^3$:

- En reprenant l'argument de dimension utilisé dans le cours ;
- En montrant concrètement que pour α donné, tout vecteur de $\mathbb{R}^3$ peut s'écrire de manière unique comme somme d'un vecteur de $\mathcal{D}_\alpha$ et d'un vecteur de $\mathcal{P}$.

Exercice de manipulation 3.16. —
On se place dans $\mathcal{M}_n(\mathbb{R})$, $n \geq 2$.
On note $\mathcal{A}_n$ l'ensemble des matrices antisymétriques et $\mathcal{T}_n$ l'ensemble des matrices triangulaires supérieures.

Vérifier que $\mathcal{A}_n$ et $\mathcal{T}_n$ sont supplémentaires dans $\mathcal{M}_n(\mathbb{R})$ de deux manières différentes :

- En montrant qu'ils sont en somme directe puis en utilisant un argument de dimension ;
- En montrant que toute matrice $n \times n$ peut s'écrire de manière unique comme somme d'une matrice antisymétrique et d'une matrice triangulaire supérieure.

Exercice de manipulation 3.17. —
On se place dans $\mathbb{K}[X]$.
On note $U(X) = X^3+1$ et E l'ensemble $\{P \in \mathbb{R}[X],\ \exists R \in \mathbb{R}[X], P(X) = U(X)R(X)\}$ des polynômes divisibles par U.

- En utilisant la division euclidienne par Q, montrer que E et $F = \mathbf{Vect}[1, X, X^2]$ sont supplémentaires dans $\mathbb{R}[X]$.
- On note $V = E \cap \mathbb{R}_4[X]$. Montrer que V et F sont supplémentaires dans $\mathbb{R}_4[X]$ en montrant d'abord qu'ils sont en somme directe puis en utilisant un argument de dimension.

Exercice de manipulation 3.18. —

1. On se place dans $\mathbb{R}_3[X]$ et on considère $F = \{P \in \mathbb{R}_3[X], P(1+X) = P(1-X)\}$.
   Montrer que la famille $(1, X^2-2X)$ est une base de F. En déduire un supplémentaire de F dans $\mathbb{R}_3[X]$.
2. On se place dans $\mathbb{R}^4$. Trouver un supplémentaire de $F = \left\{ (x,y,z,t) \in \mathbb{R}^4, \begin{array}{r} x+2y+3z+4t=0 \\ x-2y+3z-4t=0 \end{array} \right\}$.

Exercice de manipulation 3.19. —
On se place dans $\mathbb{R}[X]$. On note

- $E = \{P \in \mathbb{R}(X), P(1) = P(0) = 0\}$,
- F l'ensemble $\{P \in \mathbb{R}_3(X), P(-X) = -P(X)\}$ des polynômes impairs,
- $G = \mathbf{Vect}[(X-3)^2]$.

Montrer que E, F et G sont en somme directe.

Exercice de manipulation 3.20. —
On se place dans $\mathbb{R}^5$. Soit $U = \{(x, y,\ (x+y),\ (x-y),\ 2x), (x,y) \in \mathbb{R}^2\}$, $V = \{(x,y,z,t,s) \in \mathbb{R}^5, x=y=s=0\}$ et $W = \mathbf{Vect}[(0,0,0,0,1)]$. Montrer l'égalité $\mathbb{R}^5 = U \oplus V \oplus W$:

- En montrant que chaque vecteur de $\mathbb{R}^5$ peut s'écrire de manière unique comme somme de vecteurs de U, V, W ;
- En utilisant des bases de chacun de ces espaces.

CHAPITRE 4

APPLICATIONS LINÉAIRES

1. Définition, exemples, propriétés élémentaires

1.1. Définition, remarques pratiques et vocabulaire. —

Définition 4.1 – Application linéaire

Soient E et F deux $\mathbb{K}$-espaces vectoriels et f une application de E dans F.
On dit que f est *linéaire* lorsque

$$\forall (u,v) \in E^2, \quad f(u+v) = f(u) + f(v)$$
$$\forall u \in E, \forall \lambda \in \mathbb{K}, \quad f(\lambda \cdot u) = \lambda \cdot f(u)$$

Fixons deux $\mathbb{K}$-espaces vectoriels E et F.

Proposition 4.2 – L'image du vecteur nul par une application linéaire est le vecteur nul

Si $f : E \to F$ est linéaire, alors $f(\mathbf{0}_E) = \mathbf{0}_F$

En effet, $f(\mathbf{0}_E) = f(2\mathbf{0}_E) = 2f(\mathbf{0}_E)$, donc si on note $y = f(\mathbf{0}_E)$, c'est un vecteur de F et on a $y = 2y$, ainsi $y = \mathbf{0}_F$.

Proposition 4.3 – Comment vérifier en pratique qu'une application est linéaire

Une application $f : E \to F$ est linéaire si et seulement si elle vérifie

$$\forall (u,v) \in E^2, \forall (\lambda,\mu) \in \mathbb{K}^2, \quad f(\lambda u + \mu v) = \lambda f(u) + \mu f(v).$$

Vocabulaire : endomorphisme, forme linéaire. — Soit E un $\mathbb{K}$-espace vectoriel.

- Lorsque f est une application linéaire de E <u>dans $\mathbb{K}$</u>, on dit que c'est une *forme linéaire*. Par exemple, l'application

$$\begin{aligned}\mathrm{ev}_0 : \mathcal{F}(\mathbb{R},\mathbb{R}) &\to \mathbb{R}\\ f &\mapsto f(0)\end{aligned}$$

est linéaire [(1)], c'est une forme linéaire sur $\mathcal{F}(\mathbb{R},\mathbb{R})$.

- Lorsque f est une application *linéaire* de E <u>dans lui-même</u>, on dit que f est un *endomorphisme de* E.

  Par exemple, l'application

$$\begin{aligned}\mathrm{Id}_E : E &\to E\\ x &\mapsto x\end{aligned}$$

  est linéaire [(2)], c'est un endomorphisme de E.

1.2. Exemples d'applications linéaires. — La liste que je vais donner est longue, afin que vous puissiez prendre conscience du fait que beaucoup d'opérations mathématiques « naturelles », issues de contextes variés, sont linéaires.

(a) L'application

$$\begin{aligned}f : \mathbb{R}^2 &\to \mathbb{R}^2\\ \begin{pmatrix}x\\y\end{pmatrix} &\mapsto \begin{pmatrix}2x+3y\\7x-4y\end{pmatrix}\end{aligned}$$

est linéaire. En effet, si $u = \begin{pmatrix}x\\y\end{pmatrix}$ et $v = \begin{pmatrix}x'\\y'\end{pmatrix}$ sont deux éléments de $\mathbb{R}^2$ et si λ est un nombre réel, on a

$$f(u+v) = f\left[\begin{pmatrix}x+x'\\y+y'\end{pmatrix}\right] = \begin{pmatrix}2(x+x')+3(y+y')\\7(x+x')-4(y+y')\end{pmatrix} = \begin{pmatrix}2x+3y\\7x-4y\end{pmatrix} + \begin{pmatrix}2x'+3y'\\7x'-4y'\end{pmatrix} = f(u)+f(v)$$

et

$$f(\lambda\cdot u) = \begin{pmatrix}2(\lambda x)+3(\lambda y)\\7(\lambda x)-4(\lambda y)\end{pmatrix} = \lambda\cdot\begin{pmatrix}2x+3y\\7x-y\end{pmatrix} = \lambda\cdot f(u).$$

On aurait pu aller plus vite en introduisant la matrice $A = \begin{pmatrix}2 & 3\\7 & -4\end{pmatrix}$ et en constatant que pour tout X de $\mathbb{R}^2$, $f(X) = AX$, si bien que pour chaque (u,v) de $(\mathbb{R}^2)^2$ et chaque (λ,μ) de $\mathbb{K}^2$,

$$f(\lambda u+\mu v) = A(\lambda u+\mu v) = \lambda(Au)+\mu(Av) = \lambda f(u)+\mu f(v). \tag{1.1}$$

(b) **Application linéaire canoniquement associée à une matrice.**

La seconde stratégie que nous avons utilisée pour montrer la linéarité de l'application précédente fait apparaître un fait très général et essentiel pour la suite.

1. On a bien sûr $\forall (f,g)\in\mathcal{F}(\mathbb{R},\mathbb{R})^2, \forall(\lambda,\mu)\in\mathbb{R}^2, \mathrm{ev}_0(\lambda f+\mu g) = (\lambda f+\mu g)(0) = \lambda f+\mu g(0) = \lambda\mathrm{ev}_0(f)+\mu\mathrm{ev}_0(g)$.
2. On a bien sûr $\forall(u,v)\in E^2, \forall(\lambda,\mu)\in\mathbb{K}^2, \mathrm{Id}_E(\lambda u+\mu v) = \lambda u+\mu v = \lambda\mathrm{Id}_E(u)+\mu\mathrm{Id}_E(v)$.

Définition 4.4 – Application linéaire canoniquement associée à une matrice

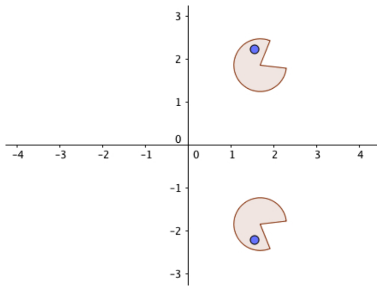
Soit $A \in \mathcal{M}_{np}(\mathbb{K})$ une matrice à n lignes et p colonnes ($n, p \in \mathbb{N}^\star$).
L'application linéaire canoniquement associée à A est l'application linéaire

$$\begin{aligned} T_A : \mathbb{K}^p &\to \mathbb{K}^n \\ X &\mapsto AX. \end{aligned}$$

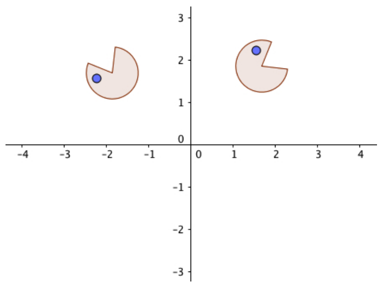
La démonstration est évidemment la même qu'en (1.1).

Beaucoup d'exemples « intéressants » d'applications linéaires sont obtenus de cette manière. Voici quelques illustrations en dimension 2, qui montrent que beaucoup des notions classiques de géométrie plane peuvent être étudiées par l'algèbre linéaire.

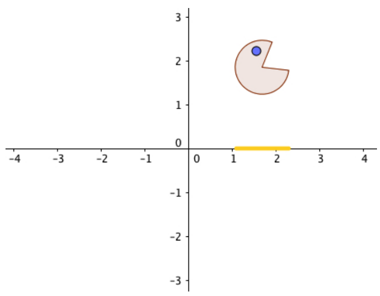
- Si A est la matrice $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$, $T_A : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} -x \\ y \end{pmatrix}$ est la *symétrie par rapport à l'axe des abscisses.*



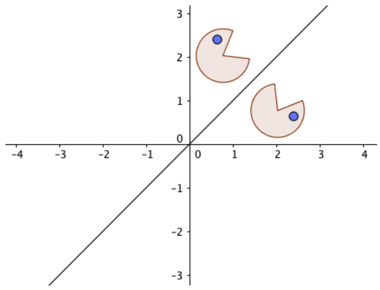
- Si θ est un nombre réel et A est la matrice $\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$, T_A est la *rotation d'angle θ autour de l'origine.*



- Si A est la matrice $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, T_A est la *projection orthogonale sur l'axe des ordonnées*,



- Si A est la matrice $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, T_A est la *transformation qui permute les coordonnées x et y*, autrement dit la *symétrie par rapport à la première bissectrice*



Nous rencontrerons très souvent des transformations de ce type, car toute application linéaire de $\mathbb{R}^p$ dans $\mathbb{R}^n$ est en fait de cette forme :

Proposition 4.5 – Toute appli. linéaire de $\mathbb{K}^p$ dans $\mathbb{K}^n$ est associée à une matrice

Soit f une application linéaire de $\mathbb{K}^p$ dans $\mathbb{K}^n$ $(n, p \in \mathbb{N}^\star)$.
Notons $(e_1, ..., e_p)$ la base canonique de $\mathbb{K}^p$.
Soit A la matrice de $\mathcal{M}_{np}(\mathbb{K})$ dont la 1^{re} colonne est le vecteur $f(e_1)$ de $\mathbb{K}^n$, dont la $2^{è}$ colonne $f(e_2)$, ... et dont la $p^{è}$ colonne est $f(e_p)$.
On a alors l'égalité $f = T_A$.

Démonstration. — Si $X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ est un vecteur de $\mathbb{K}^p$, on peut écrire $X = x_1 e_1 + x_2 e_2 + ... + x_p e_p$.
Notons $C_1 = f(e_1)$, $C_2 = f(e_2)$, etc, si bien que $(C_1, ...C_p)$ est la famille de p vecteurs de $\mathbb{K}^n$ qui donne les colonnes de la matrice A. En utilisant la linéarité de f, on obtient

$$f(X) = f(x_1e_1 + x_2e_2 + ... + x_pe_p) \overset{\text{linéarité}}{=} x_1 f(e_1) + x_2 f(e_2) + ... + x_p f(e_p) = x_1C_1 + x_2C_2 + ... + x_pC_p.$$

Mais nous avons vu page 11 que la combinaison linéaire $x_1C_1 + x_2C_2 + ... + x_pC_p$ n'est autre que le produit AX, si bien que $f(X) = AX$, ce qu'il fallait démontrer. □

Le chapitre suivant reviendra très longuement sur des idées de ce type.

(c) **Homothétie, identité, application nulle.**

Soit E un $\mathbb{K}$-espace vectoriel. Si α est un scalaire de $\mathbb{K}$, l'application

$$E \to E$$
$$x \mapsto \alpha \cdot x$$

est linéaire. On l'appelle *l'homothétie de E de rapport α.*

Lorsque $\alpha = 1$, c'est de l'application *identité id_E* que nous sommes en train de parler. Lorsque $\alpha = 0$, c'est de l'application *nulle* $x \mapsto \mathbf{0}$ que nous sommes en train de parler.
Plus généralement, si E et F sont deux espaces vectoriels, l'application de E dans F définie par $f(x) = \mathbf{0}_F$ pour tout x de E est linéaire.

(d) **Transposition et trace des matrices**

Fixons deux éléments n, p de $\mathbb{N}^\star$. L'application

$$\begin{aligned} \mathcal{M}_{np}(\mathbb{K}) &\to \mathcal{M}_{pn}(\mathbb{K}) \\ M &\mapsto M^T \end{aligned}$$

est linéaire. Voir les rappels du chapitre 1.

Si nous restreignons notre attention aux matrices carrées, l'application

$$\begin{aligned} \mathcal{M}_{n}(\mathbb{K}) &\to \mathbb{K} \\ M &\mapsto \mathrm{Tr}(M) \end{aligned}$$

est aussi linéaire, c'est une forme linéaire sur $\mathcal{M}_n(\mathbb{K})$. Voir les rappels du chapitre 1.

(e) **Dérivation de polynômes ou de fonctions.**

- L'application
$$\begin{aligned} D : \mathbb{K}[X] &\to \mathbb{K}[X] \\ P(X) &\mapsto P'(X) \end{aligned}$$
  est linéaire. C'est un endomorphisme de $\mathbb{K}[X]$.

  En effet, si P et Q sont deux polynômes de $\mathbb{K}[X]$, alors on a $D(\lambda P + \mu Q) = (\lambda P + \mu Q)' = \lambda P' + \mu Q' = \lambda D(P) + \mu D(Q)$.

- De même, si $\mathcal{D}(\mathbb{R}, \mathbb{R})$ est le sous-espace de $\mathcal{F}(\mathbb{R}, \mathbb{R})$ formé des fonctions qui sont dérivables sur $\mathbb{R}$, l'application
$$\begin{aligned} \mathcal{D}(\mathbb{R}, \mathbb{R}) &\to \mathcal{F}(\mathbb{R}, \mathbb{R}) \\ f &\mapsto f' \end{aligned}$$
  est linéaire.

(f) **Linéarité de l'intégrale**

Soient a et b deux nombres réels avec $a < b$
et $\mathcal{C}([a, b], \mathbb{R})$ le sous-espace de $\mathcal{F}([a, b], \mathbb{R})$ formé des fonctions continues sur $[a, b]$. L'application

$$\begin{aligned} \mathcal{C}([a, b], \mathbb{R}) &\to \mathbb{R} \\ f &\mapsto \int_a^b f \end{aligned}$$

est linéaire.

(g) **Translation de la variable dans une fonction**

Soit a un nombre réel. L'application

$$\begin{aligned} \tau_a : \mathcal{F}(\mathbb{R}, \mathbb{R}) &\to \mathcal{F}(\mathbb{R}, \mathbb{R}) \\ f &\mapsto \tau_a[f] = [x \mapsto f(x - a)] \end{aligned}$$

est linéaire.

(h) **Multiplication par une fonction**

Soit φ une fonction de $\mathbb{R}$ dans $\mathbb{R}$. L'application

$$\begin{aligned} \mathfrak{m}_\varphi : \mathcal{F}(\mathbb{R}, \mathbb{R}) &\to \mathcal{F}(\mathbb{R}, \mathbb{R}) \\ f &\mapsto \varphi \cdot f = [x \mapsto \varphi(x) f(x)] \end{aligned}$$

est linéaire.

(i) **Composition à droite** Soit φ une fonction de $\mathbb{R}$ dans $\mathbb{R}$. L'application

$$\begin{aligned} CD_\varphi : \mathcal{F}(\mathbb{R},\mathbb{R}) &\to \mathcal{F}(\mathbb{R},\mathbb{R}) \\ f &\mapsto f \circ \varphi \end{aligned}$$

est linéaire. Ne pas confondre avec l'exemple précédent !

Pour le démontrer, constatons que si f et g sont deux fonctions et λ et μ deux réels, $CD_\varphi(\lambda f + \mu g)$ est la fonction $x \mapsto (\lambda f + \mu g)(\varphi(x)) = \lambda f(\varphi(x)) + \mu g(\varphi(x)) = \lambda CD_\varphi(f)(x) + \mu CD_\varphi(g)(x)$, c'est-à-dire la fonction $\lambda CD_\varphi(f) + \mu CD_\varphi(g)$.

Attention, en revanche, l'application de composition à gauche

$$\begin{aligned} CG_\varphi : \mathcal{F}(\mathbb{R},\mathbb{R}) &\to \mathcal{F}(\mathbb{R},\mathbb{R}) \\ f &\mapsto \varphi \circ f \end{aligned}$$

n'est presque jamais linéaire. Par exemple, si φ est la fonction $x \mapsto x^2$, CG_φ est l'application $f \mapsto f^2$, et si f n'est pas l'application nulle, on n'a pas $CG_\varphi(2f) = 2 \cdot CG_\varphi(f)$, puisque $CG_\varphi(2f) = (2f)^2 = 4f^2 = 4 \cdot CG_\varphi(f)$.

(j) **Sommation de vecteurs**

Soit E un $\mathbb{K}$-espace vectoriel et r un élément de $\mathbb{N}^\star$.

Rappelons que le produit cartésien $E^r = \underbrace{E \times E \times ... \times E}_{r \text{ fois}}$ est un espace vectoriel. L'application

$$\begin{aligned} \Sigma : E^r &\to E \\ (u_1, ...u_r) &\mapsto u_1 + .. + u_r \end{aligned}$$

est linéaire. En effet, si $U = (u_1, ..u_r)$ et $V = (v_1, ..., v_r)$ sont deux vecteurs de E^r et si λ est un scalaire de $\mathbb{K}$, alors

$$\Sigma(U + V) = (u_1 + v_1) + ... + (u_r + v_r) = (u_1 + ... + u_r) + (v_1 + ... + v_r) = \Sigma(U) + \Sigma(V);$$

$$\Sigma(\lambda \cdot U) = \lambda \cdot u_1 + ... + \lambda \cdot u_r = \lambda \cdot (u_1 + ... + u_r) = \lambda \cdot \Sigma(U).$$

(k) **Restriction d'une application linéaire à un sous-espace vectoriel.**

Soient E et F deux $\mathbb{K}$-espaces vectoriels et $f : E \to F$ une application.

Si A est un *sous-espace vectoriel* de E et si f est linéaire, alors la restriction

$$\begin{aligned} f_{|A} : A &\to F \\ x &\mapsto f(x) \end{aligned}$$

est aussi linéaire.

(k,bis) **Inclusion d'un sous-espace vectoriel de E dans E.**
En particulier, si A est un sous-espace vectoriel de E, *l'injection canonique*

$$\begin{aligned} \iota : A &\to E \\ x &\mapsto x \end{aligned}$$

est linéaire.

1.3. Exemples d'applications non-linéaires. —

- L'application
$$f : \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ \begin{pmatrix} x \\ y \end{pmatrix} & \mapsto & x+3y+1 \end{array}$$
n'est pas linéaire. En effet, l'image du vecteur nul de $\mathbb{R}^2$ par f n'est pas nulle.
- L'application
$$g : \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ \begin{pmatrix} x \\ y \end{pmatrix} & \mapsto & xy \end{array}$$
n'est pas linéaire. En effet, $g\left[\begin{pmatrix} 2 \\ 2 \end{pmatrix}\right] = 4$, ce qui n'est pas égal à $2 \cdot g\left[\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right]$.
- Les applications
$$h : \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ \begin{pmatrix} x \\ y \end{pmatrix} & \mapsto & x^2+3y-e^x \end{array} \quad \text{et} \quad T : \begin{array}{ccc} \mathcal{F}(\mathbb{R},\mathbb{R}) & \to & \mathcal{F}(\mathbb{R},\mathbb{R}) \\ f & \mapsto & \ln(f^4+1) \end{array}$$
ne sont pas non plus linéaires. Voyez-vous pourquoi ?
- (exemple moins formel).
L'inversion des matrices n'est pas une opération linéaire. En effet :
  - Si A est une matrice inversible de $\mathcal{M}_n(\mathbb{K})$, alors $2A$ est aussi inversible, mais $(2A)^{-1}$ est égale à $0.5A^{-1}$, pas à $2A^{-1}$.
  - Au-delà des calculs, la définition de la linéarité n'a de sens que si les espaces de départ et d'arrivée sont des espaces vectoriels. Or il existe des matrices qui ne sont pas inversibles, et l'ensemble des matrices inversibles de $\mathcal{M}_n(\mathbb{K})$ n'est *pas un sous-espace vectoriel de* $\mathcal{M}_n(\mathbb{K})$ (il ne contient même pas la matrice nulle).

1.4. Opérations sur les applications linéaires. —

Définition 4.6 – Espace des applications linéaires

On note $\mathcal{L}(E,F)$ l'espace des applications linéaires de E dans F.
On note $\mathcal{L}(E)$ pour l'espace $\mathcal{L}(E,E)$ des endomorphismes de E.

Proposition 4.7 – Combinaison linéaire d'applications linéaires

Si $f : E \to F$ et $g : E \to F$ sont deux applications linéaires, et si λ, μ sont deux scalaires, alors l'application $\lambda f + \mu g$ est aussi une application linéaire.

Ainsi $\mathcal{L}(E,F)$ est un sous-espace vectoriel de $\mathcal{F}(E,F)$.

(pour la seconde phrase, on rappelle que l'application nulle de E dans F est linéaire).

Proposition 4.8 – Composition d'applications linéaires

Si $f : E \to F$ et $g : F \to G$ sont deux applications linéaires, alors la composée $g \circ f : E \to G$ est aussi une application linéaire. Si $f : E \to F$ est une application linéaire *et* bijective, *alors sa bijection réciproque $f^{-1} : F \to E$ est aussi une application linéaire.*

En combinant ces deux propriétés et les exemples (a) à (j) ci-dessus, on peut obtenir beaucoup d'applications linéaires.

Par exemple, la linéarité de l'application

$$\begin{aligned} &\mathcal{C}^\infty(\mathbb{R},\mathbb{R}) \to \mathcal{F}(\mathbb{R},\mathbb{R}) \\ &\qquad f \mapsto \Phi(f) : x \mapsto f''(x) - (x^6+8)f'(x+\pi) - \int_0^1 f(e^t)dt \end{aligned}$$

peut être obtenue en combinant les exemples (e), (g), (h), (i), (k) du numéro 4.1.2. De même, on peut obtenir la linéarité de

$$\begin{aligned} &\mathbb{K}[X]^2 \to \mathbb{K}[X] \\ &(P,Q) \mapsto XP'(X) + (X^5-3)Q(X^2+4) \end{aligned}$$

en combinant (d), (e) et les versions évidentes de (i) et (j) pour les polynômes.

Proposition 4.9 – Pour les applications linéaires T_A, composition = produit

Si $A \in \mathcal{M}_{np}(\mathbb{K})$ et $B \in \mathcal{M}_{pq}(\mathbb{K})$, on a l'égalité

$$T_A \circ T_B = T_{AB}$$

Si A est une matrice inversible de $\mathcal{M}_n(\mathbb{K})$, alors $T_A : \mathbb{K}^n \to \mathbb{K}^n$ est bijective et on a l'égalité

$$T_A^{-1} = T_{A^{-1}}$$

C'est une conséquence immédiate de la définition de T_A pour une matrice A.

Cette propriété des applications linéaires canoniquement associées à des matrices n'est pas la seule analogie qu'on puisse trouver avec la notion de produit :

Proposition 4.10 – Distributivités entre la composition et les combinaisons

Soient E, F, G trois $\mathbb{K}$-espaces vectoriels, f_1, f_2 et f des éléments de $\mathcal{L}(E,F)$, g_1, g_2 et g deux éléments de $\mathcal{L}(F,G)$, λ et μ deux scalaires. Les deux égalités suivantes sont vraies :

(i) $g \circ (\lambda f_1 + \mu f_2) = \lambda(g \circ f_1) + \mu(g \circ f_2)$;

(ii) $(\lambda g_1 + \mu g_2) \circ f = \lambda(g_1 \circ f) + \mu(g_2 \circ f)$.

Démonstration. —

(i) On constate que pour tout x de E, on a $[g \circ (\lambda f_1 + \mu f_2)](x) = g[\lambda f_1(x) + \mu f_2(x)]$, ce qui par linéarité de g vaut $\lambda g(f_1(x)) + \mu g(f_2(x)) = [\lambda(g \circ f_1) + \mu(g \circ f_2)](x)$. C'est ce qu'on voulait montrer.

(ii) La démonstration est la même que pour l'exemple (j) page 93. □

□

***Remarque 4.11* (Notation fg pour la composée $f \circ g$).** — Grâce à cette propriété, on peut manier les compositions d'applications linéaires comme s'il s'agissait de produits usuels.

Par exemple, si f est un endomorphisme d'un espace vectoriel E, on notera souvent

$$3f^6 - 7f^3 + \sqrt{2}f - 8$$

pour l'application $3\underbrace{(f \circ ... \circ f)}_{6 \text{ fois}} - 7\underbrace{(f \circ ... \circ f)}_{3 \text{ fois}} + \sqrt{2}f - 8\mathrm{Id}_E$.

De même, si f et g sont deux endomorphismes de E, on pourra parler de l'endomorphisme $f^2g - 17gf^8 + f - g + 9$.

Attention. — Les composées $f \circ g$ et $g \circ f$ sont en général différentes (songer au cas de composées $T_A \circ T_B$ si A et B sont des matrices carrées qui ne commutent pas). On prendra donc garde, dans les raccourcis d'écriture du type ci-dessus, à ne pas maltraiter l'ordre dans les produits.

Vocabulaire : puissances d'un endomorphisme, endomorphisme nilpotent.

Si E est un espace vectoriel, si f appartient à $\mathcal{L}(E)$ et si k appartient à $\mathbb{N}$, on notera ainsi f^k pour l'endomorphisme $\underbrace{(f \circ ... \circ f)}_{k \text{ fois}}$ de E (par exemple, $f^0 = \mathrm{Id}_E$, $f^1 = f$, $f^2 = f \circ f$, etc).

On dit que f est *nilpotent* s'il existe n de $\mathbb{N}^\star$ tel que f^n soit l'endomorphisme nul. Le plus petit entier p pour lequel $f^p = 0$ est appelé *l'indice de nilpotence de f*.

1.5. Effet sur les sous-espaces vectoriels. —

Dans ce paragraphe, on fixe deux espaces vectoriels E et F.

Proposition 4.12 – Image d'un sous-espace vectoriel par une application linéaire

Si $f : E \to F$ est linéaire et si A est un sous-espace vectoriel de E, alors $f(A)$ est un sous-espace vectoriel de F.

Démonstration. — Le vecteur nul $\mathbf{0}_F$ appartient à $f(A)$ puisque c'est l'image de $\mathbf{0}_E$ (qui appartient à A) par f si y, y' sont deux vecteurs de $f(A)$ et si λ et μ sont deux scalaires, il existe des vecteurs x, x' de E tels que $f(x) = y$ et $f(x') = y'$, et alors

$$\lambda y + \mu y' = \lambda f(x) + \mu f(x') \overset{\text{linéarité}}{=} f\left(\underbrace{\lambda x + \mu x'}_{\in A}\right), \quad \text{si bien que } \lambda y + \mu y' \text{ appartient à } f(A).$$

□

Proposition 4.13 – Image réciproque d'un sous-espace vectoriel par une appli. linéaire

Si $f : E \to F$ est linéaire et si B est un sous-espace vectoriel de F, alors $f^{-1}(B)$ est un sous-espace vectoriel de E.

Démonstration. — L'image du vecteur nul $\mathbf{0}_E$ par f est bien dans B puisque c'est le vecteur nul de F et B est un sous-espace de F ;
par ailleurs, si x et x' sont deux éléments de $f^{-1}(B)$ et si λ et μ sont deux scalaires,
$f(\lambda x + \mu x') = \lambda \underbrace{f(x)}_{\in B} + \mu \underbrace{f(x')}_{\in B}$ appartient à B, ce qui signifie que $\lambda x + \mu x'$ appartient encore à $f^{-1}(B)$. □

Proposition 4.14 – Graphe d'une application linéaire

Si $f : E \to F$ est linéaire, alors le graphe $\Gamma_f = \{(x,y) \in E \times F, y = f(x)\}$ est un sous-espace vectoriel de $E \times F$.

Démonstration. — On peut utiliser le résultat précédent : notons

$$\begin{aligned} u : E \times F &\to F \\ (x,y) &\mapsto y - f(x) \end{aligned}$$

et constatons que l'application u est linéaire [3]. L'ensemble $\{\mathbf{0}_F\}$ est un sous-espace vectoriel de F, donc $\Gamma_f = u^{-1}(\{\mathbf{0}_F\})$ est un sous-espace vectoriel de $E \times F$. □

La réciproque du résultat précédent est en fait vraie :

Proposition 4.15 – Graphe et linéarité

Soient E et F deux espaces vectoriels et $f : E \to F$ est une application.
Il y a équivalence entre

- *l'application f est linéaire ;*
- *le graphe $\Gamma_f = \{(x,y) \in E \times F, y = f(x)\}$ est un sous-espace vectoriel de $E \times F$.*

Démonstration. — Supposons que Γ_f soit un sous-espace vectoriel de $E \times F$. Si u, v sont deux éléments de E et λ, μ deux scalaires, les points $(u, f(u))$ et $(v, f(v))$ de $E \times F$ appartiennent au graphe Γ_f, donc le point $\lambda(u, f(u)) + \mu(v, f(v))$ appartient aussi à Γ_f. Or, il s'agit du point $(\lambda u + \mu v, \lambda f(u) + \mu f(v))$ et un point (w, z) appartient à Γ_f si et seulement si $z = f(w)$. On en déduit $f(\lambda u + \mu v) = \lambda f(u) + \mu f(v)$, et ainsi f est linéaire. □

3. Si (x_1, y_1) et (x_2, y_2) sont deux vecteurs de $E \times F$ et si λ et μ sont deux scalaires, on a $u\left[\lambda(x_1, y_1) + \mu(x_2, y_2)\right] = (\lambda y_1 + \mu y_2) - f(\lambda x_1 + \mu x_2)$, ce qui par linéarité de f vaut $\lambda y_1 + \mu y_2 - \lambda f(x_1) - \mu f(x_2) = \lambda u(x_1, y_1) + \mu u(x_2, y_2)$, cqfd.

2. Noyau et image d'une application linéaire

2.1. Noyau d'une application linéaire ou d'une matrice. —

Définition 4.16 – Noyau d'une application linéaire

Soient E et F deux $\mathbb{K}$-espaces vectoriels et $f \in \mathcal{L}(E, F)$ une application linéaire de E vers F.
Le *noyau de f*, noté $\mathrm{Ker}(f)$, est

$$\mathrm{Ker}(f) = \{x \in E,\ f(x) = \mathbf{0}_F\}.$$

C'est un sous-espace vectoriel de l'espace *de départ* E.

Vous connaissez déjà cette notion dans le cadre ensembliste : le noyau de f est l'ensemble $f^{-1}(\{\mathbf{0}_F\})$ des vecteurs dont l'image par f est $\mathbf{0}_F$. Le seul ajout de la définition ci-dessus est donc *terminologique*.

Exemple 4.17. — Considérons l'application $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ canoniquement associée à $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$. Dire qu'un vecteur $\begin{pmatrix} x \\ y \end{pmatrix}$ appartient à $\mathrm{Ker}(T_A)$, c'est dire que $T_A\left[\begin{pmatrix} x \\ y \end{pmatrix}\right] = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, c'est-à-dire $\begin{pmatrix} x + 2y \\ 3x + 6y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Ainsi, $\mathrm{Ker}(T_A)$ est la droite $\left\{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2,\ \ y = -x/2\right\}$ de $\mathbb{R}^2$.

Exemple 4.18. — Considérons l'application $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ canoniquement associée à $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.
Vérifions que $\mathrm{Ker}(T_A) = \{\mathbf{0}_{\mathbb{R}^2}\}$ de deux façons différentes :

- Dire qu'un vecteur $\begin{pmatrix} x \\ y \end{pmatrix}$ appartient à $\mathrm{Ker}(T_A)$, c'est dire que $T_A\left[\begin{pmatrix} x \\ y \end{pmatrix}\right] = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, c'est-à-dire $\begin{pmatrix} x + 2y \\ 3x + 4y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Mais le système linéaire $\begin{cases} x + 2y = 0 \\ 3x + 4y = 0 \end{cases}$ équivaut à $\begin{cases} x + 2y = 0 \\ -2 = 0 \end{cases}$, ou encore à $\begin{cases} x = 0 \\ y = 0 \end{cases}$, ce qu'il fallait démontrer.
- Un vecteur X de $\mathbb{R}^2$ est dans le noyau de T_A si et seulement si $T_A(X) = \mathbf{0}_{\mathbb{R}^2}$; or la matrice A est inversible et $T_A(X) = AX$, donc $T_A(X) = \mathbf{0}_{\mathbb{R}^2}$ équivaut à $AX = \mathbf{0}$ ou encore à $A^{-1}(AX) = A^{-1}\mathbf{0}_{\mathbb{R}^2}$, c'est-à-dire à $X = \mathbf{0}_{\mathbb{R}^2}$.

Ainsi $\mathrm{Ker}(T_A) = \{\mathbf{0}_{\mathbb{R}^2}\}$.

Exemple 4.19. — L'application

$$\begin{aligned} \Phi : \mathcal{F}(\mathbb{R}, \mathbb{R}) &\to \mathcal{F}(\mathbb{R}, \mathbb{R}) \\ f &\mapsto \Phi(f) : [x \mapsto f(x + 2\pi) - f(x)] \end{aligned}$$

est linéaire, et dire qu'une fonction f appartient à $\mathrm{Ker}(\Phi)$, c'est dire qu'on a : $\forall x \in \mathbb{R}, f(x + 2\pi) = f(x)$. Ainsi $\mathrm{Ker}(\Phi)$ est l'ensemble des fonctions 2π-périodiques.

Exemple 4.20. — L'application

$$\begin{aligned} \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}) &\to \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}) \\ y &\mapsto y' - y \end{aligned}$$

est linéaire, et son noyau est l'ensemble des fonctions $\mathcal{C}^\infty$ vérifiant l'équation différentielle linéaire $y' - y = 0$: c'est l'ensemble des fonctions proportionnelles à exp, c'est-à-dire la droite vectorielle $\mathbf{Vect}[\exp]$.

Proposition 4.21 – Noyau et injectivité

Soit $f : E \to F$ une application linéaire. Il y a équivalence entre
- *f est injective ;*
- $\mathrm{Ker}(f) = \{\mathbf{0}_E\}$.

Démonstration. — Si f est linéaire et si x_1, x_2 appartiennent à E, alors $f(x_1) - f(x_2) = f(x_1 - x_2)$ et donc

$$f(x_1) = f(x_2) \quad \text{si et seulement si} \quad x_1 - x_2 \in \mathrm{Ker}(f)$$

donc si $\mathrm{Ker}(f) = \{\mathbf{0}_E\}$, on ne peut avoir $f(x_1) = f(x_2)$ que si $x_1 = x_2$.
L'autre sens est immédiat, puisque si f est injective et si x appartient à $\mathrm{Ker}(f)$, $f(x) = f(\mathbf{0}_E)$, ce qui donne $x = \mathbf{0}_E$. □

***Remarque 4.22* (En pratique).** — Soit $f : E \to F$ une application linéaire. Puisque $\mathrm{Ker}(f)$ est un sous-espace vectoriel, on sait déjà qu'il contient $\mathbf{0}_E$. Pour vérifier qu'une application linéaire f est injective, il suffit donc de vérifier que si x est un vecteur tel que $f(x) = \mathbf{0}_F$, alors on a forcément $x = \mathbf{0}_E$.

Exemple 4.23. — Plaçons-nous dans $\mathbb{R}[X]$, considérons le sous-espace vectoriel $E_0 = \{P \in \mathbb{R}[X], P(0) = 0\}$ et

$$f : E_0 \to \mathbb{R}[X]$$
$$P(X) \mapsto XP'(X).$$

L'application f est linéaire (voir les exemples généraux ci-dessus). Si P est un polynôme appartenant à E_0 vérifiant $f(P) = 0$, alors le polynôme $XP'(X)$ est nul, donc le polynôme $P'(X)$ est nul, donc le polynôme P est constant ; puisque $P(0) = 0$, le polynôme P est nul. Ainsi $\mathrm{Ker}(f) = \{\mathbf{0}_{E_0}\}$ et f est injective.

Noyau d'une matrice. — Si A est une matrice de $\mathcal{M}_{np}(\mathbb{K})$, on appelle *noyau de A* le sous-espace $\mathrm{Ker}(T_A) = \{X \in \mathbb{K}^p, AX = \mathbf{0}_{\mathbb{R}^n}\}$ de $\mathbb{K}^p$.
On le note $\mathrm{Ker}(A)$.

Si $C_1, ...C_p$ sont les p vecteurs de $\mathbb{K}^n$ qui donnent les colonnes de la matrice A, il faut remarquer que

$$\mathrm{Ker}(A) = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathbb{K}^p,\ x_1C_1 + ... + x_pC_p = 0 \right\} :$$

un vecteur $\begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$ appartient à $\mathrm{Ker}(A)$ si et seulement s'il fournit une relation de dépendance entre les colonnes de A.

$\rightsquigarrow$ Par exemple, si A est la matrice $\begin{pmatrix} 1 & 3 & 1 & -2 \\ 2 & 5 & 3 & -6 \\ 4 & 7 & 3 & -6 \end{pmatrix}$,

le fait que la dernière colonne est « moins deux fois la troisième » se traduit par le fait que $\begin{pmatrix} 0 \\ 0 \\ 2 \\ 1 \end{pmatrix}$ appartient à $\mathrm{Ker}(A)$,

le fait que deuxième colonne est la somme de la première et de la troisième se traduit par : $\begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}$ appartient à $\mathrm{Ker}(A)$

2.2. Sous-espace image d'une application linéaire. —

Définition 4.24 – Sous-espace image d'une application linéaire

Soient E et F deux $\mathbb{K}$-espaces vectoriels et $f \in \mathcal{L}(E, F)$ une application linéaire de E vers F. Le *sous-espace image de f*, noté $\mathrm{Im}(f)$, est

$$\mathrm{Im}(f) = \{y \in F, \exists x \in E \ f(x) = y\}.$$

C'est un sous-espace vectoriel de l'espace *d'arrivée* F.

Là encore, vous connaissez déjà cette notion dans le cadre ensembliste.

Proposition 4.25 – Image et surjectivité

Soit $f : E \to F$ une application linéaire. Il y a équivalence entre

- *f est surjective ;*
- *$\mathrm{Im}(f) = F$.*

Ici il n'y a rien à montrer.

Proposition 4.26 – Détermination d'une famille génératrice de $\mathrm{Im}(f)$

Soient E et F deux $\mathbb{K}$-espaces vectoriels et $f \in \mathcal{L}(E, F)$ une application linéaire de E vers F. Si $(x_i)_{i\in I}$ est une famille de vecteurs de E, alors $f\left(\mathbf{Vect}[(x_i)_{i\in I}]\right) = \mathbf{Vect}[(f(x_i)_{i\in I}]$. En particulier, si $(x_i)_{i\in I}$ est une famille génératrice de E, on a $\mathrm{Im}(f) = \mathbf{Vect}[(f(x_i))_{i\in I}]$.

Exemple 4.27. — Fixons n dans $\mathbb{N}^\star$, plaçons-nous dans $\mathbb{K}_n[X]$ et considérons l'endomorphisme

$$\begin{aligned} D : \mathbb{K}_n[X] &\to \mathbb{K}_n[X] \\ P(X) &\mapsto P'(X). \end{aligned}$$

La famille $(1, X, ...X^n)$ est génératrice de $\mathbb{K}_n[X]$. Son image par D est la famille $(0, 1, 2X, .., nX^{n-1})$. On obtient donc : $\mathrm{Im}(D) = \mathbf{Vect}\big[0, 1, 2X, .., nX^{n-1}\big]\,\mathbf{Vect}\big[1, X, ..., X^{n-1}\big] = \mathbb{K}_{n-1}[X]$.

Ce qui précède est moins concret, mais plus rapide, que de vérifier que si $a_0 + a_1X + ... + a_{n-1}X^{n-1}$ est un polynôme arbitraire de $\mathbb{K}_{n-1}[X]$, alors il peut s'écrire $D\left(a_0X + \frac{a_1}{2}X^2 + ... + \frac{a_{n-1}}{n}X^n\right)$.

Image d'une matrice. — Si A est une matrice de $\mathcal{M}_{np}(\mathbb{K})$, on appelle *image de* A le sous-espace $\mathrm{Im}(T_A) = \{Y \in \mathbb{K}^n,\ \exists X \in \mathbb{K}^p, Y = AX\}$ de $\mathbb{K}^n$.
On le note $\mathrm{Im}(A)$. Rappelons T_A est une application de $\mathbb{K}^p$ dans $\mathbb{K}^n$ et que la définition du produit matriciel implique

Les colonnes de A sont les images par T_A des vecteurs de la base canonique de $\mathbb{K}^p$:
si e_j est le $j^{\text{ème}}$ vecteur de la base canonique de $\mathbb{K}^p$, alors Ae_i est la $j^{\text{ème}}$ colonne de A.

Comme la base canonique est une famille génératrice de $\mathbb{K}^p$, la deuxième partie de la proposition précédente implique

$\mathrm{Im}(A) = \mathbf{Vect}[C_1, ...C_p]$, où $C_1, .., C_p$ est la famille des colonnes de A.

Attention à une confusion fréquente : — si A est une matrice $n \times p$ qui n'est pas carrée, ou si f est une application linéaire de E dans F où les espaces vectoriels E et F sont différents, le noyau $\mathrm{Ker}(A)$ et l'image $\mathrm{Im}(A)$, ou le noyau $\mathrm{Ker}(f)$ et l'image $\mathrm{Im}(f)$, *ne sont pas des sous-espaces du même espace vectoriel.* Attention donc à bien vérifier la nature des objets dans $\mathrm{Ker}(f)$ ou $\mathrm{Im}(f)$ (objets abstraits, coordonnées, etc) : se tromper dessus mène à écrire des choses qui n'ont pas de sens même après des calculs justes.

Par exemple, l'application

$$\begin{aligned} \mathcal{M}_2(\mathbb{R}) &\to \mathbb{R}_2[X] \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} &\mapsto (a+d)X^2 + (b+c)X + (a-d) \end{aligned}$$

est linéaire et *son noyau est un espace de matrices*, alors que *son image est un espace de polynômes.*

2.3. Principe de superposition des solutions d'équations linéaires. — Lorsque E et F sont deux espaces vectoriels et $f : E \to F$ est une application linéaire, on dit que l'équation

$$f(x) = \mathbf{0}_F$$

d'inconnue $x \in E$, est une *équation linéaire homogène* (ou, si le contexte s'y prête, un *système d'équations linéaires homogènes*). Son ensemble de solutions est bien sûr $\mathrm{Ker}(f)$.

Une conséquence de la linéarité de f est le fait que la somme de deux solutions de l'équation $f(x) = \mathbf{0}_F$ est encore une solution. Par exemple, fixons un réel ω ; si on considère l'application linéaire

$$\begin{aligned} \Phi : \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}) &\to \mathcal{F}(\mathbb{R}, \mathbb{R}) \\ y &\mapsto y'' + \omega^2 y \end{aligned}$$

les fonctions $t \mapsto \sin(\omega t)$ et $t \mapsto \cos(\omega t)$ sont solutions de l'équation $\Phi(y) = 0$, donc pour tout α et tout β de $\mathbb{R}$, la fonction $t \mapsto \alpha\cos(\omega t) + \beta\sin(\omega t)$ en est aussi une solution.

L'abondance d'équations linéaires dans les applications des mathématiques (où, dans l'équation $f(x) = 0$, l'inconnue x peut être un objet très compliqué : fonction, matrice, image, onde, son, champ électromagnétique, état d'un système quantique...) est une des raisons de l'importance de l'algèbre linéaire dans les sciences contemporaines. Le *principe de superposition* que nous venons de discuter est, dans toutes ces applications, un outil crucial (même s'il est mathématiquement évident !).

Si un vecteur b de F est donné, l'équation

$$f(x) = b$$

est une *équation linéaire (inhomogène).*

Proposition 4.28 – Structure de l'ensemble des solutions d'une équation linéaire

Soit $f : E \to F$ une application linéaire et b un vecteur de F.
Considérons l'équation (E) : $f(x) = b$, d'inconnue $x \in E$.

- *Si b n'appartient pas à $\mathrm{Im}(f)$, cette équation n'a pas de solution.*
- *Si b appartient à $\mathrm{Im}(f)$, et si x_0 est un vecteur de E tel que $f(x_0) = b$ (*solution particulière de *(E)),*
  alors l'ensemble des solutions de (E) est l'ensemble des vecteurs de la forme $u + x_0$ avec $u \in \mathrm{Ker}(f)$.
  (les solutions de (E) sont obtenues en ajoutant, à une solution particulière, les solutions de l'équation sans second membre).

***Exemple 4.29* (Résolution d'une équation différentielle inhomogène)**

Soient b un nombre réel et a un nombre réel non nul. L'ensemble des solutions de l'équation différentielle $y' = ay + b$, dont l'inconnue est une fonction y de $\mathbb{R}$ dans $\mathbb{R}$, peut être obtenu comme suit : désignant par $\mathcal{D}(\mathbb{R}, \mathbb{R})$ l'ensemble des fonctions dérivables de $\mathbb{R}$ dans $\mathbb{R}$, par C_b la fonction constante de valeur b, et considérant l'application linéaire

$$\begin{aligned} \Phi : \mathcal{D}(\mathbb{R}, \mathbb{R}) &\to \mathcal{F}(\mathbb{R}, \mathbb{R}) \\ y &\mapsto y' - ay \end{aligned}$$

on constate qu'il s'agit de chercher les solutions de l'équation linéaire $\Phi(y) = C_b$.

La fonction constante $C_{-b/a}$ de valeur $-b/a$ est une solution particulière de cette équation, et les solutions de l'équation $\Phi(y) = 0$ sont les fonctions de la forme $x \mapsto Ke^{ax}$ où K est un nombre réel.

D'après la proposition précédente, une fonction y est solution de l'équation de départ si et seulement s'il existe un réel K tel qu'on ait : $\forall x \in \mathbb{R}, y(x) = Ke^{ax} - \frac{b}{a}$.

3. Injectivité, surjectivité d'applications linéaires

Nous avons déjà vu un critère pratique essentiel pour savoir si une application linéaire est injective, à l'aide du noyau.

En voici un autre.

Proposition 4.30 – Injectivité d'applications linéaires et familles libres

Soit $f : E \to F$ une application linéaire. Il y a équivalence entre
- *f est injective ;*
- *si $(x_i)_{i\in I}$ est une famille* libre *de E, alors la famille $(f(x_i))_{i\in I}$ est une famille libre de F.*

Démonstration. —
- Supposons que l'image par f de toute famille libre de E soit libre ; montrons que f est injective, et pour cela, vérifions que $\mathrm{Ker}(f) = \{\mathbf{0}_E\}$. Si x est un vecteur non nul de E, la famille (x) d'un seul vecteur est libre, donc avec notre hypothèse la famille (f(x)) est aussi libre, ce qui ne peut arriver que si $f(x)$ est non nul. Ainsi le seul vecteur x tel que $f(x) = 0$ est le vecteur nul de E.
- Supposons f injective. Soit $(x_i)_{i\in I}$ une famille libre de E, vérifions que $(f(x_i))_{i\in I}$ est libre. Pour cela, il faut montrer que pour tout p de $\mathbb{N}^\star$ et pour toute sous-famille à p éléments $\big(f(x_{i_1}), ..., f(x_{i_p})\big)$ de $(f(x_i))_{i\in I}$, si $\lambda_1, ..., \lambda_p$ sont des scalaires tels que $\sum\limits_{k=1}^{p} \lambda_k f(x_{i_k}) = \mathbf{0}_E$, alors tous les λ_k, $k = 1..p$ sont nuls. Mais si $\sum\limits_{k=1}^{p} \lambda_k f(x_{i_k}) = \mathbf{0}_E$, alors par linéarité de f, on a $f\left(\sum\limits_{k=1}^{p} \lambda_k x_{i_k}\right) = \mathbf{0}_F$. Comme f est injective, la combinaison $\sum\limits_{k=1}^{p} \lambda_k x_{i_k}$ doit être nulle, ce qui par liberté de $(x_i)_{i\in I}$ implique bien que tous les λ_k, $k = 1..p$, sont nuls.

□

***Exemple 4.31* (Une conséquence importante).** —

(a) Si une matrice $A \in \mathcal{M}_{np}(\mathbb{K})$ a plus de colonnes que de lignes ($p > n$), alors $T_A : \mathbb{R}^p \to \mathbb{R}^n$ n'est jamais injective. En effet, l'image par T_A de la base canonique de $\mathbb{K}^p$ est la famille $(C_1, .., C_p)$ des colonnes de A ; si f était injective cette famille devrait être libre, mais comme c'est une famille de p vecteurs de $\mathbb{K}^n$, elle ne peut pas être libre si $p > n$.

(b) Plus généralement, lorsque E et F sont de dimension finie, on constate que

si $\dim(E) > \dim(F)$, alors aucune application linéaire de E dans F ne peut être injective.

En effet, si $(e_1, ..., e_p)$ est une base de E (avec $p = \dim(E)$), alors $(f(e_1), .., f(e_p))$ est une famille de p vecteurs de f, et si f est injective elle doit être libre puisque $(e_1, ..., e_p)$ l'est, mais c'est une famille de p vecteurs de F et elle ne peut pas être libre si $p > \dim(F)$.

Proposition 4.32 – Surjectivité d'applications linéaires et familles génératrices

Soit $f : E \to F$ une application linéaire. Il y a équivalence entre

(i) f est surjective ;

(ii) l'image de toute *famille génératrice de E est une famille génératrice de F ;*

(iii) il existe une *famille génératrice de E dont l'image par f est une famille génératrice de F.*

Démonstration. —

- Supposons (i) et montrons (ii). Soit $(x_i)_{i\in I}$ une famille génératrice de E. Montrons que $(f(x_i))_{i\in I}$ est génératrice de F, c'est-à-dire que $\mathbf{Vect}[(f(x_i))_{i\in I}] = F$. La proposition 4.26 donne $\mathbf{Vect}[(f(x_i))_{i\in I}] = f\left(\mathbf{Vect}[(x_i)_{i\in I}]\right) = f(E) = F$, cqfd.
- Le fait que (ii) implique (iii) est évident.
- Supposons (iii) et montrons (i). Supposons qu'il existe une famille génératrice $(x_i)_{i\in I}$ de E dont l'image par f est une famille génératrice de F. Montrons que F est surjective, c'est-à-dire que $F = f(E)$. Puisque $(f(x_i))_{i\in I}$ est génératrice de F, on a $\mathbf{Vect}[(f(x_i))_{i\in I}] = F$; mais on a $\mathbf{Vect}[(f(x_i))_{i\in I}] = f\left(\mathbf{Vect}[(x_i)_{i\in I}]\right)$ (voir la proposition 4.26), et puisque $(x_i)_{i\in I}$ est génératrice de E, on a $\mathbf{Vect}[(x_i)_{i\in I}] = E$. Nous obtenons comme voulu $F = f(E)$.

□

***Exemple 4.33* (Une conséquence importante).** — (a) Si une matrice $A \in \mathcal{M}_{np}(\mathbb{K})$ a plus de lignes que de lignes ($p < n$), alors $T_A : \mathbb{R}^p \to \mathbb{R}^n$ n'est jamais surjective. En effet, l'image par T_A de la base canonique de $\mathbb{K}^p$ est la famille $(C_1, .., C_p)$ des colonnes de A ; si f était surjective cette famille devrait être génératrice de $\mathbb{K}^n$, mais comme c'est une famille de p vecteurs de $\mathbb{K}^n$, elle ne peut pas être génératrice si $p < n$.

(b) Plus généralement, lorsque E et F sont de dimension finie, on constate que

Si $\dim(E) < \dim(F)$, alors aucune application linéaire de E dans F ne peut être surjective.

En effet, si $(e_1, ..., e_p)$ est une base de E (avec $p = \dim(E)$), alors $(f(e_1), .., f(e_p))$ est une famille de p vecteurs de f, et si f est surjective elle doit être génératrice puisque $(e_1, ..., e_p)$ l'est, mais c'est une famille de p vecteurs de F et elle ne peut pas être génératrice si $p < \dim(F)$.

En rassemblant les deux propositionsci-dessus, nous obtenons le résultat suivant.

Proposition 4.34 – Bijectivité d'applications linéaires et bases

Soit $f : E \to F$ une application linéaire. Il y a équivalence entre

- *f est bijective ;*
- *l'image par f de* toute *base de E est une base de F ;*
- *il existe* une *base de E dont l'image par f est une base de F.*

Cas d'une application linéaire canoniquement associée à une matrice.

Soit A une matrice *carrée* de taille n. Le résultat que nous venons de montrer indique que pour savoir si $T_A : \mathbb{K}^n \to \mathbb{K}^n$ est bijective, il suffit de savoir si l'image par T_A d'une base de $\mathbb{K}^n$ est une base de $\mathbb{K}^n$. Par exemple, il suffit de savoir si l'image par T_A de la base canonique de $\mathbb{K}^n$ est une base de $\mathbb{K}^n$.

Or, nous avons vu que l'image de la base canonique de $\mathbb{K}^n$ est la famille $(C_1, ..., C_n)$ des colonnes de A. Puisque c'est une famille de n vecteurs de $\mathbb{K}^n$, pour savoir si la famille $(C_1, ..., C_n)$ est une base de $\mathbb{R}^n$ il suffit de savoir si elle est libre. Ainsi T_A est bijective si et seulement si la famille $(C_1, ..., C_n)$ des colonnes de A est libre.

Mais nous avons aussi vu page 95 que T_A est bijective si et seulement si la matrice A est inversible. Ainsi A est inversible si et seulement si la famille $(C_1, ..., C_n)$ des colonnes de A est libre.

Par ailleurs, A est inversible si et seulement si sa transposée A^T l'est (voir le cours d'algèbre du premier semestre, Proposition 8.3.12), donc on peut aussi dire que A est inversible si et seulement si T_{A^T} est inversible ; puisque les colonnes de A^T sont les lignes de A, c'est équivalent à la liberté de la famille des lignes de A.

En résumé, nous avons obtenu par un chemin abstrait deux caractérisations très pratiques de l'inversibilité d'une matrice :

A est inversible

$\Updownarrow$

T_A est bijective

$\Updownarrow$

il n'y a pas de relation linéaire entre les *colonnes* de A

$\Updownarrow$

il n'y a pas de relation linéaire entre les *lignes* de A.

On peut relever que la dernière équivalence (les colonnes sont liées si et seulement si les lignes sont aussi liées), dont on vient de voir qu'elle était vraie pour *toutes* les matrices, n'est pas du tout évidente intuitivement !

Par exemple, la matrice $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 2 & 5 & 8 \end{pmatrix}$ n'est pas inversible, car ses deux premières lignes sont les mêmes ; mais si on observe les colonnes, on pourrait être tenté à première vue de penser qu'elles sont linéairement indépendantes.

4. Isomorphismes

4.1. Définition et exemples. —

Définition 4.35 – Isomorphisme linéaire

Soient E et F deux espaces vectoriels.
Si f est une application *linéaire* et *bijective* de E dans F, on dit que c'est un *isomorphisme entre E et F*.

On dira alors que deux espaces vectoriels E et F *sont isomorphes* s'il existe un isomorphisme entre E et F.

***Exemple 4.36* (Applications linéaires canoniquement associées à des matrices, conséquences pour $\mathbb{K}^n$ et $\mathbb{K}^p$)**

Considérons une matrice A de $\mathcal{M}_{np}(\mathbb{K})$, $(n,p) \in (\mathbb{N}^\star)^2$. Nous avons vu au paragraphe précédent que $T_A \in \mathcal{L}(\mathbb{K}^p, \mathbb{K}^n)$ ne peut pas être injective si $p > n$, et que T_A ne peut pas être surjective si $p < n$. En conséquence, T_A ne peut être un isomorphisme entre $\mathbb{K}^p$ et $\mathbb{K}^n$ que si $p = n$, c'est-à-dire si la matrice A est carrée. Dans ce cas, nous avons vu que la bijectivité de T_A équivaut à l'inversibilité de A.

Par ailleurs, nous avons vu à l'exemple (a) page **??** que pour toute application linéaire f de $\mathbb{K}^p$ dans $\mathbb{K}^n$, il existe $A \in \mathcal{M}_{np}(\mathbb{K})$ pour laquelle $f = T_A$. Ce qui précède donne donc un renseignement général sur les applications linéaires entre $\mathbb{K}^p$ et $\mathbb{K}^n$:

En résumé,

> Les espaces $\mathbb{K}^p$ et $\mathbb{K}^n$ ne sont isomorphes que si $p = n$.
> Si $p > n$, aucune application linéaire de $\mathbb{K}^p$ dans $\mathbb{K}^n$ n'est injective.
> Si $p < n$, aucune application linéaire de $\mathbb{K}^p$ dans $\mathbb{K}^n$ n'est surjective.

***Exemple 4.37* (Formule de Taylor pour les polynômes).** — • Soit n un entier naturel. L'application

$$\begin{aligned} \Phi : \mathbb{K}_n[X] &\to \mathbb{K}^{n+1} \\ P &\mapsto \Big(P(0), P'(0), ...P^{(n)}(0)\Big) \end{aligned}$$

est linéaire (cela se vérifie aisément).

• Vérifions qu'elle est injective en vérifiant que $\mathrm{Ker}(\Phi) = \{0\}$: si P est un élément de $\mathbb{K}_n[X]$ vérifiant $\Phi(P) = 0$, la formule de Taylor en zéro pour le polynôme P donne

$$P(X) = \sum_{k=0}^{n} P^{(k)}(0)\frac{X^k}{k!} = \sum_{k=0}^{n} 0 = 0$$

et donc P est le polynôme nul, comme voulu.

• Vérifions qu'elle est surjective. Si $u = (a_0, ..., a_n)$ est un élément de $\mathbb{K}^n$, considérons le polynôme $P(X) = a_0 + a_1 + a_2\dfrac{X^2}{2} + ... + a_n\dfrac{X^n}{n!}$. Alors la formule de Taylor ci-dessus montre que pour tout k de $\{0, .., n\}$, $P^{(k)}(0)$ vaut a_k. Le polynôme P est donc un antécédent de u par Φ.

• Ainsi, Φ est un isomorphisme :

Si P est de degré n et si on connaît $P(0), P'(0), P''(0), ...P^{(n)}(0)$, alors on connaît P complètement.

***Exemple 4.38* (Interpolation de Lagrange).** — • Soit n un entier naturel non nul et $(a_0, a_1, ..., a_n)$ des réels *deux à deux distincts*. Considérons l'application

$$\begin{aligned} \Psi : \mathbb{R}_n[X] &\to \mathbb{R}^{n+1} \\ P &\mapsto (P(a_0), P(a_1), ...P(a_n)) . \end{aligned}$$

C'est une application linéaire. Montrons brièvement que c'est un isomorphisme.

• Si P appartient au noyau de Ψ, alors $P(a_0) = P(a_1) = ... = P(a_n) = 0$, et donc le polynôme P admet $n+1$ racines distinctes. Comme P est de degré $\leq n$, cela ne peut arriver que si P est le polynôme nul. Ainsi $\mathrm{Ker}(\Psi) = \{0\}$ et Ψ est injective.

• Expliquons pourquoi Ψ est surjective. Soit $(y_0, y_1, ..., y_n)$ un élément de $\mathbb{R}^{n+1}$. On cherche un polynôme P vérifiant $P(a_0) = y_0,\ P(a_1) = y_1,\ ...P(a_n) = y_n$.

Nous en avons trouvé un à l'exercice 11 de la feuille de TD2 : si pour chaque i de $\{0, ..n\}$ on note $L_i(X) = \prod\limits_{\substack{j=0..n \\ j\neq i}} \frac{X-a_j}{a_i-a_j}$,

on constate que $L_i(a_k) = 0$ pour tout $k \neq i$, alors que $L_i(a_i) = 1$.

Si l'on pose $P(X) = \sum\limits_{i=0}^{n} y_i L_i(X)$, on obtient donc : pour chaque k de $\{1, .., n\}$, $P(a_k) = y_k$. Ainsi $\Psi(P) = (y_1, ..., y_n)$.

• Conclusion : Ψ est injective et surjective, donc définit un isomorphisme entre $\mathbb{R}_n[X]$ et $\mathbb{R}^{n+1}$.

On notera que l'isomorphisme Ψ est très différent de l'isomorphisme Φ de l'exemple précédent !

Si P est de degré n et si on connaît $P(a_0), P(a_1), ...P(a_n)$, alors on connaît P complètement.

4.2. Résultats théoriques. —

Proposition 4.39 – Effet d'un isomorphisme sur la dimension

Soit E un espace vectoriel de dimension finie.
Si F est un espace vectoriel isomorphe à E, alors F est aussi de dimension finie et $\dim(F) = \dim(E)$.

Démonstration. — Si f est un isomorphisme entre E et F et si $(e_1, ..., e_n)$ est une base de E (avec $n = \dim(E)$), nous avons vu (lien entre bijectivité et image d'une base, page 104) que $(f(e_1), ...f(e_n))$ est une base de F. Cette famille comporte n vecteurs, donc F est de dimension $n = \dim(E)$. □

Ainsi, deux espaces (de dimension finie) ne peuvent être isomorphes que s'ils ont la même dimension. On peut préciser un peu ce constat en rassemblant les résultats des pages 103 et 104 :

Proposition 4.40 – Résultats de non-injectivité ou non-surjectivité automatique

Soient E et F des espaces vectoriels de dimension finie.
Si $\dim(E) > \dim(F)$, *aucune application linéaire de E dans F n'est injective.*
Si $\dim(E) < \dim(F)$, *aucune application linéaire de E dans F n'est surjective.*

4.3. En pratique : espaces abstraits et coordonnées. —

Le paragraphe précédent montre que deux espaces de dimension finie ne peuvent être isomorphes que s'ils ont la même dimension. Réciproquement, deux espace de même dimension (finie) sont *toujours* isomorphes.

Théorème 4.41 – Tous les espaces de dimension n sont isomorphes

Si E est un $\mathbb{K}$-espace vectoriel de dimension finie n, alors il existe un isomorphisme entre E et $\mathbb{K}^n$.

Démonstration. — soit $\mathcal{B} = (e_1, ..., e_n)$ une base de E. Considérons l'application

$$\varphi : \mathbb{K}^n \to E$$
$$(x_1, ...x_n) \mapsto x_1e_1 + ... + x_ne_n.$$

Elle est linéaire (la vérification est aisée). Vérifions que c'est un isomorphisme :

- Pour montrer l'injectivité, utilisons le fait que $\mathcal{B}$ est libre pour montrer que $\mathrm{Ker}(\varphi) = \{0\}$: si $(x_1, ..., x_n)$ est un élément de $\mathbb{K}^n$ vérifiant $\varphi(x_1, ..., x_n) = \mathbf{0}_E$, alors on a $x_1e_1 + ... + x_ne_n = \mathbf{0}_E$, et la famille $(e_1, .., e_n)$ étant libre, cela implique $(x_1, ..., x_n) = (0, ..., 0) = \mathbf{0}_{\mathbb{K}^n}$.
- Pour montrer la surjectivité, utilisons le fait que $\mathcal{B}$ est génératrice : cela signifie que pour tout x de E, il existe un n-uplet $(x_1, ...x_n)$ de scalaires vérifiant $x = x_1e_1 + ... + x_ne_n$, c'est-à-dire $\varphi(x_1, ..x_n) = x$. Tout élément x de E a donc bien un antécédent par φ.
- En conclusion, φ est bijective et c'est un isomorphisme entre $\mathbb{K}^n$ et E : φ^{-1} est un isomorphisme entre E et $\mathbb{K}^n$.

□

Ce théorème est important et mérite quelques commentaires.

1. Ainsi, un espace est de dimension n si et seulement s'il est isomorphe à $\mathbb{K}^n$. Cela donne un sens précis à la remarque « dire qu'un espace est de dimension n, c'est dire que connaître un vecteur de cet espace revient à connaître n scalaires ». La démonstration s'appuie, comme on pouvait s'en douter, sur le fait que dans un espace admettant une base $\mathcal{B}$ de cardinal n, connaître un vecteur, c'est connaître ses coordonnées dans $\mathcal{B}$.
2. Puisqu'un $\mathbb{K}$-espace vectoriel E de dimension $n \neq 0$ admet (dans le cadre de ce cours où $\mathbb{K}$ est infini) une infinité de bases différentes, on constate donc qu'*il y a une infinité d'isomorphismes différents entre E et $\mathbb{K}^n$*, donc une infinité de façons différentes de ramener la connaissance d'un vecteur de E à la connaissance de n scalaires.
   Au paragraphe précédent, nous avons par exemple croisé deux isomorphismes très différents entre $\mathbb{K}_n[X]$ et $\mathbb{K}^{n+1}$: l'un lié à la formule de Taylor, l'autre aux polynômes de Lagrange.
3. Si E et F sont deux $\mathbb{K}$-espaces vectoriels de même dimension n, il existe un isomorphisme φ entre E et $\mathbb{K}^n$ et un isomorphisme ψ entre F et $\mathbb{K}^n$; l'application $\psi^{-1} \circ \varphi$ est alors un isomorphisme entre E et F.
4. Un intérêt pratique majeur de la proposition ci-dessus est que fixer un isomorphisme entre E et $\mathbb{K}^n$ rend *complètement concrets* (donc manipulables disons par un ordinateur...) les vecteurs de E, même si E est au départ un espace qui aurait semblé abstrait (de polynômes, de matrices, d'images, de sons...). À chaque fois qu'on utilise un ordinateur pour manipuler des images ou des sons, on utilise une correspondance entre l'espace d'objets « naturels » et l'espace « accessible au calcul » des structures de données qui les encodent sur ordinateur.
5. Dans le même ordre d'idées, si un $\mathbb{R}$-espace vectoriel E est de dimension 3 mais de description un peu abstraite, fixer un isomorphisme entre E et $\mathbb{R}^3$ permet de faire des *dessins dans E*.

***Exemple 4.42* (Dessins dans l'espace des matrices symétriques).** —
Considérons l'espace $\mathcal{S}_2 = \{M \in \mathcal{M}_2(\mathbb{R}),\ M^T = M\} = \left\{\begin{pmatrix} a & b \\ b & c\end{pmatrix},\ (a,b,c) \in \mathbb{R}^3\right\}$ des matrices 2×2 symétriques.

- L'application
$$\Gamma : \mathcal{S}_2 \to \mathbb{R}^3$$
$$\begin{pmatrix} a & b \\ b & c\end{pmatrix} \mapsto \begin{pmatrix} a \\ b \\ c\end{pmatrix}$$
est linéaire.
- Elle est injective : si $\Gamma\left[\begin{pmatrix} a & b \\ b & c\end{pmatrix}\right] = \begin{pmatrix} 0 \\ 0 \\ 0\end{pmatrix}$, alors $\begin{pmatrix} a & b \\ b & c\end{pmatrix}$ est la matrice nulle.
- Elle est aussi surjective : si $u = \begin{pmatrix} x \\ y \\ z\end{pmatrix}$ est un élément de $\mathbb{R}^3$, la matrice $\begin{pmatrix} x & y \\ y & z\end{pmatrix}$ fournit un antécédent de u.
- C'est donc que Γ est un isomorphisme.

Il permet par exemple de dessiner dans $\mathbb{R}^3$ l'ensemble des matrices *de trace 3* (la condition étant $x + z = 3$, c'est un plan), l'ensemble des matrices *de déterminant* -1 (la condition étant $xz - y^2 = -1$, c'est un hyperboloïde), l'ensemble des matrices *non-inversibles, c'est-à-dire de déterminant zéro* (la condition est $xz - y^2 = 0$, donc c'est un cône)....

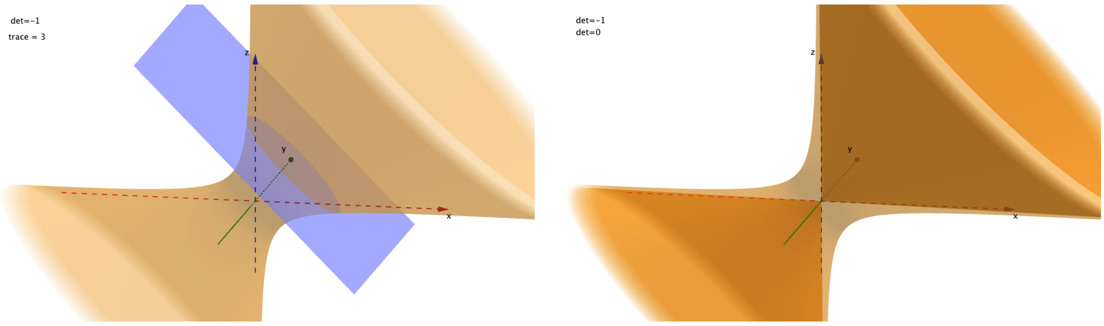


FIGURE 1. L'isomorphisme Γ révèle une correspondance inattendue entre des objets géométriques célèbres et des ensembles de matrices. Des résultats de géométrie « à l'ancienne » peuvent donc donner des renseignements sur un ensemble de matrices, et vice-versa.

5. Rang d'une application linéaire. Théorème du rang

Définition 4.43 – Rang d'une application linéaire

Soient E et F deux espaces vectoriels, f une application linéaire de E dans F.

- On dit que f est de *rang fini* si $\mathrm{Im}(f)$ est de dimension finie, sinon on dit que f est de rang infini ;
- si f est de rang fini, on appelle rang de f et on note $\mathbf{rg}[f]$ l'entier $\dim[\mathrm{Im}(f)]$.

Indiquons la cohérence avec la notion de rang d'une famille de vecteurs ou d'une matrice, vue au chapitre 2 (page 61) :
si $(x_i)_{i\in I}$ est une base de E, alors on a $\mathrm{Im}(f) = \mathbf{Vect}[(f(x_i))_{i\in I}]$, et $\mathbf{rg}[f]$ apparaît comme le rang de la famille $(f(x_i))_{i\in I}$.

En particulier, si f est l'application linéaire $T_A : \mathbb{K}^p \to \mathbb{K}^n$ associée à une matrice A de $\mathcal{M}_{np}(\mathbb{K})$ et si $(e_1, ..., e_p)$ est la base canonique de $\mathbb{K}^p$, on obtient : $\mathbf{rg}[T_A] = \dim \mathbf{Vect}[Ae_1, ..., Ae_n]$; mais pour chaque j de $\{1, .., p\}$, Ae_j est la j^e colonne C_j de A, si bien que $\mathbf{rg}[T_A]$ est la dimension de $\mathbf{Vect}[C_1, ..., C_p]$, ce qui est la définition que nous avions donnée au chapitre 2 pour $\mathbf{rg}[A]$.

Proposition 4.44 – Effet d'une application linéaire sur la dimension

Soient E et F deux espaces vectoriels, f une application linéaire de E dans F.

(a) ***Le rang est inférieur à la dimension de l'espace-but :***
Si F est de dimension finie, alors f est de rang fini et $\mathbf{rg}[f] \leq \dim(F)$.
Dans ce cas, f est surjective si et seulement si $\mathbf{rg}[f] = \dim(F)$.

(b) ***Une application linéaire ne peut « augmenter la dimension » et ne la préserve qu'en cas d'injectivité :***
si E est de dimension finie, alors $\mathbf{rg}[f] \leq \dim(E)$.
Dans ce cas, f est injective si et seulement si $\mathbf{rg}[f] = \dim(E)$.

Démonstration. —

- La première partie est évidente, puisque si F est de dimension finie, $\mathrm{Im}(f)$ est un sous-espace vectoriel de F, donc $\dim[\mathrm{Im}(f)] \leq \dim(F)$ avec égalité si et seulement si $\mathrm{Im}(f) = F$ (voir page **??**).
- C'est la seconde partie qui n'est pas évidente : montrons-la. Si E est de dimension n et si $(e_1, ..., e_n)$ est une base de E, alors $\mathrm{Im}(f) = f(\mathbf{Vect}[e_1, ..., e_n]) = \mathbf{Vect}[f(e_1), ...f(e_n)]$, donc $\mathrm{Im}(f)$ admet une famille génératrice de cardinal n et $\dim[\mathrm{Im}(f)] \leq n = \dim(E)$. Par ailleurs il y a égalité si et seulement si $(f(e_1), ...f(e_n))$ est une base de $\mathrm{Im}(f)$; dans ce cas f envoie la base $(e_1, ...e_n)$ de E sur une base de F, ce qui arrive si et seulement si f est un isomorphisme entre E et $\mathrm{Im}(f)$ (voir la proposition 4.34). Mais dire que f est un isomorphisme entre E et $\mathrm{Im}(f)$, c'est exactement dire que f est injective (pour la surjectivité il n'y a rien à vérifier).

□

Exemple 4.45. — si A est une matrice de $\mathcal{M}_{np}(\mathbb{K})$, le rang de l'application $T_A \in \mathcal{L}(\mathbb{R}^p, \mathbb{R}^n)$ est inférieur ou égal à n **et** à p, donc à $\min(n, p)$. Une matrice 12×7 a donc un rang compris entre 0 et 7.

La deuxième partie de la proposition précédente dit que lorsque E est de dimension finie, le rang est « maximal » parmi les rangs possibles lorsque f est injective, c'est-à-dire lorsque la dimension du noyau est nulle, « minimale » parmi les dimensions possibles. Autrement dit,

lorsque l'image est de dimension maximale (parmi les possibles),
le noyau est de dimension minimale (parmi les possibles).

On peut aller beaucoup plus loin : il y a un lien étroit entre la dimension du noyau et celle de l'image. C'est l'objet du théorème suivant :

Théorème 4.46 – Le théorème du rang

Soit E un espace vectoriel de dimension finie, F un espace vectoriel, f une application linéaire de E dans F.

$$\dim(E) = \dim[\mathrm{Ker}(f)] + \dim[\mathrm{Im}(f)]$$

Démonstration. — Nous allons nous appuyer sur la remarque suivante :

$$\text{Si } S \text{ est un supplémentaire de } \mathrm{Ker}(f) \text{ dans } E, \text{ alors } \dim(S) = \dim[\mathrm{Im}(f)].$$

Comment déduire le théorème de la remarque : Fixons un supplémentaire arbitraire S de $\mathrm{Ker}(f)$ dans E (on sait qu'il en existe, voir page **??**). On a alors $E = \mathrm{Ker}(f) \oplus S$, et donc $\dim(E) = \dim[\mathrm{Ker}(f)] + \dim(S)$. Si la remarque est vraie, on a $\dim(S) = \dim[\mathrm{Im}(f)]$, ce qui donne bien sûr $\dim(E) = \dim[\mathrm{Ker}(f)] + \dim[\mathrm{Im}(f)]$.

Démonstration de la remarque : soit S un supplémentaire de $\mathrm{Ker}(f)$ dans E. Considérons la restriction de f à S, et puisque $f(S) \subset f(E) = \mathrm{Im}(f)$, observons l'application

$$\begin{aligned} g : S &\to \mathrm{Im}(f) \\ x &\mapsto f(x). \end{aligned}$$

Nous allons montrer que g est un isomorphisme entre S et $\mathrm{Im}(f)$.

• Vérifions que g est injective en vérifiant que si x est un vecteur de S tel que $g(x) = 0$, alors $x = 0$. Si x appartient à S et $g(x) = 0$, alors puisque $g(x) = f(x)$ est nul, x appartient à la fois à S et à $\mathrm{Ker}(f)$, donc il appartient à $\mathrm{Ker}(f) \cap S$, qui est $\{0\}$ puisque $\mathrm{Ker}(f)$ et S sont en somme directe. Ainsi $x = 0$, comme voulu.

• Vérifions que g est surjective : si y est un élément de $\mathrm{Im}(f)$, on veut vérifier qu'il existe un x dans S vérifiant $g(x) = y$. Mais on sait que y appartient à $\mathrm{Im}(f) = f(E)$, donc il existe $\tilde{x}$ dans E tel que $g(\tilde{x}) = y$.

Puisque $E = \mathrm{Ker}(f) \oplus S$, il existe un x *de* S et un x' *de* $\mathrm{Ker}(f)$ vérifiant $\tilde{x} = x' + x$; mais alors $y = g(\tilde{x}) = g(x') + g(x) = 0 + g(x) = g(x)$, cqfd.

$\rightsquigarrow$ Conclusion : g est un isomorphisme entre S et $\mathrm{Im}(f)$, donc $\dim(S) = \dim[\mathrm{Im}(f)]$ et la remarque est démontrée. □

Attention. — Retenir « dimension *de l'espace de départ* = dimension du noyau + dimension de l'image ». Lorsque l'espace de départ et l'espace d'arrivée sont de dimensions différentes, gare à ne pas utiliser la mauvaise !

Exemple 4.47. — Soit A la matrice $\begin{pmatrix} 2 & 3 & 7 & -1 \\ 4 & 1 & 9 & 3 \end{pmatrix}$. L'application linéaire T_A va de $\mathbb{R}^4$ dans $\mathbb{R}^2$. Le théorème du rang donne

$$\dim(\mathbb{R}^4) = \dim[\mathrm{Ker}(T_A)] + \dim[\mathrm{Im}(T_A)]\,.$$

Or, $\mathrm{Im}(T_A)$ est un sous-espace vectoriel de $\mathbb{R}^2$: c'est le sous-espace $\mathbf{Vect}\left[\begin{pmatrix} 2 \\ 4 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 7 \\ 9 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \end{pmatrix}\right]$ de $\mathbb{R}^2$.

Il est de dimension 0, 1 ou 2, et comme les quatre colonnes de A ne sont pas toutes colinéaires, on sait qu'il est de dimension exactement 2.

D'après le théorème du rang, on *doit* donc avoir $4 = \dim[\mathrm{Ker}(T_A)] + 2$, donc $\dim[\mathrm{Ker}(T_A)] = 2$.
Ainsi le noyau $\mathrm{Ker}(A)$, qui est un sous-espace vectoriel de $\mathbb{R}^4$ est de dimension 2.

On peut aller plus loin et trouver concrètement $\mathrm{Ker}(A)$: rappelons qu'on peut trouver des vecteurs dans le noyau d'une matrice en observant les relations éventuelles entre ses colonnes (voir page 99). Ici on constate que $\begin{pmatrix} 7 \\ 9 \end{pmatrix} = 2\begin{pmatrix} 2 \\ 4 \end{pmatrix} + \begin{pmatrix} 3 \\ 1 \end{pmatrix}$, donc $u = \begin{pmatrix} 2 \\ 1 \\ -1 \\ 0 \end{pmatrix}$ appartient à $\mathrm{Ker}(A)$; on constate aussi que $\begin{pmatrix} -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} - \begin{pmatrix} 3 \\ 1 \end{pmatrix}$, donc $v = \begin{pmatrix} 1 \\ -1 \\ 0 \\ -1 \end{pmatrix}$ appartient aussi à $\mathrm{Ker}(A)$.

Puisque u et v appartiennent au sous-espace $\mathrm{Ker}(A)$, on a $\mathbf{Vect}[u, v] \subset \mathrm{Ker}(A)$; comme u et v ne sont pas colinéaires $\mathbf{Vect}[u, v]$ est de dimension 2, et puiqu'on a vu que $\mathrm{Ker}(A)$ est aussi de dimension 2, on obtient $\mathrm{Ker}(A) = \mathbf{Vect}[u, v]$, ce qui décrit complètement $\mathrm{Ker}(A)$.

Proposition 4.48 – Entre espaces *de même dimension*, injectivité $\iff$ surjectivité

Soient E et F deux espaces vectoriels de dimension finie.
On suppose que les dimensions $\dim(E)$ et $\dim(F)$ sont ÉGALES.
Alors si f est une application linéaire *de E dans F, les trois assertions suivantes sont équivalentes :*

- *f est injective*
- *f est surjective*
- *f est bijective*

Démonstration. — L'application f est injective si et seulement si $\mathrm{Ker}(f) = \{0\}$, et d'après le théorème du rang cela équivaut à $\dim[\mathrm{Im}(f)] = \dim(E)$; puisque E et F ont même dimension cela équivaut à $\dim[\mathrm{Im}(f)] = \dim(F)$, c'est-à-dire à la surjectivité de f. □

Ce résultat est essentiel en pratique : si on sait déjà que les espaces de départ et d'arrivée ont même dimension, on peut montrer qu'une application linéaire est bijective sans avoir à montrer la surjectivité.

Exemple 4.49. — Considérons l'application

$$\begin{aligned}\Phi : \mathbb{R}_3[X] &\to \mathbb{R}^4\\ P &\mapsto (P(0), P'(0), P(1), P'(1))\end{aligned}$$

et montrons que c'est un isomorphisme. Puisque les espaces $\mathbb{R}_3[X]$ et $\mathbb{R}^4$ sont tous les deux de dimension quatre, le résultat que nous venons de montrer dit qu'il suffit de vérifier l'injectivité.

Or, si P est un élément de $\mathbb{R}_3[X]$ vérifiant $\Phi(P) = 0$, alors $P(0) = P'(0) = 0$, donc 0 est racine double de P, et $P(1) = P'(1) = 0$, donc 1 est aussi racine double de P. Le polynôme P est de degré ≤ 3 et admet au moins deux racines doubles, donc au moins quatre racines si on les compte avec multiplicité. Ce n'est possible que si P est le polynôme nul.

Conclusion : $\mathrm{Ker}(\Phi) = \{0\}$, donc Φ est injective et le résultat précédent dit qu'elle est bijective : c'est un isomorphisme.

Attention, ce « miracle » reliant l'injectivité à la surjectivité ne vaut qu'en dimension finie

Par exemple, si E est l'espace $\mathcal{F}(\mathbb{N}, \mathbb{R})$ des suites réelles et si on considère l'application

$$\begin{aligned}\sigma : E &\to E\\ (u_0, u_1, u_2, ...) &\mapsto (u_1, u_2, u_3, ...)\end{aligned}$$

alors σ est surjective (en effet, si $v = (v_n)_{n\in\mathbb{N}} = (v_0, v_1, ...)$ est une suite réelle, la suite $u = (0, v_0, v_1, ...)$ vérifie $\sigma(u) = v$) ;

mais σ n'est pas injective : si δ est la suite $(1, 0, 0, ...)$, $\sigma(\delta)$ est la suite nulle alors que δ n'est pas la suite nulle.

Attention, ce « miracle » ne concerne que les applications linéaires

Par exemple, l'application $\exp : \mathbb{R} \to \mathbb{R}$ est injective sans être surjective...

6. Projecteurs

6.1. Définition et premiers exemples. —

Définition 4.50 – Projection sur un sous-espace parallèlement à un supplémentaire

Soient E un espace vectoriel, F et S deux sous-espaces vectoriels *supplémentaires* de E. Rappelons que pour tout x de E, il existe un *unique* couple (x_F, x_S) de $F \times S$ vérifiant $x = x_F + x_S$. La *projection sur F parallèlement à S* est l'application

$$\begin{aligned} E &\to E \\ x &\mapsto x_F. \end{aligned}$$

C'est un endomorphisme de E.

Exemple 4.51. — Dans $\mathbb{R}^3$, considérons le plan $F = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3,\ z = 0 \right\}$.

- La droite $S_1 = \mathbf{Vect}\left[\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right]$ est un supplémentaire de F (voir le chapitre 3). Notons π_1 la projection sur F parallèlement à S_1. Alors on a
$$\pi_1 = \left[\begin{pmatrix} 0.2 \\ -0.1 \\ 3 \end{pmatrix}\right] = \begin{pmatrix} 0.2 \\ -0.1 \\ 0 \end{pmatrix}, \quad \text{car} \quad \begin{pmatrix} 0.2 \\ -0.1 \\ 3 \end{pmatrix} = \underbrace{\begin{pmatrix} 0.2 \\ -0.1 \\ 0 \end{pmatrix}}_{\in F} + \underbrace{\begin{pmatrix} 0 \\ -0 \\ 3 \end{pmatrix}}_{\in S_1}.$$
Plus généralement, pour tout $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ de $\mathbb{R}^3$, on a $\pi_1 \left[\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right] = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix}$.
- Cela dit, la droite $S_2 = \mathbf{Vect}\left[\begin{pmatrix} -2 \\ 0 \\ 3 \end{pmatrix}\right]$ est un autre supplémentaire de F. Notons π_2 la projection sur F parallèlement à S_1. Alors on a
$$\pi_2 \left[\begin{pmatrix} 0.2 \\ -0.1 \\ 3 \end{pmatrix}\right] = \begin{pmatrix} 2.2 \\ -0.1 \\ 0 \end{pmatrix}, \quad \text{car} \quad \begin{pmatrix} 0.2 \\ -0.1 \\ 3 \end{pmatrix} = \underbrace{\begin{pmatrix} 2.2 \\ -0.1 \\ 0 \end{pmatrix}}_{\in F} + \underbrace{\begin{pmatrix} -2 \\ 0 \\ 3 \end{pmatrix}}_{\in S_2}.$$
Plus généralement, pour tout $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ de $\mathbb{R}^3$, on a $\pi_2 \left[\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right] = \begin{pmatrix} x - 2z/3 \\ y \\ 0 \end{pmatrix} + \begin{pmatrix} 2z/3 \\ 0 \\ z \end{pmatrix}$.

***Exemple 4.52* (Parties paire et impaire d'une fonction).** — Rappelons que dans $\mathcal{F}(\mathbb{R}, \mathbb{R})$, les sous-espaces $\mathcal{P}$ et $\mathcal{I}$ des fonctions paires et impaires sont supplémentaires (voir page 79). Notons p la projection sur $\mathcal{P}$ parallèlement à $\mathcal{I}$. Alors on a [(4)]

$$p(\exp) = \cosh, \qquad \text{car} \quad \exp = \underbrace{\cosh}_{\text{paire}} + \underbrace{\sinh}_{\text{impaire}}.$$

4. On rappelle que $\cosh(x) = \dfrac{e^x + e^{-x}}{2}$ et $\sinh(x) = \dfrac{e^x - e^{-x}}{2}$ pour tout x de $\mathbb{R}$.

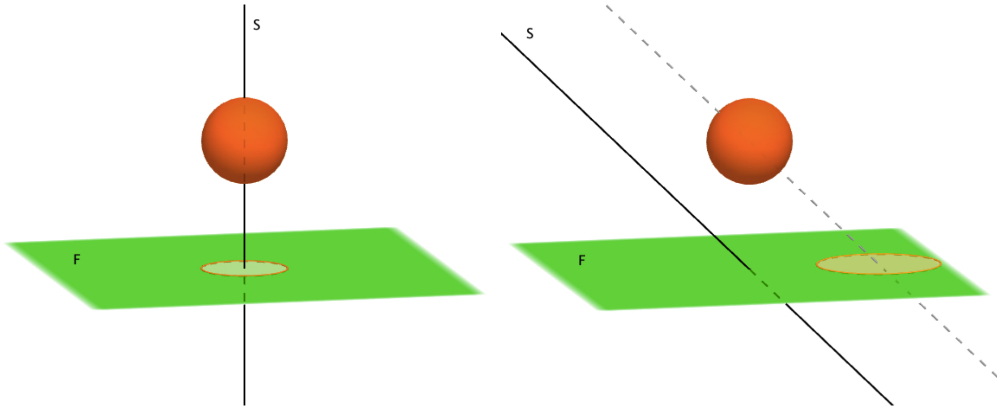


FIGURE 2. Un plan F de $\mathbb{R}^3$, deux supplémentaires S différents et ce que devient la boule centrée en $(0,0,3)$ et de rayon 1 lorsqu'on la projette sur F parallèlement à chacun de ces deux supplémentaires.

Si f est une fonction de $\mathbb{R}$ dans $\mathbb{R}$, on appelle souvent $p(f)$ la *partie paire de* f.

On a vu page 79 qu'il est possible de donner une expression explicite de $p(f)$: on a

$$\forall x \in \mathbb{R}, f(x) = \frac{f(x)+f(-x)}{2} + \frac{f(x)-f(-x)}{2}$$

et les fonctions $x \mapsto \frac{f(x)+f(-x)}{2}$ et $x \mapsto \frac{f(x)-f(-x)}{2}$ sont respectivement paire et impaire ; ainsi, $p(f)$ n'est autre que la fonction $x \mapsto \frac{f(x)+f(-x)}{2}$.

Exemple 4.53 **(Deux projections différentes sur l'espace des polynômes de degré 1)**

Dans l'espace $\mathbb{R}_4[X]$,

- les sous-espaces $\mathcal{P}_2 = \mathbf{Vect}[1, X]$ et $\mathcal{R} = \mathbf{Vect}\left[X^2, X^3, X^4\right]$ sont supplémentaires. La projection sur $\mathcal{P}_2$ parallèlement à $\mathcal{R}$ est l'application qui, à un polynôme $aX^4 + bX^3 + cX^2 + dX + e$ de $\mathbb{R}_4[X]$, associe le polynôme $dX + e$.
- les sous-espaces $\mathcal{P}_2 = \mathbf{Vect}[1, X]$ et $\mathcal{R} = \mathbf{Vect}\left[1 + X^2, X + X^3, X^4\right]$ sont supplémentaires (savez-vous le montrer ?). Si on note π la projection sur $\mathcal{P}_2 = \mathbf{Vect}[1, X]$ parallèlement à $\mathcal{R}$, on a par exemple $\pi(1 + X^2 + X^3) = -X$,
car la décomposition de $(1+X^2+X^3)$ selon $\mathcal{P}_2 \oplus \mathcal{R}$ est $(1+X^2+X^3) = \underbrace{[-X]}_{\in \mathcal{P}_2} + \underbrace{\left[(1+X^2) + (X+X^3)\right]}_{\in \mathcal{R}}$.

Proposition 4.54 – Le noyau = le supplémentaire, l'image = ce sur quoi on projette

Soient E un espace vectoriel, F et S deux sous-espaces vectoriels supplémentaires *de E, p la projection sur F parallèlement à S. On a*
- $\mathrm{Ker}(p) = S$
- $\mathrm{Im}(p) = F$
- *Un vecteur x appartient à $\mathrm{Im}(p)$ si et seulement si $p(x) = x$.*

Démonstration. — Soit x un vecteur de E. Il existe un unique couple (x_F, x_S) vérifiant : $x_F \in F$, $x_S \in S$ et $x = x_F + x_S$.

- Dire que $p(x) = 0$, c'est dire que $x_F = 0$, ou encore $x = x_S$, ce qui équivaut à $x \in S$.

- Visiblement $\mathrm{Im}(p) \subset F$, puisque $p(x)$ appartient à F pour tout x de E. Réciproquement, si x appartient à F, l'unique décomposition de x en somme d'un vecteur de F et d'un vecteur de S est $x = x+0$, donc $p(x) = x$ et x admet un antécédent par p (lui-même), donc x appartient à $\mathrm{Im}(p)$
- Vu le point précédent, il suffit de montrer que x appartient à F si et seulement si $p(x) = x$. Mais dire que $p(x) = x$, c'est dire que $x = x_F$, ce qui équivaut à $x \in F$.

□

Vous pouvez reprendre les expressions explicites trouvées dans les trois exemples ci-dessus et vous convaincre que cette proposition rend bien compte de la situation.

Il est important de retenir la dernière partie de la proposition précédente :

Pour un projecteur, le sous-espace image est l'ensemble des points fixes.

Proposition 4.55 – Un projecteur est idempotent

Soient E un espace vectoriel, F et S deux sous-espaces vectoriels supplémentaires *de E. Notons p la* projection sur F parallèlement à S. *Alors on a l'égalité d'endomorphismes*

$$p^2 = p$$

(qui signifie : $\forall x \in E, p(p(x)) = p(x)$.)

6.2. Caractérisation algébrique des projecteurs. —

Proposition 4.56 – Caractérisation algébrique des projecteurs

Soit E un espace vectoriel. Si

$$\begin{cases} p \text{ est une application linéaire de } E \text{ dans } E \\ p^2 = p, \end{cases}$$

alors il existe deux sous-espaces supplémentaires F et S vérifiant : p est la projection sur F parallèlement à S.

Vocabulaire. — On dit souvent qu'une application de E dans E est un *projecteur* lorsqu'il existe il existe deux sous-espaces supplémentaires F et S vérifiant : p est la projection sur F parallèlement à S. On peut donc résumer les deux propositions qu'on vient de lire de la manière suivante :

$$p \text{ est un projecteur} \iff p \text{ est linéaire et } p^2 = p.$$

Proposition 4.57 – Comment retrouver les sous-espaces F et S

i $p \in \mathcal{L}(E)$ est un projecteur, alors
- $\mathrm{Ker}(p)$ *et* $\mathrm{Im}(p)$ *sont supplémentaires dans E*
- *p est la projection sur* $\mathrm{Im}(p) = \{x \in E, p(x) = x\}$ *parallèlement à* $\mathrm{Ker}(p)$.

***Exemple 4.58* (Exemple avec un endomorphisme de $\mathbb{R}^3$).** — Considérons la matrice $A = \frac{1}{2}\begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix}$ et l'application linéaire T_A de $\mathbb{R}^3$ dans $\mathbb{R}^3$.

- La transformation T_A vérifie : $T_A^2 = T_A \circ T_A = T_{A^2}$ (voir page 95). Or, on constate l'égalité $A^2 = A$ (calcul !).
- Par conséquent, l'application linéaire T_A est un projecteur. Pour la caractériser géométriquement, la proposition précédente dit que nous devons trouver son noyau et l'ensemble de ses points fixes.
- Soit $u = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ un vecteur de $\mathbb{R}^2$. On constate que

$$T_A u = u \text{ équivaut à } \begin{cases} x - y + z = 2x \\ -x + y + z = 2y \\ 2z = 2z \end{cases} \quad \text{ou encore à :} z = x + y$$

tandis que

$$T_A u = \mathbf{0}_{\mathbb{R}^3} \text{ équivaut à } \begin{cases} x - y + z = 0 \\ -x + y + z = 0 \\ z = 0 \end{cases} \quad \text{ou encore à :} \begin{cases} x = y \\ z = 0 \end{cases}, \text{ c'est-à-dire à } u \in \mathbf{Vect}\left[\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\right].$$

- Ainsi, on sait que l'ensemble des points fixes de T_A est le plan $\mathcal{P}$ d'équation $z = x + y$ et que le noyau de T_A est la droite $\mathcal{D} = \mathbf{Vect}\left[\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\right]$. D'après la proposition précédente, T_A est la projection sur $\mathcal{P}$ parallèlement à $\mathcal{D}$ (et en plus, on a automatiquement $\mathrm{Im}(T_A) = \mathcal{P}$).

***Exemple 4.59* (Matrices symétriques et antisymétriques).** — Plaçons-nous dans l'espace $E = \mathcal{M}_n(\mathbb{R})$ et considérons l'application

$$\begin{aligned} \sigma : \mathcal{M}_n(\mathbb{R}) &\to \mathcal{M}_n(\mathbb{R}) \\ M &\mapsto \frac{M + M^T}{2}. \end{aligned}$$

- C'est un endomorphisme de E (car la transposition l'est et qu'une combinaison linéaire d'applications linéaires est linéaire) ; pour toute matrice M de E, on a :

$$\sigma(\sigma(M)) = \sigma\left(\frac{M + M^T}{2}\right) = \frac{\frac{M + M^T}{2} + \left(\frac{M + M^T}{2}\right)^T}{2} = \frac{(M + M^T) + (M^T + M)}{4} = \frac{2M + 2M^T}{2} = \frac{M + M^T}{2}$$

si bien que $\sigma^2(M) = \sigma(M)$. Ainsi $\sigma^2 = \sigma$ et σ est un projecteur.
- Sur quoi et dans quelle direction ? Observons les points fixes et le noyau de σ.
L'équation $\sigma(M) = M$ équivaut à $\frac{M+M^T}{2} = M$, c'est-à-dire à $M + M^T = 2M$, autrement dit $M^T = M$. Ainsi $\sigma(M) = M$ si et seulement si la matrice M est symétrique.
Par ailleurs, $\sigma(M) = 0$ équivaut à $\frac{M+M^T}{2} = 0$ ou encore à $M^T = -M$. Ainsi $\mathrm{Ker}(\sigma)$ est l'ensemble des matrices antisymétriques.
- Conclusion : σ est la projection sur l'espace des matrices symétriques parallèlement à celui des matrices antisymétriques.

***Exemple 4.60* (Développement de Taylor à l'ordre 2).** — Plaçons-nous dans $\mathbb{R}[X]$ et considérons l'application

$$\tau : \mathbb{R}[X] \to \mathbb{R}[X]$$
$$P \mapsto P(7) + P'(7)(X-7) + \frac{P''(7)}{2}(X-7)^2.$$

- Montrons que τ est un projecteur. Soiti P est un polynôme de $\mathbb{R}[X]$; nous devons montrer que $\tau(\tau(P)) = \tau(P)$. Notons $Q = \tau(P)$; nous devons montrer que $\tau(Q) = Q$, c'est-à-dire qu'on a $Q(X) = Q(7) + Q'(7)(X-7) + \dfrac{Q''(7)}{2}(X-7)^2$. Or, Q est de degré inférieur ou égal à 2, puisque $Q = \tau(P) = P(7) + P'(7)(X-7) + \dfrac{P''(7)}{2}(X-7)^2$. La formule de Taylor pour le polynôme Q dit alors exactement que $Q(X) = Q(7) + Q'(7)(X-7) + \dfrac{Q''(7)}{2}(X-7)^2$, comme voulu.

  Ainsi $\tau^2 = \tau$ et τ est un projecteur
- Sur quoi, dans quelle direction ? Quels sont l'ensemble des points fixes et le noyau de τ ?

  Dire que $\tau(P) = P$, c'est dire qu'on a $P(X) = P(7) + P'(7)(X-7) + \dfrac{P''(7)}{2}(X-7)^2$: c'est équivalent au fait que P soit de degré inférieur ou égal à 2 (en effet, si c'est le cas on a visiblement $\deg(P) \leq 2$, et si $\deg(P) \leq 2$ la formule de Taylor garantit $\tau(P) = P$ comme précédemment).

  Dire que $\tau(P) = 0$, c'est dire que $P(7) = P'(7) = P''(7) = 0$, autrement dit que 7 est racine (au moins) triple de P, autrement dit que $(X-7)^3$ divise P.
- Conclusion : τ est la projection sur l'espace des polynômes de degré inférieur ou égal à 2, parallèlement à l'espace des polynômes divisibles par $(X-7)^3$.

7. Formes linéaires, hyperplans, dualité

7.1. Formes linéaires, espace dual. —

Définition 4.61 – Forme linéaire, espace dual

Soit E un $\mathbb{K}$-espace vectoriel.
Une *forme linéaire sur* E est une application linéaire de E dans $\mathbb{K}$.
L'ensemble des formes linéaires sur E est appelé *dual de* E ; on le note souvent $E^\star$ plutôt que $\mathcal{L}(E, \mathbb{K})$.

Rappelons que $\mathcal{L}(E, \mathbb{K})$ est un $\mathbb{K}$-espace vectoriel.
Pour chaque $\mathbb{K}$-espace vectoriel E, on dispose donc d'un « autre » $\mathbb{K}$-espace vectoriel, son dual.

Exemple 4.62. — Si $E = \mathbb{R}^3$ et si (a, b, c) est un triplet de nombres réels, alors l'application

$$\begin{aligned} \varphi_{a,b,c} : \mathbb{R}^3 &\to \mathbb{R} \\ (x, y, z) &\mapsto ax + by + cz \end{aligned}$$

est une forme linéaire sur E. Rappelons par ailleurs que toute forme linéaire sur $\mathbb{R}^3$ peut s'écrire $\varphi_{a,b,c}$ pour un certain (a, b, c) de $\mathbb{R}^3$ (voir page **??**).

Exemple 4.63. — De même si $E = \mathbb{R}^n$ et si $\vec{a} = (a_1, ..., a_n)$ est un n-uplet de nombres réels, alors l'application

$$\begin{aligned} \varphi_{\vec{a}} : \mathbb{R}^n &\to \mathbb{R} \\ (x_1, .., x_n) &\mapsto a_1x_1 + ... + a_nx_n \end{aligned}$$

est une forme linéaire sur E. Grâce au résultat page **??**, tout élément de $E^\star$ est de la forme $\varphi_{\vec{a}}$ pour un certain a de $\mathbb{R}^n$.

Exemple 4.64. — Si $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$ et si a est un nombre réel, alors l'application

$$\begin{aligned} \mathrm{ev}_a : E &\to \mathbb{R} \\ f &\mapsto f(a) \end{aligned}$$

est une forme linéaire sur E.
Si $E = \mathcal{C}(\mathbb{R}, \mathbb{R})$, alors l'application

$$\begin{aligned} \mathrm{Int}_{[0,1]} : E &\to \mathbb{R} \\ f &\mapsto \int_0^1 f(t)dt \end{aligned}$$

est une forme linéaire sur E.

Exemple 4.65. — Si $E = \mathbb{R}[X]$, l'application

$$\begin{aligned} \mathbb{R}[X] &\to \mathbb{R} \\ P &\mapsto P(1) - P(2) \end{aligned}$$

est une forme linéaire sur E (en effet, c'est la différence des formes linéaires $\mathrm{ev}_1 : P \mapsto P(1)$ et $\mathrm{ev}_2 : P \mapsto P(2)$ sur E.

Proposition 4.66 – Surjectivité automatique pour les formes linéaires non nulles

Si une forme linéaire $\varphi : E \to \mathbb{K}$ est non nulle, alors φ est surjective.

Démonstration. — L'espace $\mathrm{Im}(\varphi)$ est un sous-espace vectoriel du $\mathbb{K}$-espace vectoriel $\mathbb{K}$ et ce dernier est de dimension 1, donc $\mathrm{Im}(\varphi)$ est de dimension 0 ou 1, donc égal soit à $\{0\}$ soit à $\mathbb{K}$ tout entier. Dire que φ est non nulle, c'est dire que $\mathrm{Im}(\varphi) \neq \{0\}$, ainsi $\mathrm{Im}(\varphi) = \mathbb{K}$ et φ est surjective. ☐

7.2. Hyperplans. —
Rappelons que si a, b, c sont trois réels et qu'ils ne sont pas tous nuls, l'ensemble $\mathcal{P} = \left\{(x,y,z) \in \mathbb{R}^3,\ ax+by+cz = 0\right\}$ est un plan de $\mathbb{R}^3$. On peut reformuler l'équation de $\mathcal{P}$ à l'aide d'une forme linéaire :

$$\mathcal{P} = \mathrm{Ker}(\varphi),\ \text{où } \varphi \text{ est la forme linéaire } (x,y,z) \mapsto ax+by+cz \text{ sur } \mathbb{R}^3.$$

Cela motive la définition suivante, qui donne un nom géométriquement suggestif à l'« ensemble des solutions d'une seule équation linéaire ».

Définition 4.67 – Hyperplan dans un espace vectoriel quelconque

Soit E un espace vectoriel. On dit qu'une partie H de E est un *hyperplan de E* lorsqu'il existe une *forme linéaire $\varphi \in E^\star$ non nulle* qui vérifie :

$$H = \mathrm{Ker}(\varphi).$$

Attention. — Si $\varphi = \mathbf{0}_{E^\star}$ est la forme linéaire nulle, l'espace $E = \mathrm{Ker}(\mathbf{0}_{E^\star})$ ne mérite pas d'être appelé un hyperplan
(pas plus que l'espace $\mathbb{R}^3$ tout entier ne mérite d'être appelé un plan).

Exemple 4.68. — Reprenons les exemples de formes linéaires des exemples du §7.1

(a) Un hyperplan de $\mathbb{R}^3$ n'est rien d'autre qu'un plan au sens usuel (passant par l'origine).

(b) Dans $\mathbb{R}^n$, un hyperplan est un sous-ensemble de la forme $\{(x_1, ..., x_n), a_1x_1 + ... + a_nx_n = 0\}$ où $a_1, ..., a_n$ sont des nombres réels non tous nuls.

(c) Dans $\mathcal{F}(\mathbb{R}, \mathbb{R})$, si a est un nombre réel, alors $\{f \in \mathcal{F}(\mathbb{R}, \mathbb{R}),\ f(a) = 0\}$ est un hyperplan.

(d) Dans $\mathcal{C}(\mathbb{R}, \mathbb{R})$, l'ensemble des fonctions vérifiant $\int_0^1 f(t)dt = 0$ est un hyperplan.

(e) Dans $\mathbb{R}[X]$, l'ensemble des polynômes P vérifiant $P(1) = P(2)$ est un hyperplan.

Pour se faire une idée plus concrète de ce qu'est un hyperplan, le résultat qui vient est très utile.

Proposition 4.69 – Supplémentaires et dimensions des hyperplans

Soient E un espace vectoriel et H un sous-espace vectoriel de E. Alors

$$H \text{ est un hyperplan} \iff \text{il existe un vecteur non nul } u \text{ de } E \text{ vérifiant : } E = H \oplus \textbf{Vect}[u].$$

Corollaire 4.70 – Hyperplan : cas de la dimension finie

Si E est un espace de dimension finie n, alors

$$H \text{ est un hyperplan} \iff \dim(H) = n - 1.$$

Démonstration de la proposition. —

- Soit H un hyperplan de E : c'est le noyau d'une forme linéaire non nulle φ sur E. Si u est un vecteur n'appartenant pas à H (il est en particulier non-nul), alors $\mathbf{Vect}[u]$ et H sont en somme directe (nous l'avons vu page **??**, exemple 1(c)). De plus, pour tout x de E, on a $x = (x - \varphi(x)u) + \varphi(x)u$, donc x appartient à $H + \mathbf{Vect}[u]$, si bien que $H + \mathbf{Vect}[u] = E$. Nous avons montré : $E = H \oplus \mathbf{Vect}[u]$.
- Réciproquement, s'il existe un vecteur non nul u de E vérifiant : $E = H \oplus \mathbf{Vect}[u]$, pour chaque x de E, il existe un unique couple (h_x, t_x) de $H \times \mathbb{K}$ vérifiant : $x = h_x + t_x u$. Comme t_x est un scalaire entièrement déterminé par x, on peut considérer l'application $\varphi : x \mapsto t_x$ de E dans $\mathbb{K}$. Il est facile de vérifier qu'elle est linéaire [5] : par conséquent, φ est une forme linéaire sur E et puisque $\varphi(x) = 0$ si et seulement si $x = h_x$ appartient à H, on a $H = \mathrm{Ker}(\varphi)$, donc H est un hyperplan.

□

Remarque 4.71. — En dimension finie, le corollaire découle immédiatement de la proposition. Cependant, on peut aller plus vite pour montrer que si H est un hyperplan alors $\dim(H) = n - 1$. En effet, si H est un hyperplan, il existe une forme linéaire non nulle φ vérifiant $H = \mathrm{Ker}(\varphi)$; mais on a vu que dans ce cas φ est surjective et $\mathrm{Im}(\varphi)$ est de dimension 1, le théorème du rang donne alors immédiatement $\dim(H) + 1 = \dim(E)$.

Proposition 4.72 – Les hyperplans sont les plus grands sous-espaces vectoriels stricts

- *si H est un hyperplan et $u \notin H$, alors $E = H \oplus \mathbf{Vect}[u]$.*
- *si H est un hyperplan et F est un sous-espace vectoriel de E qui contient H, alors soit $F = H$, soit $F = E$.*

Démonstration. — Nous avons vu le premier point au cours de la démonstration de la proposition ci-dessus, et le second est une conséquence immédiate du premier (si F contient H et n'est pas égal à H, il contient un vecteur u qui n'appartient pas à H, et alors il contient $H + \mathbf{Vect}[u]$ qui est E tout entier). □

On pourra comparer ce qui précède aux exemples de la page 78 : le corollaire ci-dessus étend aux espaces vectoriels quelconques l'observation « si $\mathcal{P}$ est un plan de $\mathbb{R}^3$, toute droite vectorielle non contenue dans $\mathcal{P}$ en donne un supplémentaire ».

Proposition 4.73 – Toutes les équations d'un hyperplan donné sont proportionnelles

Si φ et ψ sont deux formes linéaires non nulles sur E et si les hyperplans $\mathrm{Ker}(\varphi)$ et $\mathrm{Ker}(\psi)$ sont identiques,
alors il existe un scalaire λ vérifiant $\psi = \lambda\varphi$.

5. Si x et y sont deux vecteurs et α, β deux scalaires, alors $\alpha x + \beta y = (\alpha h_x + \beta h_y) + (\alpha t_x + \beta t_y)u$, mais il est aussi égal à $h_{\alpha x+\beta y} + t_{\alpha x+\beta y}u$; par unicité des décompositions des vecteurs de $H \oplus \mathbf{Vect}[u]$, on en déduit $t_{\alpha x+\beta y}u = (\alpha t_x + \beta t_y)u$, et comme u est non nul, $t_{\alpha x+\beta y} = \alpha t_x + \beta t_y$, cqfd.

Démonstration. — Soient φ et ψ sont deux formes linéaires non nulles sur E ayant le même noyau, noté H. Fixons un vecteur u arbitraire qui n'appartienne pas à H ; compte tenu de la proposition précédente, on a $E = H \oplus \mathbf{Vect}[u]$.

- Si $\varphi(u)$ valait zéro, la forme linéaire φ serait nulle [6], donc on a $\varphi(u) \neq 0$. Notons λ le réel $\dfrac{\psi(u)}{\varphi(u)}$.
- Si maintenant x est un vecteur quelconque de E et $x = h_x + t_x u$ est sa décomposition selon $H \oplus \mathbf{Vect}[u]$ (où $h_x \in H$ et $t_x \in \mathbb{R}$), on a

$$\psi(x) = \psi(h_x) + t_x\psi(u) = t_x\psi(u) = t_x\left(\lambda\varphi(u)\right) = \lambda\left(t_x\varphi(u)\right) = \lambda\left(0 + t_x\varphi(u)\right) = \lambda\varphi(x)$$

  donc $\psi = \lambda\varphi$, comme annoncé.

□

7.3. Base duale. —

Définition 4.74 – Forme linéaire extrayant une coordonnée

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, $\mathcal{B} = (e_1, ..., e_n)$ une base de E.
Pour chaque j de $\{1, .., n\}$, la *forme « $j^{\text{ème}}$ coordonnée dans $\mathcal{B}$ »* est l'application

$$\begin{aligned} e_j^\star : E &\to \mathbb{K} \\ x = \sum_{k=1}^{n} \lambda_k e_k &\mapsto \lambda_j. \end{aligned}$$

C'est une forme linéaire sur E.

Rappelons qu'écrire « $x = \sum_{k=1}^{n} \lambda_k e_k$ » signifie que les coordonnées de x dans la base $\mathcal{B}$ sont données par le n-uplet $(\lambda_1, ..., \lambda_n)$: ainsi les réels $\lambda_1, ..., \lambda_n$ sont entièrement déterminés par x, ce qui justifie la définition ci-dessus.

Exemple 4.75. — Si $E = \mathbb{R}^3$, la forme « $2^{\text{ème}}$ coordonnée dans la base canonique de E » est l'application $(x, y, z) \mapsto y$ de $\mathbb{R}^3$ dans $\mathbb{R}$.

Exemple 4.76. — Toujours si $E = \mathbb{R}^3$, la famille $\mathcal{B} = \left(\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\2\\0\end{pmatrix}, \begin{pmatrix}1\\1\\3\end{pmatrix}\right)$ est une base de E.

Comme $\begin{pmatrix}x\\y\\z\end{pmatrix} = (x - y - \frac{z}{6})\begin{pmatrix}1\\0\\0\end{pmatrix} + (y - \frac{z}{6})\begin{pmatrix}1\\2\\0\end{pmatrix} + \frac{z}{3}\begin{pmatrix}1\\1\\3\end{pmatrix}$ pour tout $\begin{pmatrix}x\\y\\z\end{pmatrix}$ de $\mathbb{R}^3$,
la forme « $1^{\text{ère}}$ coordonnée dans la base canonique de E » est l'application $(x, y, z) \mapsto x - y - \frac{z}{6}$ de $\mathbb{R}^3$ dans $\mathbb{R}$.
Si $E = \mathbb{R}_3[X]$, la forme « $2^{\text{ème}}$ coordonnée dans la base canonique de E » est l'application $P \mapsto P'(0)$ de $\mathbb{R}_3[X]$ dans $\mathbb{R}$: en effet, la formule de Taylor donne, pour chaque P de $\mathbb{R}_3[X]$, $P(X) = P(0) + P'(0)X + \frac{P''(0)}{2}X^2 + \frac{P^{(3)}(0)}{6}X^3$.

6. En effet, tout vecteur x peut s'écrire $h_x + t_x u$ avec $h_x \in H = \mathrm{Ker}(\varphi)$ et $t_x \in \mathbb{R}$; si $\varphi(u)$ était nul on aurait toujours $\varphi(x) = \varphi(h_x) + t_x\varphi(u) = 0 + 0$.

Proposition 4.77 – Notion de base duale

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie, $\mathcal{B} = (e_1, ..., e_n)$ une base de E.
La famille $\mathcal{B}^\star = (e_1^\star, ..., e_n^\star)$ est une base de l'espace dual $E^\star$.
On dit que c'est la base duale *de $\mathcal{B}$.*

Démonstration. — • Vérifions que $\mathcal{B}^\star$ est libre dans $E^\star$. Si $\lambda_1, .., \lambda_n$ sont des scalaires et si $\lambda_1 e_1^\star + ... + \lambda_n e_n^\star$ est la forme linéaire nulle, alors $[\lambda_1 e_1^\star + ... + \lambda_n e_n^\star](e_1) = 0$; or, $[\lambda_1 e_1^\star + ... + \lambda_n e_n^\star](e_1)$ vaut $\lambda_1 \cdot 1 + \lambda_2 \cdot 0 + .. + \lambda_n \cdot 0 = \lambda_1$, donc $\lambda_1 = 0$. En évaluant sur chacun des vecteurs de $\mathcal{B}$, on vérifie de même que les λ_i sont tous nuls, et $\mathcal{B}^\star$ est libre.

- Vérifions que $\mathcal{B}^\star$ est génératrice de $E^\star$. Soit φ une forme linéaire sur E. Si x est un vecteur de E, ses coordonnées $(x_1, .., x_n)$ dans la base $\mathcal{B}$ sont données par $e_1^\star(x)$, ..., $e_n^\star(x)$; ainsi

$$\varphi(x) = \varphi(x_1 e_1 + ... + x_n e_n) = x_1 \varphi(e_1) + ... + x_n \varphi(e_n) = \varphi(e_1) e_1^\star(x) + ... + \varphi(e_n) e_n^\star(x).$$

Comme cette égalité est vraie pour tout x, on obtient l'égalité de formes linéaires : $\varphi = \varphi(e_1)e_1^\star + ... + \varphi(e_n)e_n^\star$. Elle exprime φ comme combinaison linéaire des éléments de $\mathcal{B}^\star$: ainsi $\varphi \in \mathbf{Vect}[\mathcal{B}^\star]$. Nous avons montré que $E^\star = \mathbf{Vect}[\mathcal{B}^\star]$.

□

Exemple 4.78. —
- Si $E = \mathbb{R}^3$ et si $\mathcal{B}$ est la base canonique de E, alors la base duale de $\mathcal{B}$ est la famille $(e_1^\star, e_2^\star, e_3^\star)$ de formes linéaires sur E : ici $e_1^\star$ est l'application $(x, y, z) \mapsto x$, $e_2^\star$ l'application $(x, y, z) \mapsto y$ et $e_3^\star$ l'application $(x, y, z) \mapsto z$. .
- Si $E = \mathbb{R}^3$ et si $\mathcal{B}$ est la base canonique de E, alors la base duale de $\mathcal{B}$ est la famille $(e_1^\star, e_2^\star, e_3^\star)$ de formes linéaires sur E : ici $e_1^\star$ est l'application $(x, y, z) \mapsto x$, $e_2^\star$ l'application $(x, y, z) \mapsto y$ et $e_3^\star$ l'application $(x, y, z) \mapsto z$.
- Si $E = \mathbb{R}_3[X]$, et si $\mathcal{B} = (1, X, X^2, X^3)$ est la base canonique de E, la base duale $\mathcal{B}^\star$ est donnée par les formes linéaires $P \mapsto P(0)$, $P \mapsto P'(0)$, $P \mapsto \frac{P''(0)}{2}$ et $P \mapsto \frac{P^{(3)}(0)}{6}$.

Corollaire 4.79 – Dimension de l'espace dual $E^\star$ si E est de dimension finie

Si E est de dimension finie, *alors l'espace $E^\star$ est aussi de dimension finie et* $\dim(E^\star)$ *n'est autre que* $\dim(E)$.

Remarque 4.80. — On peut montrer qu'aucun analogue ne peut être vrai en dimension infinie (le « théorème d'Erdós-Kaplansky » dit que E et $E^\star$ ne sont jamais isomorphes). De manière générale, pour beaucoup d' espaces de dimension infinie « intéressants » (par exemple des espaces de fonctions...), la notion d'espace dual est, avec quelques variantes, à la fois essentielle (y compris pour les applications les plus concrètes) et délicate à manier.

7.4. ($\star$) Complément : base antéduale. —
Dans le numéro précédent, nous avons vu comment construire une base de $E^\star$ à partir d'une base de E. Toutes les bases de $E^\star$ peuvent en fait être obtenues de cette manière :

Proposition 4.81 – Notion de base antéduale

soit E un $\mathbb{K}$-espace vectoriel de dimension finie, $E^\star$ son dual, $\Phi = (\varphi_1, ..., \varphi_n)$ une base de $E^\star$.
Il existe une unique base $\mathcal{B}$ de E dont Φ soit la base duale, c'est-à-dire vérifiant $\Phi = \mathcal{B}^\star$.
On dit que c'est la base antéduale de Φ.

Exemple 4.82. — fixons un entier $n \geq 1$, plaçons-nous dans $E = \mathbb{R}_n[X]$. Fixons $n+1$ réels $a_0, .., a_n$ deux à deux distincts. Pour chaque i de $\{0, .., n\}$, notons ev_{a_i} la forme linéaire $P \mapsto P(a_i)$ sur E. La famille $(\mathrm{ev}_{a_0}, \mathrm{ev}_{a_1}, .., \mathrm{ev}_{a_n})$ est alors une base de $E^\star$; sa base antéduale, une famille de polynômes de $\mathbb{K}_n[X]$, est la famille $(L_0, .., L_n)$ des *polynômes interpolateurs de Lagrange* mentionnée page 106 et étudiée dans l'exercice 11 de la feuille de TD3.

Démonstration de la proposition. —

- Nous allons nous appuyer sur l'observation suivante : si $x \in E$ et si $\varphi(x) = 0$ pour tout φ de $E^\star$, alors $x = \mathbf{0}_E$.
  En effet, si $\mathcal{B}(e_1, .., e_n)$ est une base de E et si $(e_1^\star, ..., e_n^\star)$ est sa base duale, on doit avoir $e_i^\star(x) = 0$ pour tout i. Or, on peut écrire $x = x_1 e_1 + ... + x_n e_n$ pour un certain n-uplet $(x_1, .., x_n) \in \mathbb{K}^n$ (qui donne les coordonnées de x dans la base $\mathcal{B}$). Mais alors $e_i^\star(x) = x_i$ pour tout $i \in \{1, ..n\}$; ainsi toutes les coordonnées x_i doivent être nulles et $x = \mathbf{0}_E$.
- Prouvons la partie « existence » de la proposition. Considérons l'application linéaire

$$\begin{aligned} u : E &\to \mathbb{K}^n \\ x &\mapsto (\varphi_1(x), ..., \varphi_n(x)) \,. \end{aligned}$$

- Montrons qu'elle est injective en montrant que $\mathrm{Ker}(u) = \{\mathbf{0}_E\}$. Si x appartient à $\mathrm{Ker}(u)$, on a $\varphi_i(x) = 0$ pour tout x de E ; mais alors dès que φ est combinaison linéaire des φ_i, on a $\varphi(x) = 0$, et puisque $\mathbf{Vect}[\varphi_1, ..., \varphi_n] = E^\star$, on doit avoir $\varphi(x) = 0$ pour tout φ de $E^\star$; l'observation donne alors $x = \mathbf{0}_E$.
- Ainsi u est linéaire et injective ; comme E et $\mathbb{K}^n$ ont la même dimension (voir la proposition **??**), elle est automatiquement surjective (voir la proposition 4.48) et c'est un isomorphisme.

- Notons $(\varepsilon_1, ..., \varepsilon_n)$ la base canonique de $\mathbb{K}^n$; on peut obtenir une famille de n vecteurs de E en posant $(b_1, ..., b_n) = \big(u^{-1}(\varepsilon_1), ..., u^{-1}(\varepsilon_n)\big)$. Comme u est un isomorphisme, la famille $\mathcal{B} = (b_1, ..., b_n)$ est une base de E (proposition 4.34).

- Nous allons montrer que la base duale $\mathcal{B}^\star$ n'est autre que Φ.
  Notons $(\ell_1, ..., \ell_n)$ pour $\mathcal{B}^\star$ et montrons que chaque tout i de $\{1, .., n\}$ on a $\ell_i = \varphi_i$.
  Soit x un vecteur de E. Écrivons $x = x_1 b_1 + ... + x_n b_n$ où $(x_1, .., x_n)$ désigne les coordonnées de x dans la base $\mathcal{B}$.
  Pour chaque i de $\{1, .., n\}$, on a $\ell_i(x) = x_i$ par définition de la base duale.
  Or on constate que $u(x) = x_1 u(e_1) + ... + x_n u(e_n) = x_1 \varepsilon_1 + ... + x_n \varepsilon_n$ est le vecteur $(x_1, .., x_n)$ de $\mathbb{K}^n$.
  Mais la définition de $u(x)$ est $u(x) = (\varphi_1(x), ..., \varphi_n(x))$; ainsi pour tout $j \in \{1, ..n\}$ on a $\varphi_i(x) = \ell_i(x)$, cqfd.
- Prouvons la partie « unicité » de la proposition en nous appuyant sur l'isomorphisme u utilisé pour la partie « existence ». Soit $\mathcal{B} = (b_1, ..., b_n)$ une base de E vérifiant $\mathcal{B}^\star = \Phi$. Par définition de la base duale, $\varphi_1(b_1) = 1$ et pour tout i de $\{2, .., n\}$, $\varphi_i(b_1) = 0$; formulé avec l'application u, ce constat signifie $u(b_1) = (1, 0, ..0) = \varepsilon_1$. De même on constate que $u(b_i) = \varepsilon_i$ pour tout i de $\{1, .., n\}$: cela signifie que $\mathcal{B}$ est *nécessairement* la base de E dont l'image par u est la base canonique de $\mathbb{K}^n$. La base construite pour la partie « existence » est la seule possible !

□

7.5. Application : nombre d'équations indépendantes et dimension de l'espace des solutions. — Fixons deux entiers naturels non nuls n et p. Considérons un système linéaire de d équations à n inconnues

$$(S) : \begin{cases} a_{11}x_1 + ... + a_{1n}x_n = & 0 \\ \vdots & \vdots \\ a_{dn}x_1 + ... + a_{dn}x_n = & 0. \end{cases}$$

On peut réécrire (S) sous la forme

$$(S) : \begin{cases} \varphi_1(x_1, ..., x_n) = & 0 \\ \vdots & \vdots \\ \varphi_d(x_1, ..., x_n) = & 0. \end{cases}$$

où pour chaque i de $\{1, .., d\}$, $\varphi_i : \mathbb{R}^n \to \mathbb{R}$ est la fonction $(x_1, ..., x_n) \mapsto a_{i1}x_1 + a_{i2}x_2 + ... + a_{in}x_n$. Nous avons vu (exemple 4.5.1(b) page 118) que φ_1, φ_2, ..., φ_d sont des *formes linéaires sur* $\mathbb{R}^p$.

Les exercices sur la notion de dimension vous ont sans doute habitués à l'idée suivante :

Si les équations $\varphi_1(x) = 0$, $\varphi_2(x) = 0$, ..., $\varphi_d(x) = 0$ sont indépendantes,
alors l'espace des solutions de (S) est de dimension $n - d$.

Mais que signifie l'expression « équations indépendantes », que nous utilisons régulièrement sans jamais l'avoir définie ?

Nous l'utilisons d'habitude pour dire qu'il n'est *pas possible d'exprimer l'une des équations comme combinaison linéaire des autres.* Mais cela signifie que la famille $(\varphi_1, \varphi_2, ...\varphi_d)$ est *libre en tant que famille de vecteurs de l'espace $E^\star$.*

Au contraire, si la famille $(\varphi_1, \varphi_2, \varphi_3)$ est liée en tant que famille de vecteurs de l'espace $E^\star$, alors on peut obtenir un système équivalent à (S) en *éliminant les redondances*, c'est-à-dire en cherchant à *extraire de $(\varphi_1, \varphi_2, .., \varphi_d)$ une famille libre maximale.* Mais cela revient à chercher la dimension de $\mathbf{Vect}[\varphi_1, .., \varphi_d]$ de $E^\star$, c'est-à-dire *le rang de $(\varphi_1, \varphi_2, .., \varphi_d)$.*

Grâce aux observations précédentes, il est raisonnable de penser qu'étant donné un espace vectoriel E de dimension finie, les notions de liberté et de rang pour les familles de formes linéaires sur E permettent de trouver la dimension de l'espace des solutions d'un système linéaire dont l'inconnue est un vecteur de E.

Proposition 4.83 – Lien entre rang de la famille d'équations et espace des solutions

1. *Soit F un sous-espace vectoriel de E de dimension d. Il existe une famille libre $(\varphi_1, .., \varphi_{n-d})$ de $(n-d)$ formes linéaires sur E vérifiant : $F = \bigcap_{k=1}^{n-d} \mathrm{Ker}(\varphi_i)$.*
2. *Si p un entier naturel non nul et si $(\varphi_1, .., \varphi_p)$ est une famille de formes linéaires sur E, alors la dimension de $\bigcap_{k=1}^{p} \mathrm{Ker}(\varphi_i)$ est $(n-r)$, où r est le rang $(\varphi_1, .., \varphi_p)$ comme famille de vecteurs de $E^\star$.*

Démonstration. —

1. Fixons une base $(e_1, ..., e_d)$ de F et complétons-la en une base $\mathcal{B} = (e_1, ..., e_d, e_{d+1}, ..., e_n)$ de E. Notons $\Phi = (\psi_1, .., \psi_d, \psi_{d+1}, .., \psi_n)$ la base duale de $\mathcal{B}$.

Un vecteur $x = x_1e_1 + ... + x_de_d + x_{d+1}e_{d+1} + ... + x_ne_n$ appartient à F si et seulement s'il est combinaison linéaire de $e_1, .., e_d$, c'est-à-dire si et seulement si on a $x_{d+1} = x_{d+2} = ... = x_n = 0$. Or par définition, la base duale Φ rassemble les formes « coordonnées dans la base $\mathcal{B}$ », si bien que $x_{d+1} = \psi_{d+1}(x)$, que $x_{d+2} = \psi_{d+2}(x)$, etc.

Nous constatons donc les équivalences suivantes :

$$x \in F \iff \psi_{d+1}(x) = \psi_{d+2}(x) = ... = \psi_n(x) = 0 \iff x \in \bigcap_{k=d+1}^{n} \operatorname{Ker}(\psi)$$

En posant $\varphi_1 = \psi_{d+1}$, $\varphi_2 = \psi_{d+2}$, ..., $\varphi_{n-d} = \psi_n$ et en constatant que $(\varphi_1, .., \varphi_{n-d})$ est libre dans $E^\star$ puisqu'extraite de la base Φ de $E^\star$, on obtient le résultat voulu.

2. Soit $(\varphi_1, .., \varphi_p)$ est une famille de formes linéaires sur E ; notons r son rang. On sait que parmi ces p formes linéaires il y en a r qui forment une famille libre de $E^\star$; quitte à renuméroter, supposons qu'il s'agit de $(\varphi_1, .., \varphi_r)$.

   Rappelons que nous devons trouver la dimension du sous-espace $F = \bigcap_{k=1}^{p} \operatorname{Ker}(\varphi_i)$. Par définition du rang, on a $\mathbf{Vect}[\varphi_1, .., \varphi_d] = \mathbf{Vect}[\varphi_1, .., \varphi_r]$, donc on sait que φ_{r+1}, φ_{r+2}, ..., φ_d peuvent s'obtenir comme combinaison linéaire de $(\varphi_1, .., \varphi_r)$. Ainsi on peut réécrire $F = \bigcap_{k=1}^{r} \operatorname{Ker}(\varphi_i)$.

   Complétons maintenant $(\varphi_1, .., \varphi_n)$ en une base $\Phi = (\varphi_1, .., \varphi_n)$ de $E^\star$, et considérons la base de E qui lui est antéduale ; notons-la $\mathcal{B} = (e_1, .., e_n)$. Ainsi, tout vecteur x de E peut s'écrire sous la forme

$$x = \varphi_1(x)e_1 + ... + \varphi_r(x)e_r + \varphi_{r+1}(x)e_{r+1} ... + \varphi_n(x)e_n.$$

   Mais alors

$$x \in F \iff \varphi_1(x) = ... = \varphi_d(x) = 0 \iff x = \varphi_{r+1}(x)e_{r+1} ... + \varphi_n(x)e_n \implies x \in \mathbf{Vect}[e_{r+1}, .., e_n]\,.$$

   On obtient donc l'inclusion : $F \subset \mathbf{Vect}[e_{r+1}, .., e_n]$. Or, par définition de la base duale, on sait que pour tout $i \leq r$ la forme linéaire φ_i vaut zéro sur $e_{r+1}, .., e_n$, donc $e_{r+1}, .., e_n$ appartiennent à F et en fait $\mathbf{Vect}[e_{r+1}, .., e_n] \subset F$.

   Nous avons trouvé une base de F formée de $n - r$ vecteurs et montré, comme voulu, l'égalité $\dim(F) = n - r$.

□

CHAPITRE 5

REPRÉSENTATION MATRICIELLE DES APPLICATIONS LINÉAIRES

1. Une application linéaire est déterminée par l'image d'une base

Soient E et F deux espaces vectoriels et f une application linéaire de E vers F. Nous avons vu que si $(e_i)_{i\in I}$ est une base de E, alors certains renseignements sur la famille $(f(e_i))_{i\in I}$ donnent des informations précieuses sur l'application f : par exemple si $(f(e_i))_{i\in I}$ est génératrice de F alors f est surjective, tandis que si $(f(e_i))_{i\in I}$ est libre alors f est injective.

Nous allons voir maintenant qu'on peut aller beaucoup plus loin : si $(e_i)_{i\in I}$ est une base de E et si on connaît la famille $(f(e_i))_{i\in I}$, alors on connaît *complètement* l'application f.

Théorème 5.1 – Connaître une application linéaire $\iff$ connaître l'image d'une base

Soient
- *E et F deux espaces vectoriels,*
- *$\mathcal{B} = (e_i)_{i\in I}$ une base de E,*
- *$(v_i)_{i\in I}$ une famille de vecteurs de F indexée par I.*

Il existe une unique application linéaire f de E dans F vérifiant : $\forall i \in I,\ f(e_i) = v_i$.

Exemple 5.2. — Cherchons à décrire une application *linéaire* f de $\mathbb{R}^3$ dans $\mathbb{R}^3$ qui, sur la base canonique, se comporte comme suit :

$$f\left[\begin{pmatrix}1\\0\\0\end{pmatrix}\right] = \begin{pmatrix}1\\2\\3\end{pmatrix}, \qquad f\left[\begin{pmatrix}0\\1\\0\end{pmatrix}\right] = \begin{pmatrix}4\\5\\6\end{pmatrix}, \qquad f\left[\begin{pmatrix}0\\0\\1\end{pmatrix}\right] = \begin{pmatrix}7\\8\\9\end{pmatrix}.$$

Si $\begin{pmatrix}x\\y\\z\end{pmatrix}$ est un vecteur de $\mathbb{R}^3$, on a $\begin{pmatrix}x\\y\\z\end{pmatrix} = x\begin{pmatrix}1\\0\\0\end{pmatrix} + y\begin{pmatrix}0\\1\\0\end{pmatrix} + z\begin{pmatrix}0\\0\\1\end{pmatrix}$. Comme f est linéaire, on doit avoir

$$f\left[\begin{pmatrix}x\\y\\z\end{pmatrix}\right] = f\left[x\begin{pmatrix}1\\0\\0\end{pmatrix} + y\begin{pmatrix}0\\1\\0\end{pmatrix} + z\begin{pmatrix}0\\0\\1\end{pmatrix}\right] = x\begin{pmatrix}1\\2\\3\end{pmatrix} + y\begin{pmatrix}4\\5\\6\end{pmatrix} + z\begin{pmatrix}7\\8\\9\end{pmatrix} = \begin{pmatrix}x+4y+7z\\2x+5y+8z\\3x+6y+9z\end{pmatrix}$$

ce qui donne « la formule pour calculer » $f\left[\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right]$ lorsque x,y et z sont quelconques :

$$f\left[\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right] = \begin{pmatrix} x+4y+7z \\ 2x+5y+8z \\ 3x+6y+9z \end{pmatrix}.$$

Cela décrit bien sûr complètement f, et on constate que $f = T_A$, où A est la matrice $\begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}$.

Démonstration du théorème. — Il n'y a aucune idée dans le cas général qui n'apparaisse pas déjà sur l'exemple précédent.

Rappelons que tout vecteur x de E peut s'écrire de manière unique comme $\sum\limits_{i\in I} x_i e_i$, où $(x_i)_{i\in I}$ est la famille des coordonnées de x dans $\mathcal{B}$ (c'est une famille de scalaires, qui est à support fini). Si f vérifie les hypothèses du théorème, on doit avoir

$$f(x) = f\left(\sum_{i\in I} x_i e_i\right) = \sum_{i\in I} x_i f(e_i) = \sum_{i\in I} x_i v_i.$$

Ainsi f coïncide nécessairement avec l'application $\varphi : \sum\limits_{i\in I} x_i e_i \mapsto \sum\limits_{i\in I} x_i v_i$, qui est complètement déterminée par les familles $(e_i)_{i\in I}$ et $(v_i)_{i\in I}$.

L'application φ est linéaire et vérifie bien $\forall i \in I,\ \varphi(e_i) = v_i$: c'est donc bien une solution du problème initial, et nous avons vu que c'est la seule solution. $\square$

Voici d'autres exemples d'utilisation du théorème ci-dessus.

Exemple 5.3. — Soit Φ une application de $\mathbb{R}[X]$ dans lui-même. Si Φ est linéaire et vérifie : $\Phi(1) = 0$, $\Phi(X) = 1$, $\Phi(X^2) = 2X$, ... et plus généralement $\Phi(X^n) = nX^{n-1}$ pour tout n de $\mathbb{N}$, alors Φ est *nécessairement* la dérivation $P \mapsto P'$.

Exemple 5.4. — Soit Ψ une forme linéaire sur $\mathcal{M}_2(\mathbb{R})$. Si $\Psi\left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right] = \Psi\left[\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right] = \Psi\left[\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\right] = \Psi\left[\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right] = 1$, alors Ψ est nécessairement l'application $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto a+b+c+d$.

2. Matrice d'une application linéaire

2.1. Définition et exemples. —

Définition 5.5 – Matrice d'une application linéaire dans des bases

Soient
- E et F deux $\mathbb{K}$-espaces vectoriels de dimension finie,
- f une application linéaire de E dans F,
- $\mathcal{B} = (e_1, .., e_p)$ une base de E,
- $\mathcal{C} = (u_1, .., u_n)$ une base de F.

On appelle *matrice de f dans les bases $\mathcal{B}$ et $\mathcal{C}$*, et on note

$$\mathcal{M}_{\mathcal{C},\mathcal{B}}\,[f]$$

la matrice de $\mathcal{M}_{np}(\mathbb{K})$ dont la j-ème colonne (pour $j \in \{1, .., p\}$) recense les coordonnées de $f(e_j)$ dans la base $\mathcal{C}$.

***Remarque 5.6* (Sur les notations).** —
- **Attention à l'ordre qui est la première source d'erreurs sur ce chapitre.** Dans ce cours on note les bases de départ et d'arrivée *de droite à gauche*. Lorsqu'il y a risque de confusion, on n'hésitera pas à noter $\mathcal{M}_{\text{arr}=\mathcal{C},\ \text{dép}=\mathcal{B}}\,[f]$ cette matrice.
- Si l'espace *d'arrivée* est de dimension n et l'espace *de départ* de dimension p, alors $\mathcal{M}_{\mathcal{C},\mathcal{B}}\,[f]$ comporte n lignes et p colonnes.
- Lorsque $E = F$ et $\mathcal{C} = \mathcal{B}$ (même base au départ et à l'arrivée), on notera $\mathcal{M}_{\mathcal{B}}\,[f]$ plutôt que $\mathcal{M}_{\mathcal{B},\mathcal{B}}\,[f]$.

Nous donnons maintenant une série d'exemples significatifs.

Exemple 5.7. — Considérons l'application (évidemment linéaire)

$$\begin{aligned} f : \mathbb{R}_3[X] &\to \mathbb{R}^2 \\ P &\mapsto (P(2), P'(2)), \end{aligned}$$

notons $\mathcal{B}$ la base canonique $(1, X, X^2, X^3)$ de l'espace de départ $\mathbb{R}_3[X]$, $\mathcal{C}$ la base canonique de l'espace d'arrivée $\mathbb{R}^2$.

La matrice $\mathcal{M}_{\mathcal{C}\mathcal{B}}(f)$ est une matrice à 2 lignes et 4 colonnes.

Pour la trouver, il faut observer $f(1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $f(X) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $f(X^2) = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$ et $f(X^3) = \begin{pmatrix} 8 \\ 12 \end{pmatrix}$: on a

$$\mathcal{M}_{\mathcal{C}\mathcal{B}}(f) = \begin{pmatrix} 1 & 2 & 4 & 8 \\ 0 & 1 & 4 & 12 \end{pmatrix}.$$

Exemple 5.8. — Soit f un endomorphisme de $\mathbb{R}^3$. Dire que la matrice de f dans la base canonique [1] est $\begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 8 \\ 5 & 5 & 5 \end{pmatrix}$,

c'est dire que $f\left[\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right] = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}$, $f\left[\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right] = \begin{pmatrix} 2 \\ -1 \\ 5 \end{pmatrix}$ et $f\left[\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right] = \begin{pmatrix} 0 \\ 8 \\ 5 \end{pmatrix}$.

1. Autrement dit la matrice $\mathcal{M}_{\mathcal{C}\mathcal{C}}(f)$ où $\mathcal{C}$ est la base canonique de $\mathbb{R}^3$.

Dans ce cas f est l'application $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x+2y \\ 3x-y+8z \\ 5x+5y+5z \end{pmatrix}$. Plus généralement, si A est une matrice de $\mathcal{M}_{np}(\mathbb{K})$, et si f est une application linéaire de $\mathbb{K}^p$ dans $\mathbb{K}^n$, dire que la matrice de f dans les bases canoniques de $\mathbb{K}^p$ et $\mathbb{K}^n$ est la matrice A, c'est dire que f est la transformation T_A canoniquement associée à A.

$f : \mathbb{K}^p \to \mathbb{K}^n$ est la transformation T_A $\iff$ f a pour matrice A dans les bases canoniques de $\mathbb{K}^p$ et $\mathbb{K}^n$.

Exemple 5.9. — Considérons l'espace $E = \mathbb{R}_2[X]$ et l'endomorphisme $\Phi : P(X) \mapsto P(X+3)$ de E. Pour trouver la matrice de Φ dans la base $\mathcal{B} = (1, X, X^2)$, il suffit d'exprimer $\Phi(1)$, $\Phi(X) = X+3$ et $\Phi(X^2)$ à l'aide de 1, X et X^2 :

on constate que $\mathcal{M}_{\mathcal{B}}(\Phi) = \begin{pmatrix} 1 & 3 & 9 \\ 0 & 1 & 6 \\ 0 & 0 & 1 \end{pmatrix}$.

Exemple 5.10. — Considérons toujours l'espace $E = \mathbb{R}_2[X]$ et l'endomorphisme $\Phi : P(X) \mapsto P(X+3)$ de E. Cette fois notons $\mathcal{B}$ la base canonique et $\mathcal{C}$ la base (U, V, W) où $U(X) = X^2+1, V(X) = X^2$ et $W(X) = (X+3)^2$ (pourquoi est-ce une base ?).

Pour trouver la matrice $\mathcal{M}_{\mathcal{CB}}(\Phi)$, il faut exprimer les vecteurs $\Phi(1)$, $\Phi(X)$ et $\Phi(X^2)$ (images des vecteurs de $\mathcal{B}$), mais cette fois *en fonction des vecteurs de* $\mathcal{C}$, donc en fonction de X^2+1, X^2 et $(X+1)^2$.

On constate que $\Phi(1) = U(X) - V(X)$, que $\Phi(X) = \frac{3}{2}U(X) + \frac{-5}{3}V(X) + \frac{1}{6}W(X) =$ et que $\Phi(X^2) = W(X)$.

Ainsi $\mathcal{M}_{\mathcal{CB}}(\Phi) = \begin{pmatrix} 1 & 3/2 & 0 \\ -1 & -5/3 & 0 \\ 0 & 1/6 & 1 \end{pmatrix}$.

Exemple 5.11. — Voici un exemple qui montre qu'une application linéaire peut avoir une matrice *compliquée* ou *qui rend l'application difficile à comprendre* dans certaines bases et une matrice *très simple* ou *qui rend l'application facile à comprendre* dans d'autres bases. Considérons l'application

$$f : \mathbb{R}^3 \to \mathbb{R}^3$$
$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} 2x-y \\ -x+2z \\ -x-2y+4z \end{pmatrix}$$

et notons $\mathcal{B}$ la base canonique de $\mathbb{R}^3$, $\mathcal{B}'$ la base (u, v, w) du même espace où $u = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $v = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$, $w = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$.

La matrice $\mathcal{M}_{\mathcal{B}}(f)$ de f dans la base canonique est $A = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 0 & 2 \\ -1 & -2 & 4 \end{pmatrix}$, puisque $f = T_A$ (voir l'ex. 2bis ci-dessus).

Par ailleurs, on constate que $f(u) = u$, $f(v) = 2v$ et $f(w) = 3w$:
en conséquence, dans la base $\mathcal{B}' = (u, v, w)$, la matrice de f est la matrice diagonale $\mathcal{M}_{\mathcal{B}'}(f) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.

Voici à présent une remarque théorique importante. Fixons deux bases $\mathcal{B} = (e_1, .., e_p)$, $\mathcal{C} = (u_1, .., u_n)$ de E et F. Nous avons vu qu'à toute application linéaire f de $\mathcal{L}(E, F)$, on peut associer sa matrice dans les bases $\mathcal{B}$ et $\mathcal{C}$: on dispose donc d'une application

$$\begin{aligned} \mathbf{Mat} : \mathcal{L}(E, F) &\to \mathcal{M}_{np}(\mathbb{K}) \\ f &\mapsto \mathcal{M}_{\mathcal{C},\mathcal{B}}[f]. \end{aligned}$$

L'application **Mat** est bijective : connaître f et connaître sa matrice dans les bases $\mathcal{B}$ et $\mathcal{C}$, c'est « la même chose ».

Ce fait est d'une importance pratique immense, car il permet de manipuler *par le calcul* (et d'automatiser sur un ordinateur) des applications linéaires abstraites, donc qui relient des objets « compliqués » (par exemple susceptibles d'intervenir dans le traitement de données, de sons, d'images...).

Démonstration. — Il faut montrer que pour toute matrice A de $\mathcal{M}_{np}(\mathbb{K})$, il existe une unique application $f \in \mathcal{L}(E, F)$ vérifiant $\mathbf{Mat}(f) = A$. Or, notons $A = (a_{ij})$; dire que $\mathbf{Mat}(f) = A$, c'est dire que

- que $f(e_1)$ est la combinaison linéaire $v_1 = \sum_{i=1}^{n} a_{i1} u_i$ des vecteurs de $\mathcal{C}$ (les coefficients qui apparaissent ici sont ceux de la première colonne de A).
- et plus généralement, que pour chaque j de $\{1, .., p\}$, $f(e_j)$ est la combinaison linéaire $v_j = \sum_{i=1}^{n} a_{ij} u_i$ des vecteurs de $\mathcal{C}$ (les coefficients qui apparaissent ici sont ceux de la j-ème colonne de A).

Nous sommes donc en train d'essayer de montrer qu'il existe une unique application f vérifiant : pour j de $\{1, .., p\}$, $f(e_j) = v_j$. Mais c'est vrai : c'est le théorème 5.1. □

2.2. Comment les opérations sur les applications linéaires se traduisent sur leurs matrices. —

Proposition 5.12 – Matrice d'une combinaison d'applications linéaires

Soient

- *E, F, deux espaces vectoriels de dimension finie,*
- *f et g deux applications linéaires de E dans F,*
- *α et β deux réels,*
- *$\mathcal{B}$, $\mathcal{C}$ des bases de E et F.*

On a l'égalité

$$\mathcal{M}_{\mathcal{C},\mathcal{B}}[\alpha f + \beta g] = \alpha\, \mathcal{M}_{\mathcal{C},\mathcal{B}}[f] + \beta\, \mathcal{M}_{\mathcal{C},\mathcal{B}}[g]$$

Démonstration. — Il suffit d'écrire la définition. Si $\mathcal{B} = (e_1, .., e_p)$ et $\mathcal{C} = (u_1, .., u_n)$, alors pour chaque j de $\{1, .., p\}$, la j-ème colonne de $\mathcal{M}_{\mathcal{C},\mathcal{B}}[\alpha f + \beta g]$ rassemble les coordonnées de $(\alpha f + \beta g)(e_j)$ dans $\mathcal{C}$. Mais puisque $(\alpha f + \beta g)(e_j) = \alpha f(e_j) + \beta g(e_j)$, cette colonne est la combinaison linéaire formée avec les coefficients α et β à partir des colonnes donnant les coordonnées de $f(e_j)$ et de $g(e_j)$ dans $\mathcal{C}$, c'est-à-dire des j-èmes colonnes de $\mathcal{M}_{\mathcal{C},\mathcal{B}}[f]$ et $\mathcal{M}_{\mathcal{C},\mathcal{B}}[g]$. Les deux matrices $\mathcal{M}_{\mathcal{C},\mathcal{B}}[\alpha f + \beta g]$ et $\alpha\, \mathcal{M}_{\mathcal{C},\mathcal{B}}[f] + \beta\, \mathcal{M}_{\mathcal{C},\mathcal{B}}[g]$ ont ainsi les mêmes colonnes, elles sont donc identiques. □

Proposition 5.13 – Matrice d'une composée = produit des matrices

Soient
- E, F, G *trois espaces vectoriels de dimension finie,*
- $f \in \mathcal{L}(E, F)$ *et* $g \in \mathcal{L}(E, G)$,
- $\mathcal{B}$, $\mathcal{C}$ *et* $\mathcal{D}$ *des bases de* E, F *et* G *respectivement.*

On a alors

$$\mathcal{M}_{\mathcal{D},\mathcal{B}}\left[g \circ f\right] = \mathcal{M}_{\mathcal{D},\mathcal{C}}\left[g\right] \mathcal{M}_{\mathcal{C},\mathcal{B}}\left[f\right]$$

C'est une conséquence de l'observation suivante :

Lemme 5.14 – Comment décrire l'image par f d'un vecteur à l'aide de $\mathcal{M}_{\mathcal{C}\mathcal{B}}[f]$

Soit v *un vecteur de* E.
Notons $X^v_{\mathcal{B}}$ *la matrice-colonne (de* $\mathcal{M}_{p1}(\mathbb{K})$, *où* $p = \dim(E)$*) rassemblant les coordonnées de* v *dans la base* $\mathcal{B}$.
Alors les coordonnées de $f(v)$ *dans la base* $\mathcal{C}$ *sont données par le produit matriciel*

$$X^{f(v)}_{\mathcal{C}} = \mathcal{M}_{\mathcal{C},\mathcal{B}}\left[f\right] \cdot X^v_{\mathcal{B}}.$$

Démonstration du lemme. — Notons $(\alpha_1, .., \alpha_p)$ le p-uplet des coordonnées de v dans $\mathcal{B}$. Rappelons que le produit matrice-vecteur $\mathcal{M}_{\mathcal{C},\mathcal{B}}\left[f\right] \cdot X^v_{\mathcal{B}}$ est une combinaison linéaire des colonnes de $\mathcal{M}_{\mathcal{C},\mathcal{B}}\left[f\right]$: si on note $A_1, .., A_p$ les colonnes de cette matrice, on a

$$\mathcal{M}_{\mathcal{C},\mathcal{B}}\left[f\right] \cdot X^v_{\mathcal{B}} = \alpha_1 A_1 + ... + \alpha_p A_p$$

Par ailleurs, $X^{f(v)}_{\mathcal{C}}$ est la colonne des coordonnées de $f(v)$ dans $\mathcal{C}$.

Mais on a $f(v) = f\left(\sum_{j=1}^{p} \alpha_j e_j\right) = \sum_{j=1}^{p} \alpha_j f(e_j)$; ainsi, pour chaque i de $\{1, .., n\}$, la coordonnée de chaque $f(e_j)$ selon le i-ème vecteur de $\mathcal{C}$ est le coefficient en i-ème position dans la colonne A_j : ainsi le i-ème coefficient de $X^{f(v)}_{\mathcal{C}}$ est le même que le i-ème coefficient de $\alpha_1 A_1 + ... + \alpha_p A_p$, ces deux colonnes sont donc identiques comme voulu. □

Démonstration de la proposition.. — Pour tout u de E, on peut décrire les coordonnées dans la base $\mathcal{D}$ du vecteur $(g \circ f)(u)$ de deux façons différentes : on a d'une part $X^{(g \circ f)(u)}_{\mathcal{D}} = \mathcal{M}_{\mathcal{D},\mathcal{B}}\left[g \circ f\right] \cdot X^u_{\mathcal{B}}$, mais d'autre part $X^{(g \circ f)(u)}_{\mathcal{D}} = X^{g[f(u)]}_{\mathcal{D}} = \mathcal{M}_{\mathcal{D},\mathcal{C}}\left[g\right] \cdot X^{f(u)}_{\mathcal{C}} = \mathcal{M}_{\mathcal{D},\mathcal{C}}\left[g\right] \mathcal{M}_{\mathcal{C},\mathcal{B}}\left[f\right] \cdot X^u_{\mathcal{B}}$. Ainsi :

$$\forall u \in E, \ (\mathcal{M}_{\mathcal{D},\mathcal{B}}\left[g \circ f\right]) \cdot X^u_{\mathcal{B}} = (\mathcal{M}_{\mathcal{D},\mathcal{C}}\left[g\right] \mathcal{M}_{\mathcal{C},\mathcal{B}}\left[f\right]) \cdot X^u_{\mathcal{B}}. \qquad (\star)$$

Notons maintenant $\mathcal{B} = (e_1, .., e_p)$. Comme l'égalité précédente est vraie pour tout u, on peut observer ce qu'elle donne pour $u = e_1$: dans ce cas $X^u_{\mathcal{B}}$ est le vecteur $(1, 0, .., 0)$ de la base canonique de $\mathbb{K}^p$, et les deux membres de $(\star)$ sont les premières colonnes des matrices $(\mathcal{M}_{\mathcal{D},\mathcal{B}}\left[g \circ f\right])$ et $(\mathcal{M}_{\mathcal{D},\mathcal{C}}\left[g\right] \mathcal{M}_{\mathcal{C},\mathcal{B}}\left[f\right])$... qui ont donc la même première colonne. Pour chaque j de $\{1, .., p\}$, appliquer l'égalité $(\star)$ à $u = e_j$ montre de même que ces deux matrices ont la même j-ème colonne, elles sont donc identiques. □

Voici quelques conséquences très utiles de la proposition 5.13.

(a) Si f est une application linéaire bijective, alors $\mathcal{M}_{\mathcal{C},\mathcal{B}}\left[f\right]$ est une matrice inversible et $\mathcal{M}_{\mathcal{C},\mathcal{B}}\left[f\right]^{-1} = \mathcal{M}_{\mathcal{B},\mathcal{C}}\left[f^{-1}\right]$.

(b) Soient E est un espace vectoriel de dimension finie et f un endomorphisme de E. S'il existe une base $\mathcal{B}$ de E dans laquelle la matrice $A = \mathcal{M}_{\mathcal{B}}(f)$ vérifie $A^2 = A$, alors f est un projecteur.

En effet, on a alors $\mathcal{M}_{\mathcal{B}}(f) = A = A^2 = \mathcal{M}_{\mathcal{B}}(f^2)$, et nous avons vu que deux endomorphismes dont les matrices dans $\mathcal{B}$ coïncident sont identiques, donc $f^2 = f$.

Par ailleurs, si f est un projecteur, alors pour toute base $\mathcal{B}$ de E, la matrice $A = \mathcal{M}_{\mathcal{B}}(f)$ vérifie $A^2 = A$ (pourquoi ?).

Relevons par ailleurs un renseignement théorique :

Proposition 5.15 – Dimension de l'espace des applications linéaires

oient E et F deux espaces vectoriels de dimension finie. La dimension de $\mathcal{L}(E, F)$ est donnée par

$$\dim [\mathcal{L}(E, F)] = \dim(E) \cdot \dim(F).$$

Démonstration. — Nous avons vu qu'une fois fixées deux bases $\mathcal{B}$ et $\mathcal{C}$ de E et F, on dispose d'une application

$$\begin{aligned} \mathbf{Mat} : \mathcal{L}(E, F) &\to \mathcal{M}_{np}(\mathbb{K}) \\ f &\mapsto \mathcal{M}_{\mathcal{C},\mathcal{B}}[f]. \end{aligned}$$

On observe alors que

- Une phrase résumant la proposition ci-dessus est : « **Mat** est linéaire ».
- Nous avons vu à la fin du §**??** que **Mat** est bijective.
- Ainsi **Mat** est un isomorphisme, et donc $\dim [\mathcal{L}(E, F)] = \dim [\mathcal{M}_{np}(\mathbb{K})] = np = pn = \dim(E) \dim(F)$.

$\square$

3. Changements de coordonnées, changements de base

3.1. Matrice de changement de base. —

Définition 5.16 – Matrice de passage ou de changement de base

Soient
- E un espace vectoriel de dimension finie n
- $\mathcal{B} = (e_1, .., e_n)$ et $\mathcal{C} = (u_1, .., u_n)$ deux bases de E.

On appelle *matrice de passage de* $\mathcal{B}$ *à* $\mathcal{C}$, et on note

$$P_{\mathcal{B} \text{ vers } \mathcal{C}} \qquad \text{ou} \qquad P_{\mathcal{B}\to\mathcal{C}}$$

la matrice de $\mathcal{M}_n(\mathbb{K})$ dont la j-ème colonne (pour $j \in \{1, .., n\}$) recense les coordonnées de u_j dans la base $\mathcal{B}$.

Exemple 5.17. — Dans $\mathbb{R}^3$,

pour trouver la matrice de passage de la base $(u, v, w) = \left(\begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}\right)$ à la base $(u', v', w') = \left(\begin{pmatrix}2\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}5\\5\\5\end{pmatrix}\right)$,

il suffit d'exprimer u', v' et w' en fonction de u, v et w :
or $u' = 1 \cdot u + 1 \cdot v + 0 \cdot w$, $v' = (-1) \cdot u + 1 \cdot v + 0 \cdot w$ et $w' = 0 \cdot u + 0 \cdot v + 5 \cdot w$, donc

$$P_{(u,v,w) \text{ vers } (u',v',w')} = \begin{pmatrix}1 & -1 & 0\\1 & 1 & 0\\0 & 0 & 5\end{pmatrix}.$$

Exemple 5.18 **(Passage de la base canonique à une autre).** — Si $\mathcal{C} = \left[\begin{pmatrix}a_{11}\\ \vdots \\ a_{n1}\end{pmatrix}, .. \begin{pmatrix}a_{1n}\\ \vdots \\ a_{nn}\end{pmatrix}\right]$ est une base de $\mathbb{R}^n$, et si $\mathcal{B}$ est la base canonique de $\mathbb{R}^n$, alors la matrice de passage de $\mathcal{B}$ à $\mathcal{C}$ est $\begin{pmatrix}a_{11} & \dots & a_{1n}\\ \vdots & & \vdots\\ a_{n1} & \dots & a_{nn}\end{pmatrix}$: ses colonnes donnent les coordonnées dans la base canonique des vecteurs de $\mathcal{C}$.

Exemple 5.19. — Si $E = \mathbb{R}_2[X]$,

- la matrice de passage de la base $(1, X, X^2)$ à la base $(X^2, X, 1)$ est $\begin{pmatrix}0 & 0 & 1\\0 & 1 & 0\\1 & 0 & 0\end{pmatrix}$

• la matrice de passage de la base $(1, X+1, X^2+X+1)$ à la base $(3, 2X-1, X^2)$ est $\begin{pmatrix} 3 & -4 & 0 \\ 0 & 2 & -1 \\ 0 & 0 & 1 \end{pmatrix}$,

puisque

$$\begin{cases} 3 & = \ \ 3 \cdot 1 + 0 \cdot (X+1) + 0 \cdot (X^2+X+1), \\ 2X-1 & = -4 \cdot 1 + 2 \cdot (X+1) + 0 \cdot (X^2+X+1), \\ X^2 & = \ \ 0 \cdot 1 + -1 \cdot (X+1) + 1 \cdot (X^2+X+1). \end{cases}$$

Proposition 5.20 – Interprétation théorique de la matrice de changement de base

Avec les notations de la définition ci-dessus, on a l'égalité

$$P_{\mathcal{B} \text{ vers } \mathcal{C}} = \mathcal{M}_{arr=\mathcal{B},\ d\acute{e}p=\mathcal{C}}\,[id_E]$$

Attention, l'ordre n'est probablement pas celui qu'on imagine en premier !

Démonstration. — Il suffit d'écrire les définitions. Notons $\mathcal{C} = (u_1, .., u_n)$; alors :

• Par définition, $\mathcal{M}_{\text{arr}=\mathcal{B},\ \text{dép}=\mathcal{C}}\,[\text{id}_E]$ est la matrice dont la j-ème colonne rassemble les coordonnées de l'image du j-ème vecteur de la base $\mathcal{C}$, donc de $\text{id}_E(u_j)$, dans la base $\mathcal{B}$.

• Par définition, $P_{\mathcal{B} \text{ vers } \mathcal{C}}$ est la matrice dont la j-ème colonne rassemble les coordonnées de u_j dans $\mathcal{B}$. □

Remarque 5.21 **(Inversibilité des matrices de passage).** — (a) Dans le contexte ci-dessus, $P_{\mathcal{B} \text{ vers } \mathcal{B}'}$ est toujours inversible : l'interprétation théorique la fait apparaître comme matrice dans des bases de l'application linéaire bijective id_E. Comme $\text{id}_E^{-1} = \text{id}_E$, on a d'ailleurs (remarque (a) page 131) :

$$P^{-1}_{\mathcal{B} \text{ vers } \mathcal{B}'} = P_{\mathcal{B}' \text{ vers } \mathcal{B}}$$

(b) Réciproquement, si P est une matrice inversible, alors c'est une matrice de passage : ses colonnes $C_1, .., C_n$ forment une base $\mathcal{C}$ de $\mathbb{K}^n$ (voir page 104), et si on note $\mathcal{B}$ la base canonique de $\mathbb{K}^n$, nous avons vu dans le premier exemple page 133 que $P = P_{\mathcal{B} \text{ vers } \mathcal{C}}$.

Si P est une matrice inversible, on a : $\quad P \quad = \quad P_{\text{(base canonique)} \quad \text{vers} \quad \text{(base des colonnes de } P)}$.

3.2. Changement de coordonnées pour les vecteurs. —

Proposition 5.22 – Lien entre les coordonnées d'un vecteur dans deux bases différentes

Soient
- *E un espace vectoriel de dimension finie ;*
- *$\mathcal{B}$ et $\mathcal{B}'$ deux bases de E et $P_{\mathcal{B} \text{ vers } \mathcal{B}'}$ la matrice de passage associée ;*
- *u un vecteur de E ;*
- *$X^u_{\mathcal{B}'}$ la colonne rassemblant les coordonnées de u dans $\mathcal{B}'$, $X^u_{\mathcal{B}}$ celle qui rassemble ses coordonnées dans $\mathcal{B}$.*

On a l'égalité

$$X^u_{\mathcal{B}} = P_{\mathcal{B} \text{ vers } \mathcal{B}'} \cdot X^u_{\mathcal{B}'}.$$

Attention, les erreurs sur ce résultat sont très fréquentes : à gauche il y a les « anciennes » coordonnées.

Démonstration. — C'est une conséquence immédiate du lemme **??** et de l'interprétation qu'on vient de donner de $P_{\mathcal{B} \text{ vers } \mathcal{B}'}$. □

Exemple 5.23. — On se place dans $\mathbb{R}^2$. Fixons un réel θ et considérons les deux vecteurs $u_\theta = \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$, $v_\theta = \begin{pmatrix} -\sin(\theta) \\ \cos(\theta) \end{pmatrix}$.

La famille (u_θ, v_θ) est formée de deux vecteurs non colinéaires (ils sont orthogonaux !), donc c'est une base de $\mathbb{R}^2$.

Si $\begin{pmatrix} x \\ y \end{pmatrix}$ est un vecteur de $\mathbb{R}^2$, ses coordonnées dans la base canonique (e_1, e_2) sont bien sûr x et y ; pour trouver ses coordonnées (x', y') dans la base (u_θ, v_θ), la proposition précédente dit qu'on a

$$\begin{pmatrix} x \\ y \end{pmatrix} = P_{(e_1,e_2) \text{ vers } (u_\theta, v_\theta)} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

autrement dit

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}.$$

N'oubliez pas qu'ici les nouvelles coordonnées sont *à droite* : pour les trouver, il convient donc d'écrire [2]

$$\begin{aligned} \begin{pmatrix} x' \\ y' \end{pmatrix} &= \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta) & +\sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \\ &= \begin{pmatrix} x\cos(\theta) + y\sin(\theta) \\ -x\sin(\theta) + y\cos(\theta) \end{pmatrix}. \end{aligned}$$

Exemple 5.24. — Plaçons-nous dans l'espace $\mathbb{R}_4[X]$ et considérons le polynôme $U(X) = X^3 - 1$. Notons
- $\mathcal{B}$ la base canonique $(1, X, X^2, X^3, X^4)$ de E,
- $\mathcal{B}'$ la base $(1, (X+1), (X+1)^2, (X+1)^3, (X+1)^4)$.

2. Le fait que $\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}^{-1} = \begin{pmatrix} \cos(\theta) & +\sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}$ peut se vérifier à la main.

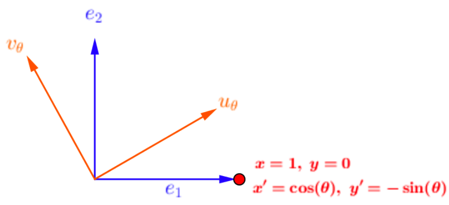


FIGURE 1. Coordonnées du point rouge dans (e_1, e_2) et (u_θ, v_θ).

Quelles sont les coordonnées de U dans $\mathcal{B}'$?

Si on utilise la proposition ci-dessus, on doit commencer par constater que $P_{\mathcal{B} \text{ vers } \mathcal{B}'} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$.

Ensuite, la colonne des coordonnées de U dans la base canonique est $X_{\mathcal{B}}^U = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$; si on note $X_{\mathcal{B}'}^U = \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \\ \varepsilon \end{pmatrix}$ celle des coordonnées de U dans $\mathcal{B}'$, on doit donc avoir

$$X_{\mathcal{B}}^U = P_{\mathcal{B} \text{ vers } \mathcal{B}'} X_{\mathcal{B}'}^U, \qquad \text{c'est-à-dire} \qquad \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \\ \varepsilon \end{pmatrix} = \begin{pmatrix} \alpha + \beta + \gamma + \delta + \varepsilon \\ \beta + 2\gamma + 3\delta + 4\varepsilon \\ \gamma + 3\delta + 6\varepsilon \\ \delta + 4\varepsilon \\ \varepsilon \end{pmatrix} ;$$

on est ramené à un système triangulaire dont on trouve aisément la solution : $\varepsilon = 0$, $\delta = 1$, $\gamma = -3$, $\beta = 3$, $\alpha = -2$.

On peut se convaincre qu'on n'a pas fait d'erreur en se rappelant la formule de Taylor pour les polynômes (p. 45) :

on sait que pour chaque réel a, on a $U(X) = \sum_{k=1}^{4} \frac{U^{(k)}(a)}{k!}(X-a)^k$. C'est pour $a = -1$ qu'on obtient les coordonnées de U dans $\mathcal{B}'$, et on constate bien que $U(-1) = -2$, $U'(-1) = 3(-1)^2 = 3$, etc.

3.3. Changement de base pour les matrices d'applications linéaires. —

Théorème 5.25 – Changement de base pour la matrice d'une application linéaire

Soient

- *E et F deux espaces vectoriels de dimension finie,*
- *f une application linéaire de E dans F,*
- *$\mathcal{B}$ et $\mathcal{B}'$ deux bases de E,*
- *$\mathcal{C}$ et $\mathcal{C}'$ deux bases de F.*

On a l'égalité

$$\mathcal{M}_{\mathcal{C}',\mathcal{B}'}[f] = P_{\mathcal{C}' \text{ vers } \mathcal{C}} \cdot \mathcal{M}_{\mathcal{C},\mathcal{B}}[f] \cdot P_{\mathcal{B} \text{ vers } \mathcal{B}'}.$$

Démonstration. —

$$\mathcal{M}_{\mathcal{C}'\mathcal{B}'}[f] = \mathcal{M}_{\mathcal{C}'\mathcal{C}}[\mathrm{id}_F]\ \mathcal{M}_{\mathcal{C}\mathcal{B}}[f]\ \mathcal{M}_{\mathcal{B}\mathcal{B}'}[\mathrm{id}_E]$$

et les interprétations théoriques des matrices de passage (proposition 5.3.1) font le reste. □

Corollaire 5.26 – Cas particulier : chang^{t} de base pour les matrices d'endomorphismes

Soient
- *E un espace vectoriel de dimension finie,*
- *f un endomorphisme de E,*
- *$\mathcal{B}$ et $\mathcal{B}'$ deux bases de E.*

On a l'égalité

$$\mathcal{M}_{\mathcal{B}'}[f] = P^{-1}_{\mathcal{B} \text{ vers } \mathcal{B}'} \cdot \mathcal{M}_{\mathcal{B}}[f] \cdot P_{\mathcal{B} \text{ vers } \mathcal{B}'}.$$

Exemple 5.27. — Nous pouvons voir fonctionner ce résultat sur l'exemple 5.11.

- Notons $\mathcal{B}$ la base canonique de $\mathbb{R}^3$ et $\mathcal{B}'$ la base (u, v, w) où $u = \begin{pmatrix}1\\1\\1\end{pmatrix}$, $v = \begin{pmatrix}2\\0\\1\end{pmatrix}$, $w = \begin{pmatrix}1\\-1\\-1\end{pmatrix}$.

  La matrice de passage $\mathcal{B}$ à $\mathcal{B}'$ est $P_{\mathcal{B} \text{ vers } \mathcal{B}'} = \begin{pmatrix}1/2 & 3/2 & -1\\0 & -1 & 1\\1/2 & -1/2 & -1\end{pmatrix}$ (voir (b) page 133).

  Pour trouver son inverse, deux possibilités :
  — soit on cherche l'inverse par la méthode « du pivot » ;
  — soit on utilise le fait que $P^{-1}_{\mathcal{B} \text{ vers } \mathcal{B}'} = P_{\mathcal{B}' \text{ vers } \mathcal{B}}$ et on cherche la matrice de passage de (u, v, w) vers la base canonique. Il s'agit dans ce cas d'exprimer les vecteurs de la base canonique à l'aide de u, v et w : or on constate que $\begin{pmatrix}1\\0\\0\end{pmatrix} = \frac{u+w}{2}$, $\begin{pmatrix}0\\1\\0\end{pmatrix} = \frac{3}{2}u - v - \frac{1}{2}w$ et $\begin{pmatrix}0\\0\\1\end{pmatrix} = -u + v - w$, si bien que $P_{\mathcal{B}' \text{ vers } \mathcal{B}} = \begin{pmatrix}1/2 & 3/2 & -1\\0 & -1 & 1\\1/2 & -1/2 & -1\end{pmatrix}$.

- Considérons l'endomorphisme f dont la matrice dans la base canonique est $\begin{pmatrix}2 & -1 & 0\\-1 & 0 & 2\\-1 & -2 & 4\end{pmatrix}$.

  Le théorème ci-dessus dit donc que $\mathcal{M}_{\mathcal{B}'}(f) = \begin{pmatrix}1/2 & 3/2 & -1\\0 & -1 & 1\\1/2 & -1/2 & -1\end{pmatrix}\begin{pmatrix}2 & -1 & 0\\-1 & 0 & 2\\-1 & -2 & 4\end{pmatrix}\begin{pmatrix}1/2 & 3/2 & -1\\0 & -1 & 1\\1/2 & -1/2 & -1\end{pmatrix}$...

  le calcul donne $\begin{pmatrix}1 & 0 & 0\\0 & 2 & 0\\0 & 0 & 3\end{pmatrix}$, comme prévu compte tenu des discussions antérieures de cet exemple.

Exemple 5.28. — Reprenons à présent l'exemple 5.7 : considérons

$$f : \mathbb{R}_3[X] \to \mathbb{R}^2$$
$$P \mapsto (P(2), P'(2)),$$

Notons $\mathcal{B}$ la base canonique de $\mathbb{R}_3[X]$, $\mathcal{B}'$ la base $(1, X-2, (X-2)^2, (X-2)^3)$. Notons $\mathcal{C}$ la base canonique de $\mathbb{R}^2$.

Nous avons vu que

$$\mathcal{M}_{\mathcal{CB}}(f) = \begin{pmatrix}1 & 2 & 4 & 8\\0 & 1 & 4 & 12\end{pmatrix}.$$

Que vaut $\mathcal{M}_{\mathcal{CB}'}(f)$?

En développant les vecteurs de $\mathcal{B}'$on constate que $P_{\mathcal{B} \text{ vers } \mathcal{B}'} = \begin{pmatrix} 1 & -2 & 4 & -8 \\ 0 & 1 & -4 & 12 \\ 0 & 0 & 1 & -6 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

La proposition sur le changement de base donne

$$\mathcal{M}_{\mathcal{C}\mathcal{B}'}(f) = \mathcal{M}_{\mathcal{C}\mathcal{B}}(f) P_{\mathcal{B} \text{ vers } \mathcal{B}'} = \begin{pmatrix} 1 & 2 & 4 & 8 \\ 0 & 1 & 4 & 12 \end{pmatrix} \begin{pmatrix} 1 & -2 & 4 & -8 \\ 0 & 1 & -4 & 12 \\ 0 & 0 & 1 & -6 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

et le calcul donne « miraculeusement »

$$\mathcal{M}_{\mathcal{C}\mathcal{B}'}(f) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Doit-on s'en étonner ? Ici il était particulièrement facile d'observer les images des vecteurs de $\mathcal{B}'$ dans la base canonique $\mathcal{C}$ de $\mathbb{R}^2$: on a $f(1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $f(X-2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $f((X-2)^2) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ et $f((X-3)^2) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

4. Matrices semblables

Définition 5.29 – Matrices semblables

Soient n un entier naturel non nul et M, N deux matrices carrées de $\mathcal{M}_n(\mathbb{K})$.
On dit que M et N sont semblables lorsqu'il existe une matrice P de $\mathcal{M}_n(\mathbb{K})$ qui soit inversible et qui vérifie

$$N = P^{-1}MP.$$

La notion est évidemment liée à ce qui précède :

Proposition 5.30 – Semblables = représentent la même appli. dans des bases différentes

Soit n un entier naturel non nul et M, N deux matrices carrées de $\mathcal{M}_n(\mathbb{K})$. Les conditions suivantes sont équivalentes :

(i) Les matrices M et N sont semblables

(ii) Il existe un endomorphisme f de $\mathbb{R}^n$ et deux bases $\mathcal{B}$ et $\mathcal{C}$ de $\mathbb{R}^n$ vérifiant : $M = \mathcal{M}_{\mathcal{B}}[f]$ et $N = \mathcal{M}_{\mathcal{C}}[f]$.

Démonstration. — Le fait que (ii) implique (i) est une conséquence immédiate de la dernière proposition du paragraphe **??**. Pour montrer que (i) implique (ii), supposons $N = P^{-1}MP$; notons $\mathcal{B}$ la base canonique de $\mathbb{K}^n$, $\mathcal{C}$ la famille des colonnes de P. Nous avons vu au numéro **??**, remarque (b), que $\mathcal{C}$ est une base de $\mathbb{K}^n$ et que $P = P_{\mathcal{B}\text{ vers}\mathcal{C}}$. Par ailleurs si f est l'application linéaire T_M canoniquement associée à M, on sait que $M = \mathcal{M}_{\mathcal{B}}[f]$, si bien que

$$N = P^{-1}MP = P^{-1}_{\mathcal{B}\text{ vers}\mathcal{C}} \cdot \mathcal{M}_{\mathcal{B}}[f] \cdot P_{\mathcal{B}\text{ vers}\mathcal{C}} = P_{\mathcal{C}\text{ vers}\mathcal{B}} \cdot \mathcal{M}_{\mathcal{B}}[f] \cdot P_{\mathcal{B}\text{ vers}\mathcal{C}} = \mathcal{M}_{\mathcal{C}}[\mathrm{id} \circ f \circ id]$$

ce qui montre (ii). □

Voici une série d'exemples.

Exemple 5.31. — Reprenons l'exemple 5.11 : nous avons montré que si A est la matrice $\begin{pmatrix} 2 & -1 & 0 \\ -1 & 0 & 2 \\ -1 & -2 & 4 \end{pmatrix}$, alors

- la matrice de T_A dans la base canonique est A
- dans la base $\mathcal{B}' = \left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \right)$, on a vu que : $\mathcal{M}_{\mathcal{B}}[T_A] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.

Ainsi A et $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ sont semblables.

Exemple 5.32. — Soit A une matrice de $\mathcal{M}_2(\mathbb{R})$ vérifiant $A^2 = 0$ mais $A \neq 0$. Montrons que A est semblable à $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$.

Fixons un vecteur X de $\mathbb{R}^2$ vérifiant $AX \neq 0$ (il en existe, puisque $A \neq 0$) ; considérons la famille $(u, v) = (X, AX)$.

C'est une famille libre : en effet, si $\alpha X + \beta AX = 0$ avec $(\alpha, \beta) \in \mathbb{R}^2$, alors $A(\alpha X + \beta AX) = 0$, donc $\alpha(AX) + \beta(A^2 X) = 0$, c'est-à-dire $\alpha(AX) = 0$; comme AX n'est pas nul, on obtient $\alpha = 0$ puis $\beta(AX) = 0$ et $\beta = 0$.

Ainsi (u, v) est une famille libre de deux vecteurs de $\mathbb{R}^2$; c'en est une base. Elle vérifie $T_A(u) = v$ et $T_A(v) = A^2 X = 0$.

On a donc bien $\mathcal{M}_{(u,v)}(T_A) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, comme annoncé.

Exemple 5.33. — Soit A une matrice de $\mathcal{M}_3(\mathbb{R})$ vérifiant $A^2 = A$. Montrons que si A n'est ni la matrice nulle ni la matrice identité, alors A est semblable soit à $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, soit à $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

On sait que l'application T_A est un projecteur ; rappelons que si $F = \operatorname{Im}(A)$ et $S = \operatorname{Ker}(A)$, alors $F = \{X \in \mathbb{R}^n, AX = X\}$ et que T_A est le projecteur sur F parallèlement à S. Rappelons également que si $\mathcal{B}_F$ est une base de F et $\mathcal{B}_S$ une base de S, alors la réunion $\mathcal{B}_F \cup \mathcal{B}_S$ est une base de $\mathbb{R}^3$ (voir le numéro 3.3, page 77) .

- Si $\dim(F) = 0$, alors $\mathbb{R}^3 = F \oplus S = \{0\} \oplus S = S = \operatorname{Ker}(T_A)$ et T_A est l'application nulle, donc A est la matrice nulle.
- Si $\dim(F) = 1$, alors $\dim(S) = 2$; choisissons une base (u) de F et une base (v, w) de S, alors $\mathcal{B} = (u, v, w)$ est une base de $\mathbb{R}^3$ et $T_A(u) = u$, $T_A(v) = T_A(w) = 0$, si bien que $\mathcal{M}_{\mathcal{B}}[f] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.
- Si $\dim(F) = 2$, alors $\dim(S) = 1$; choisissons une base (u, v) de F et une base (w) de S, alors $\mathcal{B} = (u, v, w)$ est une base de $\mathbb{R}^3$ et $T_A(u) = u$, $T_A(v) = v$, $T_A(w) = 0$, si bien que $\mathcal{M}_{\mathcal{B}}[f] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.
- Si $\dim(F) = 3$, alors $\mathbb{R}^3 = F = \{X \in \mathbb{R}^n, \quad AX = X\}$, donc A est en fait la matrice identité.

***Remarque 5.34* (Utilisation pratique de la notion).** — Les exemples ci-dessus devraient montrer clairement que le fait qu'une matrice A soit semblable à une matrice « simple » révèle que la transformation T_A est en fait facile à décrire géométriquement, mais à condition d'utiliser un repère différent de la base canonique de $\mathbb{R}^n$. Si $A = PBP^{-1}$ avec B « agréable » et P inversible, alors la matrice P permet de passer de la base canonique à un repère dans lequel la transformation T_A est « plus facile à comprendre » que si on s'en tient à la base canonique de $\mathbb{R}^n$.

Nous venons de voir que pour montrer que deux matrices *sont* semblables, il est souvent nécessaire de comprendre les transformations géométriques sous-jaçentes et d'en révéler une parenté cachée. Ce n'est pas un problème évident, comme nous commencerons à le voir au chapitre 7.

En général, montrer que deux matrices sont semblables, c'est difficile.

En revanche, quelques indices permettent parfois de voir très vite que deux matrices *ne sont pas* semblables. En voici deux :

Proposition 5.35 – Trace et rang de deux matrices semblables

Soient n un entier naturel non nul et M, N deux matrices carrées de $\mathcal{M}_n(\mathbb{K})$.
Si M et N sont semblables, alors on a $Tr(M) = Tr(N)$ et $\boldsymbol{rg}[M] = \boldsymbol{rg}[N]$.

Démonstration. — (i) Pour la trace, rappelons que si A et B sont deux matrices de $\mathcal{M}_n(\mathbb{K})$, alors $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$. Par conséquent, si M est une matrice quelconque de $\mathcal{M}_n(\mathbb{K})$ et si $N = P^{-1}MP$ est une matrice semblable à M, alors

$$\operatorname{Tr}(N) = \operatorname{Tr}\left(P^{-1}MP\right) = \operatorname{Tr}\left([P^{-1}][MP]\right) = \operatorname{Tr}\left([MP][P^{-1}]\right) = \operatorname{Tr}\left(MPP^{-1}\right) = \operatorname{Tr}(M).$$

(ii) Pour le rang, il faut remarquer $M = P^{-1}NP$, alors les sous-espaces $\operatorname{Im}(M)$ et $\operatorname{Im}(N)$ sont liés par la relation suivante : $\operatorname{Im}(N) = T_P\left[\operatorname{Im}(N)\right]$. En effet, un vecteur Y de de $\mathbb{R}^n$ appartient à $\operatorname{Im}(N)$ si et seulement si PY appartient à $\operatorname{Im}(M)$:

- Si Y appartient à $\operatorname{Im}(N)$, alors il existe un vecteur X vérifiant : $Y = NX$, ce qui équivaut à $Y = P^{-1}MPX$ ou encore :
  $(PY) = M(PX)$; ainsi PY appartient à $\operatorname{Im}(M)$.
- Réciproquement, si PY appartient à $\operatorname{Im}(M)$, alors il existe un vecteur X vérifiant $PY = MU$, mais alors :
  $Y = P^{-1}MU = P^{-1}MP(PU) = N(PU)$, si bien que Y appartient à $\operatorname{Im}(N)$.

Mais P étant une matrice inversible, l'application linéaire T_P est un isomorphisme, donc préserve la dimension (voir page 107). Les dimensions de $\operatorname{Im}(M)$ et $\operatorname{Im}(N)$, qui sont les rangs de M et N, sont donc identiques.

□

La façon la plus courante d'utiliser ce résultat est la remarque suivante :

Si deux matrices n'ont pas la même trace ou pas le même rang, elles n'ont aucune chance d'être semblables.

Exemple 5.36. —

(a) Les matrices $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ et $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$ ne sont pas semblables : leurs traces sont différentes.

(b) Les matrices $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ et $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ ont la même trace (zéro) mais elles ont des rangs différents : la première est de rang 1, la seconde de rang 2. Elles ne sont donc pas semblables.

Attention, il est très fréquent que deux matrices aient la même trace et le même rang sans être semblables.

Exemple 5.37. — (a) Les matrices $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ et $H = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ ne sont pas semblables. En effet, H est le double $2I_2$ de la matrice identité, donc pour *toute* matrice inversible P, on a $P^{-1}HP = H$ et on ne peut jamais avoir $P^{-1}IP = A$.

(b) Les matrices $A = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$ et $H = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$ ont même trace (5) et même rang (2), mais le chapitre qui vient montrera facilement qu'elles ne sont pas semblables.

CHAPITRE 6

DÉTERMINANT

1. Introduction

- Vous avez déjà résolu des systèmes linéaires. Beaucoup de ces systèmes étaient *à coefficients entiers* ; vous avez par ailleurs souvent rencontré le cas où il y a une unique solution.

  Voici quelques exemples, avec des coefficients que j'ai choisis « au hasard » :

  Le système $\begin{cases} 5x+3y=1 \\ x+4y=2 \end{cases}$, d'inconnue $(x,y)\in\mathbb{R}^2$, a une unique solution, donnée par : $x=\dfrac{-2}{17}$,
  $$y=\frac{9}{17}$$
  Le système $\begin{cases} x+3y-6z=0 \\ x+4y+5z=-2 \\ 2x+y+z=1 \end{cases}$, d'inconnue $(x,y,z)\in\mathbb{R}^3$, a une unique solution, donnée par :
  $$x=\frac{57}{68},\ y=\frac{-37}{68},\ z=\frac{-9}{68}$$
  Le système $\begin{cases} x+2y+3z+4t=1 \\ x-y-z+t=3 \\ 4x+3y+2z+t=1 \\ x+3y-3z-t=2 \end{cases}$, d'inconnue $(x,y,z,t)\in\mathbb{R}^4$, a une unique solution : $x=\dfrac{29}{40}$,
  $$y=\frac{-11}{40},\ z=\frac{-41}{40},\ t=\frac{39}{40}.$$

- On peut trouver banal que les solutions soient toujours données par des nombres rationnels : après tout, la méthode « du pivot » que vous avez apprise pour résoudre ces systèmes revient à une sucession d'additions/soustractions et multiplications/divisions faisant intervenir les coefficients du système.

  Il peut sembler en revanche plus surprenant que dans tous les cas ci-dessus, les valeurs numériques des coordonnées de la solution aient *toutes le même dénominateur* dans leur écriture sous forme de fraction irréductible. Est-ce que les exemples ci-dessus ne sont pas représentatifs, ou cela cache-t-il un fait général ?

- Pour les systèmes 2×2, vous savez déjà que c'est un fait général. En effet, si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est une matrice 2×2 à coefficients entiers et si (u, v) est un couple d'entiers, le système

$$\begin{cases} ax + by = u \\ cx + dy = v \end{cases}$$

a une unique solution si et seulement si la matrice A est inversible (voir page **??**), et dans ce cas beaucoup d'entre vous ont appris en Terminale que

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

si bien que la solution est donnée par $\begin{pmatrix} x \\ y \end{pmatrix} = A^{-1} \begin{pmatrix} u \\ v \end{pmatrix}$, autrement dit $x = \dfrac{ud - bv}{ad - bc}$, $y = \dfrac{au - cv}{ad - bc}$.

Dans les écritures en fraction irréductible des deux rationnels x et y, les dénominateurs sont donc *toujours* des diviseurs du *déterminant* **Det** $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$. Si $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ est inversible, ce nombre est bien sûr non nul.

Rappelons par ailleurs que le déterminant permet également de « tester » l'inversibilité de A : dire que A n'est pas inversible, c'est dire que la famille des colonnes de A est liée (voir page 131) : ici il n'y a que deux colonnes, donc A est inversible si et seulement si $\begin{pmatrix} a \\ c \end{pmatrix}$ et $\begin{pmatrix} b \\ d \end{pmatrix}$ ne sont pas colinéaires, ce qui signifie [1] $ad \neq bc$, ou encore $ad - bc \neq 0$.

- Les exemples qui ont ouvert ce chapitre suggèrent que l'existence d'un « dénominateur commun naturel » n'est pas spécifique aux systèmes 2×2.

- Afin de voir où nous allons, observons le cas des systèmes 3×3. Considérons un système linéaire

$$\begin{pmatrix} \alpha & a & d \\ \beta & b & e \\ \gamma & c & f \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} u \\ v \\ w \end{pmatrix}, \qquad \text{c'est-à-dire} \qquad \begin{cases} \alpha x + ay + dz = u \\ \beta x + by + ez = v \\ \gamma x + cy + fz = w \end{cases}$$

où l'inconnue est le triplet $(x, y, z) \in \mathbb{R}^3$ et où les données $\alpha, \beta, \gamma, a, b, c, d, e, f, u, v, w$ sont des entiers. [2]

Cherchons à résoudre ce système par la méthode habituelle « du pivot », nous allons bien sûr « chasser x et y de l'une des équations pour trouver z ». Sauf incident [3], le système initial est équivalent à

$$\begin{pmatrix} \alpha & a & d \\ 0 & \alpha b - \beta a & \alpha e - \beta d \\ 0 & \alpha c - \gamma a & \alpha f - \gamma d \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} u \\ \alpha v - \beta u \\ \alpha w - \gamma u \end{pmatrix}$$

ce qui, sauf incident [4] et malgré la laideur des formules à venir, équivaut à

1. Songer que si $a \neq 0$ et $c \neq 0$, la dernière égalité dit simplement que les rapports $\frac{a}{b}$ et $\frac{c}{d}$ ne sont pas égaux...
2. Je choisis les notations simplement pour que la suite soit plus facile à lire.
3. Un incident ne se produit que si l'opération élémentaire $L_2 \leftarrow \alpha L_2 - \beta L_1$ ne donne pas un système équivalent, c'est-à-dire si $\alpha = 0$
4. Un incident ne se produit que si $\alpha b - \beta a = 0$, pour la même raison qu'à la note précédente.

$$\begin{pmatrix}\alpha & a & d\\ 0 & \alpha b-\beta a & \alpha e-\beta d\\ 0 & 0 & (\alpha b-\beta a)(\alpha f-\gamma d)-(\alpha c-\gamma a)(\alpha e-\beta d)\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}=\begin{pmatrix}u\\ \alpha v-\beta u\\ (\alpha b-\beta a)(\alpha w-\gamma u)-(\alpha c-\gamma a)(\alpha v-\beta u)\end{pmatrix}$$

N'ayez pas peur, nous avons trouvé z ! Ou plutôt, nous avons trouvé z à condition que le nombre

$$\begin{aligned}(\alpha b-\beta a)(\alpha f-\gamma d)-(\alpha c-\gamma a)(\alpha e-\beta d) &= \alpha^2 bf-\alpha b\gamma d+-\beta\alpha af+\beta a\gamma d-\alpha^2 ce+\alpha\beta cd+\alpha\gamma ae-\beta\gamma ad\\ &= \alpha\cdot[\alpha\,(bf-ce)-\beta\,(af-dc)+\gamma\,(ae-bd)]\end{aligned}$$

soit non-nul. Si c'est le cas, alors on trouve, après une simplification analogue de la troisième coordonnée du membre de droite,

$$z=\frac{\alpha\cdot[\alpha\,(bw-cv)-\beta\,(aw-cu)+\gamma\,(av-bu)]}{\alpha\cdot[\alpha\,(bf-ce)-\beta\,(af-dc)+\gamma\,(ae-bd)]}.$$

Bien sûr on peut trouver x et y en procédant de façon analogue, et apparaît aux dénominateurs la même « somme alternée »

$$\alpha\cdot\mathrm{Det}\begin{pmatrix}\cdot & \cdot & \cdot\\ \cdot & b & e\\ \cdot & c & f\end{pmatrix}-\beta\cdot\mathrm{Det}\begin{pmatrix}\cdot & a & d\\ \cdot & \cdot & \cdot\\ \cdot & c & f\end{pmatrix}+\gamma\cdot\mathrm{Det}\begin{pmatrix}\cdot & a & d\\ \cdot & b & e\\ \cdot & \cdot & \cdot\end{pmatrix}$$

(ici j'ai essayé d'utiliser une écriture suggestive : on reconnaît dans les termes $bf-ce$, etc des déterminants de matrices 2×2 extraites de A en rayant la première colonne et l'une des lignes).

C'est cette somme alternée qu'on appelle *déterminant* de la matrice $\begin{pmatrix}\alpha & a & d\\ \beta & b & e\\ \gamma & c & f\end{pmatrix}$.

2. Déterminant d'une matrice carrée

2.1. Définition. — En remarquant que la définition du déterminant 2×2 est aussi $\mathbf{Det}\begin{pmatrix} \alpha & a \\ \beta & b \end{pmatrix} = \alpha \cdot \mathbf{Det} \cdot \begin{pmatrix} \cdot & \cdot \\ \cdot & b \end{pmatrix} - \beta \cdot \mathbf{Det}\begin{pmatrix} \cdot & a \\ \cdot & \cdot \end{pmatrix}$, on peut imaginer à partir des cas 2×2 et 3×3, une généralisation « naturelle » du déterminant aux matrices $n \times n$

Notation pour la matrice obtenue en rayant une ligne et une colonne.

Si A est une matrice $n \times n$ et si i et j sont deux indices entre 1 et n, on note $A^\star_{ij}$ la matrice $(n-1) \times (n-1)$ obtenue en rayant la i-ème ligne et la j-ème colonne de A.

Par exemple, si A est la matrice $\begin{pmatrix} 1 & 0 & 9 \\ 0 & 5 & 0 \\ -4 & 0 & -6 \end{pmatrix}$, on a $A^\star_{11} = \begin{pmatrix} \cdot & \cdot & \cdot \\ \cdot & 5 & 0 \\ \cdot & 0 & -6 \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & -6 \end{pmatrix}$, tandis que $A^\star_{32} = \begin{pmatrix} 1 & \cdot & 9 \\ 0 & \cdot & 0 \\ \cdot & \cdot & \cdot \end{pmatrix} = \begin{pmatrix} 1 & 9 \\ 0 & 0 \end{pmatrix}$.

Définition 6.1 – Déterminant d'une matrice carrée

Pour définir le déterminant d'une matrice carrée, utilisons une récurrence sur la taille de la matrice :

- Pour les matrices 1×1, le déterminant de $A = (x)$ est le scalaire x.
- Supposons la notion de déterminant définie pour les matrices $(n-1) \times (n-1)$.
  Alors si $A = (a_{ij})_{1 \le i,j \le n}$ est une matrice $n \times n$, on pose

$$\mathbf{Det}(A) = \sum_{i=1}^{n} (-1)^{i+1}\, a_{i1}\, \mathbf{Det}\left(A^\star_{i1}\right).$$

Si A appartient à $\mathcal{M}_n(\mathbb{K})$, $\mathbf{Det}(A)$ est ainsi un scalaire de $\mathbb{K}$.

Notation à barres verticales. — Si $A = (a_{ij})$ appartient à $\mathcal{M}_n(\mathbb{K})$, on note souvent $\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}$ pour $\mathbf{Det}\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$.

Exemple 6.2. —

$$\begin{vmatrix} 1 & 0 & 9 \\ 0 & 5 & 0 \\ -4 & 0 & -6 \end{vmatrix} = 1 \cdot \begin{vmatrix} 5 & 0 \\ 0 & -6 \end{vmatrix} - 0 \cdot \begin{vmatrix} 0 & 9 \\ 0 & -6 \end{vmatrix} + (-4) \begin{vmatrix} 0 & 9 \\ 5 & 0 \end{vmatrix} = -30 + 0 + (-4) \cdot 45 = -30 + 150 = 120$$

Exemple 6.3. —

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \\ 2 & 3 & 2 & 3 \\ 1 & -1 & 2 & -2 \end{vmatrix} = 1 \begin{vmatrix} 3 & 2 & 1 \\ 3 & 2 & 3 \\ -1 & 2 & -2 \end{vmatrix} - 4 \begin{vmatrix} 2 & 3 & 4 \\ 3 & 2 & 3 \\ -1 & 2 & -2 \end{vmatrix} + 2 \begin{vmatrix} 2 & 3 & 4 \\ 3 & 2 & 1 \\ -1 & 2 & -2 \end{vmatrix} - 1 \begin{vmatrix} 2 & 3 & 4 \\ 3 & 2 & 1 \\ 3 & 2 & 3 \end{vmatrix}$$

(n'oubliez pas l'alternance des signes dans le développement !).

Pour poursuivre, on doit ensuite calculer $\begin{vmatrix} 3 & 2 & 1 \\ 3 & 2 & 3 \\ -1 & 2 & -2 \end{vmatrix}$, $\begin{vmatrix} 2 & 3 & 4 \\ 3 & 2 & 3 \\ -1 & 2 & -2 \end{vmatrix}$, $\begin{vmatrix} 2 & 3 & 4 \\ 3 & 2 & 1 \\ -1 & 2 & -2 \end{vmatrix}$ et $\begin{vmatrix} 2 & 3 & 4 \\ 3 & 2 & 1 \\ 3 & 2 & 3 \end{vmatrix}$.

On voit ici que le calcul d'un déterminant peut être une affaire de patience... et que la moindre inattention peut fausser le résultat.

2.2. Cas des matrices triangulaires supérieures. — Ce cas où le calcul est immédiat est crucial et sert souvent, comme nous le verrons, de support aux calculs généraux.

Considérons une matrice triangulaire supérieure

$$A = \begin{pmatrix} a_{11} & \star & \dots & \star \\ 0 & a_{22} & \star & \vdots \\ 0 & 0 & \ddots & \star \\ 0 & 0 & 0 & a_{nn} \end{pmatrix}$$

où les étoiles au-dessus de la diagonale figurent des coefficients arbitraires, *a priori* non nuls (et bien sûr pas forcément tous égaux). Vu la définition, on constate les égalités

$$\mathbf{Det}(A) = a_{11} \begin{vmatrix} a_{22} & \star & \star \\ 0 & \ddots & \star \\ 0 & 0 & a_{nn} \end{vmatrix} + 0 \cdot \begin{vmatrix} \star & \dots & \star \\ 0 & \ddots & \star \\ 0 & 0 & a_{nn} \end{vmatrix} + \dots + 0 \cdot \begin{vmatrix} \star & \dots & \star \\ a_{22} & \star & \vdots \\ 0 & \ddots & \star \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & \star & \star \\ 0 & \ddots & \star \\ 0 & 0 & a_{nn} \end{vmatrix}$$

Nous obtenons donc, par une récurrence immédiate :

Proposition 6.4 – Déterminant d'une matrice triangulaire supérieure

Le déterminant d'une matrice triangulaire supérieure est le produit de ses coefficients diagonaux.

2.3. La « formule tout-développée ». — Vu la définition, le déterminant est une somme alternée de produits de coefficients de la matrice. Compte tenu de l'exemple (b) ci-dessus, on peut penser que c'est une somme qui comporte beaucoup de termes, et on peut n'avoir pas très envie d'écrire ce que cela donne pour une matrice $n \times n$ quelconque.

Il n'est cependant pas complètement inutile d'écrire la forme générale que prend le développement.

Nous obtenons donc le résultat suivant : l'énoncer de manière si théorique n'a pas grand intérêt pour calculer en pratique des déterminants, mais nous en tirerons par la suite des renseignements qui eux, sont très pratiques pour les calculs concrets.

Proposition 6.5 – La « formule affreuse avec $n!$ termes ».

Notons $\mathfrak{S}_n$ l'ensemble des permutations de $\{1, .., n\}$, c'est-à-dire des bijections de $\{1, .., n\}$ dans lui-même. Rappelons que $\mathfrak{S}_n$ est un ensemble fini, de cardinal $(n!)$.
Il existe une fonction $\varepsilon : \mathfrak{S}_n \to \{\pm 1\}$ vérifiant : pour toute matrice $A = (a_{ij})$ de $\mathcal{M}_n(\mathbb{K})$,

$$\boldsymbol{Det}(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdot a_{\sigma(2)2} \cdot \dots \cdot a_{\sigma(n)n}.$$

Cette formule étrange permet de soupçonner que le déterminant a des propriétés de « symétrie » cachées : par exemple la première colonne, qui était le support de la définition par récurrence donnée plus haut, n'y joue pas de rôle particulier. Nous complèterons cette remarque au numéro 6.3.

Interprétation. — Dans chaque terme du type $a_{\sigma(1)1} \cdot a_{\sigma(2)2} \cdot \ldots \cdot a_{\sigma(n)n}$ figure un coefficient de chaque colonne de A, et le fait que σ soit bijective dit que chaque ligne est représentée exactement une fois dans $\sigma(1), \sigma(2), \ldots, \sigma(n)$. Par exemple, dans le déterminant $\begin{vmatrix} \alpha & a & u \\ \beta & b & v \\ \gamma & c & w \end{vmatrix}$ vont apparaître les produits $\alpha b w$, $\alpha v c$, mais pas $\alpha a b$ (ce terme fait apparaître deux coefficients d'une même ligne), ni $\beta a v$... Pour le reste, la moitié des termes est affectée d'un signe $+$ et l'autre d'un signe $-$, mais la façon dont se répartissent les signes est subtile et il ne semble pas prioritaire de la retenir.

Voici deux exemples de conséquences de la formule tout-développée.

Le cas des matrices* 3×3 *: « règle de Sarrus » . — C'est peut-être le seul cas pratique utile à retenir :

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{13}a_{12}a_{23} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32};$$

formule qui évidemment n'a aucune chance de tenir dans une boîte cranienne : n'essayez surtout pas de l'apprendre.

On doit à Pierre-Frédéric Sarrus (1798-1861) une béquille mnémotechnique : si on réécrit les deux premières lignes de A « sous » A, comme sur la figure de gauche ci-dessous, les produits qui apparaissent avec un signe $+$ sont ceux qui sont indiqués en bleu sur la figure centrale, ceux qui apparaissent avec un signe $-$ sont ceux qui sont indiqués en bleu sur la figure de droite.

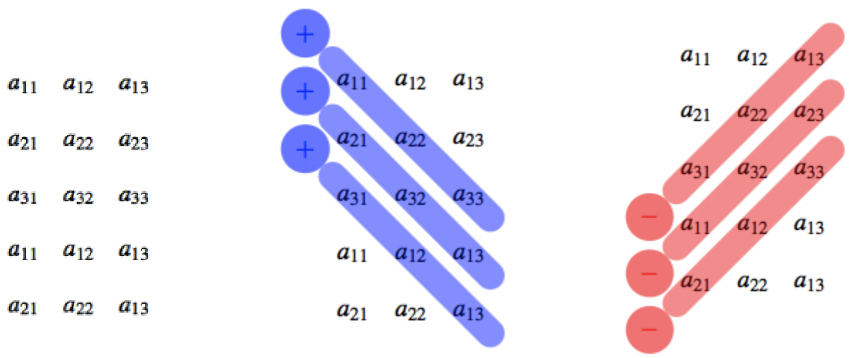


FIGURE 1. Illustration de la règle de Sarrus

Le déterminant est une fonction polynomiale des coefficients de la matrice

Revenant au cas de matrices $n \times n$ quelconque, On constate ici que le déterminant ne fait intervenir qu'une (grosse) somme de produits de n des n^2 coefficients de la matrice. On dit souvent pour le résumer que c'est une *fonction polynomiale homogène de degré n* des coefficients de la matrice.

Voici une conséquence importante : si A et B sont deux matrices quelconques de $\mathcal{M}_n(\mathbb{K})$, la fonction

$$\begin{aligned} \mathbb{K} &\to \mathbb{K} \\ t &\mapsto \det(A + tB) \end{aligned}$$

est polynomiale de degré au plus n. Nous nous servirons abondamment de cette remarque au chapitre 7.

3. Formes multilinéaires alternées et naturalité du déterminant

3.1. Propriétés pratiques du déterminant. —

3.1.1. *Vocabulaire : déterminant d'une famille de n vecteurs de $\mathbb{K}^n$.* —

Si $u_1, ..., u_n$ sont n vecteurs de $\mathbb{R}^n$ (autant de vecteurs que de coordonnées dans les vecteurs !), on définit $\mathbf{Det}(u_1, .., u_n)$ comme le déterminant de la matrice $n \times n$ dont les lignes sont $u_1, ..., u_n$: par exemple, si $u_1 = (1, 2, 3)$, $u_2 = (4, 5, 6)$ et $u_3 = (7, 8, 9)$, on note $\mathbf{Det}(u_1, u_2, u_3)$ le nombre $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$.

On peut donc voir le déterminant comme une application de $\underbrace{(\mathbb{K}^n) \times ... \times (\mathbb{K}^n)}_{n \text{ facteurs}}$ dans $\mathbb{K}$.

3.1.2. *Déterminant et opérations élémentaires.* —

Proposition 6.6 – Trois propriétés théoriques du déterminant

(i) ***Linéarité par rapport à chaque ligne :***

Si i_0 est un entier de $\{1, .., n\}$, si $u_1, ..u_{i_0-1}, u_{i_0+1}, ..., u_n$, v, w sont des vecteurs de $\mathbb{R}^n$ et λ, μ sont deux scalaires,

$$\mathbf{Det}(u_1, .., u_{i_0-1}, (\lambda v + \mu w), u_{i_0+1}, .., u_n) =$$
$$\lambda \cdot \mathbf{Det}(u_1, .., u_{i\_0m1}, v, u_{i_0+1}, .., u_n) + \mu \cdot \mathbf{Det}(u_1, .., u_{i_0-1}, w, u_{i_0+1}, .., u_n)$$

(ii) ***Si deux lignes sont identiques, alors le déterminant est nul :***

Si deux des vecteurs $(u_1, ..u_n)$ sont identiques, alors $\Phi(u_1, .., u_n) = 0$

(iii) ***Déterminant de la matrice identité : $\mathbf{Det}(I_n) = 1$.***

Démonstration. —

(i) Notons $A = (a_{ij})$ la matrice dont les lignes sont $u_1, .., u_{i_0-1}, (\lambda v + \mu w), u_{i+1}, .., u_n$. Notons $B = (b_{ij})$ la matrice dont les lignes sont $u_1, .., u_{i_0-1}, v, u_{i_0+1}, .., u_n$ et $C = (c_{ij})$ la matrice dont les lignes sont $u_1, .., u_{i_0-1}, v, u_{i_0+1}, .., u_n$. Les matrices A, B et C ont toutes leurs lignes identiques à l'exception de la ligne numéro i_0, et la i_0-ème ligne de A est égale à « λ fois celle de B plus μ fois celle de C ».

C'est $\mathbf{Det}(A) = \mathbf{Det}(u_1, .., u_{i-1}, (\lambda v + \mu w), u_{i+1}, .., u_n)$ que nous devons calculer : la formule tout-développée donne
$$\mathbf{Det}(A) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdot a_{\sigma(2)2} \cdot ... \cdot a_{\sigma(n)n}.$$
Fixons $\sigma \in \mathfrak{S}_n$ et observons le terme de la somme qui correspond à σ, c'est-à-dire $a_{\sigma(1)1} \cdot a_{\sigma(2)2} \cdot ... \cdot a_{\sigma(n)n}$.. Dans ce produit, un seul des termes est situé sur la i_0-ème ligne, puisqu'il existe un seul j tel que $\sigma(j)$ soit égal à i_0 ; notons j_0 cet indice de colonne. On a alors $a_{\sigma(k)k} = b_{\sigma(k)k} = c_{\sigma(k)k}$ pour tout $k \neq j_0$, tandis qu'on a $a_{\sigma(j_0)j_0} = a_{i_0j_0} = \lambda b_{i_0j_0} + \mu c_{i_0j_0} = \lambda b_{\sigma(j_0)j_0} + \mu c_{\sigma(j_0)j_0}$. Lorsqu'on effectue le produit $a_{\sigma(1)1} \cdot a_{\sigma(2)2} \cdot ... \cdot a_{\sigma(n)n}$, il y a « un seul terme à séparer en $b_{\sigma(j_0)j_0} + \mu c_{\sigma(j_0)j_0}$ », on obtient donc
$$\forall \sigma \in \mathfrak{S}_n, \quad a_{\sigma(1)1} \cdot a_{\sigma(2)2} \cdot ... \cdot a_{\sigma(n)n} = \lambda \left(a_{\sigma(1)1} \cdot a_{\sigma(2)2} \cdot ... \cdot a_{\sigma(n)n}\right) + \mu \left(c_{\sigma(1)1} \cdot c_{\sigma(2)2} \cdot ... \cdot c_{\sigma(n)n}\right)$$
d'où le résultat.

(ii) *(Démonstration volontairement incomplète.)* Raisonnons par récurrence sur la taille n de la matrice. Pour $n = 1$, il n'y a rien à démontrer. Soit n un entier naturel non nul fixé ; supposons que le déterminant de toute matrice $n \times n$ comportant deux lignes identiques soit nul. Soit $(u_1, .., u_{n+1})$ une famille de $n+1$ vecteurs de $\mathbb{K}^{n+1}$; supposons qu'il existe deux indices i_0 et i_1 qui vérifient $u_{i_0} = u_{i_1}$. En notant $(a_{ij})_{1 \leq i,j \leq n+1}$ la matrice $(n+1) \times (n+1)$ dont les lignes sont les u_i, la définition du déterminant donne

$$\begin{aligned}\mathbf{Det}(u_1, .., u_{n+1}) &= \sum_{i=1}^{n+1} (-1)^{i+1}\, a_{i1}\, \mathbf{Det}\left(A^\star_{i1}\right) \\ &= a_{i_0 1}(-1)^{i_0+1}\mathbf{Det}\left(A^\star_{i_0 1}\right) + a_{i_1 1}(-1)^{i_1+1}\mathbf{Det}\left(A^\star_{i_1 1}\right) + \sum_{i \notin \{i_0, i_1\}} (-1)^{i+1}(-1)^{i+1}\, a_{i1}\, \mathbf{Det}\left(A^\star_{i1}\right).\end{aligned}$$

Dans chacune des matrices $n \times n$ de la forme $A^\star_{i1}$ avec $i \notin \{i_0, i_1\}$, il y a deux lignes identiques aux positions qui proviennent de i_0 et i_1 ; chacun des termes $\mathbf{Det}\left(A^\star_{i1}\right)$ avec $i \notin \{i_0, i_1\}$ vaut donc zéro. Il reste à montrer que $(-1)^{i_1+1}\mathbf{Det}\left(A^\star_{i_1 1}\right)$ et $(-1)^{i_0+1}\mathbf{Det}\left(A^\star_{i_0 1}\right)$ sont opposés. Nous ne le ferons pas complètement, mais voici quelques indications : on remarquera que les matrices $A^\star_{i_1 1}$ et $A^\star_{i_0 1}$ ont les mêmes lignes, mais pas dans le même ordre. Pour comprendre comment leurs déterminants sont reliés l'un à l'autre, on peut utiliser la formule tout-développée du §2.3 ; on peut relier les déterminants des deux matrices en observant les deux formules et en utilisant certaines des propriétés de l'application $\varepsilon : \mathfrak{S}_n \to \{\pm 1\}$: comme nous avons évité d'entrer dans le détail des propriétés de cette application, nous n'irons pas jusqu'à écrire la démonstration complète.

(iii) La matrice I_n est une matrice diagonale (donc triangulaire) dont la diagonale ne comporte que des 1 : on obtient son déterminant grâce à la proposition 6.2.2.

□

Soulignons quelques conséquences pratiques très importantes.

Proposition 6.7 – Déterminant et opérations de la méthode du pivot

(a) ***Ajout d'un multiple d'une ligne à une autre :*** *Si $(u_1, ..u_n)$ est une famille de n vecteurs de $\mathbb{K}^n$, si i et j sont deux entiers de $\{1, .., n\}$ avec $j \neq i$ et et si α est un scalaire quelconque, alors*

$$\boldsymbol{Det}(u_1, .., u_{j-1}, (u_j + \alpha u_i), u_{j+1}, .., u_n) = \boldsymbol{Det}(u_1, .., u_n).$$

Autrement dit, une opération du type $L_j \leftarrow L_j + \alpha L_i$ $(i \neq j)$ ne change pas le déterminant.

(b) ***Changement d'échelle sur une seule ligne :***
Si une seule ligne d'une matrice A est multipliée par un scalaire λ, alors $\boldsymbol{Det}(A)$ est multiplié par λ.

(c) ***Échange de deux lignes :***
Si $n \geq 2$, si $u_1, .., u_n$ sont deux vecteurs de $\mathbb{R}^n$ et $i < j$ sont deux entiers distincts entre 1 et n, alors

$$\boldsymbol{Det}(u_1, .., u_{i-1}, u_j, u_{i+1}, .., u_{j-1}, u_i, u_{j+1}, .., u_n) = (-1) \cdot \boldsymbol{Det}(u_1, .., u_n).$$

Lorsqu'on échange deux lignes d'une matrice A, le déterminant est donc changé en son opposé.

Démonstration. — (a) est une conséquence immédiate de (i) et (ii), (b) est un cas particulier de (i). Pour (c) : écrire

$$\begin{aligned}
0 &= \mathbf{Det}\,(u_1, .., u_{i-1}, u_i + u_j, u_{i+1}, .., u_{j-1}, u_i + u_j, u_{j+1}, .., u_n) \\
&= \mathbf{Det}\,(..., u_i, ..., u_i + u_j, ...) + \mathbf{Det}\,(..., u_j, ..., u_i + u_j, ...) \\
&= \mathbf{Det}\,(..., u_i, ..., u_j, ...) + \mathbf{Det}\,(..., u_i, ..., u_i, ...) + \mathbf{Det}\,(..., u_j, ..., u_i, ...) + \mathbf{Det}\,(..., u_j, ..., u_j, ...) \\
&= \mathbf{Det}\,(..., u_i, ..., u_j, ...) + 0 + \mathbf{Det}\,(..., u_j, ..., u_i, ...) + 0
\end{aligned}$$

où la première ligne est justifiée par la propriété (ii) ci-dessus, la deuxième et la troisième par (i), la quatrième par (ii). □

Les opérations (a), (b) et (c) sont celles qui interviennent dans la « méthode du pivot » à laquelle vous êtes habitués. Cela mène au constat suivant :

Théorème 6.8 – Calcul du déterminant à l'aide de la « méthode du pivot »

Par une suite d'opérations des types (a) à (c) ci-dessus, on peut ramener le calcul du déterminant d'une matrice quelconque à celui d'une matrice triangulaire supérieure.

Exemple 6.9. — Utilisons cette méthode pour calculer le déterminant de $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 4 \\ 3 & 2 & 1 \end{pmatrix}$.

Les résultats précédents permettent de suivre ce que devient le déterminant lorsqu'on « réduit » la matrice :

$$\begin{aligned}
\begin{vmatrix} 1 & 2 & 3 \\ 2 & 4 & 4 \\ 3 & 2 & 1 \end{vmatrix} &= \begin{vmatrix} 1 & 2 & 3 \\ 0 & 0 & -2 \\ 0 & -4 & -8 \end{vmatrix} && (L_2 \leftarrow L_2 - 2L_1 \quad \text{et} \quad L_3 \leftarrow L_3 - 3L_1) \\
&= (-2) \cdot (-4) \begin{vmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 1 & 2 \end{vmatrix} && (\text{le } (-2) \text{ vient de la deuxième ligne et le } (-4) \text{ de la troisième}) \\
&= 8 \begin{vmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \\ 0 & 1 & 2 \end{vmatrix} && (\text{ici je recopie simplement}) \\
&= 8 \cdot (-1) \cdot \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{vmatrix} && (\text{échange de lignes}) \\
&= -8 \cdot 1 = -8 && (\text{matrice triangulaire})
\end{aligned}$$

3.1.3. *Conséquences pratiques.* — La proposition 6.7 a plusieurs conséquences spectaculaires. En voici une :

Proposition 6.10 – Une relation de dépendance entre les lignes donne un dét. nul

Si la famille $u_1, ..u_n$ est liée, alors $\boldsymbol{Det}(u_1, .., u_n) = 0$.
Ainsi, si les lignes d'une matrice A forment une famille liée, alors le déterminant de A est nul.

Démonstration. — Si $u_1, ..u_n$ est liée, alors il existe i_0 et des scalaires α_j, $j \in \{1, ..n\} - \{i_0\}$, vérifiant $u_{i_0} = \sum_{\substack{i=1 \\ i \neq i_0}}^{n} \alpha_j \cdot u_j$. Soit A la matrice dont les lignes sont les u_i et B la matrice obtenue à partir de A par l'opération élémentaire $L_{i_0} \leftarrow L_{i_0} - \sum_{\substack{i=1 \\ i \neq i_0}}^{n} \alpha_j \cdot L_j$. Par la propriété (a) ci-dessus, on a donc $\mathbf{Det}(A) = \mathbf{Det}(B)$.

Mais la matrice B comporte une ligne de zéros : or, la formule tout-développée **??** fait apparaître $\mathbf{Det}(B)$ comme une somme de termes dont chacun est un produit de coefficients de B où chaque ligne est représentée une fois, donc où l'un des coefficients est nul. Chacun des termes du développement est donc nul et $\mathbf{Det}(B)$ est nul. □

Voici un cas particulier facile à repérer :

Si deux des vecteurs $u_1, ..u_n$ sont colinéaires, alors $\mathbf{Det}(u_1, .., u_n) = 0$.

Ainsi, si deux lignes d'une matrice A sont proportionnelles, alors le déterminant de A est nul. Par exemple $\begin{vmatrix} 1 & 2 & 3 \\ 2 & 0 & 7 \\ 3 & 6 & 9 \end{vmatrix} = 0$.

Exemple 6.11. — Le déterminant $\begin{vmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 3 & 2 & 1 \end{vmatrix}$ est nul, puisque la deuxième ligne est combinaison linéaire des lignes 1 et 3 (c'est le quart de leur somme).

Voici une deuxième conséquence très utile de la proposition 6.7

Proposition 6.12 – Changement d'échelle sur toutes les lignes en même temps

Pour toute matrice A de $\mathcal{M}_n(\mathbb{K})$ et pour tout λ de $\{1, .., n\}$, on a

$$\boldsymbol{Det}(\lambda A) = \lambda^n \cdot \boldsymbol{Det}(A)$$

Ne pas oublier la puissance, qui dépend de la taille de la matrice!!!

Exemple 6.13. — $\begin{vmatrix} 3 & 3 \\ 3 & 0 \end{vmatrix} = 9 \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix}$, tandis que $\begin{vmatrix} 3 & 3 & 0 \\ 0 & 0 & 3 \\ 3 & 6 & -3 \end{vmatrix} = 27 \cdot \begin{vmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 2 & -1 \end{vmatrix}$.

3.2. Pourquoi ces propriétés pratiques caractérisent le déterminant. — Nous allons maintenant démontrer que les propriétés apparues dans la proposition 6.7 expliquent la forme très particulière du déterminant. Pour cela, il nous faut les reformuler en termes très abstraits.

Définition 6.14 – Forme k-linéaire alternée

Soit E un $\mathbb{K}$ espace vectoriel, k un entier naturel non nul et

$$\Phi : E^k \to \mathbb{K}$$
$$(u_1, ..u_k) \mapsto \Phi(u_1, ..., u_k)$$

une application qui, à partir de k vecteurs de E, produit un nombre.

(a) On dit que Φ est une **forme k-linéaire** si lorsqu'on fixe un i de $\{1, ..k\}$ et des vecteurs $u_1, ..u_{i-1}, u_{i+1}, u_k$ de E, l'application

$$E \to \mathbb{K}$$
$$v \mapsto \Phi(u_1, .., u_{i-1}, v, u_{i+1}, .., u_k)$$

est linéaire.

(b) On dit que Φ est **alternée** lorsque si deux des vecteurs $(u_1, ..u_k)$ sont identiques, alors $\Phi(u_1, .., u_k) = 0$.

Par exemple, l'application $(u_1, .., u_n) \mapsto \mathbf{Det}(u_1, .., u_n)$ est une forme n-linéaire alternée sur l'espace $E = \mathbb{K}^n$.

On vérifie aisément que si Φ_1 et Φ_2 sont deux formes k-linéaires alternées sur un $\mathbb{K}$-espace vectoriel E et si λ_1, λ_2 sont deux scalaires de $\mathbb{K}$, alors l'application $\lambda_1\Phi_1 + \lambda_2\Phi_2$ est une forme k-linéaire alternée. Ainsi, l'ensemble des formes k-linéaires alternées sur E est un sous-espace vectoriel de $\mathcal{F}(E^k, \mathbb{K})$. On le note traditionnellement $\bigwedge^k(E)$.

Théorème 6.15 – Les formes n-linéaires alternées, en dimension n, sont proportionnelles

Soit E un espace vectoriel de dimension finie. Notons n la dimension de E. Alors l'espace $\bigwedge^n(E)$ est de dimension 1.
*En conséquence, si Φ est une forme n-linéaire alternée sur $\mathbb{K}^n$, alors $\Phi = \lambda \cdot$ **Det** pour un certain scalaire λ.*

Ce théorème peut paraître abstrait, mais nous allons voir que les propriétés les plus concrètes du déterminant en découlent.

Démonstration du théorème. — Nous commencerons par montrer le théorème lorsque $E = \mathbb{K}^n$, puis nous en déduirons le résultat pour E quelconque.

- Soit $\Phi : (\mathbb{K}^n)^n \to \mathbb{K}$ une forme n-linéaire alternée sur K^n. Nous allons donner une formule complètement développée pour $\Phi(u_1, ..., u_n)$ lorsque $u_1, .., u_n$ sont des vecteurs de E, et nous la comparerons à la formule tout-développée du numéro 2.3.

  Soient $\mathbf{u}_1, .., \mathbf{u}_n$ des vecteurs de E. Notons $(\mathbf{e}_1, .., \mathbf{e}_n)$ la base canonique de $\mathbb{K}^n$ et écrivons

$$\mathbf{u}_1 = a_{11}\mathbf{e}_1 + a_{12}\mathbf{e}_2 + ... + a_{1n}\mathbf{e}_n,$$
$$\mathbf{u}_2 = a_{21}\mathbf{e}_1 + a_{22}\mathbf{e}_2 + ... + a_{2n}\mathbf{e}_n,$$

etc.

En utilisant la multilinéarité, on peut alors développer

$$\begin{aligned}
\Phi(u_1,..,u_n) &= \Phi\left(a_{11}\mathbf{e}_1 + a_{12}\mathbf{e}_2 + ... + a_{1n}, \quad \mathbf{u}_2, \quad .., \quad \mathbf{u}_n\right) \\
&= \sum_{i=1}^{n} a_{1i} \cdot \Phi(\mathbf{e}_i, \mathbf{u}_2, .., \mathbf{u}_n) \\
&= \sum_{i=1}^{n} a_{1i} \cdot \Phi(\ \mathbf{e}_i, \quad a_{21}\mathbf{e}_1 + a_{22}\mathbf{e}_2 + ... + a_{2n}\mathbf{e}_n, \quad \mathbf{u}_3, .., \quad \mathbf{u}_n) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{1i}a_{2j}\Phi(\ \mathbf{e}_i, \mathbf{e}_j, \mathbf{u}_3, .., \quad \mathbf{u}_n), \\
&= \ \text{etc} \\
&= \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} \cdots \sum_{i_{n-1}=1}^{n}\sum_{i_n=1}^{n} a_{1i_1}a_{2i_2}...a_{(n-1)i_{n-1}}a_{ni_n}\Phi(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, ..., \mathbf{e}_{i_n}).
\end{aligned}$$

En remarquant que si deux des i_k, $k \in \{1, ..n\}$, sont identiques, alors on a $\Phi(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, ..., \mathbf{e}_{i_n})$, on fait apparaître $\Phi(u_1, .., u_n)$ comme une combinaison de termes de la forme $a_{1i_1}a_{2i_2}...a_{(n-1)i_{n-1}}a_{ni_n}$, où l'application $(1, 2, ..., n) \to (i_1, i_2, .., i_n)$ est bijective. On peut donc réécrire

$$\Phi(u_1, .., u_n) = \sum_{\sigma \in \mathfrak{S}_n} \varphi(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} ... a_{n\sigma(n)}$$

où nous avons noté $\varphi(\sigma) = \Phi(\mathbf{e}_{\sigma(1)}, \mathbf{e}_{\sigma(2)}, ..., \mathbf{e}_{\sigma(n)})$.

Si nous avions étudié plus en détail l'application ε : qui intervient dans la formule tout-développée du §2.3, nous pourrions analyser plus finement l'application φ et constater que $\varphi = \lambda\varepsilon$, où λ est le scalaire $\Phi(\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n)$. Nous n'écrirons pas les détails.

- Revenons à un espace vectoriel E quelconque, supposé de dimension n. On sait qu'il existe un isomorphisme $\iota : E \to \mathbb{K}^n$ entre E et $\mathbb{K}^n$. Si Φ est une forme n-linéaire alternée sur $\mathbb{K}^n$, alors on vérifie sans peine que l'application

$$\begin{aligned}
\tilde{\Phi} : E^n &\to \mathbb{K} \\
(u_1, ...u_n) &\mapsto \Phi(\iota(u_1), \iota(u_2), ..., \iota(u_n))
\end{aligned}$$

est une forme n-linéaire alternée sur E.

Nous venons de décrire une opération produisant un élément de l'espace $\Lambda^n(E)$ à partir d'un élément de l'espace $\Lambda^n(\mathbb{K}^n)$, donc une application

$$\begin{aligned}
\Lambda^n(\mathbb{K}^n) &\to \Lambda^n(E) \\
\Phi &\mapsto \tilde{\Phi}.
\end{aligned}$$

En observant les définitions, on s'aperçoit que cette application est linéaire. De plus, elle est bijective : sa bijection réciproque est l'application qui, à une forme linéaire Ψ sur E, associe la forme linéaire $(v_1, ...v_n) \mapsto \Phi(\iota^{-1}(v_1), \iota^{-1}(v_2), ..., \iota^{-1}(v_n))$ sur $\mathbb{R}^n$.

Ainsi, nous avons obtenu un isomorphisme entre les espaces $\Lambda^n(E)$ et $\Lambda^n(\mathbb{K}^n)$; ils ont donc la même dimension. Cette dimension commune est 1 puisqu'on a vu que $\Lambda^n(\mathbb{K}^n)$ est de dimension 1.

$\square$

3.3. Premières conséquences pratiques. —

Proposition 6.16 – Développement par rapport à une autre colonne que la première

Soit n un entier de $\mathbb{N}^\star$. Fixons un entier j entre 1 et n. Alors pour toute matrice A de $\mathcal{M}_n(\mathbb{K})$,

$$\boldsymbol{Det}(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \boldsymbol{Det}(A^\star_{ij}).$$

Démonstration. — L'application

$$\begin{aligned} \Phi : (\mathbb{K}^n) &\to \mathbb{K} \\ (u_1, .., u_n) &\mapsto \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \mathbf{Det}(A^\star_{ij}), \text{ où } A \text{ est la matrice dont les lignes sont les } u_i, \end{aligned}$$

est une forme n-linéaire alternée sur $\mathbb{K}^n$ (la démonstration est identique à celle des propriétés (i) et (ii) du déterminant). Elle peut donc s'écrire $\lambda \cdot \mathbf{Det}$ où λ est un scalaire. Mais si $(u_1, ..u_n)$ est la famille donnant les lignes de la matrice identité, $\Phi(u_1, ..u_n) = \mathbf{Det}(u_1, ..u_n) = 1$, donc λ doit être égal à 1 et $\Phi = \mathbf{Det}$. □

Remarque 6.17. — Comment retenir l'alternance de signes.

Vous pouvez retenir que le développement par rapport à la k-ème colonne est la somme alternée des termes $(a_{ik} A^\star_{ik})$ avec les signes comme dans les tableaux suivants :

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix}, \quad \begin{vmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{vmatrix}, \text{ etc}$$

Exemple 6.18. — Pour calculer le déterminant $\begin{vmatrix} 1 & 0 & 3 \\ 3 & 0 & 5 \\ 2 & 2 & 2 \end{vmatrix}$, on peut s'appuyer sur la première colonne, ce qui donne

$$\begin{vmatrix} 1 & 0 & 3 \\ 3 & 0 & 5 \\ 2 & 2 & 2 \end{vmatrix} = 1 \cdot \begin{vmatrix} 0 & 5 \\ 2 & 2 \end{vmatrix} - 3 \cdot \begin{vmatrix} 0 & 3 \\ 2 & 2 \end{vmatrix} + 2 \begin{vmatrix} 0 & 3 \\ 0 & 5 \end{vmatrix}$$

mais on peut aussi s'appuyer sur la deuxième colonne, ce qui donne (n'oubliez pas qu'il faut faire attention à l'alternance des signes) :

$$\begin{vmatrix} 1 & 0 & 3 \\ 3 & 0 & 5 \\ 2 & 2 & 2 \end{vmatrix} = -0 \cdot \begin{vmatrix} 3 & 5 \\ 2 & 2 \end{vmatrix} + 0 \cdot \begin{vmatrix} 1 & 3 \\ 2 & 2 \end{vmatrix} - 2 \begin{vmatrix} 1 & 3 \\ 3 & 5 \end{vmatrix}.$$

Le deuxième chemin nécessite moins de calculs, n'est-ce pas ?

Proposition 6.19 – Déterminant de la transposée

Si A est une matrice carrée de $\mathcal{M}_n(\mathbb{K})$ et A^T sa transposée, alors

$$\boldsymbol{Det}(A^T) = \boldsymbol{Det}(A).$$

Démonstration. — Notons $A = (a_{ij})_{1\le i,j\le n}$. La définition du déterminant est

$$\mathbf{Det}(A) = \sum_{i=1}^{n} (-1)^{i+1} \cdot a_{i1} \cdot A^{\star}_{i1}$$

mais on peut aussi calculer le déterminant en développant par rapport à la deuxième colonne :

$$\mathbf{Det}(A) = \sum_{i=1}^{n} (-1)^{i+2} \cdot a_{i2} \cdot A^{\star}_{i2}$$

ou par rapport à d'autres colonnes :

$$\vdots$$

$$\mathbf{Det}(A) = \sum_{i=1}^{n} (-1)^{i+n} \cdot a_{in} \cdot A^{\star}_{in}$$

On obtient ainsi n descriptions du déterminant (une par colonne). En les sommant, on constate que

$$n \cdot \mathbf{Det}(A) = \sum_{j=1}^{n} \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot A^{\star}_{ij}.$$

Dans cette formule les lignes et colonnes jouent des rôles parfaitement symétriques. En écrivant l'égalité analogue pour $\mathbf{Det}(A^T)$, on constate que $n \cdot \mathbf{Det}(A^T)$ est donnée par la *même* formule... et donc $\mathbf{Det}(A) = \mathbf{Det}(A^T)$. □

Corollaire 6.20 – Développement par rapport à une ligne

Soit n un entier de $\mathbb{N}^{\star}$. Fixons un entier i entre 1 et n. Alors pour toute matrice A de $\mathcal{M}_n(\mathbb{K})$,

$$\boldsymbol{Det}(A) = \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \boldsymbol{Det}(A^{\star}_{ij}).$$

Exemple 6.21. — Pour calculer le déterminant $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 9 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & -2 & -3 & 1 \end{vmatrix}$, on a tout intérêt à s'appuyer sur la deuxième ligne : en veillant aux signes (par exemple en ayant en tête la répartition dessinée page 154) on constate que

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 9 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & -2 & -3 & 1 \end{vmatrix} = -9 \begin{vmatrix} 2 & 3 & 4 \\ 2 & 2 & 2 \\ -2 & 3 & 1 \end{vmatrix} = -9 \begin{vmatrix} 2 & 3 & 4 \\ 0 & -1 & -1 \\ 0 & 0 & 5 \end{vmatrix} = (-9) \cdot (2 \cdot (-1) \cdot 5) = 90$$

Remarque 6.22. — Une autre conséquence de l'égalité $\mathbf{Det}(A) = \mathbf{Det}(A^T)$ est que les propriétés pratiques du paragraphe 3.1 restent vraies si on effectue des combinaisons ou échanges de colonnes plutôt que de lignes. Par exemple, si deux colonnes sont identiques alors le déterminant est nul, etc.

4. Déterminant d'un produit

Voici un théorème fondamental, très surprenant si on songe aux formules compliquées qui définissent le déterminant, bien moins surprenant si on se souvient de notre motivation pour étudier ces sommes étranges (l'étude du dénominateur commun aux coordonnées d'un vecteur $X \in \mathbb{K}^n$ solution d'un système du type $AX = Y$ avec $A \in \mathcal{M}_n(\mathbb{K})$ et $Y \in \mathbb{K}^n$ à coefficients entiers).

Théorème 6.23 – Le déterminant d'un produit, c'est le produit des déterminants

Si A et B sont deux matrices de $\mathcal{M}_n(\mathbb{K})$, alors

$$\boldsymbol{Det}(AB) = \boldsymbol{Det}(A)\boldsymbol{Det}(B).$$

Démonstration. — Fixons une matrice B de $\mathcal{M}_n(\mathbb{K})$. Considérons l'application

$$\begin{aligned}\Phi : (\mathbb{K}^n)^n &\to \mathbb{K}\\ (u_1, .., u_n) &\mapsto \mathbf{Det}(MB), \text{ où } M \text{ est la matrice dont les lignes sont les } u_i.\end{aligned}$$

Constatons que c'est une forme n-linéaire alternée sur $\mathbb{K}^n$:

- La linéarité par rapport à chaque variable se vérifie sans peine, car le coefficient en position (i, j) dans MA est le produit scalaire $u_i \cdot C_j$, où C_j est la j-ème colonne de B.
- Si deux des u_i sont identiques, deux des lignes de la matrice M dont les lignes sont les u_i sont identiques, mais en observant les lignes de MB, on constate alors que deux aussi sont identiques : $\mathbf{Det}(MB) = 0$, donc $\Phi(u_1, ..., u_n) = 0$.

Par le théorème fondamental sur les formes multilinéaires alternées (théorème 6.15), Φ est proportionnelle au déterminant : ainsi, il existe un scalaire λ vérifiant $\forall M \in \mathcal{M}_n(\mathbb{K}),\ \mathbf{Det}(MB) = \lambda\mathbf{Det}(M)$. En prenant $M = I_n$, on obtient $\lambda \cdot 1 = \mathbf{Det}(B)$.

Nous avons bien démontré : $\forall B \in \mathcal{M}_n(\mathbb{K}), \forall M \in \mathcal{M}_n(\mathbb{K}),\ \mathbf{Det}(MB) = \mathbf{Det}(B)\mathbf{Det}(M)$. □

Le théorème ci-dessus est probablement le plus important de la théorie des déterminants. Parmi ses conséquences qui sont omniprésentes en mathématiques, citons :

Théorème 6.24 – Déterminant et inversibilité

Une matrice A est inversible si et seulement si $\boldsymbol{Det}(A) \neq 0$. Dans ce cas $\boldsymbol{Det}(A^{-1}) = \frac{1}{\boldsymbol{Det}(A)}$.

Démonstration. —

- Si A est inversible, alors $A(A^{-1}) = I_n$, et alors $\mathbf{Det}(A)\mathbf{Det}(A^{-1}) = \mathbf{Det}(I_n) = 1$, donc $\mathbf{Det}(A)$ ne peut pas être nul et on a bien $\mathbf{Det}(A^{-1}) = \frac{1}{\mathbf{Det}(A)}$;
- Si A n'est pas inversible, alors il existe une relation de liaison entre ses lignes (voir page 105), et nous avons vu que le déterminant de A est nul dans ce cas (voir la proposition 6.10).

□

Théorème 6.25 – Deux matrices semblables ont même déterminant

Si A et B sont deux matrices semblables, alors $\boldsymbol{Det}(A) = \boldsymbol{Det}(B)$.

Démonstration. — Dire que A et B sont semblables, c'est dire qu'il existe une matrice inversible P vérifiant $A = PBP^{-1}$, mais alors $\mathbf{Det}(A) = \mathbf{Det}(P^{-1})\mathbf{Det}(B)\mathbf{Det}(P)$, mais on vient de voir que $\mathbf{Det}(P^{-1}) = \frac{1}{\mathbf{Det}(P)}$ et le produit $\frac{1}{\mathbf{Det}(P)}\mathbf{Det}(B)\mathbf{Det}(P)$ est un produit de nombre réels, il vaut donc bien sûr $\mathbf{Det}(B)$. □

Définition 6.26 – Déterminant d'un endomorphisme abstrait

Soient E un espace vectoriel de dimension finie et f un endomorphisme de E.
Si $\mathcal{B}$ et $\mathcal{B}'$ sont deux bases de E, alors les matrices $\mathcal{M}_{\mathcal{B}}(f)$ et $\mathcal{M}_{\mathcal{B}'}(f)$ ont le même déterminant.
La valeur commune à toutes les matrices de la forme $\mathcal{M}_{\mathcal{B}}(f)$ ($\mathcal{B}$ base de E) est appellée *déterminant de* f, notée $\mathbf{Det}(f)$.

5. Déterminant et volume

5.1. Déterminants de familles de vecteurs et volumes de parallélépipèdes. — Si u et v sont deux vecteurs de $\mathbb{R}^2$, le *parallélogramme porté par u et v* est l'ensemble

$$\{\alpha u + \beta v,\ \alpha \in [0,1],\ \beta \in [0,1]\}.$$

s Si u, v et w sont deux vecteurs de $\mathbb{R}^3$, le *parallélépipède porté par u, v et w* est l'ensemble

$$\{\alpha u + \beta v + \gamma w,\ \alpha \in [0,1],\ \beta \in [0,1],\ \gamma \in [0,1]\}.$$

Plaçons-nous dans $\mathbb{R}^2$ et fixons deux vecteurs u et v. Nous allons montrer qu'il y a un lien entre l'aire du parallélogramme porté par u et v et le déterminant de u et v.

En utilisant les coordonnées dans la base canonique, notons $u = \begin{pmatrix} a \\ b \end{pmatrix}$ et $v = \begin{pmatrix} c \\ d \end{pmatrix}$ et représentons ce parallélogramme.

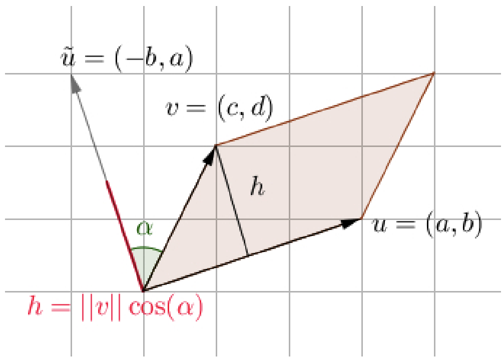


FIGURE 2. Parallélogramme porté par deux vecteurs de $\mathbb{R}^2$.

Rappelons que son aire peut être calculée (ou définie !) par la formule « base × hauteur » : avec les notations de la figure ci-dessus (où je vous fais confiance pour vous rappeler la trigonométrie nécessaire à montrer que la longueur rouge est $\|v\|\,|\cos(\alpha)|$), l'aire du parallélogramme porté par u et v est $\|u\| \times h = \|u\|\,(\|v\|\,|\cos(\alpha)|)$.

Or, on rappelle que le produit scalaire entre les vecteurs $\tilde{u} = (-b, a)$ et $v = (c, d)$ peut être décrit soit comme $\|\tilde{u}\|\,\|v\|\cos(\alpha)$, soit comme $-bc + ad$, qui n'est autre que $\mathbf{Det}(u, v)$. Passant aux valeurs absolues, on obtient

$$|\mathbf{Det}(u,v)| = \|\tilde{u}\|\,\|v\|\,|\cos(\alpha)| = \|u\|\,\|v\|\,|\cos(\alpha)| = \|u\| \times h = \text{aire du parallélogramme porté par } u \text{ et } v.$$

Nous avons montré la partie (i) du théorème suivant, très surprenant si on songe à nos motivations initiales sur les dénominateurs apparaissant dans les solutions de systèmes linéaires à coefficients entiers :

Théorème 6.27 – Déterminant et volume dans l'espace à deux ou trois dimensions

(i) Soient u et v deux vecteurs de $\mathbb{R}^2$. L'aire du parallélogramme porté par u et v est $|\boldsymbol{Det}(u,v)|$.

(ii) Soient u, v et w trois vecteurs de $\mathbb{R}^3$. Le volume du parallélépipède porté par u,v et w est $|\boldsymbol{Det}(u,v,w)|$.

Je ne démontrerai pas (ii) ici (ce n'est pas vraiment plus difficile, mais c'est long).

Le résultat ci-dessus est un bon complément au fait, déjà montré, que

- dans le plan $\mathbf{Det}(u,v)$ est nul si et seulement si u et v sont colinéaires (parallélogramme aplati),
- dans l'espace, $\mathbf{Det}(u,v,w)$ est nul si et seulement si u, v et w sont coplanaires (parallélépipède aplati).

Ce théorème permet même d'imaginer une définition du « volume d'un parallélépipède en dimensions supérieures » :

Définition 6.28 – Une définition du volume d'un parallélotope dans l'espace $\mathbb{R}^n$, $n > 3$

Si $(u_1, .., u_n)$ est une famille de n vecteurs de $\mathbb{R}^n$ et P est l'ensemble $\left\{ \sum_{i=1}^{n} \alpha_i v_i,\ (\alpha_1, .., \alpha_n) \in [0,1]^n \right\}$, alors on dira que P est le parallélotope porté par $u_1, .., u_n$ et on appellera volume de P, noté $\mathbf{Vol}(P)$, le nombre $|\mathbf{Det}(u_1, .., u_n)|$.

Les objets qui ne sont pas des parallélotopes ont un volume, naturellement ; il est lié au déterminant, mais expliquer comment dépasserait largement le cadre de ce cours. La *théorie de la mesure* et la *formule de changement de variables dans les intégrales multiples* vous l'expliqueront dans la suite de vos études (en L2 et surtout L3).

5.2. Déterminant d'un endomorphisme et transformation des volumes. —

Proposition 6.29 – Déterminant et facteur de transformation du volume

Soient f un endomorphisme de $\mathbb{R}^n$ et $\mathcal{U} = (u_1, .., u_n)$ une famille de n vecteurs de $\mathbb{R}^n$.
Le volume du parallélotope porté par $\mathcal{U}$ et le volume du parallélotope porté par $f(\mathcal{U}) = (f(u_1), .., f(u_n))$ sont reliés par

$$\mathbf{Det}(f(u_1), .., f(u_n)) = \mathbf{Det}(f) \cdot \mathbf{Det}(u_1, .., u_n)$$

autrement dit,

$$\mathbf{Vol}[f(\mathcal{U})] = \mathbf{Det}(f) \cdot \mathbf{Vol}(\mathcal{U}).$$

Démonstration. — Notons A la matrice de f dans la base canonique, si bien que f est l'application linéaire T_A associée à A. Nous devons calculer le déterminant de la famille $(f(u_1), .., f(u_n)) = (Au_1, .., Au_n)$. C'est (en combinant la définition du déterminant d'une famille de vecteurs, page 148 et la propriété sur le déterminant d'une transposée, proposition 6.19) le déterminant de la matrice B dont les colonnes sont $Au_1, .., Au_n$. Or, en observant la définition du produit matriciel, on s'aperçoit que pour chaque j, la j-ème colonne du produit AB est Au_j.

Ainsi

$$\mathbf{Det}((f(u_1), .., f(u_n))) = \mathbf{Det}(Au_1, .., Au_n) = \mathbf{Det}(AB) = \mathbf{Det}(A)\mathbf{Det}(B) = \mathbf{Det}(f)\mathbf{Det}(u_1, .., u_n)$$

la dernière égalité vient du fait que A est la matrice de f dans une certaine base, donc $\mathbf{Det}(f) = \mathbf{Det}(A)$ par définition de $\mathbf{Det}(f)$; dans le même temps $\mathbf{Det}(B) = \mathbf{Det}(u_1, .., u_n)$ par définition de B. □

6. (⋆) Complément : retour aux systèmes linéaires

6.1. Les formules de Cramer. —

Théorème 6.30 – Formules de Cramer pour la résolution des systèmes linéaires

Soient A une matrice de $\mathcal{M}_n(\mathbb{K})$ et Y un vecteur de $\mathbb{R}^n$.
Considérons le système linéaire $AX = Y$, d'inconnue le vecteur $X = (x_1, .., x_n)$ de $\mathbb{R}^n$.
Notons C_1, ..., C_n les colonnes de A.
Si A est inversible, alors l'unique solution $X = A^{-1}Y$ est donnée par :

$$x_1 = \frac{\mathbf{Det}(Y, C_2, .., C_n)}{\mathbf{Det}(A)}, \qquad x_2 = \frac{\mathbf{Det}(C_1, Y, C_3, .., C_n)}{\mathbf{Det}(A)}, \qquad x_n = \frac{\mathbf{Det}(C_1, .., C_{n-1}, Y)}{\mathbf{Det}(A)}.$$

Ces formules montrent notamment que si A et Y sont à coefficients entiers et si $AX = Y$ a une unique solution, les coordonnées de l'unique solution du système sont des nombres rationnels (puisque le déterminant d'une matrice à coefficients entiers est évidemment à coefficients entiers compte tenu de la définition ou de la formule tout-développée), et que $\mathbf{Det}(A)$ « apparaît au dénominateur de chacune de ces coordonnées ».

Remarque 6.31. — En général, il est *beaucoup plus long* de résoudre un système de Cramer à l'aide des formules ci-dessus qu'avec la méthode de Gauss. Pour les systèmes de petite taille, elles peuvent cependant être utiles.

6.2. Inversion des matrices à l'aide de déterminants. —

Vocabulaire : mineurs, cofacteurs, comatrice. — Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$

Si i et j sont des entiers de $\{1, .., n\}$, on appelle *mineur principal de A d'indice (i, j)* le déterminant $\mathbf{Det}(A^\star_{ij})$,

et *cofacteur de A d'indice (i, j)* le nombre $\mathrm{Cof}_{ij}[A] = (-1)^{i+j}\mathbf{Det}(A^\star_{ij})$ (un cofacteur est donc simplement le déterminant d'une sous-matrice, mais au signe près). On note $\mathbf{Com}(A)$ la matrice $(\mathrm{Cof}_{ij}[A])_{ij}$ de $\mathcal{M}_n(\mathbb{K})$, souvent appelée *matrice des cofacteurs*, ou *comatrice*, de A.

Proposition 6.32 – Inverse d'une matrice et matrice des cofacteurs

Si A est inversible, alors

$$A^{-1} = \frac{1}{\mathbf{Det}(A)} \cdot \mathbf{Com}(A)^T.$$

Deux conséquences classiques. — (a) Si A est à coefficients entiers et inversible, alors A^{-1} est à coefficients rationnels.

(b) Si A est à coefficients entiers et inversible, alors A^{-1} est à coefficients entiers si et seulement [5] si $\mathbf{Det}(A) = \pm 1$.

5. La proposition précédente montre l'implication « si $\mathbf{Det}(A) = \pm 1$, alors A^{-1} est à coefficients entiers ; pour l'autre implication, si A^{-1} est à coefficients entiers alors $\mathbf{Det}(A)\mathbf{Det}(A^{-1}) = \mathbf{Det}(I_n) = 1$, donc $\mathrm{Det}(A)$ est un entier dont le produit avec un autre entier vaut 1, ce qui n'est le cas que de 1 et -1.

Démonstration. — La matrice A^{-1} est l'unique solution de l'équation $AB = I_n$, où l'inconnue B est une *matrice* de $\mathcal{M}_n(\mathbb{K})$.

Notons $C_1, ...C_n$ les colonnes de A^{-1}. Rappelons que la première colonne d'une matrice est donnée par le produit entre cette matrice et le premier vecteur $e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ de la base canonique. On a ainsi $ABe_1 = AC_1$ et $I_n e_1 = e_1$, donc C_1 est l'unique solution de l'équation $AC_1 = e_1$.

De même, pour chaque j, C_j est l'unique solution du système linéaire $AC_j = e_j$, où e_j est le j-ème vecteur de la base canonique de $\mathbb{K}^n$.

On peut ensuite utiliser les formules de Cramer pour trouver C_j : la i-ème coordonnée de C_j, donc est donnée par $\dfrac{\mathbf{Det}\,(D_1, .., D_{i-1}, e_j, D_{i+1}, .., D_n)}{\mathbf{Det}(A)}$, où $D_1, ..D_n$ sont les colonnes de A.

En développant $\mathbf{Det}\,(D_1, .., D_{i-1}, e_j, D_{i+1}, .., D_n)$, par rapport à la i-ème colonne, on constate qu'il est égal à $(-1)^{i+j}\mathbf{Det}(A^{\star}_{ji})$, qui est le coefficient en position (i, j) dans $\mathbf{Com}(A)^T$. □

CHAPITRE 7

VALEURS PROPRES ET VECTEURS PROPRES

1. Valeurs propres, vecteurs propres, espaces propres

1.1. Valeurs propres d'un endomorphisme. —

Définition 7.1 – Valeurs propres et vecteurs propres d'un endomorphisme

Soient E un $\mathbb{K}$-espace vectoriel et f un endomorphisme de E.

(i) Un scalaire λ de $\mathbb{K}$ est une *valeur propre de f* s'il existe un vecteur *non nul* x de E vérifiant $f(x) = \lambda x$.

(ii) Si x est un vecteur *non nul* de E vérifiant $f(x) = \lambda x$, on dit que x est *vecteur propre de f pour la valeur propre λ*.

Vocabulaire pour les matrices. — Si A est une matrice de $\mathcal{M}_n(\mathbb{K})$, on parlera des valeurs propres et vecteurs propres de A pour les valeurs propres et vecteurs propres de l'endomorphisme T_A de $\mathbb{K}^n$.

***Exemple 7.2* (Un exemple ave des matrices).** — Si A est la matrice $\begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$,

- on a $T_A\left[\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right] = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$, donc 4 est une valeur propre de T_A et $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ est vecteur propre de T_A pour la valeur propre 4 ;
- on a $T_A\left[\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right] = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, donc 0 est valeur propre de T_A et $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ est vecteur propre de T_A pour la valeur propre 0.

***Exemple 7.3* (Fonction sinus et opérateur dérivée seconde).** — Plaçons-nous dans $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$. Pour chaque ω de $\mathbb{R}$, notons s_ω la fonction $x \mapsto \sin(\omega x)$. Considérons l'endomorphisme $\Phi : f \mapsto f''$ de E. Soit ω un réel. Les résultats usuels sur les dérivées de sinus et cosinus impliquent $s''_\omega = -\omega^2 s_\omega$: ainsi, s_ω est un vecteur propre de Φ pour la valeur propre $-\omega^2$.

***Exemple 7.4* (Valeurs propres sur $\mathbb{R}$ et valeurs propres sur $\mathbb{C}$).** — La matrice $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ n'a pas de valeur propre réelle.

En effet, si λ était valeur propre de A, il existerait un vecteur non nul $u = \begin{pmatrix}\alpha\\ \beta\end{pmatrix}$ de $\mathbb{R}^2$ vérifiant $Au = \lambda u$, mais alors on aurait $\begin{pmatrix}-\beta\\ \alpha\end{pmatrix} = \lambda\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$, ce qui signifie $\alpha = \lambda\beta$ et $\beta = -\lambda\alpha$... mais alors on aurait $\alpha = -\lambda^2\alpha$, et les deux côtés de cette égalité sont de signe contraire, sauf si $\alpha = 0$, et on aurait aussi $\beta = -\lambda^2\beta$, ce qui implique $\beta = 0$. Le vecteur u serait alors nul.

Nous avons montré que si λ est un réel quelconque, il n'existe aucun vecteur *non nul* de $\mathbb{R}^2$ vérifiant $Au = \lambda u$, et ainsi λ ne peut être valeur propre de A.

(c, bis) En revanche, si on voit $A = \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$ comme une matrice de $\mathcal{M}_n(\mathbb{C})$ et si on considère l'endomorphisme T_A de $\mathbb{C}^n$, alors A *admet des valeurs propres complexes non réelles.* En effet, si λ est un nombre *complexe* et si on cherche un vecteur $u = \begin{pmatrix}\alpha\\ \beta\end{pmatrix}$ de $\mathbb{C}^2$ (non plus de $\mathbb{R}^2$) vérifiant $Au = \lambda u$, alors les considérations de l'exemple (c) montrent que pour que α ou β puisse être non nul, il est nécessaire que $-\lambda^2$ soit égal à 1, c'est-à-dire que λ soit égal à i ou à $-i$. Se pourrait-il que i et $-i$ soient des valeurs propres complexes de A ?

C'est effectivement le cas : on constate que $A\begin{pmatrix}1\\ -i\end{pmatrix} = \begin{pmatrix}i\\ 1\end{pmatrix} = i\begin{pmatrix}1\\ -i\end{pmatrix}$ et que $A\begin{pmatrix}1\\ i\end{pmatrix} = \begin{pmatrix}-i\\ 1\end{pmatrix} = (-i)\begin{pmatrix}1\\ i\end{pmatrix}$, ainsi $\begin{pmatrix}1\\ -i\end{pmatrix}$ est un vecteur propre de A pour la valeur propre i et $\begin{pmatrix}1\\ i\end{pmatrix}$ en est un vecteur propre pour la valeur propre $-i$.

1.2. Sous-espaces propres d'un endomorphisme. —

Définition 7.5 – Sous-espace propre associé à une valeur propre

Soient E un $\mathbb{K}$-espace vectoriel et f un endomorphisme de E ; soit λ un scalaire de $\mathbb{K}$.
Les assertions suivantes sont équivalentes :

(i) le nombre λ est valeur propre de f,

(ii) l'endomorphisme $(f - \lambda\mathrm{id}_E)$ n'est pas injectif,

(iii) le noyau $\mathrm{Ker}(f - \lambda\mathrm{id}_E)$ n'est pas réduit à $\{0_E\}$.

Dans ce cas, le sous-espace vectoriel $V_\lambda = \mathrm{Ker}(f - \lambda\mathrm{id}_E)$ est appelé *sous-espace propre de f pour la valeur propre λ.*

Vocabulaire pour les matrices. — Si A est une matrice, on parlera des espaces propres de A pour désigner les espaces propres de T_A ; concrètement si λ est valeur propre de A, alors le sous-espace propre de A pour la valeur propre λ est $\mathrm{Ker}(A - \lambda I_n)$.

Exemple 7.6 **(Noyau et valeur propre 0).** — Dire que 0 est valeur propre de f revient donc à dire que f n'est pas injectif ; dans ce cas le sous-espace propre V_0 pour la valeur propre 0 n'est rien d'autre que le noyau $\mathrm{Ker}(f)$ de f.

Exemple 7.7. — Soit A la matrice $\begin{pmatrix}2 & 1 & 1\\ 1 & 2 & 1\\ 1 & 1 & 2\end{pmatrix}$. On constate que $T_A\left[\begin{pmatrix}1\\ -1\\ 0\end{pmatrix}\right] = \begin{pmatrix}1\\ -1\\ 0\end{pmatrix}$, donc 1 est valeur propre de A.

- Le sous-espace propre de A pour la valeur propre 1 est $V_1 = \mathrm{Ker}(A - I_n)$, c'est-à-dire $\mathrm{Ker}(B)$ où $B = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.
- La matrice B est visiblement de rang 1 et c'est une matrice 3×3, donc son noyau est de dimension 2. En observant les relations entre les colonnes de B, on constate que $u = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ et $v = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$ appartiennent à $\mathrm{Ker}(B)$; puisqu'on sait déjà que $\mathrm{Ker}(B)$ est de dimension 2, on en déduit $V_1 = \mathrm{Ker}(B) = \mathbf{Vect}[u, v]$.

Exemple 7.8. — Reprenant l'exemple 7.3 ci-dessus, plaçons-nous dans $E = \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ et considérons l'endomorphisme $\Phi : f \mapsto f''$ de E. Nous avons vu que sin est vecteur propre de Φ pour la valeur propre -1 ; l'espace propre V_{-1} de Φ pour la valeur propre -1 est l'ensemble des fonctions $\mathcal{C}^\infty$ qui vérifient $f'' = -f$, donc l'ensemble des solutions de l'équation différentielle $f'' + f = 0$: c'est $\mathbf{Vect}[\sin, \cos]$ (vu dans les TD du chapitre 4 et démontré dans un document mis en ligne à cette occasion).

1.3. Relation entre les espaces propres. —

Attention, pour ce qui suit il est important d'être au point sur la notion de somme directe de plus de deux sous-espaces vectoriels : voir page 82.

Proposition 7.9 – Les sous-espaces propres sont en somme directe

Soient E un $\mathbb{K}$-espace vectoriel et f un endomorphisme de E.
Soit k un entier naturel non nul. Soient $\lambda_1, ..., \lambda_k$ des valeurs propres de f deux à deux distinctes.
Les sous-espaces propres associés à $\lambda_1, ..., \lambda_k$ sont en somme directe : on a $V_{\lambda_1} + ... + V_{\lambda_k} = V_{\lambda_1} \oplus ... \oplus V_{\lambda_k}$.

Démonstration. — Fixons un $\mathbb{K}$-espace vectoriel E et un endomorphisme f de E. Nous allons raisonner par récurrence. Pour chaque k de $\mathbb{N}^\star$, notons $\mathcal{P}(k)$ l'énoncé :

> « Si $\lambda_1, .., \lambda_k$ sont des valeurs propres deux à deux distinctes de f et si $x_1, .., x_k$ sont des éléments de $V_{\lambda_1}, .., V_{\lambda_k}$ respectivement, alors l'égalité $x_1 + ... + x_k = \mathbf{0}_E$ implique $x_1 = ... = x_k = \mathbf{0}_E$ »

qui n'est que la traduction de l'énoncé de la proposition où on a inséré la définition de « somme directe de k sous-espaces ».

- L'énoncé $\mathcal{P}(1)$ est évidemment vrai (il n'y a rien à montrer...)
- Soit k un entier de $\mathbb{N}^\star$. Supposons $\mathcal{P}(k)$ vrai. Soient $\lambda_1, .., \lambda_{k+1}$ des valeurs propres deux à deux distinctes de f et $x_1, .., x_{k+1}$ des éléments de $V_{\lambda_1}, .., V_{\lambda_{k+1}}$ respectivement.

  Supposons l'égalité $x_1 + ... + x_{k+1} = \mathbf{0}_E$ et montrons que tous les x_i, $i = 1, .., (k+1)$, sont nuls.

  On constate que
  $$x_{k+1} = -x_1 - ... - x_k \tag{1.1}$$
  donc par linéarité de f
  $$f(x_{k+1}) = -f(x_1) - ... - f(x_k).$$
  Or nous partons d'éléments d'espaces propres de f, donc pour chaque i de $\{1, .., k+1\}$, on a $f(x_i) = \lambda_i x_i$. Ainsi
  $$\lambda_{k+1} x_{k+1} = -\lambda_1 x_1 - ... - \lambda_k x_k. \tag{1.2}$$

En retranchant λ_{k+1} fois l'équation (1.1) à l'équation (1.2), il vient

$$\mathbf{0}_E = (\lambda_{k+1} - \lambda_1)x_1 + ... + (\lambda_{k+1} - \lambda_k)x_k.$$

Mais alors l'énoncé $\mathcal{P}(k)$ donne $\forall i \in \{1, .., k\}$, $(\lambda_{k+1} - \lambda_i)x_i = \mathbf{0}_E$ et comme tous les $(\lambda_{k+1} - \lambda_i)$ sont non-nuls (valeurs propres deux à deux distinctes), tous les x_i, $i = 1, .., k$ sont nuls. Puisque $x_1 + ... + x_k + x_{k+1} = \mathbf{0}_E$, x_{k+1} est nul également, et nous avons montré que tous les x_i, $i = 1, .., (k+1)$ sont nuls.

□

Corollaire 7.10 – Nombre de valeurs propres $\leq$ dimension de l'espace

Soit E un $\mathbb{K}$-espace vectoriel ***de dimension finie*** *et f un endomorphisme de E.*
Si E est de dimension n, alors f a au plus n valeurs propres distinctes.

Démonstration. — Dire que λ est valeur propre de f, c'est dire que $V_\lambda \neq \{0\}$, autrement dit que $\dim(V_\lambda) \geq 1$. Par conséquent si f admet k valeurs propres distinctes, la dimension de $V_{\lambda_1} \oplus ... \oplus V_{\lambda_k}$ est au moins [1] k. Mais $V_{\lambda_1} \oplus ... \oplus V_{\lambda_k}$ est un sous-espace vectoriel de E qui est de dimension n, sa dimension ne peut donc pas dépasser n et on ne peut pas avoir $k > n$ sous peine de contradiction. □

Définition 7.11 – Spectre d'un endomorphisme

L'ensemble des valeurs propres de f est appelé le *spectre* de f, noté **Spectre**(f).

La proposition ci-dessus dit donc que si E est un espace de dimension finie et f appartient à $\mathcal{L}(E)$, alors **Spectre**(f) est un ensemble fini dont le cardinal ne peut dépasser la dimension de E. De même, lorsque A est une matrice carrée à coefficients dans $\mathbb{K}$, on note **Spectre**(A) le spectre de $f = T_A$; comme c'est sur $\mathbb{K}^n$ qu'agit T_A, nous venons de voir que :

Une matrice $n \times n$ admet au plus n valeurs propres.

Exemple 7.12. — Nous avons vu à l'exemple 7.7 que la matrice $A = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$ admet 4 et 0 pour valeurs propres ; elle ne peut en admettre d'autres. Ainsi **Spectre**$(A) = \{0; 4\}$.

1. Ici on utilise le fait qu'en situation de somme directe, $\dim(V_{\lambda_1} \oplus ... \oplus V_{\lambda_k}) = \dim(V_{\lambda_1}) + ... + \dim(V_{\lambda_k})$.

2. Le polynôme caractéristique

Dans tout ce paragraphe (et jusqu'à la fin du cours), on suppose E de dimension finie et $\dim(E) \neq 0$.

2.1. L'équation polynomiale qui permet de trouver les valeurs propres. — Nous avons vu dans la définition 7.5 que λ est valeur propre de f si et seulement si $(f - \lambda \mathrm{id}_E)$ n'est pas injectif. Or, si E est de dimension finie, dire que $(f - \lambda \mathrm{id}_E)$ est non-injectif est équivalent à dire qu'il est non-surjectif, ou encore à dire que $\mathbf{Det}(\lambda \mathrm{id}_E - f) = 0$. Ainsi

$$\lambda \text{ est valeur propre de } f \iff \mathbf{Det}(\lambda \mathrm{id}_E - f) = 0.$$

Trouver les valeurs propres revient donc à résoudre l'*équation caractéristique* $\mathbf{Det}(\lambda \mathrm{id}_E - f) = 0$, d'inconnue $\lambda \in \mathbb{K}$. Nous avons démontré au chapitre 6 que c'est une équation polynomiale (voir page 147) de degré n (voir TD6, exercice 9).

Définition 7.13 – Polynôme caractéristique d'un endomorphisme

Soient E un $\mathbb{K}$-espace vectoriel *de dimension finie* et f un endomorphisme de E.
Pour tout x de $\mathbb{K}$, on pose $\chi_f(x) = \mathbf{Det}(x\, \mathrm{Id}_E - f)$.
Cette formule définit [2] un polynôme χ_f de $\mathbb{K}[X]$, le *polynôme caractéristique de* f. Il a les deux propriétés suivantes :

- Il s'agit d'un polynôme unitaire de degré n.
- Un scalaire λ est racine de χ_f si et seulement si c'est une valeur propre de f.

Cas des matrices. — Si A est une matrice carrée de taille n, on notera χ_A pour χ_{T_A}, où T_A est l'endomorphisme de $\mathbb{K}^n$ canoniquement associé à A : ainsi $\chi_A(x) = \mathbf{Det}(xI_n - A)$ pour tout x de $\mathbb{K}$.

***Exemple 7.14* (Cas des matrices 2×2).** — Si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ avec a, b, c, d dans $\mathbb{K}$, alors le polynôme caractéristique de A est donné par

$$\chi_A(\lambda) = \begin{vmatrix} \lambda - a & -b \\ -c & \lambda - d \end{vmatrix} = (\lambda - a)(\lambda - d) - bc = \lambda^2 - (a+d)\lambda + (ad - bc). \text{ Ainsi :}$$

$$\text{Si } A \text{ est une matrice } 2 \times 2, \text{ alors } \chi_A(X) = X^2 - \mathbf{Tr}(A) \cdot X + \mathbf{Det}(A).$$

Cas des matrices triangulaires. —

$$\text{Si} \quad A = \begin{pmatrix} a_1 & \star & \dots & \star \\ 0 & a_2 & & \\ \vdots & & \ddots & \star \\ 0 & \dots & 0 & a_n \end{pmatrix}, \quad \text{alors :} \quad \forall \lambda \in \mathbb{K},\ \chi_A(\lambda) = \begin{vmatrix} \lambda - a_1 & \star & \dots & \star \\ 0 & \lambda - a_2 & & \\ \vdots & & \ddots & \star \\ 0 & \dots & 0 & \lambda - a_n \end{vmatrix}. \quad \text{En conséquence :}$$

2. Lorsque le corps de base $\mathbb{K}$ est infini, ce qui est toujours le cas dans ce cours, une fonction polynomiale $x \mapsto P(x)$ de $\mathbb{K}$ dans $\mathbb{K}$ est la même chose qu'un polynôme $P \in \mathbb{K}[X]$.

Si A est une matrice triangulaire $n \times n$ de diagonale $a_1, ..., a_n$, alors $\chi_A(X) = (X - a_1)...(X - a_n)$. Les valeurs propres d'une matrice triangulaire sont donc données par ses coefficients diagonaux.

***Exemple 7.15* (Un calcul typique pour une matrice 3×3).** — Notons A la matrice $\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$.

Pour chaque λ de $\mathbb{K}$, on a :

$$\begin{aligned}
\chi_A(\lambda) = & \begin{vmatrix} \lambda-2 & -1 & -1 \\ -1 & \lambda-2 & -1 \\ -1 & -1 & \lambda-2 \end{vmatrix} \\
= & \begin{vmatrix} \lambda-4 & -1 & -1 \\ \lambda-4 & \lambda-2 & -1 \\ \lambda-4 & -1 & \lambda-2 \end{vmatrix} \quad C_1 \leftarrow C_1 + C_2 + C_3 \\
= & (\lambda-4) \begin{vmatrix} 1 & -1 & -1 \\ 1 & \lambda-2 & -1 \\ 1 & -1 & \lambda-2 \end{vmatrix} \\
= & (\lambda-4) \begin{vmatrix} 1 & -1 & -1 \\ 0 & \lambda-1 & 0 \\ 0 & 0 & \lambda-1 \end{vmatrix} L_2 \leftarrow L_2 - L_1, \text{ idem } L_3 \\
= & (\lambda-4)(\lambda-1)^2.
\end{aligned}$$

La matrice A admet donc exactement deux valeurs propres : $\mathbf{Spectre}(A) = \{1, 4\}$.

***Exemple 7.16* (Un cas plus abstrait).** — Soit E l'espace $\mathcal{M}_2(\mathbb{R})$ des matrices 2×2 à coefficients réels ; considérons la matrice $U = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$ et l'endomorphisme

$$\begin{aligned}
\Phi : E &\to E \\
M &\mapsto UM - MU.
\end{aligned}$$

Pour trouver le polynôme caractéristique de Φ, il faut calculer, pour chaque λ de $\mathbb{R}$, le déterminant de l'endomorphisme $(\lambda \mathrm{id}_E - \Phi)$ de $E = \mathcal{M}_2(\mathbb{R})$. Compte tenu de la définition du déterminant d'un endomorphisme d'espace abstrait, définition 6.26, il faut que nous écrivions la matrice de $(\lambda \mathrm{Id}_E - \Phi)$ dans une certaine base de $\mathcal{M}_2(\mathbb{R})$ et que nous calculions son déterminant.

En cherchant la matrice de $\Psi_\lambda = (\lambda \mathrm{Id}_E - \Phi)$ dans la base canonique $\mathcal{B} = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right)$ de $\mathcal{M}_2(\mathbb{R})$, c'est-à-dire en exprimant l'image par Ψ_λ de chacune des matrices de $\mathcal{B}$ en fonction des matrices de $\mathcal{B}$, on trouve

$$\mathcal{M}_{\mathcal{B}}(\lambda \mathrm{Id}_E - \Phi) = \begin{pmatrix} \lambda & 0 & -2 & 0 \\ 2 & \lambda+2 & 0 & -2 \\ 0 & 0 & \lambda-2 & 0 \\ 0 & 0 & 2 & \lambda \end{pmatrix}.$$

Pour calculer le polynôme caractéristique χ_Φ, c'est le déterminant de cette matrice 4×4 que nous devons calculer.

En développant par rapport à la troisième ligne, on trouve : $\chi_\Phi(\lambda) = (\lambda - 2) \begin{vmatrix} \lambda & 0 & 0 \\ 2 & \lambda+2 & -2 \\ 0 & 0 & \lambda \end{vmatrix}$;

en développant cette fois par rapport à la première ligne on obtient

$$\forall \lambda \in \mathbb{R}, \quad \chi_\Phi(\lambda) = (\lambda - 2)(\lambda) \begin{vmatrix} \lambda + 2 & -2 \\ 0 & \lambda \end{vmatrix} = \lambda^2(\lambda - 2)(\lambda + 2).$$

On constate donc que les valeurs propres de Φ sont $0, 2$ et -2.

Remarques sur le calcul du polynôme caractéristique. — Si A est une matrice $n \times n$, pour calculer le polynôme caractéristique χ_A, il s'agit de calculer le déterminant $\mathbf{Det}(X \cdot I_n - A)$... donc toutes les méthodes exposées au chapitre 6 sont susceptibles de fournir le résultat.

Cependant, deux remarques imposent de faire ce calcul soigneusement.

- On ne s'intéresse au polynôme caractéristique que parce que ses racines sont les valeurs propres de A,
- Si A n'est pas une matrice 2×2, χ_A est un polynôme de degré supérieur ou égal à 3 ; sauf dans de rares cas, il n'est pas facile [3] de trouver les racines d'un polynôme de degré n si $n \geq 3$.

Il serait donc imprudent de ne pas essayer de mener le calcul d'une façon qui facilite la factorisation du polynôme caractéristique. Il est important de s'entraîner à ce type de calcul pour ne pas obtenir une expression de χ_f qui soit juste mais trop difficile à exploiter. Dans l'exemple 7.15 ci-dessus, il aurait été imprudent de calculer trop brutalement $\chi_A(X)$.

Existence de valeurs propres sur $\mathbb{C}$, mais pas nécessairement sur $\mathbb{R}$.

- Si E est un $\mathbb{C}$-espace vectoriel et si f est un endomorphisme de E, alors χ_f est un polynôme à coefficients complexes ; il admet donc [4] au moins une racine, et f admet au moins une valeur propre.
- Par contre, si E est un $\mathbb{R}$-espace vectoriel et f un endomorphisme de E, alors χ_f est un polynôme à coefficients réels. Il admet bien des racines complexes, mais rien ne garantit qu'il admette des racines réelles, il existe donc des endomorphismes n'ayant pas de valeurs propres réelles.
- Par exemple, fixons un réel θ et considérons la matrice $A = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ (rappelons que T_A est la rotation de $\mathbb{R}^2$ d'angle θ dans le sens direct). Le polynôme caractéristique de A est $\chi_A(X) = X^2 - (2\cos(\theta))X + 1$ (voyez-vous pourquoi ?). C'est un polynôme du second degré dont le discriminant est $4\cos^2(\theta) - 4 = 4(\cos^2(\theta) - 1)$... qui est strictement négatif, sauf si θ est congru à zéro modulo π. Par conséquent A n'a *aucune valeur propre réelle*, sauf si $\theta \equiv 0\ [\pi]$, ce qui correspond au cas où A est égale à $\pm\, I_2$ (rotation d'angle 0^o ou 180^o). Bien sûr, le polynôme χ_A admet toujours deux racines *complexes* ; on peut vérifier qu'il s'agit des nombres $e^{i\theta}$ et $e^{-i\theta}$.

2.2. Deux matrices semblables ont les mêmes valeurs propres. —

Proposition 7.17 – Deux matrices semblables ont le même polynôme caractéristique

Soient A et B deux matrices semblables. Alors les polynômes caractéristiques χ_A et χ_B sont identiques.
Par conséquent, si A et B sont semblables, alors elles ont les mêmes valeurs propres.

Démonstration. — Si $B = P^{-1}AP$ avec P inversible, alors pour chaque x on a $xI_n - B = P^{-1}(xI_n - A)P$, donc les matrices $xI_n - A$ et $xI_n - B$ sont semblables, elles ont donc même déterminant et $\chi_A(x) = \chi_B(x)$. □

3. Si $n = 3$ ou $n = 4$, il existe des formules mais elles sont très peu commodes et il est hors de question de les retenir ; si $n \geq 5$, il n'est même pas possible d'écrire une formule « explicite » permettant de trouver les valeurs exactes des racines.

4. Le théorème de d'Alembert-Gauss assure l'existence d'au moins une racine complexe de χ_f dès qu'il n'est pas constant ; or χ_f ne peut pas être constant, puisqu'il est de degré $n = \dim(E)$ et qu'on a supposé au début de ce paragraphe que la dimension de E n'est pas nulle.

Dimensions des espaces propres de deux matrices semblables. — Soient A et B deux matrices carrées de même taille. Dans ce paragraphe, on notera $V_{A,\lambda}$ et$V_{B,\lambda}$ les sous-espaces propres de A et B pour la valeur propre λ.

Proposition 7.18 – Matrices semblables et dimension des espaces propres

Si A et B sont semblables et si λ est valeur propre de A (donc de B), les dimensions $\dim(V_{A,\lambda})$ *et* $\dim(V_{B,\lambda})$ *sont égales.*

Démonstration. — Supposons que B soit égale à $P^{-1}AP$, où P est une matrice inversible. Alors dire qu'un vecteur u appartient à $V_{B,\lambda}$, c'est dire que $Bv = \lambda v$, ou encore $P^{-1}BPv = \lambda v$, ce qui équivaut à $A(Pv) = \lambda(Pv)$. Par conséquent, v appartient à $V_{B,\lambda}$ si et seulement si Pv appartient à $V_{A,\lambda}$. Comme P est inversible, si $(e_1, .., e_{m_\lambda})$ est une base de $V_{P,\lambda}$, alors $(Pe_1, .., Pe_{m_\lambda})$ est une base de $V_{A,\lambda}$. □

2.3. Multiplicité comme racine du polynôme caractéristique et dimension du sous-espace propre. —

Proposition 7.19 – Dimension du sous-espace propre et multiplicité comme racine

Soient E un $\mathbb{K}$-espace vectoriel de dimension finie *et f un endomorphisme de E.*
Si λ est une valeur propre et si m_λ est sa multiplicité comme racine du polynôme χ_f, alors on a

$$1 \leq \dim(V_\lambda) \leq m_\lambda$$

En particulier si λ est racine *simple* du polynôme caractéristique, alors V_λ est de dimension 1.

Démonstration de la proposition. — Soit $(e_1, .., e_{d_\lambda})$ une base de V_λ. Complétons-la en une base $\mathcal{B}$ de E. Si on écrit la matrice de f dans la base $\mathcal{B}$, on constate que les d_λ premières colonnes ont des zéros partout, sauf sur la diagonale où il y a des λ :

$$\mathcal{M}_{\mathcal{B}}(f) = \begin{pmatrix} \lambda I_{d_\lambda} & \star \dots \star \\ 0 & \tilde{M} \end{pmatrix}$$

où $\tilde{M}$ est une matrice $(n-1) \times (n-1)$.

On peut donc écrire $\chi_f(X) = \mathbf{Det} \begin{pmatrix} (X-\lambda) I_{d_\lambda} & \star \dots \star \\ 0 & XI_{n-d_\lambda} - \tilde{M} \end{pmatrix}$; en développant par rapport à la première colonne et en répétant cela m_λ fois, on constate que $\chi_f(X) = (X-\lambda)^{d_\lambda}\mathbf{Det}(XI_{n-d_\lambda} - \tilde{M})$. Cela prouve que $(X-\lambda)^{d_\lambda}$ divise $\chi_f(X)$. En conséquence, la multiplicité m_λ de la racine λ est supérieure à d_λ, c'est ce qu'on voulait démontrer. □

Exemple 7.20. — Reprenons la matrice $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ des exemples précédents. Nous avons vu que son polynôme caractéristique est $\chi_A(X) = (X-4)(X-1)^2$.

- Les valeurs propres de A sont 1 et 4. Grâce à la proposition ci-dessus, nous obtenons les renseignements suivants :
  — l'espace propre V_4 est de dimension 1 ;
  — l'espace propre V_1 est soit de dimension 1, soit de dimension 2.

- Pour trouver une base de V_4, il suffit donc de trouver un vecteur propre pour la valeur propre 4 ; or on peut constater que la somme des coefficients sur chaque ligne est 4, ainsi $A\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}4\\4\\4\end{pmatrix}$: le vecteur $\begin{pmatrix}1\\1\\1\end{pmatrix}$ appartient à V_4 et forme à lui tout seul une base de V_4
- Pour trouver la dimension de V_1, on peut soit s'appuyer sur le fait que c'est l'ensemble des vecteurs X de $\mathbb{R}^3$ qui vérifient le système linéaire $AX = 1 \cdot X$ et chercher une base de l'ensemble des solutions, soit s'appuyer sur le fait que V_1 est le noyau de la matrice $A - 1I_3$; nous avons vu à l'exemple (c) page 163 qu'il est de dimension 2 et qu'une base en est $\left(\begin{pmatrix}1\\-1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\-1\end{pmatrix}\right)$.

Exemple 7.21. — Soit B la matrice $\begin{pmatrix}5 & 1 & 0\\0 & 5 & 0\\0 & 0 & -2\end{pmatrix}$.

- C'est une matrice triangulaire, donc son polynôme caractéristique est $\chi_B(X) = (X+2)(X-5)^2$ et ses valeurs propres sont -2 et 5. En s'appuyant sur les multiplicités dans la factorisation de χ_B, on obtient les renseignements suivants :
  — l'espace propre V_{-2} est de dimension 1 ;
  — l'espace propre V_5 est soit de dimension 1, soit de dimension 2.
- Déterminons l'espace propre V_{-2}. Il s'agit du noyau de la matrice $B + 2I = \begin{pmatrix}7 & 1 & 0\\0 & 7 & 0\\0 & 0 & 0\end{pmatrix}$. Or la dernière colonne de cette matrice est nulle, on constate donc que le dernier vecteur de la base canonique, le vecteur $e_3 = \begin{pmatrix}0\\0\\1\end{pmatrix}$, appartient à V_2. Puisque nous savons déjà que V_2 est de dimension 1, le vecteur e_3 forme à lui tout seul une base de V_{-2}.

- L'espace propre V_5 est de dimension 1 ou 2. Comment trancher ?

  On constate qu'un vecteur $\begin{pmatrix}x\\y\\z\end{pmatrix}$ appartient à V_5 si et seulement si $B\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}5x\\5y\\5z\end{pmatrix}$, ce qui équivaut à $\begin{cases}5x+y &= 5x\\5y &= 5y\\-2z &= 5z\end{cases}$, ou encore à $\begin{cases}y=0\\z=0\end{cases}$ (sans condition sur x), autrement dit $\begin{pmatrix}x\\y\\z\end{pmatrix} \in \textbf{Vect}\left[\begin{pmatrix}1\\0\\0\end{pmatrix}\right]$.

  On conclut que V_5 est de dimension 1 et que $\begin{pmatrix}1\\0\\0\end{pmatrix}$ en est une base [5].

---

5. On aurait aussi pu le voir en écrivant $V_5 = \mathrm{Ker}(B - 5I_3) = \mathrm{Ker}\begin{pmatrix}0 & 1 & 0\\0 & 0 & 0\\0 & 0 & -7\end{pmatrix}$... n'est-ce pas ?

2.4. En pratique : trouver valeurs propres et vecteurs propres d'une matrice. —

- Pour trouver les valeurs propres de A, on peut calculer le polynôme caractéristique de A ; c'est un simple calcul de déterminant, où il faut veiller à ne pas faire d'erreur et à ne pas utiliser des stratégies de calcul qui rendent difficile le fait de trouver les racines (voir page 168).
- Une fois connues les valeurs propres de A, la recherche des sous-espaces propres revient à la résolution d'un système linéaire pour chaque valeur propre ou à la recherche du noyau d'une matrice pour chaque valeur propre. En effet, pour chaque valeur propre λ, l'espace propre V_λ est par définition le noyau $\operatorname{Ker}(A - \lambda I_n)$, c'est-à-dire l'ensemble des vecteurs X de $\mathbb{K}^n$ qui vérifient $AX = \lambda X$. À λ fixé, il s'agit donc simplement de résoudre le système linéaire $AX = \lambda X$, ou de chercher le noyau de la matrice $A - \lambda I_n$. Les TD vous donneront de nombreux exemples.
- Il peut être utile, avant de se lancer dans le calcul de χ_A, de chercher à savoir s'il y a des valeurs propres ou vecteurs propres « évidents », notamment dans les cas de grande dimension.

3. Endomorphismes diagonalisables

3.1. Définition et exemples. —

Définition 7.22 – Endomorphisme diagonalisable

Soient E un $\mathbb{K}$-espace vectoriel *de dimension finie* et f un endomorphisme de E.
On dit que f est *diagonalisable* s'il existe une base de E formée de vecteurs propres de f.

Vocabulaire pour les matrices. — Si A est une matrice de $\mathcal{M}_n(\mathbb{K})$, on dira que A est diagonalisable si l'endomorphisme T_A de $\mathbb{K}^n$ est diagonalisable.

Exemple 7.23. — Pour $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$, nous avons trouvé les valeurs propres (1 et 4) et décrit les espaces propres :

$V_4 = \mathbf{Vect}[u]$ et $V_1 = \mathbf{Vect}[v, w]$, où $u = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $v = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ et $w = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$.

La famille (u, v, w) est donc formée de vecteurs propres de A ; de plus, c'est la réunion de bases des espaces V_4 et V_1 ; comme ces espaces sont en somme directe, on en déduit (proposition 3.3, page 77) que notre famille est libre : comme c'est une famille libre de trois vecteurs de $\mathbb{R}^3$, c'est une base de $\mathbb{R}^3$.

Nous constatons donc que A est diagonalisable.

Exemple 7.24. — Pour $B = \begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -2 \end{pmatrix}$, nous avions aussi trouvé les valeurs propres (-2 et 5) et les espaces propres : $V_{-2} = \mathbf{Vect}[e_3]$ et $V_5 = \mathbf{Vect}[e_1]$ où $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ et $e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Tout vecteur propre de B est donc proportionnel soit à e_1, soit à e_3 ; en conséquence, il est impossible de trouver un vecteur w qui soit un vecteur propre de B et pour lequel la famille (e_1, e_3, w) serait libre. Nous constatons donc que B n'est pas diagonalisable.

Proposition 7.25 – Explication du terme « diagonalisable »

Soit E un $\mathbb{K}$-espace vectoriel ***de dimension finie*** *et f un endomorphisme de E.*
Les deux conditions suivantes sont équivalentes :

(i) f est diagonalisable ;

(ii) il existe une base de E dans laquelle la matrice de f est diagonale.

Démonstration. — Si $\mathcal{U} = (u_1, .., u_n)$ est une base de E formée de vecteurs propres de f associés aux valeurs propres $\lambda_1, ..., \lambda_n$ respectivement, alors la matrice de f dans la base $\mathcal{U}$ n'est autre que la matrice diagonale

$$\mathbf{Diag}(\lambda_1, .., \lambda_n) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

Réciproquement, si $\mathcal{U} = (u_1, .., u_n)$ est une base de E, dire que la matrice $\mathcal{M}_{\mathcal{U}}(f)$ est la matrice diagonale $\begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix}$ pour des scalaires $\alpha_1, .., \alpha_n$, c'est dire que $f(u_1) = \alpha_1 u_1$, $f(u_2) = \alpha_2 u_2$, etc ; dans ce cas tous les vecteurs de la base $\mathcal{U}$ sont des vecteurs propres de f, et f est diagonalisable. □

Proposition 7.26 – Diagonalisabilité : cas des matrices

Soit A une matrice de $\mathcal{M}_n(\mathbb{K})$. Alors

- *La matrice A est diagonalisable si et seulement si elle est semblable à une matrice diagonale.*
- *Dans ce cas, en notant*
  - *$\mathcal{U} = (u_1, .., u_n)$ est une base de $\mathbb{R}^n$ formée de vecteurs propres de A associés aux valeurs propres $\lambda_1, .., \lambda_n$,*
  - *$\mathcal{B}$ est la base canonique et P est la matrice de passage $P_{\mathcal{B} \text{ vers } \mathcal{U}}$ (celle dont les colonnes sont les coordonnées des u_i),*
  - *D est la matrice diagonale $\boldsymbol{Diag}(\lambda_1, .., \lambda_n)$,*

  on a l'égalité $A = P \cdot D \cdot P^{-1}$.

Démonstration. — Dire que A est diagonalisable, c'est dire qu'il existe une base $\mathcal{U} = (u_1, .., u_n)$ de $\mathbb{R}^n$ formée de vecteurs propres de A. Notons $\lambda_1, .., \lambda_n$ les valeurs propres correspondantes, alors on a toujours $\mathcal{M}_{\mathcal{B}}(T_A) = A$ (exemple 2bis page 129), alors qu'on vient de voir à la proposition précédente que $\mathcal{M}_{\mathcal{U}}(T_A) = D = \mathbf{Diag}(\lambda_1, .., \lambda_n)$ est diagonale ; par conséquent A et D sont semblables. Le reste de la proposition est une application immédiate du cours du chapitre 5. □

Exemple 7.27. — Nous avons vu que la matrice $A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$ est diagonalisable, que ses valeurs propres sont 1 et 4 et que les espaces propres sont $V_4 = \mathbf{Vect}[u]$ et $V_1 = \mathbf{Vect}[v, w]$, où $u = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, $v = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ et $w = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$.

La famille $(u, v, w) = \left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right)$ est une base de $\mathbb{R}^3$ formée de vecteurs propres de A associés, pour le premier à la valeur propre 4, pour les deux derniers à la valeur propre 1. La proposition que nous venons d'écrire donne :

$$A = PDP^{-1}, \quad \text{avec} \quad D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ et } P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix}$$

3.2. Lien avec la dimension des espaces propres. — Supposons que E soit un espace de dimension n. Pour qu'il existe une base formée de vecteurs propres de f, il faut qu'on puisse trouver n vecteurs propres linéairement indépendants.

Théorème 7.28 – Diagonalisabilité et dimension des espaces propres

Soient E un espace vectoriel de dimension finie et f un endomorphisme de E.
Il y a équivalence entre

$$f \text{ est diagonalisable} \iff \begin{cases} \text{Le polynôme } \chi_f \text{ est scindé, } \chi_f(X) = (X-\lambda_1)^{m_1}...(X-\lambda_k)^{m_k} \\ \textbf{et} \\ \text{pour chaque valeur propre } \lambda_i \text{ de } f, \text{ on a } \dim(V_\lambda) = m_\lambda \end{cases}$$

En particulier, si χ_f est *scindé à racines simples*, alors f est diagonalisable.

Exemple 7.29. —

(a) Soit N une matrice triangulaire supérieure $\begin{pmatrix} a_1 & \star & \dots & \star \\ 0 & a_2 & & \vdots \\ \vdots & & \ddots & \star \\ 0 & \dots & 0 & a_n \end{pmatrix}$ *dont tous les coefficients diagonaux sont distincts*. Alors $\chi_A(X) = (X-a_1)...(X-a_n)$ est scindé à racines simples, donc N est diagonalisable.

(b) Nous avons vu que si A est la matrice $\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$, alors le polynôme caractéristique de A est $(X-4)(X-1)^2$, et que les espaces propres V_4 et V_1 sont bien de dimension 1 et 2, respectivement ; par conséquent A est diagonalisable.

(c) Nous avons vu que si B est la matrice $\begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & -2 \end{pmatrix}$, le polynôme caractéristique de B est $(X-5)^2(X+2)$, mais que l'espace propre V_5 n'est pas de dimension 2 : par conséquent, la matrice B n'est pas diagonalisable.

3.3. En pratique : comment étudier la diagonalisabilité d'une matrice. —

1. Calculer le polynôme caractéristique de A, trouver ses racines et une factorisation.
2. Pour chaque valeur propre λ, décrire l'espace propre associé et voir si sa dimension est égale ou pas à la puissance de $(X-\lambda)$ qui apparaît dans la factorisation du polynôme caractéristique. Si c'est le cas pour *toutes* les valeurs propres alors A est diagonalisable, s'il y a *une* valeur propre pour laquelle ce n'est pas le cas alors A n'est pas diagonalisable.
3. Si A est diagonalisable et si $\mathcal{U} = (u_1, .., u_n)$ est une base de E formée de vecteurs propres de f pour les valeurs propres $\lambda_1, .., \lambda_n$ respectivement, alors la matrice P dont les colonnes sont les u_i vérifie
$$A = P \cdot \mathrm{Diag}(\lambda_1, ..., \lambda_n) \cdot P^{-1}.$$

3.4. Un exemple classique d'application : puissances d'une matrice. — Soit A une matrice diagonalisable ; si P est une matrice inversible et $D = \mathbf{Diag}(\lambda_1, .., \lambda_n)$ est une matrice diagonale vérifiant $A = PDP^{-1}$, alors pour tout k de $\mathbb{N}^\star$, $A^n = (PDP^{-1})(PDP^{-1})...(PDP^{-1}) = P \cdot D \cdot (P^{-1}P) \cdot D... \cdot D \cdot (P^{-1}P) \cdot D \cdot P^{-1}$ n'est autre que $P \cdot (D^k) \cdot P^{-1}$.

Or, D^k est très facile à calculer : si $D = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$, alors $D^k = \begin{pmatrix} \lambda_1^k & & \\ & \ddots & \\ & & \lambda_n^k \end{pmatrix}$ pour chaque k de $\mathbb{N}$.

Si on connaît la matrice P explicitement, on connaît donc facilement A^k.

Exemple 7.30 **(Expression explicite de la suite de Fibonacci).** —

Soit $(u_n)_{n\in\mathbb{N}}$ la suite réelle définie par $u_0 = 0$, $u_1 = 1$ et pour chaque n de $\mathbb{N}$, $u_{n+2} = u_{n+1} + u_n$.
Alors on a, pour chaque entier naturel n,

$$\begin{pmatrix} u_{n+2} \\ u_{n+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} u_n \\ u_{n-1} \end{pmatrix} = \ldots = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{n+1} \begin{pmatrix} u_1 \\ u_0 \end{pmatrix}.$$

Notant A la matrice $\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$, on a donc pour chaque n de $\mathbb{N}$, $\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = A^n \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Si nous parvenons à calculer explicitement A^n pour chaque n, nous obtiendrons une expression explicite de u_n en fonction de n. Compte tenu de la discussion précédente, cherchons à diagonaliser A.

- On constate que $\chi_A(\lambda) = \lambda^2 - \lambda - 1$ pour chaque λ de $\mathbb{R}$; par conséquent les valeurs propres de A sont les solutions de l'équation $\lambda^2 = \lambda + 1$: le nombre d'or $\varphi = \dfrac{1+\sqrt{5}}{2}$ ainsi que $\tilde{\varphi} = \dfrac{1-\sqrt{5}}{2}$.
- L'espace propre pour la valeur propre φ est :
  $V_\varphi = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2,\ A \begin{pmatrix} x \\ y \end{pmatrix} = \varphi \begin{pmatrix} x \\ y \end{pmatrix} \right\} = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2, \begin{pmatrix} x+y \\ x \end{pmatrix} = \varphi \begin{pmatrix} x \\ y \end{pmatrix} \right\} = \mathbf{Vect}\left[\begin{pmatrix} \varphi \\ 1 \end{pmatrix} \right]$.[6]
- On vérifie de même que l'espace propre pour la valeur propre $\tilde{\varphi}$ est $V_{\tilde{\varphi}} = \mathbf{Vect}\left[\begin{pmatrix} \tilde{\varphi} \\ 1 \end{pmatrix} \right]$.
- Ainsi, on a $A = PDP^{-1}$ avec $P = \begin{pmatrix} \varphi & \tilde{\varphi} \\ 1 & 1 \end{pmatrix}$ et $D = \begin{pmatrix} \varphi & 0 \\ 0 & \varphi \end{pmatrix}$.

Nous avons donc, pour chaque n de $\mathbb{N}$,

$$\begin{aligned} A^n &= \begin{pmatrix} \varphi & \tilde{\varphi} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \varphi^n & 0 \\ 0 & \tilde{\varphi}^n \end{pmatrix} \begin{pmatrix} \varphi & \tilde{\varphi} \\ 1 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \varphi & \tilde{\varphi} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \varphi^n & 0 \\ 0 & \tilde{\varphi}^n \end{pmatrix} \cdot \left[\frac{1}{\sqrt{5}} \begin{pmatrix} 1 & -\tilde{\varphi} \\ -1 & \varphi \end{pmatrix} \right] \\ &= \begin{pmatrix} \varphi^{n+1} & \tilde{\varphi}^{n+1} \\ \varphi^n & \tilde{\varphi}^n \end{pmatrix} \cdot \left[\frac{1}{\sqrt{5}} \begin{pmatrix} 1 & -\tilde{\varphi} \\ -1 & \varphi \end{pmatrix} \right] \\ &= \frac{1}{\sqrt{5}} \begin{pmatrix} \varphi^{n+1} - \tilde{\varphi}^{n+1} & \varphi\tilde{\varphi}^{n+1} - \tilde{\varphi}\varphi^{n+1} \\ \varphi^n - \tilde{\varphi}^n & \varphi\tilde{\varphi}^n - \tilde{\varphi}\varphi^n \end{pmatrix} \end{aligned}$$

et enfin

$$\begin{pmatrix} u_{n+1} \\ u_n \end{pmatrix} = A^n \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \text{ première colonne de } A^n = \frac{1}{\sqrt{5}} \begin{pmatrix} \varphi^{n+1} - \tilde{\varphi}^{n+1} \\ \varphi^n - \tilde{\varphi}^n \end{pmatrix},$$

ce qui donne u_n explicitement :

$$u_n = \frac{\varphi^n - \tilde{\varphi}^n}{\sqrt{5}} = \frac{(1+\sqrt{5})^n - (1-\sqrt{5})^n}{2^n \cdot \sqrt{5}}.$$

6. Pour la dernière égalité, on sait que A admet deux valeurs propres distinctes et c'est une matrice 2×2, donc V_φ est de dimension 1, et on constate que $A \begin{pmatrix} \varphi \\ 1 \end{pmatrix} = \begin{pmatrix} \varphi + 1 \\ \varphi \end{pmatrix} = \begin{pmatrix} \varphi^2 \\ \varphi \end{pmatrix} = \varphi \begin{pmatrix} \varphi \\ 1 \end{pmatrix}$, en se rappelant que $\varphi^2 = \varphi + 1$.

4. (⋆) Complément : trigonalisation sur ℂ et théorème de Cayley-Hamilton

4.1. Notion d'endomorphisme trigonalisable et cas des espaces vectoriels complexes. —

Définition 7.31 – Endomorphisme trigonalisable

Soient E un $\mathbb{K}$-espace vectoriel de dimension finie et f un endomorphisme de E. On dit que f est *trigonalisable* s'il existe une base $\mathcal{B}$ de E dans laquelle la matrice $\mathcal{M}_{\mathcal{B}}$ de f est triangulaire supérieure.

Remarque 7.32. — Dans ce cas, les valeurs propres de f sont données par les coefficients diagonaux de $\mathcal{M}_{\mathcal{B}}$ (puisque le spectre de f est le même que celui de la matrice triangulaire $\mathcal{M}_{\mathcal{B}}$, dont les valeurs propres se lisent sur la diagonale).

Théorème 7.33 – La trigonalisation est toujours possible sur ℂ

Si E est un $\mathbb{C}$-espace vectoriel de dimension finie ***(attention, il est essentiel ici que le corps de base soit ℂ)****,*
alors tout endomorphisme de E est trigonalisable.

Démonstration. — Raisonnons par récurrence sur la dimension de E.

Pour chaque n de $\mathbb{N}^\star$, notons $\mathcal{P}(n)$ l'énoncé

« Pour tout $\mathbb{C}$-espace vectoriel V de dimension n, pour tout endomorphisme φ de V, il existe une base de V dans laquelle la matrice de φ est triangulaire supérieure. »

- Si $n = 1$, l'énoncé est évidemment vrai.
- Soit n un entier de $\mathbb{N}^\star$. Supposons $\mathcal{P}(n)$ vrai. Soit E un $\mathbb{C}$-espace vectoriel de dimension $n+1$ et f un endomorphisme de E. Le polynôme caractéristique χ_f est à coefficients complexes, donc il admet au moins une racine, si bien que f admet au moins une valeur propre λ. Notons u un vecteur propre associé (il est donc non nul). Soit $\mathcal{B} = (u, u_2, .., u_{n+1})$ une base de E dont le premier vecteur est u (vous rappelez-vous pourquoi une telle base existe ?). Comme $f(u) = \lambda u$, la matrice de f dans la base $\mathcal{B}$ est de la forme
$$\begin{pmatrix} \lambda & \star & \dots & \star \\ 0 & & & \\ \vdots & & M & \\ 0 & & & \end{pmatrix}$$
où M est une matrice $n \times n$. Soit V le sous-espace $\mathbf{Vect}[u_2, .., u_n]$ de E. On a $\mathbf{Vect}[u] \oplus V = E$; notons π la projection sur V parallèlement à $\mathbf{Vect}[u]$. L'application $\pi \circ f$ définit un endomorphisme de V ; d'après l'énoncé $\mathcal{P}(n)$, il existe une base $(v_2, .., v_{n+1})$ de V dans laquelle la matrice N de φ est triangulaire supérieure.

  La famille $\mathcal{B}'$ est alors une base de E (voyez-vous pourquoi ?), et en observant la matrice $\mathcal{M}_{\mathcal{B}}(f)$, on constate qu'elle est de la forme
$$\begin{pmatrix} \lambda & \star' & \dots & \star' \\ 0 & & & \\ \vdots & & N & \\ 0 & & & \end{pmatrix}$$
donc triangulaire supérieure.

□

4.2. Conséquences pratiques sur les valeurs propres. —

Proposition 7.34 – Lien entre valeurs propres, trace et déterminant

Soit E un $\mathbb{C}$-espace vectoriel de dimension finie (on suppose toujours ici que le corps de base est $\mathbb{C}$).
Soit f un endomorphisme de E.
Notons $\chi_f(X) = (X-\lambda_1)^{m_1}..(X-\lambda_k)^{m_k}$ la factorisation de χ_f sur $\mathbb{C}$. Alors

$$\textbf{\textit{Tr}}(f) = m_1 \cdot \lambda_1 + ... + m_k\lambda_k$$
$$\textbf{\textit{Det}}(f) = \lambda_1^{m_1} \cdot ... \cdot \lambda_k^{m_k}$$

Remarque pratique. — Si vous devez calculer et factoriser le polynôme caractéristique d'un endomorphisme ou d'une matrice, ce résultat permet de détecter facilement une éventuelle erreur : si A est une matrice 3×3 si vous obtenez comme polynôme caractéristique $(X-1)^2(X+3)$, alors la trace de A est *nécessairement* égale à $1+1+(-3)$; dans le cas où la matrice A serait explicite sous vos yeux et où sa trace ne serait pas égale à -1, c'est qu'il y a une erreur dans votre calcul ou votre factorisation du polynôme caractéristique.

Démonstration de la proposition. — C'est une conséquence immédiate du théorème précédent. En effet, si f est un endomorphisme de E, il existe une base $\mathcal{B}$ de E dans laquelle la matrice M de f est triangulaire ; comme on a alors $\chi_f = \chi_M = (X-\lambda_1)^{m_1}..(X-\lambda_k)^{m_k}$, la diagonale de M comporte nécessairement m_1 fois le nombre λ_1, m_2 fois le nombre λ_2, ..., et m_k fois le nombre λ_k. Comme $\mathrm{Tr}(f) = \mathrm{Tr}(M)$ et $\mathrm{Det}(f) = \mathbf{Det}(M)$, on obtient tout de suite la proposition. □

Cas des matrices réelles. — Si A est une matrice à coefficients réels, on peut toujours factoriser χ_A sur $\mathbb{C}$ en $(X-\lambda_1)^{m_1}..(X-\lambda_k)^{m_k}$, où les nombres $\lambda_1,..,\lambda_n$ sont les racines complexes de χ_A (certaines peuvent ne pas être réelles).

En voyant T_A comme l'endomorphisme $\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \mapsto A \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$ de $\mathbb{C}^n$ et en appliquant la proposition ci-dessus, on constate qu'on a toujours $\mathbf{Tr}(f) = m_1 \cdot \lambda_1 + ... + m_k\lambda_k$ et $\mathbf{Det}(f) = \lambda_1^{m_1} \cdot ... \cdot \lambda_k^{m_k}$, il faut simplement faire intervenir dans la formule toutes les racines complexes de χ_A avec leur multiplicités, pas seulement les valeurs propres réelles de A.

***Exemple 7.35* (Illustration avec une matrice réelle).** —
Soit A la matrice $\begin{pmatrix} 2 & 3 & -1 \\ 1 & 2 & 1 \\ 1 & -1 & 4 \end{pmatrix}$. En observant que la somme des coefficients est la même sur chacune des lignes, on constate que $A\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ 4 \end{pmatrix}$. Ainsi l'une des valeurs propres de A est 4.

Notons $\chi_A(X) = (X-4)(X-\alpha)(X-\beta)$ la factorisation sur $\mathbb{C}$ du polynôme caractéristique de A (à ce stade les nombres α et β ne sont pas forcément distincts, pas forcéments différents de 4).

Alors le résultat précédent donne $4+\alpha+\beta = \mathbf{Tr}(A) = 8$ et $4\alpha\beta = \mathbf{Det}(A) = 16$, donc on doit avoir

$$\begin{cases} \alpha+\beta & = 4 \\ \alpha\beta & = 4, \end{cases}$$

ce qui équivaut à $\alpha = \beta = 2$. Nous constatons donc que $\chi_A(X) = (X-4)(X-2)^2$ et nous avons trouvé toutes les valeurs propres de A : $\mathbf{Spectre}(A) = \{2, 4\}$.

4.3. (⋆) Complément : théorème de Cayley-Hamilton. — Le théorème de Cayley-Hamilton est une identité remarquable (et surprenante) vérifiée par tout endomorphisme en dimension finie. C'est un résultat théorique fondamental pour poursuivre l'étude des valeurs propres et vecteurs propres au-delà des premiers éléments présentés dans ce chapitre ; nous n'aurons pas le temps de le faire, mais nous achevons ce cours en énonçant le théorème et en proposant dans les exercices quelques illustrations pratiques.

Théorème 7.36 – Théorème de Cayley-Hamilton

Soit E un $\mathbb{K}$-espace vectoriel de dimension finie n ***(ici le corps de base peut être $\mathbb{R}$ ou $\mathbb{C}$)****.*
Soit f un endomorphisme de E.
Notons $\chi_f(X) = X^n + a_{n-1}X^{n-1} + ... + a_1X + a_0$, alors on a l'égalité

$$f^n + a_{n-1}f^{n-1} + ... + a_1 f + a_0 id_E = 0.$$

Conséquence pour les matrices à coefficients réels. — Si A est une matrice à coefficients réels, on peut voir T_A comme l'endomorphisme $\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \mapsto A \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$ de $\mathbb{C}^n$. Le théorème de Cayley-Hamilton donne $\chi_A(A) = 0$. Or $\chi_A(\lambda) = \mathbf{Det}(\lambda I_n - A)$ est un polynôme à coefficients réels dès que A l'est, et peut être calculé sans référence aux nombres complexes. Ainsi la relation $\chi_A(A) = 0$ est vraie pour une matrice à coefficients réels et ne fait pas intervenir de nombres complexes, même si nous sommes passés par une extension à $\mathbb{C}^n$ pour la démontrer.

Démonstration. — [7]

Notons $\chi_f(X) = (X-\lambda_1)(X-\lambda_2)...(X-\lambda_n)$, quitte à ce que certains des λ_n soient égaux dans le cas où les racines ne sont pas toutes simples. Soit $\mathcal{B}$ une base de E dans laquelle la matrice de f est triangulaire supérieure : elle est de la forme

$$\begin{pmatrix} \lambda_1 & \star & \dots & \star \\ 0 & \lambda_2 & & \\ \vdots & & \ddots & \star \\ 0 & \dots & 0 & \lambda_n \end{pmatrix}$$

L'endomorphisme $g = f^n + a_{n-1}f^{n-1} + ... + a_1 f + a_0 \mathrm{id}_E$ n'est autre que $(f - \lambda_1 \mathrm{id}_E)(f - \lambda_2 \mathrm{id}_E)...(f - \lambda_n \mathrm{id}_E) = (f - \lambda_n \mathrm{id}_E)...(f - \lambda_2 \mathrm{id}_E)...(f - \lambda_1 \mathrm{id}_E)$. Sa matrice dans la base $\mathcal{B}$ est donc le produit

$$\mathcal{M}_{\mathcal{B}}(g) = \begin{pmatrix} \lambda_1 - \lambda_n & \star & \dots & \star \\ 0 & \lambda_2 - \lambda_n & & \\ \vdots & & \ddots & \star \\ 0 & \dots & 0 & 0 \end{pmatrix} \cdots \begin{pmatrix} \lambda_1 - \lambda_2 & \star & \dots & \star \\ 0 & 0 & & \\ \vdots & & \ddots & \star \\ 0 & \dots & 0 & \lambda_n - \lambda_2 \end{pmatrix} \begin{pmatrix} 0 & \star & \dots & \star \\ 0 & \lambda_2 - \lambda_1 & & \\ \vdots & & \ddots & \star \\ 0 & \dots & 0 & \lambda_n - \lambda_1 \end{pmatrix}$$

7. **Une démonstration qui ne marche pas.** Comme l'égalité à montrer peut s'écrire $\chi_f(f) = 0$, il est facile de sous-estimer la difficulté de ce résulat et de croire que la démonstration se réduit à écrire que $\chi_f(\lambda) = \mathbf{Det}(\lambda \mathrm{id} - f)$, « donc » $\chi_f(f) = \mathbf{Det}(f \mathrm{id} - f) = 0$. Mais écrire $\chi_f(\lambda) = \mathbf{Det}(\lambda \mathrm{id} - f)$ n'a de sens que si λ est un scalaire, on ne peut pas se contenter de substituer l'endomorphisme f à la place du scalaire λ et de former « (f id -f) ». Même si la démonstration que je vous donne ici n'est pas très difficile, elle repose sur le théorème de trigonalisation, qui utilise lui-même des notions venant de *tous* les chapitres de notre cours : cela porte sur un endomorphisme et sur sa représentation matricielle, la notion de dimension finie y est essentielle ; dans la démonstration on a croisé le polynôme caractéristique, donc un déterminant ; on s'y est appuyé sur un projecteur, ce qui nécessite de manipuler les notions de somme directe et de supplémentaire, etc.

En observant attentivement le produit des deux dernières matrices, on constate que les deux premières colonnes sont nulles, et en effectuant tranquillement les produits (en commençant par la droite), on observe que le produit des k dernières matrices commence par k colonnes de zéros pour chaque k de $\{1, ..n\}$. Finalement, on constate donc que $\mathcal{M}_{\mathcal{B}}(g)$ est la matrice nulle, donc g est l'endomorphisme nul, CQFD. □

Exemple d'application : retrouver la formule pour l'inverse d'une matrice* 2×2*.

Soit $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ une matrice 2×2. On rappelle que $\chi_A(X) = X^2 - \alpha X + \beta$ avec $\alpha = a + d = \mathbf{Tr}(A)$ et $\beta = ad - bc = \mathbf{Det}(A)$.

Supposons mainenant que A est inversible. Le théorème de Cayley-Hamilton donne

$$A^2 - \alpha A + \beta I_2 = 0$$

autrement dit

$$A(A - \alpha I_2) = -\beta I_2$$

ou encore, puisque $\beta = \mathbf{Det}(A)$ n'est pas nul,

$$A \cdot \frac{\alpha I_2 - A}{\beta} = I_2$$

ce qui montre que A^{-1} n'est autre que la matrice $\dfrac{\alpha I_2 - A}{\beta} = \dfrac{\mathbf{Tr}(A) \cdot I_2 - A}{\mathbf{Det}(A)} = \frac{1}{ad-bc} \begin{pmatrix} (a+d) - a & -b \\ -c & (a+d) - d \end{pmatrix}$.

On retrouve bien $A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$; c'est la formule habituelle (qui peut ainsi être démontrée de multiples façons : résolution d'un système linéaire, « pivot », formule utilisant la comatrice au chapitre 6, théorème de Cayley-Hamilton...).

INDEX