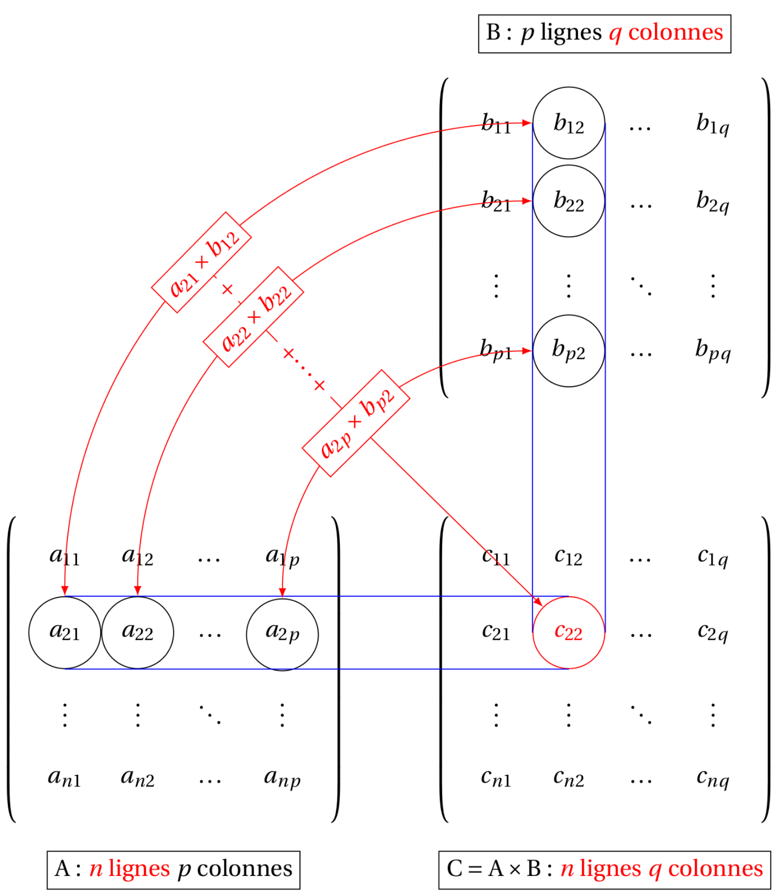
Intégrales Multiples et Algèbre Linéaire

François DE MARÇAY
Département de Mathématiques d'Orsay
Université Paris-Sud, France

« Celui qui enseigne une chose la connaît rarement à fond, car s'il l'étudiait à fond afin de l'enseigner, il n'aurait alors plus assez de temps disponible pour l'enseigner. »

Jacques-Henri D'AGUESSEAU.



Méthodologie de travail
Mathématiques 170 pour PolyTech U-PSUD PEIP1 S2

Joël MERKER alias François DE MARÇAY
Département de Mathématiques d'Orsay
Université Paris-Sud, France

• **Notes de cours.** Des notes de cours seront transmises par courriel sous forme pdf.

• **Devoirs à la maison.** Quatre devoirs à la maison seront à rendre. Ils seront établis par le professeur responsable de l'UE Math-170, puis corrigés par les chargés de travaux dirigés. Chaque devoir non rendu se verra attribuer une note de $\frac{\mathbf{0}}{\mathbf{20}}$ qui contribuera à hauteur de **5 %** de la note finale.

• **Modalités de contrôle.** Les **100 %** de la note finale complète comprendront :

☐ **20 %** contrôle continu = les **4** devoirs à la maison.

☐ **20 %** examen 1, durée 1h30, le 12 février 2019.

☐ **20 %** examen 2, durée 1h30, le 14 mars 2019.

☐ **20 %** examen 3, durée 1h30, le 3 avril 2019.

☐ **20 %** examen 4, durée 1h30, fin mai 2019.

• **Examens.** Les sujets des **4** examens seront établis par le professeur responsable de l'UE Math-170. Les copies seront intégralement corrigées par ledit professeur.

• **Règle d'or pendant les cours :**

Interdiction absolue d'utiliser et de consulter
smartphones, téléphones et ordinateurs portables
et tous autres gadgets électroniques contraires au travail

• **Modalités d'application de cette règle d'or.** Les étudiants qui contreviendront à cette règle seront exclus sur le champ de la salle de cours. Le cours ne reprendra que lorsque les étudiants en question seront sortis de la salle de cours.

• **Lecture régulière du cours.** Chaque étudiant s'imposera de lire, relire et étudier régulièrement le cours. Ce travail s'effectuera occasionnellement, même sur des courtes périodes d'une dizaine de minutes, à la maison, à la bibliothèque ou dans les transports en commun. ***C'est en lisant qu'on développe son intelligence***, car on absorbe les intelligences variées d'autres personnes sans rester confiné en soi-même, voire infiniment pire : confiné à l'abrutissement total du tripotage crétinisant de smartphone !

• **Assiduité au cours.** C'est principalement le cours oral au tableau qui permettra de transmettre les idées informelles et les intuitions importantes. Aussi, lecture du cours et présence au cours seront-elles *deux activités complémentaires et indispensables pour une préparation optimale au métier de scientifique*. De plus, *on lit beaucoup plus facilement les notes de cours après avoir écouté le professeur*. ***De toute façon, une bonne prise de notes manuscrites personnelles a plus de valeur que les polycopiés.***

• **Prise de notes pendant les séances de cours.** ***L'existence de documents écrits transmis par les professeurs ne dispense absolument pas de prendre des notes manuscrites complètes et soignées***.

Table des matières

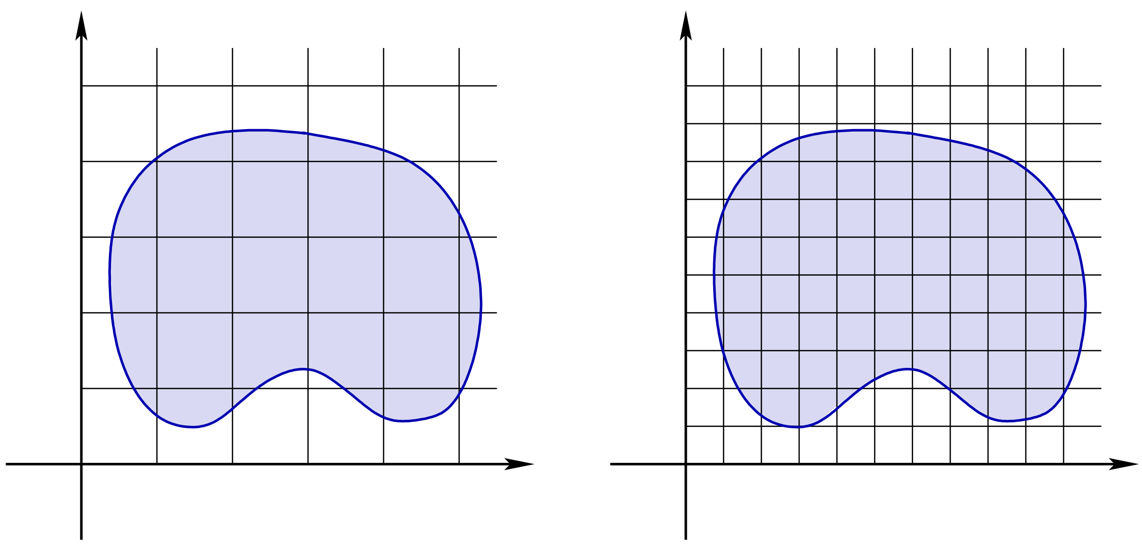
Intégrales doubles

François DE MARÇAY
Département de Mathématiques d'Orsay
Université Paris-Sud, France

1. Introduction

2. Pavages du plan $\mathbb{R}^2$

Afin de déterminer l'aire — la mesure de surface — d'une partie $A \subset \mathbb{R}^2$ du plan euclidien, une idée simple et naturelle consiste à utiliser des quadrillages de plus en plus fins.



Soit donc $\{0, \mathbf{i}, \mathbf{j}\}$ le repère orthonormé canonique du plan euclidien $\mathbb{R}^2$ muni des coordonnées cartésiennes (x, y), où $\mathbf{i} = (1, 0)$ et $\mathbf{j} = (0, 1)$ sont les deux vecteurs de base. Pour deux paires quelconques de nombres réels $a < b$ et $c < d$, le produit d'intervalles semi-ouverts :

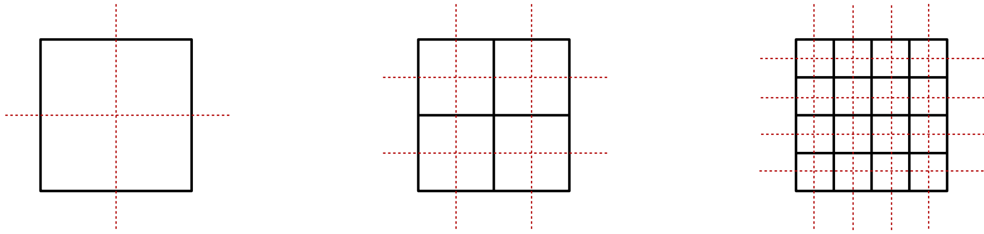
$$[a, b[\times [c, d[$$

est un rectangle élémentaire, ou *pavé*, semi-ouvert.

Lorsque a, b, et c, d sont des paires d'entiers consécutifs, la réunion de tels *pavés entiers :*

$$\boldsymbol{P}_0 := \bigcup_{m \in \mathbb{Z}} \bigcup_{p \in \mathbb{Z}} \big[m,\, m+1\big[\times \big[p,\, p+1\big[\;= \mathbb{R}^2$$

recouvre tout le plan, sans intersections. L'ensemble de ces pavés $[m, m+1[\times [p, p+1[$ constitue alors un *pavage de profondeur* 0, où de manière plus imagée *carrelage*, qui sera noté $\boldsymbol{P}_0$.



Ensuite, découpons chaque carreau (pavé) de $\boldsymbol{P}_0$ en son milieu, verticalement et horizontalement, pour obtenir le *pavage de profondeur* 1 :

$$\boldsymbol{P}_1 := \bigcup_{m\in\mathbb{Z}} \bigcup_{p\in\mathbb{Z}} \left[\frac{m}{2}, \frac{m+1}{2}\right[\times \left[\frac{p}{2}, \frac{p+1}{2}\right[= \mathbb{R}^2.$$

Ainsi, chaque pavé se fragmente en $2 \times 2 = 4$ sous-pavés. En itérant les coups de sabre, pour tout entier $k \geqslant 0$, on produit le *pavage de profondeur* k *:*

$$\boldsymbol{P}_k := \bigcup_{m\in\mathbb{Z}} \bigcup_{p\in\mathbb{Z}} \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right[\times \left[\frac{p}{2^k}, \frac{p+1}{2^k}\right[= \mathbb{R}^2,$$

qui recouvre le plan par une réunion *disjointe* de carreaux d'aires de plus en plus petites :

$$\frac{1}{2^k} \times \frac{1}{2^k} = \frac{1}{2^{2k}} \underset{k\to\infty}{\longrightarrow} 0.$$

On s'imagine alors aisément qu'en augmentant la profondeur, on a de plus en plus de chance de bien approximer l'« *aire* » d'une partie donnée $A \subset \mathbb{R}^2$.

3. Aire inférieure et aire supérieure d'un ensemble borné $A \subset \mathbb{R}^2$

Rappelons qu'un ensemble $A \subset \mathbb{R}^2$ est dit *borné* s'il est contenu dans un disque de rayon assez grand — A ne s'évade pas à l'infini. Cela équivaut à dire que le *diamètre* de A, défini comme le supremum de la distance entre deux quelconques de ses points :

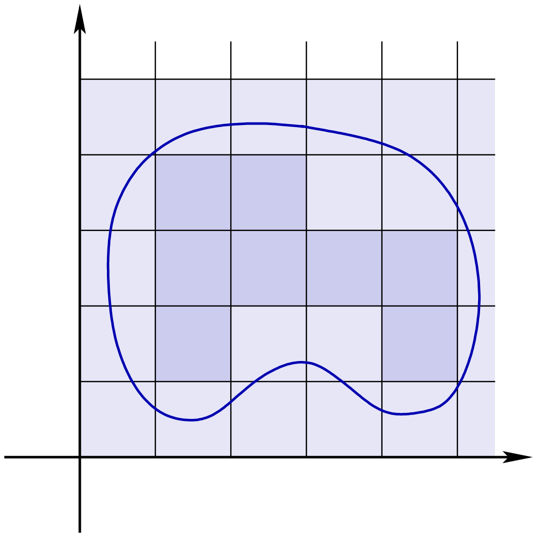
$$\begin{aligned} \operatorname{diam} A &:= \sup_{\substack{(x_1,y_1)\in A \\ (x_2,y_2)\in A}} \sqrt{(x_1-x_2)^2+(y_1-y_2)^2} \\ &\in \mathbb{R}_+ \cup \{\infty\} \end{aligned}$$

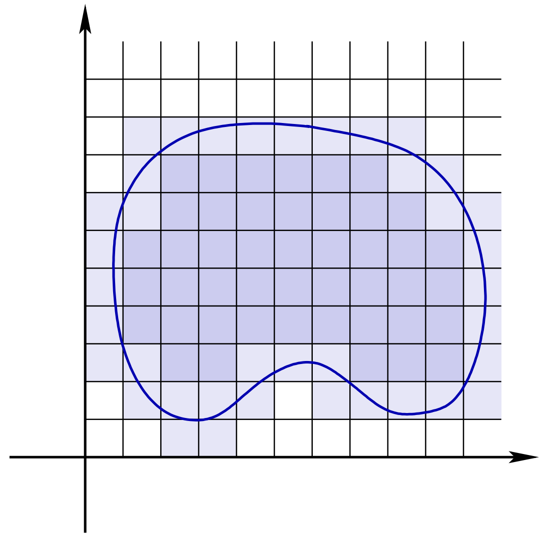
est *fini :*

$$\operatorname{diam} A < \infty \qquad \Longleftrightarrow \qquad A \text{ borné.}$$

Nous supposerons donc A borné, et notre objectif est de donner un sens mathématique rigoureux à la notion de « *mesure* » de l'étendue — aire, surface — de A. Évidemment, on convient que l'unité de mesure est :

$$\operatorname{aire}\big([0,1[\times[0,1[\big) = 1 \cdot 1 = 1.$$





L'idée spontanée est de compter les pavés entièrement contenus dans l'ensemble :

$$\mathrm{N}_0^- \;:=\; \mathsf{Card}\,\big\{\text{pavés de } \boldsymbol{P}_0 \text{ contenus dans } A\big\},$$

puis ceux, plus nombreux, qui le rencontrent :

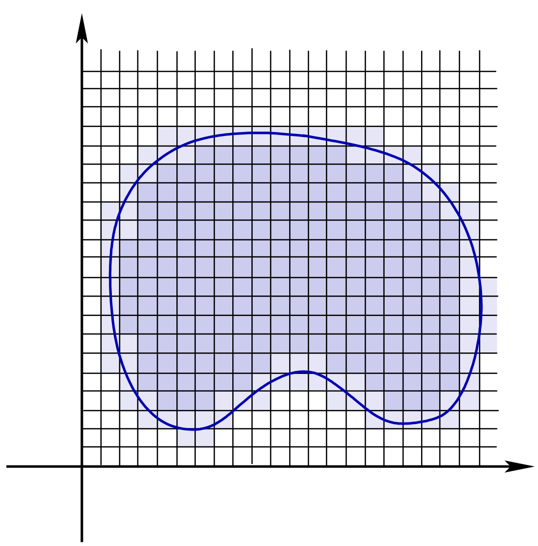
$$\mathrm{N}_0^+ \;:=\; \mathsf{Card}\,\big\{\text{pavés de } \boldsymbol{P}_0 \text{ intersectant } A\big\}.$$

Évidemment, comme chaque pavé de $\boldsymbol{P}_0$ est d'aire 1^2, on a :

$$1^2 \cdot \mathrm{N}_0^- \;\leqslant\; \overset{?}{\mathsf{aire}}(A) \;\leqslant\; 1^2 \cdot \mathrm{N}_0^+,$$

mais on ne sait pas encore calculer l'aire de A, ce qu'on signifie en la chapeautant par un point d'interrogation. En tout cas, puisque A est borné, on a :

$$\mathrm{N}_0^+ \;<\; \infty.$$



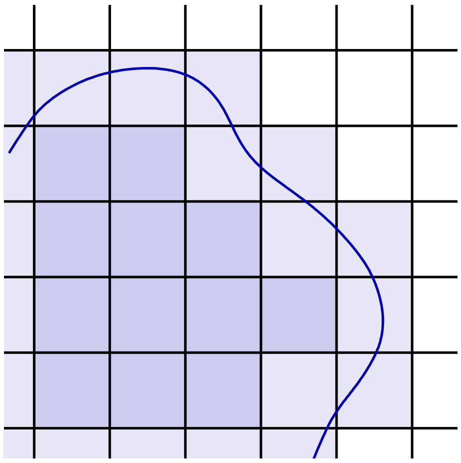
Passons ensuite au pavage suivant $\boldsymbol{P}_1$, plus fin que $\boldsymbol{P}_0$, et comptons de manière similaire :

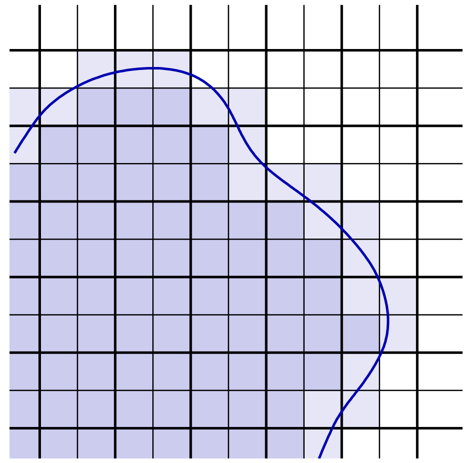
$$\mathrm{N}_1^- \;:=\; \mathsf{Card}\,\big\{\text{pavés de } \boldsymbol{P}_1 \text{ contenus dans } A\big\},$$

$$\mathrm{N}_1^+ \;:=\; \mathsf{Card}\,\big\{\text{pavés de } \boldsymbol{P}_1 \text{ intersectant } A\big\}.$$

Pour la même raison, et comme chaque pavé de $\boldsymbol{P}_1$ est d'aire $\frac{1}{2^2}$:

$$\frac{1}{2^2}\,\mathrm{N}_1^- \ \leqslant\ \overset{?}{\mathsf{aire}}(A) \ \leqslant\ \frac{1}{2^2}\,\mathrm{N}_1^+.$$





Mais comme tous les sous-pavés d'un pavé contenu dans A sont toujours *aussi* contenus dans A, et comme les sous-pavés d'un pavé intersectant A n'intersectent *pas toujours* A, on déduit deux inégalités cruciales :

$$\mathrm{N}_0^- \ \leqslant\ \frac{1}{2^2}\,\mathrm{N}_1^- \qquad\qquad \text{et} \qquad\qquad \frac{1}{2^2}\,\mathrm{N}_1^+ \ \leqslant\ \mathrm{N}_0^+,$$

ce qui confirme l'intuition géométrique d'après laquelle la génération 1 approxime mieux l'« *aire* » de A que la génération 0 :

$$\mathrm{N}_0^- \ \leqslant\ \frac{1}{2^2}\,\mathrm{N}_1^- \ \leqslant\ \overset{?}{\mathsf{aire}}(A) \ \leqslant\ \frac{1}{2^2}\,\mathrm{N}_1^+ \ \leqslant\ \mathrm{N}_0^+,$$

Sur cette pauvre aire (en question), l'étau se resserre si on itère. En effet, pour tout entier $k \in \mathbb{N}$, avec les deux comptages :

$$\begin{aligned} \mathrm{N}_k^- &:= \mathsf{Card}\,\big\{\text{pavés de } \boldsymbol{P}_k \text{ contenus dans } A\big\}, \\ \mathrm{N}_k^+ &:= \mathsf{Card}\,\big\{\text{pavés de } \boldsymbol{P}_k \text{ intersectant } A\big\}, \end{aligned}$$

le même raisonnement montre que lors du passage de $\boldsymbol{P}_{k-1}$ à $\boldsymbol{P}_k$, l'approximation s'améliore nécessairement :

$$\frac{1}{2^{2(k-1)}}\mathrm{N}_{k-1}^- \ \leqslant\ \frac{1}{2^{2k}}\,\mathrm{N}_k^- \ \leqslant\ \overset{?}{\mathsf{aire}}(A) \ \leqslant\ \frac{1}{2^{2k}}\,\mathrm{N}_k^+ \ \leqslant\ \frac{1}{2^{2(k-1)}}\mathrm{N}_{k-1}^+,$$

puisque tout pavé de $\boldsymbol{P}_{k-1}$ qui est inclus dans A donne 2^2 pavés de $\boldsymbol{P}_k$ *aussi* inclus dans A — ce qui démontre la première inégalité —, et puisque tout pavé de $\boldsymbol{P}_{k-1}$ qui rencontre A donne *au plus* 2^2 pavés de $\boldsymbol{P}_k$ qui rencontrent A — ce qui démontre la dernière inégalité.

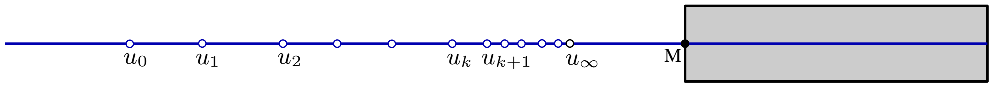
On obtient par conséquent à gauche une première suite de valeurs approchées de l'« *aire* » de A, première suite qui est *croissante et majorée* (par $\mathrm{N}_0^+ < \infty$) :

$$\frac{1}{2^0}\,\mathrm{N}_0^- \ \leqslant\ \frac{1}{2^2}\,\mathrm{N}_1^- \ \leqslant\ \cdots\ \leqslant\ \frac{1}{2^{2k}}\,\mathrm{N}_k^- \ \leqslant\ \cdots \ \underset{k\to\infty}{\longrightarrow}\ \boldsymbol{?},$$

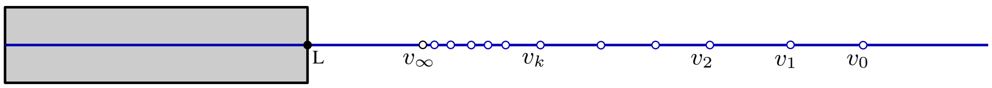
et à droite, on obtient une deuxième suite de valeurs approchées par excès qui est *décroissante et minorée* (par 0, puisque tout est positif !) :

$$\boldsymbol{?}\ \underset{k\to\infty}{\longleftarrow}\ \cdots\ \frac{1}{2^{2k}}\,\mathrm{N}_k^+ \ \leqslant\ \cdots\ \leqslant\ \frac{1}{2^2}\,\mathrm{N}_1^+ \ \leqslant\ \frac{1}{2^0}\,\mathrm{N}_0^+.$$

Or le Cours d'Analyse a démontré qu'une suite numérique $(u_k)_{k=0}^{\infty}$ croissante $u_k \leqslant u_{k+1}$ et majorée $u_k \leqslant \mathrm{M} < \infty$ possède toujours une limite $u_\infty \leqslant \mathrm{M}$.



Inversement, on sait aussi qu'une suite numérique $(v_k)_{k=0}^{\infty}$ décroissante $v_{k+1} \leqslant v_k$ et minorée $v_k \geqslant \mathrm{L} > -\infty$ possède toujours une limite $v_\infty \geqslant \mathrm{L}$.

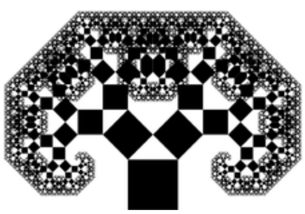
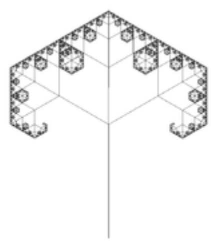


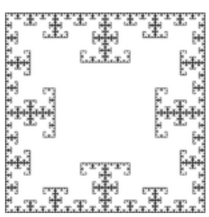
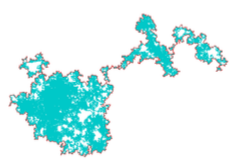
Grâce à ces deux théorèmes fondamentaux dont la démonstration rigoureuse remonte au début du 19ème siècle, nous déduisons que nos deux suites approximantes ont une limite :

$$\lim_{k\to\infty} \frac{1}{2^{2k}}\,\mathrm{N}_k^- =: \mathsf{aire}^-(A) \leqslant \overset{?}{\mathsf{aire}}(A) \leqslant \mathsf{aire}^+(A) := \lim_{k\to\infty} \frac{1}{2^{2k}}\,\mathrm{N}_k^+,$$

et s'il doit exister une « *aire* » ou « *mesure de surface* » de A, elle est nécessairement *comprise* — pour des raisons géométriques évidentes — entre ces deux limites.

Terminologie 3.1. La limite croissante à gauche $\mathsf{aire}^-(A)$ sera appelée *aire inférieure* de A, et la limite décroissante à droite $\mathsf{aire}^+(A)$ sera appelée *aire supérieure* de A.

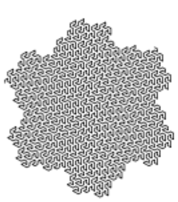
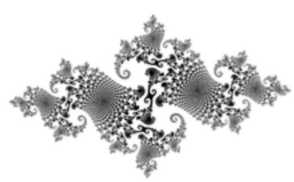
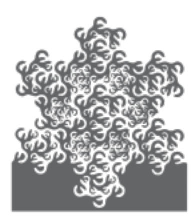



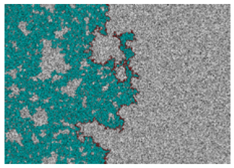
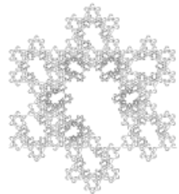
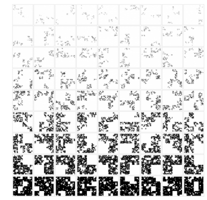




Malheureusement, il existe certains sous-ensembles bizarroïdes $A \subset \mathbb{R}^2$ — penser à des fractals, ou à des taches d'encre transpercées d'acide à toutes les échelles microscropiques — pour lesquels ces deux aires inférieure et supérieure diffèrent :

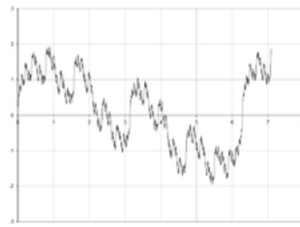
$$\mathsf{aire}^-(A) \overset{\text{strictement}}{<} \mathsf{aire}^+(A).$$

L'étude de tels ensembles va bien au-delà du niveau de ce cours, dans lequel tous les ensembles A considérés auront des frontières très simples et très régulières, et pour lesquels il est assez aisé de vérifier que ces deux aires inférieure et supérieure *coïncident*.

En tout état de cause, nous avons l'intuition que la plupart du temps, les approximations convergent vers une unique valeur.



Oublions, donc, ces figures fantomatiques qui semblent mystérieuses.

Définition 3.2. On dit qu'une partie $A \subset \mathbb{R}^2$ du plan est *quarrable* si son aire inférieure est égale à son aire supérieure :

$$\mathsf{aire}^-(A) \,=\, \mathsf{aire}^+(A).$$

Terminologie 3.3. La valeur commune s'appelle alors l'*aire* de A :

$$\mathsf{aire}(A) \,:=\, \mathsf{aire}^-(A) \,=\, \mathsf{aire}^+(A).$$

Évidemment, l'ensemble vide $\emptyset \subset \mathbb{R}^2$ est quarrable, d'aire nulle :

$$\mathsf{aire}(\emptyset) \,=\, 0.$$

Par des arguments mathématiques rigoureux — et arides —, on démontre des propriétés naturelles et intuitivement évidentes, que nous admettrons.

Lemme 3.4. *Pour tout entier $k \geqslant 0$, tout pavé de profondeur k est quarrable, et a pour mesure de surface :*

$$\mathsf{aire}\bigg(\Big[\frac{m}{2^k},\,\frac{m+1}{2^k}\Big[\,\times\,\Big[\frac{p}{2^k},\,\frac{p+1}{2^k}\Big[\bigg) \,=\, \frac{1}{2^k}\cdot\frac{1}{2^k} \,=\, \frac{1}{2^{2k}}. \qquad \square$$

Plus généralement, considérons un rectangle semi-ouvert quelconque.

Lemme 3.5. *Pour tous nombres réels :*

$$-\infty \,<\, a \,<\, b \,<\, \infty \qquad\qquad \textit{et} \qquad\qquad -\infty \,<\, c \,<\, d \,<\, \infty,$$

le rectangle $[a,b[\times[c,d[$ est quarrable, et a pour mesure de surface :

$$\mathsf{aire}\big([a,b[\,\times[c,d[\big) \,=\, (b-a)\,(d-c). \qquad \square$$

Il en va de même pour le rectangle fermé $[a,b]\times[c,d]$.

Grâce à ces énoncés élémentaires, l'ensemble :

$$\mathscr{Q} \,:=\, \big\{A \subset \mathbb{R}^2 \text{ partie quarrable}\big\} \,\neq\, \emptyset$$

est non vide. Par construction, il existe alors une application bien définie :

$$\begin{aligned} \mathscr{Q} &\longrightarrow \mathbb{R}_+ \\ A &\longmapsto \mathsf{aire}(A), \end{aligned}$$

qui attribue la valeur 0 à $\emptyset$, et la valeur 1 à tout pavé $[m,m+1[\,\times[p,p+1[$ de profondeur 1.

Proposition 3.6. *Si $A, B \in \mathcal{Q}$ sont deux parties quarrables du plan, alors :*

(1) $A \cup B \in \mathcal{Q}$;

(2) $A \cap B \in \mathcal{Q}$;

(3) $A \cap B = \emptyset \Longrightarrow \mathsf{aire}(A \cup B) = \mathsf{aire}(A) + \mathsf{aire}(B)$;

(4) $A \subset B \Longrightarrow \mathsf{aire}(A) \leqslant \mathsf{aire}(B)$;

(5) $\mathsf{aire}\big(\Phi(A)\big) = \mathsf{aire}(A)$ *pour toute isométrie euclidienne* Φ *de* $\mathbb{R}^2$. □

Rappelons que les isométries euclidiennes sont tout simplement des compositions de translations et de rotations dans le plan.

(3') Plus généralement, sans l'hypothèse $A \cap B = \emptyset$, on a :

$$\mathsf{aire}\big(A \cup B\big) \;=\; \mathsf{aire}(A) + \mathsf{aire}(B) - \mathsf{aire}\big(A \cap B\big).$$

4. Domaines de $\mathbb{R}^2$ à bords réguliers par morceaux

Dans les exercices mathématiques de calcul intégral, et dans les applications à la physique, les sous-ensembles $A \subset \mathbb{R}^2$ seront des *domaines :*

$$D \,\subset\, \mathbb{R}^2,$$

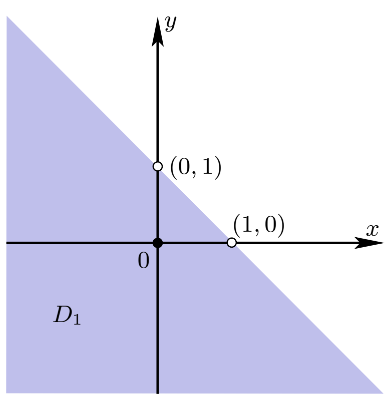
c'est-à-dire des ouverts connexes, toujours représentés par un nombre fini d'inégalités explicites :

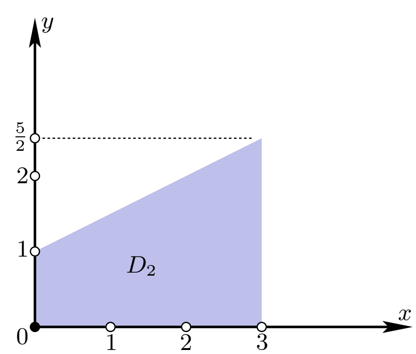
$$D \;=\; \big\{(x,y) \in \mathbb{R}^2 \colon\; d_1(x,y) < 0,\; d_2(x,y) < 0,\; \ldots,\; d_{\text{K}}(x,y) < 0\big\},$$

où les fonctions $d_k(x,y)$ pour $1 \leqslant k \leqslant \text{K}$ seront concrètes et compréhensibles. Donnons deux premiers exemples.

$$\begin{aligned} D_1 \;&:=\; \big\{(x,y) \in \mathbb{R}^2 \colon\; x + y < 1\big\}, \\ D_2 \;&:=\; \big\{(x,y) \in \mathbb{R}^2 \colon\; 0 < y,\; 0 < x < 3,\; 2y - 2 < x\big\}. \end{aligned}$$

Ce premier domaine D_1 est un demi-plan, il n'est pas borné, donc nous l'écartons.





En pratique, pour dessiner un domaine de cette espèce :

$$D \;=\; \big\{(x,y) \in \mathbb{R}^2 \colon\; d_k(x,y) < 0,\; k = 1, \ldots, \text{K}\big\},$$

on trace toutes les courbes :

$$\gamma_k \;:=\; \big\{(x,y) \in \mathbb{R}^2 \colon\; d_k(x,y) = 0\big\} \qquad\qquad (1 \leqslant k \leqslant \text{K}),$$

dans le plan. Pour D_1, on a une unique droite $\{x + y - 1 = 0\}$, et pour D_2, on a 4 droites :

$$\big\{y = 0\big\}, \qquad \big\{x = 0\big\}, \qquad \big\{x - 3 = 0\big\}, \qquad \big\{-x + 2y - 2 = 0\big\}.$$

Rappelons que pour représenter une droite d'équation $ax + by + c = 0$ avec $a \neq 0$, on marque son point $(0, -\frac{c}{b})$ d'intersection avec l'axe des ordonnées $\{x = 0\}$, son point $(-\frac{c}{a}, 0)$ d'intersection avec l'axe des abscisses $\{y = 0\}$, puis on utilise une règle pour tracer la droite entre ces deux points.

Ensuite, chaque courbe $\{d_k(x, y) = 0\}$ délimite, comme frontière commune, deux parties du plan :

$$\big\{d_k(x, y) < 0\big\} \qquad \text{et} \qquad \big\{d_k(x, y) > 0\big\},$$

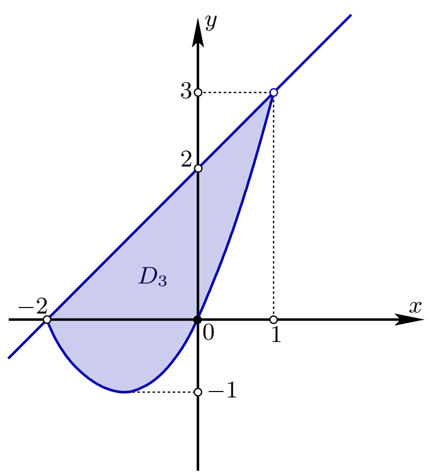
puis on repère la partie où la fonction d_k est (strictement) négative, et enfin, on *hachure* ce lieu. Le domaine $D = \bigcap_k \{d_k < 0\}$ est alors l'*intersection* des régions hachurées.

Évidemment, il faut voir/ré-écrire :

$$D_2 = \Big\{(x, y) \in \mathbb{R}^2 \colon\ -y < 0,\ -x < 0,\ x - 3 < 0,\ -x + 2y - 2 < 0\Big\}.$$

D'ailleurs, le sens des inégalités $d_k < 0$ n'a pas vraiment d'importance, on peut toujours s'y ramener, par exemple :

$$f - 5\,g > h \qquad \Longleftrightarrow \qquad -f + 5\,g + h < 0.$$



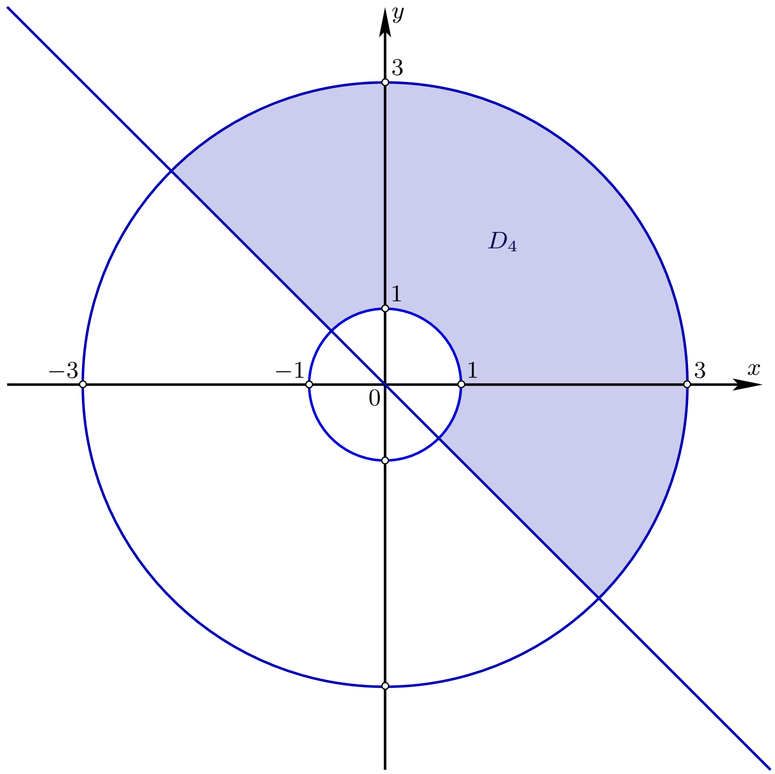
Voici un troisième exemple :

$$D_3 := \big\{(x, y) \in \mathbb{R}^2 \colon\ x^2 + 2x < y < x + 2\big\},$$

que l'on peut ré-écrire :

$$\big\{(x + 1)^2 - 1 < y\big\} \bigcap \big\{y < x + 2\big\},$$

comme intersection de la partie au-dessus d'une parabole d'axe la droite $\{x = -1\}$ et au-dessous d'une droite de pente égale à 1. Ici, le bord ou la frontière ∂D_3 de ce secteur parabolique est constitué d'un segment et d'un arc de parabole.



Encore un autre exemple, un anneau scié :

$$D_4 := \big\{(x,y) \in \mathbb{R}^2 \colon\ 1 < x^2 + y^2 < 3,\ x + y > 0\big\}.$$

Le bord ∂D_4 de ce demi-anneau est constitué de deux demi-cercles concentriques, et de deux segments de droites alignés.

Généralement, étant donné un domaine $D = \{d_1 < 0, \dots, d_{\mathrm{K}} < 0\}$ comme ci-dessus, son *bord* ∂D est contenu dans la réunion finie des courbes :

$$\gamma_k := \big\{(x,y) \in \mathbb{R}^2 \colon\ d_k(x,y) = 0\big\} \qquad (1 \leqslant k \leqslant \mathrm{K}).$$

Théorème 4.1. *Tout domaine borné du plan de la forme :*

$$D = \big\{(x,y) \in \mathbb{R}^2 \colon\ d_1(x,y) < 0,\ d_2(x,y) < 0,\ \dots,\ d_{\mathrm{K}}(x,y) < 0\big\},$$

où $d_1, d_2, \dots, d_{\mathrm{K}}$ sont des fonctions continûment différentiables, est quarrable, avec :

$$\mathsf{aire}(D) = \mathsf{aire}\big(D \cup \partial D\big). \qquad \square$$

Autrement dit, le bord compte pour du beurre — miam ! Bien que la démonstration ne soit pas accessible avec les outils présentés dans ce cours, expliquons l'hypothèse concernant les fonctions $d_1, d_2, \dots, d_{\mathrm{K}}$.

Une fonction :

$$d \colon \quad \mathbb{R}^2 \longrightarrow \mathbb{R}$$

est dite *continûment différentiable* si ses deux dérivées partielles :

$$\frac{\partial d}{\partial x}(x,y) := \lim_{\varepsilon \to 0} \frac{d(x+\varepsilon,\, y) - d(x,y)}{\varepsilon} \qquad \text{et} \qquad \frac{\partial d}{\partial y}(x,y) := \lim_{\varepsilon \to 0} \frac{d(x,\, y+\varepsilon) - d(x,y)}{\varepsilon}$$

existent en tout point $(x,y) \in \mathbb{R}^2$, et sont continues sur $\mathbb{R}^2$. En pratique, cela sera toujours le cas. Plus bas, la Définition 7.1 explicitera précisément ce qu'on entend par *continuité* d'une fonction de deux variables.

5. μ-partition d'un ensemble quarrable

Si, comme plus haut, $\mathscr{Q} \subset \mathscr{P}(\mathbb{R}^2)$ désigne l'ensemble des parties quarrables du plan euclidien $\mathbb{R}^2$, c'est-à-dire celles auxquelles on peut attribuer une aire, nous abrégerons dorénavant :

$$\mu := \text{aire},$$

la lettre grecque 'μ' étant l'initiale du mot grec $\mu\epsilon\tau\rho o\nu$, signifiant 'mesure'.

Rappelons que tout pavé fermé $[a, a+1] \times [c, c+1]$ de côté 1 a pour mesure :

$$\big(a+1-1\big)\,\big(c+1-1\big) = 1 \cdot 1 = 1,$$

et que les parties quarrables $A, B \in \mathscr{Q}$ satisfont la formule :

$$\mu\big(A \cup B\big) = \mu(A) + \mu(B) - \mu\big(A \cap B\big).$$

Terminologie 5.1. On dira que deux parties quarrables $A, B \in \mathscr{Q}$ sont μ-*disjointes* quand leur intersection est d'aire nulle :

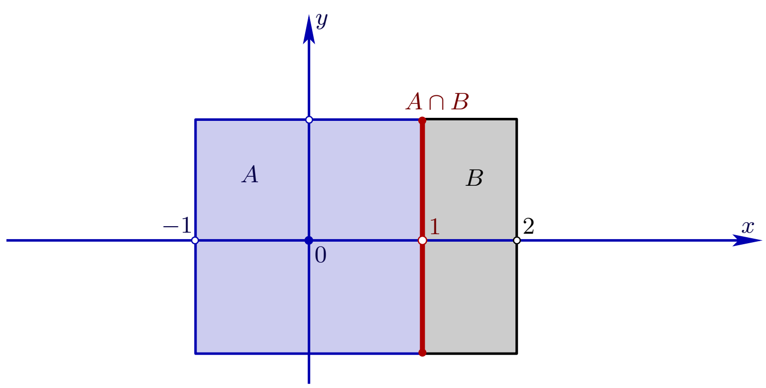
$$0 = \mu\big(A \cap B\big).$$

Cette propriété sera parfois notée :

$$\emptyset_\mu = A \cap B.$$

Dans cette circonstance, on a μ-additivité simple :

$$A \cap B = \emptyset_\mu \qquad \Longrightarrow \qquad \begin{aligned} \mu\big(A \cup B\big) &= \mu(A) + \mu(B) - \underline{\mu(A \cap B)} \\ &= \mu(A) + \mu(B) - 0. \end{aligned}$$



Puisqu'on peut vérifier que tout segment de droite est d'aire nulle (heureusement !), deux pavés fermés n'ayant qu'un côté en commun satisfont cela, par exemple :

$$A := [-1, 1] \times [-1, 1] \qquad \text{et} \qquad B := [1, 2] \times [-1, 1],$$
$$A \cap B = \{1\} \times [-1, 1].$$

Définition 5.2. On appelle μ-*partition* d'une partie quarrable $A \in \mathscr{Q}$ toute famille finie $(A_i)_{1 \leqslant i \leqslant n}$ de sous-ensembles $A_i \subset A$ satisfaisant :

(1) $A_i \in \mathscr{Q}$ pour tout indice $1 \leqslant i \leqslant n$;

(2) $A_1 \cup \cdots \cup A_i \cup \cdots \cup A_n = A$;

(3) $A_{i_1} \cap A_{i_2} = \emptyset_\mu$, pour tous indices $1 \leqslant i_1 < i_2 \leqslant n$.

En termes plus 'littéraires', une μ-partition de A est constituée d'une famille finie de parties quarrables, deux à deux μ-disjointes, dont la réunion donne A. On a alors évidemment :

$$\mu(A) = \mu(A_1) + \cdots + \mu(A_i) + \cdots + \mu(A_n).$$

Définition 5.3. On dira qu'une μ-partition $(A'_{i'})_{1\leqslant i\leqslant n'}$ de $A \in \mathscr{Q}$ est *plus fine* qu'une autre μ-partition $\big(A_i\big)_{1\leqslant i\leqslant n}$ si tout A_i est réunion d'une sous-famille extraite de $(A_{i'})_{1\leqslant i'\leqslant n'}$, à savoir si pour tout i, il existe $A'_{i'_1}, \dots, A'_{i'_p}$ avec $1 \leqslant i'_1 < \cdots < i'_p \leqslant n'$ tels que :

$$A_i \,=\, A'_{i'_1} \cup \cdots \cup A'_{i'_p}.$$

Autrement dit, une μ-partition *plus fine* possède *plus* de morceaux de puzzle.

Dès qu'on connaît *deux* μ-partitions quelconques de A :

$$\big(A_i\big)_{1\leqslant i\leqslant n} \qquad\qquad \text{et} \qquad\qquad \big(B_j\big)_{1\leqslant j\leqslant m},$$

la famille des intersections :

$$\big(A_i \cap B_j\big)_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant m}}$$

est une nouvelle μ-partition de A à $n \cdot m$ morceaux de puzzle, qui est plus fine que chacune des deux μ-partitions données.

Notation 5.4. On notera $\mathscr{P}_\mu(A)$ l'ensemble des μ-partitions de A, et une μ-partition donnée $(A_i)_{1\leqslant i\leqslant n}$ sera notée en abrégé :

$$p \,:=\, \big(A_i\big)_{1\leqslant i\leqslant n}.$$

6. Intégrale double d'une fonction définie sur un ensemble quarrable

Considérons une partie quarrable $A \subset \mathbb{R}^2$, ainsi qu'une fonction *bornée :*

$$f\colon \quad A \longrightarrow \mathbb{R}.$$

D'après un théorème d'Analyse, l'infimum et le supremum de f sur A existent, et sont finis :

$$-\infty \,<\, m := \inf_{(x,y)\in A} f(x,y) \,\leqslant\, \sup_{(x,y)\in A} f(x,y) =: M \,<\, \infty.$$

Étant donné une μ-partition $p = (A_i)_{1\leqslant i\leqslant n}$ de A, introduisons aussi les infima et les suprema de f sur les A_i :

$$m_i \,:=\, \inf_{(x,y)\in A_i} f(x,y) \qquad\qquad \text{et} \qquad\qquad M_i \,:=\, \sup_{(x,y)\in A_i} f(x,y).$$

Ensuite, formons les deux sommes de Darboux 2-dimensionnelles :

$$\begin{aligned} s(p) \,&=\, s\big(A_\bullet\big) \,:=\, m_1\,\mu(A_1) + m_2\,\mu(A_2) + \cdots + m_n\,\mu(A_n), \\ S(p) \,&=\, S\big(A_\bullet\big) \,:=\, M_1\,\mu(A_1) + M_2\,\mu(A_2) + \cdots + M_n\,\mu(A_n). \end{aligned}$$

Il est clair que :

$$s(p) \,\leqslant\, S(p). \tag{6.1}$$

On peut alors développer la même théorie que celle de l'intégrale de Riemann en dimension 1, via les sommes de Darboux, dont voici, sans démonstrations, les passages importants. Deux propositions principales interviennent.

Proposition 6.2. *Si une μ-partition $(A'_{i'})_{1\leqslant i'\leqslant n'}$ d'un ensemble quarrable $A \subset \mathbb{R}^2$ est plus fine qu'une autre μ-partition $\big(A_i\big)_{1\leqslant i\leqslant n}$, alors :*

$$s\big(A_i\big) \,\leqslant\, s\big(A'_{i'}\big) \qquad\qquad \textit{et} \qquad\qquad S\big(A'_{i'}\big) \,\leqslant\, S\big(A_i\big). \qquad \square$$

Cette proposition est absolument cruciale, car elle exprime qu'en raffinant les μ-partitions :

$$s\big(A_i\big) \leqslant s\big(A'_{i'}\big) \leqslant s\big(A''_{i''}\big) \leqslant \cdots \leqslant \int\!\!\int_A f(\overset{?}{x},y)\,dxdy \leqslant \cdots \leqslant S\big(A''_{i''}\big) \leqslant S\big(A'_{i'}\big) \leqslant S\big(A_i\big),$$

on a de plus en plus de chances d'approcher la valeur d'une intégrale double de f sur A :

$$\int\!\!\int_A f(\overset{?}{x},y) \approx \sum_{i=1}^{n} \mu(A_i)\cdot\Big(\text{valeur de } f \text{ sur } A_i\Big),$$

en s'imaginant que si tous les A_i sont assez petits, les valeurs de f sur A_i sont presque constantes :

$$\inf_{A_i} f \approx \text{valeur(s) de } f \text{ sur } A_i \approx \sup_{A_i} f.$$

Bien entendu, ce (misérable) raisonnement 'à la physicienne' sera autoritairement exclu de notre '*mathematics select club*' ! Soyons rigoureux !

La deuxième proposition principale est encore plus cruciale.

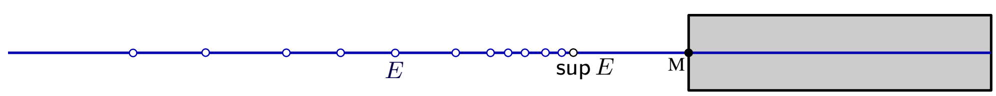
Proposition 6.3. *Étant donné deux μ-partitions* quelconques $p \in \mathscr{P}_\mu(A)$ *et* $q \in \mathscr{P}_\mu(A)$*, on a toujours :*

$$s(p) \leqslant S(q).$$

Preuve. La démonstration est très simple : la μ-partition intersection $p \cap q$ étant à la fois plus fine que p et plus fine que q, la Proposition 6.2 et l'inégalité (6.1) donnent :

$$s(p) \leqslant s\big(p\cap q\big) \leqslant S\big(p\cap q\big) \leqslant S(q). \qquad \square$$

Grâce à cela, l'ensemble des nombres $s(p)$ quand p parcourt $\mathscr{P}_\mu(A)$ est *majoré* par $S(q_0)$, où $q_0 \in \mathscr{P}_\mu(A)$ est fixée quelconque. À présent, rappelons le

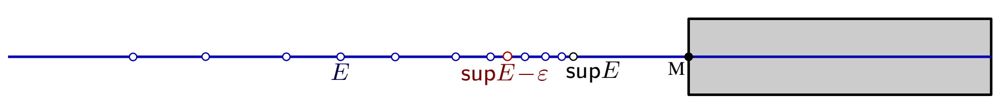


Lemme 6.4. [Borne supérieure] *Soit $E \subset \mathbb{R}$ un sous-ensemble de nombres réels qui est borné supérieurement au sens où il existe une constante* $\mathrm{M} < \infty$ *majorant tous ses éléments :*

$$x \leqslant \mathrm{M} \qquad (\forall\, x \in E).$$

Alors E admet un supremum :

$$\sup E := \sup_{x\in E} x \leqslant \mathrm{M}. \qquad \square$$



Par définition, ce supremum est l'unique nombre réel $\sup E \in \mathbb{R}$ majorant *au mieux* tous les éléments :

$$x \leqslant \sup E \qquad (\forall\, x \in E),$$

au sens où dès qu'on descend un tout petit peu la barre à droite, à savoir dès qu'on soustrait un $\varepsilon > 0$ arbitrairement petit, il n'est plus vrai que :

$$x \overset{\textbf{faux}}{\leqslant} \sup E - \varepsilon \qquad (\forall\, x \in E).$$

De manière équivalente, pour tout $\varepsilon > 0$, il existe $x \in E$ avec $x > \sup E - \varepsilon$. Attention ! En général, $\sup E < \mathrm{M}$ strictement : le 'mur' M peut *a priori* être très loin du supremum $\sup E$, comme sur les figures.

Par conséquent, on peut introduire l'*intégrale inférieure de f sur A :*

$$\sigma_f(A) := \sup_{p \in \mathscr{P}_\mu(A)} s(p).$$

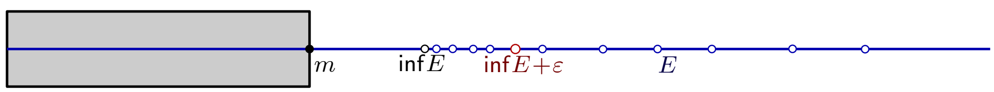
Ensuite, de manière symétrique, l'ensemble des nombres $S(q)$ quand q parcourt $\mathscr{P}_\mu(A)$ est *minoré* par $s(p_0)$, où $p_0 \in \mathscr{P}_\mu(A)$ est fixée quelconque.

Lemme 6.5. [Borne inférieure] *Si $E \subset \mathbb{R}$ est borné inférieurement au sens où il existe une borne finie $m > -\infty$ minorant tous ses éléments :*

$$m \leqslant x \qquad (\forall\, x \in E),$$

alors E admet un infimum :

$$m \leqslant \inf_{x \in E} x =: \inf E. \qquad \square$$



Par définition, cet infimum est l'unique nombre réel $\inf E \in \mathbb{R}$ minorant *au mieux* tous les éléments :

$$\inf E \leqslant x \qquad (\forall\, x \in E),$$

au sens où dès qu'on remonte un tout petit peu la barre à gauche, à savoir dès qu'on additionne un $\varepsilon > 0$ arbitrairement petit, il n'est plus vrai que :

$$\inf E + \varepsilon \overset{\textbf{faux}}{\leqslant} x \qquad (\forall\, x \in E).$$

Par conséquent, on peut introduire l'*intégrale supérieure de f sur A :*

$$\Sigma_f(A) := \inf_{p \in \mathscr{P}_\mu(A)} S(p).$$

Par construction, on a évidemment :

$$\sigma_f(A) \leqslant \Sigma_f(A).$$

Définition 6.6. On dit qu'une fonction bornée $f \colon A \longrightarrow \mathbb{R}$ définie sur un ensemble quarrable $A \subset \mathbb{R}^2$ est *intégrable au sens de Darboux sur A* lorsque son intégrale inférieure coïncide avec son intégrale supérieure :

$$\sigma_f(A) = \Sigma_f(A).$$

On note alors cette valeur commune :

$$\iint_A f(x,y)\, dxdy,$$

et on la nomme *intégrale double de f sur la partie quarrable A.*

Notation 6.7. L'ensemble des fonctions intégrables sur A sera noté :

$$\mathrm{Int}\,(A, \mathbb{R}) := \big\{ f \colon A \longrightarrow \mathbb{R} \colon \text{ bornées intégrables} \big\}.$$

Comme dans la théorie des intégrales de Riemann sur des segments de $\mathbb{R}$, on démontre plusieurs théorèmes fondamentaux.

Théorème 6.8. *Soient A et B deux parties quarrables du plan qui sont μ-disjointes :*

$$0 = \mu\big(A \cap B\big).$$

Pour toute fonction bornée $f \colon A \cup B \longrightarrow \mathbb{R}$, on a l'équivalence :

$$f \text{ est intégrable sur } A \text{ et sur } B \qquad \Longleftrightarrow \qquad f \text{ est intégrable sur } A \cup B.$$

Dans ce cas :

$$\int\!\!\int_{A\cup B} f(x,y)\,dxdy = \int\!\!\int_A f(x,y)\,dxdy + \int\!\!\int_B f(x,y)\,dxdy. \qquad \square$$

En élaguant les notations, on peut récrire cela :

$$\int\!\!\int_{A\cup B} f = \int\!\!\int_A f + \int\!\!\int_B f.$$

Ensuite, le théorème suivant montre que l'espace $\mathsf{Int}\,(A, \mathbb{R})$ est un $\mathbb{R}$-espace vectoriel.

Théorème 6.9. *Si f et g sont deux fonctions bornées intégrables sur une partie quarrable $A \subset \mathbb{R}^2$, alors pour tous réels $\lambda, \mu \in \mathbb{R}$, la combinaison linéaire :*

$$\lambda\, f + \mu\, g \in \mathsf{Int}\,(A, \mathbb{R})$$

est aussi intégrable sur A, avec :

$$\int\!\!\int_A \Big(\lambda\, f(x,y) + \mu\, g(x,y)\Big)\,dxdy = \lambda \int\!\!\int_A f(x,y)\,dxdy + \mu \int\!\!\int_A g(x,y)\,dxdy. \qquad \square$$

Avant d'aller plus loin, exprimons un critère d'intégrabilité. Par construction, sur une partie quarrable $A \subset \mathbb{R}^2$, pour toute fonction bornée $f \colon A \longrightarrow \mathbb{R}$, pour toute paire de μ-partitions $p = (A_i)_{1 \leqslant i \leqslant n}$ et $q = (B_j)_{1 \leqslant j \leqslant m}$ de A, on a des inégalités :

$$s(p) \leqslant \sigma_f(A) \leqslant \overset{?}{\int\!\!\int_A} f \leqslant \Sigma_f(A) \leqslant S(q),$$

et pour que $\iint_A f$ existe, c'est-à-dire pour que $\sigma_f(A) = \Sigma_f(A)$, il faut et il suffit que l'on puisse rendre la différence :

$$S(q) - s(p) \text{ arbitrairement petite.}$$

En remplaçant p et q par la partition plus fine $p \cap q$ qui satisfait :

$$s(p) \leqslant s\big(p \cap q\big) \leqslant \overset{?}{\int\!\!\int_A} f \leqslant S\big(p \cap q\big) \leqslant S(q),$$

et en considérant $p \cap q$, on déduit le

Lemme 6.10. [Critère d'intégrabilité] *Pour que f soit intégrable (au sens de Darboux) sur A, il faut et il suffit que pour tout $\varepsilon > 0$, il existe une μ-partition p de A telle que :*

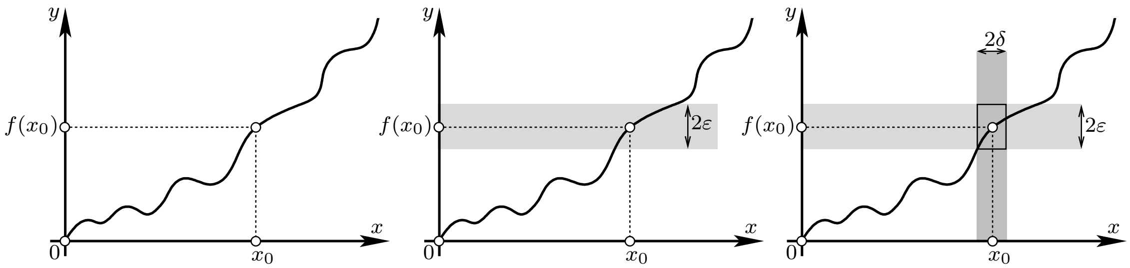
$$S(p) - s(p) \leqslant \varepsilon. \qquad \square$$

7. Fonctions continues sur $D \cup \partial D$

Comme dans la théorie de l'intégrale de Riemann sur un intervalle $[a, b] \subset \mathbb{R}$, il existe des fonctions bornées sur des ensembles quarrables $A \subset \mathbb{R}^2$ qui ne sont *pas* intégrables, *i.e.* pour lesquelles $\sigma_f(A) < \Sigma_f(A)$. Heureusement, dans pratiquement tous les cas existant en mathématiques et en physique, les fonctions considérées sont *continues* — quitte à redécouper A en un nombre fini de morceaux —, et nous allons démontrer que les fonctions continues *sont* intégrables. Avec cela, on peut élaborer des milliers d'exercices excitants !

Rappelons tout d'abord la définition des fonctions continues d'une variable $f\colon [a,b] \longrightarrow \mathbb{R}$. La définition classique due à Weierstrass stipule que f est continue en un point $x_0 \in]a,b[$ lorsque :

$$\forall\,\varepsilon > 0 \quad \exists\,\delta = \delta(\varepsilon) > 0 \quad \Big(\forall\,x \quad |x - x_0| \leqslant \delta \implies \big|f(x) - f(x_0)\big| \leqslant \varepsilon\Big).$$



Géométriquement, quelle que soit la finesse $2\,\varepsilon > 0$ d'une bande horizontale centrée autour de la droite horizontale $\{y = f(x_0)\}$, il existe une bande verticale de largeur $2\,\delta$ assez petite centrée autour de la droite verticale $\{x = x_0\}$ telle que toute la partie du graphe correspondante *reste entièrement enfermée dans la bande horizontale choisie à l'avance*.

Définition 7.1. Sur un sous-ensemble du plan $E \subset \mathbb{R}^2$, une fonction $f\colon E \longrightarrow \mathbb{R}$ est dite *continue en un point fixé* $(x_0, y_0) \in E$ lorsque :

$$(x,y) \longrightarrow (x_0,y_0) \qquad\qquad \text{entraîne} \qquad\qquad f(x,y) \longrightarrow f(x_0,y_0).$$

Ici, x tend vers x_0 et simultanément, y tend vers y_0. Plus précisément, avec la norme euclidienne :

$$\big\|(x,y) - (x_0,y_0)\big\| \;:=\; \sqrt{(x-x_0)^2 + (y-y_0)^2},$$

cela s'exprime rigoureusement par :

$$\forall\,\varepsilon > 0 \quad \exists\,\delta = \delta(x_0,y_0,\varepsilon) \quad \text{tel que} \quad \big\|(x,y) - (x_0,y_0)\big\| \leqslant \delta \implies \big|f(x,y) - f(x_0,y_0)\big| \leqslant \varepsilon.$$

Définition 7.2. Une fonction $f\colon E \longrightarrow \mathbb{R}$ est dite *continue sur* E quand elle est continue en *tout point* $(x_0, y_0) \in E$.

La plupart du temps, nous aurons affaire à des ensembles $E = D \cup \partial D$ qui sont des domaines $D \subset \mathbb{R}^2$ bordés par un nombre fini de courbes $\{d_k(x,y) = 0\}$, et à des fonctions :

$$f\colon \quad D \cup \partial D \longrightarrow \mathbb{R},$$

continues dans D union son bord ∂D. La classe des fonctions continue est stable par les opérations usuelles. Nous admettrons la

Proposition 7.3. **(1)** *Les fonctions polynomiales :*

$$\sum_{\text{finie}} a_{k,\ell}\, x^k\, y^\ell \qquad\qquad (a_{k,\ell} \in \mathbb{R}),$$

sont continues sur $\mathbb{R}^2$ *tout entier.*

(2) *Si* $f\colon D \cup \partial D \longrightarrow \mathbb{R}$ *et* $g\colon D \cup \partial D \longrightarrow \mathbb{R}$ *sont continues, et si* $\lambda, \mu \in \mathbb{R}$ *sont des constantes, alors* $\lambda\, f + \mu\, g$ *et* $f \cdot g$ *sont aussi continues sur* $D \cup \partial D$.

(3) *Si de plus* $g\colon D \cup \partial D \longrightarrow \mathbb{R}^*$ *ne s'annule jamais, alors* $\frac{f}{g}$ *est continue sur* $D \cup \partial D$.

(4) *Si* $f\colon D \cup \partial D \longrightarrow I \subset \mathbb{R}$ *est continue et prend ses valeurs dans un certain sous-intervalle* I *de* $\mathbb{R}$, *et si* $\phi\colon I \longrightarrow \mathbb{R}$ *est une fonction continue d'une variable, alors* $\phi \circ f$ *est continue sur* $D \cup \partial D$. $\square$

Par exemple :

$$(x,y) \longmapsto e^{xy} - 5\sin(x+y) + 3\sqrt{x^2+y^2}$$

est continue sur $\mathbb{R}^2 \backslash \{(0,0)\}$, et n'est *pas* continue en $(0,0)$.

Dans les exemples concrets, pendant les séances TD et aussi lors des examens, les fonctions à étudier seront fabriquées comme sommes, produits, quotients non singuliers, compositions, de fonctions classiques élémentaires connues :

$$x^2, \quad y^\ell, \quad e^x, \quad \log y, \quad \sin x, \quad \cos y, \quad \cosh x, \quad \sinh y, \quad \sqrt{x}, \quad \sqrt[7]{y}, \quad \textit{etc.},$$

et seront continues dans les domaines $D \cup \partial D$ jusqu'à leur bord ∂D. La plupart du temps, on se contentera de vérifier qu'elles sont bien définies, *e.g.* qu'on ne divise pas par 0, qu'on ne prend pas la racine carrée ou le logarithme d'un nombre < 0, *etc.*

Notre objectif, maintenant, est de démontrer que les fonctions continues $f\colon D \cup \partial D \longrightarrow \mathbb{R}$ définies sur des domaines-types sont intégrables. Une section préliminaire est nécessaire.

8. Sous-ensembles compacts du plan $\mathbb{R}^2$

En cours de Topologie Générale (niveau L2 ou L3), on démontre que les domaines *bornés* $D \cup \partial D$ avec leur bord ∂D comme ci-dessus sont des ensembles *fermés*. On peut prendre comme définition le

Théorème 8.1. [Admis] *Pour un sous-ensemble $E \subset \mathbb{R}^2$, on a l'équivalence :*

$$E \text{ est compact} \qquad \Longleftrightarrow \qquad E \text{ est fermé et borné}. \qquad \square$$

Le concept d'ensemble compact est très important dans toutes les mathématiques actuelles. Mais qu'entend-on par « fermé » ?

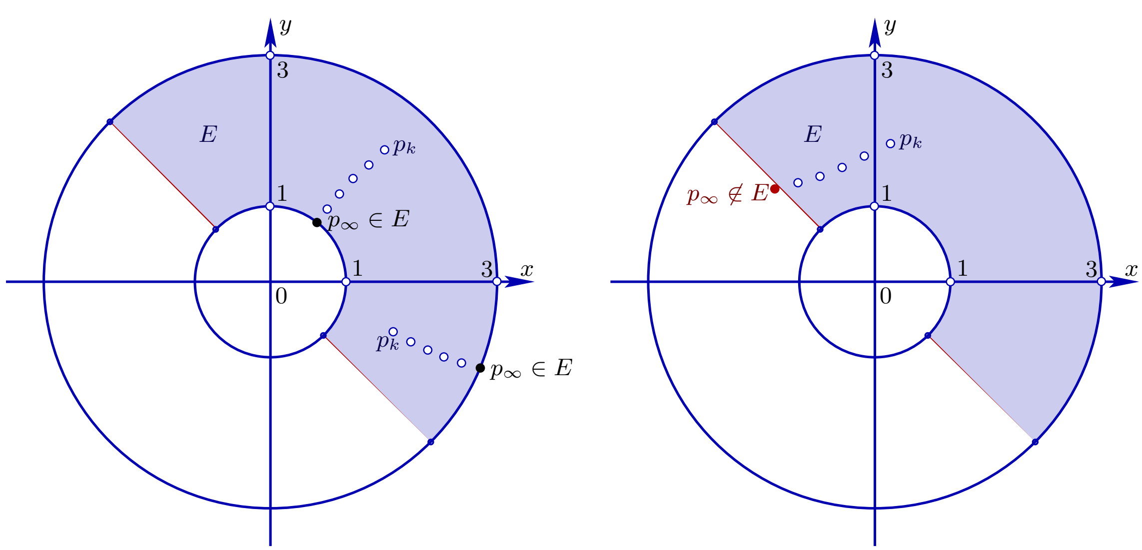
Définition 8.2. Un sous-ensemble $E \subset \mathbb{R}^2$ est dit *fermé* si, pour toute suite de points :

$$(p_k)_{k=1}^{\infty}, \qquad\qquad p_k \in E,$$

qui converge vers un certain point $p_\infty \in \mathbb{R}^2$, on a encore :

$$p_\infty \in E.$$

Autrement dit, les suites de points de E ne peuvent s'échapper de E : caserne !



Par (contre-)exemple, le demi-anneau semi-fermé et semi-ouvert :

$$E := \big\{(x,y) \in \mathbb{R}^2 \colon\ 1 \leqslant x^2 + y^2 \leqslant 3,\ x + y > 0\big\}$$

n'est *pas* fermé (partout) — noter les deux inégalités faibles '$\leqslant$', distinctes de l'inégalité stricte '$>$'. En effet, si une suite $(p_k)_{k=1}^{\infty}$ tend vers un point p_∞ d'un des deux demi-cercles (figure de gauche), on a encore $p_\infty \in E$, car l'inégalité est faible. Mais si p_∞ est sur l'un des deux segments (figure de droite), on a :

$$p_\infty \notin E,$$

car l'inégalité $x+y>0$ est stricte, donc la droite $\{x+y=0\} \cap E = \emptyset$.

Définition 8.3. La *fermeture* (ou *adhérence*) $\overline{E}$ d'un sous-ensemble $E \subset \mathbb{R}^2$ est la réunion de tous les points $p_\infty \in \mathbb{R}^2$ qui sont limites :

$$p_\infty = \lim_{k\to\infty} p_k$$

de suites quelconques de points $p_k \in E$ convergeant dans $\mathbb{R}^2$.

Intuitivement, on ajoute à E tous les points qui se trouvent au bord. Par exemple — noter le changement pour la dernière inégalité — :

$$\overline{E} = \big\{(x,y) \in \mathbb{R}^2 \colon\ 1 \leqslant x^2 + y^2 \leqslant 3,\ x + y \geqslant 0\big\}.$$

En général, il suffit de remplacer les inégalités strictes par des inégalités faibles. Pour nous, l'énoncé utile sera le suivant.

Théorème 8.4. [Admis] *La fermeture d'un domaine borné défini par des fonctions continûment différentiables :*

$$D = \big\{(x,y) \in \mathbb{R}^2 \colon\ d_1(x,y) < 0,\ \dots,\ d_{\mathrm{K}}(x,y) < 0\big\},$$

est :

$$\overline{D} = \big\{(x,y) \in \mathbb{R}^2 \colon\ d_1(x,y) \leqslant 0,\ \dots,\ d_{\mathrm{K}}(x,y) \leqslant 0\big\}.$$

De plus son bord est contenu — souvent strictement — dans la réunion des courbes :

$$\partial D := \overline{D} \backslash D \subset \bigcup_{1 \leqslant k \leqslant \mathrm{K}} \big\{d_k(x,y) = 0\big\}. \qquad \square$$

Ainsi, nous étudierons des domaines bornés $D \subset \mathbb{R}^2$ dont la fermeture $\overline{D} = D \cup \partial D$ est un sous-ensemble *compact* de $\mathbb{R}^2$. Ici, le mot « *compact* » suggère l'idée d'objet bien délimité, sans passoire, empoignable, visible, concret.

Théorème 8.5. [Admis] *Soit $A \subset \mathbb{R}^2$ un sous-ensemble compact,* i.e. *fermé et borné. Alors toute fonction continue $f \in \mathscr{C}^0(A, \mathbb{R})$ sur A est en fait uniformément continue sur A.*

Ce concept de *continuité uniforme* est aussi très important en Analyse. Il exprime que le $\delta = \delta(x_0, y_0, \varepsilon)$ dans la condition de continuité de f en un point $(x_0, y_0) \in A$:

$$\forall\, \varepsilon > 0 \quad \exists\, \delta = \delta(x_0, y_0, \varepsilon) > 0 \quad \text{tel que} \quad \big\|(x,y) - (x_0,y_0)\big\| \leqslant \delta \implies \big|f(x,y) - f(x_0,y_0)\big| \leqslant \varepsilon,$$

ne dépend en fait *pas* de $(x_0, y_0) \in A$, au sens de la

Définition 8.6. Une fonction $f \colon A \longrightarrow \mathbb{R}$ définie sur une sous-ensemble $A \subset \mathbb{R}^2$ est dite *uniformément continue* si :

$$\forall\, \varepsilon > 0 \quad \exists\, \delta = \delta(\varepsilon) > 0 \quad \text{tel que} \quad \begin{pmatrix} \forall\, (x,y) \in A \quad \forall\, (x',y') \in A \\ \big\|(x,y) - (x',y')\big\| \leqslant \delta \implies \big|f(x,y) - f(x',y')\big| \leqslant \varepsilon \end{pmatrix}.$$

Toute fonction uniformément continue est continue, mais l'implication générale inverse :

$$f \quad \text{continue} \qquad \Longrightarrow \qquad f \quad \text{uniformément continue}$$

n'est vraie que pour des fonctions définies sur des sous-ensembles compacts $A \subset \mathbb{R}^2$, ce que dit le théorème ci-dessus.

Un deuxième théorème préliminaire est nécessaire.

Théorème 8.7. *Soit $A \subset \mathbb{R}^2$ un sous-ensemble compact, et soit $f \in \mathscr{C}^0(A, \mathbb{R})$ une fonction continue. Alors :*

(1) *L'image $f(A) \subset \mathbb{R}$ est aussi un sous-ensemble compact —* i.e. *fermé et borné ;*

(2) *Il existe $p_- \in A$ et $p_+ \in A$ tels que :*

$$f(p_-) = \inf_{q \in A} f(q) \qquad \text{et} \qquad f(p_+) = \sup_{q \in A} f(q). \qquad \square$$

Cette deuxième propriété exprime que les bornes inférieure et supérieure d'une fonction continue sont *atteintes* — nul besoin de prendre une limite — en certains points du compact.

9. Intégrale double d'une fonction continue sur un compact quarrable

Nous pouvons enfin énoncer et démontrer le théorème fondamental de ce chapitre.

Théorème 9.1. *Toute fonction continue sur un compact quarrable y est intégrable.*

Démonstration. À tout $\varepsilon > 0$, la continuité uniforme de f sur A fait correspondre un $\delta = \delta(\varepsilon)$ tel que :

$$\big\|(x, y) - (x', y')\big\| \leqslant \delta \implies \big|f(x, y) - f(x', y')\big| \leqslant \varepsilon.$$

Puisque A est quarrable, il existe une μ-partition $p = (A_i)_{1 \leqslant i \leqslant n}$ de A dont les constituants A_i sont tous de diamètre inférieur à δ. Il suffit en effet de prendre les intersections de A avec un pavage $\boldsymbol{P}_k$ de profondeur k assez grande pour que la diagonale de ces pavés carrés satisfasse $\sqrt{2}\,\frac{1}{2^k} \leqslant \delta$. Quitte à ajouter la frontière de ces pavés, qui est d'aire nulle, nous pouvons supposer que tous les A_i sont fermés, donc compacts.

Alors sur chaque A_i pour $1 \leqslant i \leqslant n$, le Théorème 8.7 garantit que f atteint ses bornes inférieure et supérieure :

$$\begin{aligned} &\exists\,(x_i^-, y_i^-) \in A_i, \quad m_i := f\big(x_i^-, y_i^-\big) = \inf_{A_i} f, \\ &\exists\,(x_i^+, y_i^+) \in A_i, \quad M_i := f\big(x_i^+, y_i^+\big) = \sup_{A_i} f. \end{aligned}$$

Or puisque le diamètre de A_i est inférieur à δ, on a :

$$\big\|(x_i^-, y_i^-) - (x_i^+, y_i^+)\big\| \leqslant \delta \qquad \Longrightarrow \qquad (0 \leqslant) \quad M_i - m_i \leqslant \varepsilon.$$

Il en résulte, pour cette μ-partition p :

$$\begin{aligned} S(p) - s(p) &= \sum_{i=1}^{n} M_i\,\mu(A_i) - \sum_{i=1}^{n} m_i\,\mu(A_i) \\ &\leqslant \varepsilon \sum_{i=1}^{n} \mu(A_i) \\ &= \varepsilon\,\mu\big(A_1 + \cdots + A_i + \cdots + A_n\big) \quad = \varepsilon\,\mu(A), \end{aligned}$$

et comme $\varepsilon\,\mu(A)$ est arbitrairement petit si $\varepsilon > 0$ est arbitrairement petit, puisque $\mu(A)$ est une constante, le Critère d'intégrabilité 6.10 est vérifié. Ceci conclut la démonstration de ce théorème fondamental. $\square$

10. Valeur absolue d'une intégrale double et formule de la moyenne

En théorie de l'intégration, les inégalités sont un outil très fréquemment utile.

Théorème 10.1. *Si $f\colon A \longrightarrow \mathbb{R}$ est intégrable sur une partie quarrable $A \subset \mathbb{R}^2$, et si $f \geqslant 0$, alors :*

$$\iint_A f(x,y)\,dxdy \;\geqslant\; 0.$$

Démonstration. En effet, pour une fonction bornée positive $f\colon A \longrightarrow \mathbb{R}_+$, pour toute μ-partition $p = (A_i)_{1\leqslant i\leqslant n}$ de A, les bornes inférieures $m_i \geqslant 0$ de f sur les A_i sont évidemment toutes positives, donc $s(p) \geqslant 0$, et enfin en prenant le supremum, le résultat reste positif :

$$\iint_A f(x,y)\,dxdy \;=\; \sigma_f(A) \;=\; \sup_{p\in\mathscr{P}_\mu(A)} s(p) \;\geqslant\; 0. \qquad \square$$

Corollaire 10.2. *Si f et g sont intégrables sur A, alors :*

$$f \;\leqslant\; g \qquad\Longrightarrow\qquad \iint_A f(x,y)\,dxdy \;\leqslant\; \iint_A g(x,y)\,dxdy.$$

Démonstration. En effet, il suffit d'appliquer le théorème précédent à la fonction positive $g - f \geqslant 0$. □

De la même façon qu'en théorie de l'intégrale de Riemann 1-dimensionnelle, on établit le résultat suivant, exceptionnellement important, en Analyse, en Théorie des nombres, et en Géométrie.

Théorème 10.3. *Si une fonction bornée $f\colon A \longrightarrow \mathbb{R}$ est intégrable sur une partie quarrable $A \subset \mathbb{R}^2$, alors :*

$$\left|\iint_A f(x,y)\,dxdy\right| \;\leqslant\; \iint_A \big|f(x,y)\big|\,dxdy. \qquad \square$$

Toutefois, dans ce cours de L1 majoritairement concentré sur des calculs exacts, nous l'utiliserons peu souvent.

Voici maintenant ce qu'on appelle la *Première formule de la moyenne*.

Théorème 10.4. *Soit $f \in \mathsf{Int}\,(A,\mathbb{R})$ une fonction intégrable sur $A \subset \mathbb{R}^2$ quarrable. Alors en posant :*

$$m \;:=\; \inf_A f \qquad\qquad et \qquad\qquad M \;:=\; \inf_A f,$$

on a :

$$m\cdot\mu(A) \;\leqslant\; \iint_A f(x,y)\,dxdy \;\leqslant\; M\cdot\mu(A).$$

Démonstration. En effet, $m\,\mu(A)$ et $M\,\mu(A)$ sont respectivement les $s(p)$ et $S(p)$ correspondant à la μ-partition $p = A$ réduite à un seul élément, la partie A elle-même tout entière. □

Définition 10.5. Le nombre :

$$\lambda \;:=\; \frac{1}{\mu(A)}\iint_A f(x,y)\,dxdy$$

se nomme *valeur moyenne* de f sur A.

Si A est une partie compacte du plan, et si f est continue sur A, alors il existe $(x_0,y_0) \in A$ tel que $f(x_0,y_0) = \lambda$. On obtient ainsi le

Théorème 10.6. *Si $A \subset \mathbb{R}^2$ est quarrable et si $f\colon A \longrightarrow \mathbb{R}$ est continue, il existe un $(x_0,y_0) \in A$ tel que :*

$$\frac{1}{\mu(A)}\iint_A f(x,y)\,dxdy \;=\; f(x_0,y_0). \qquad \square$$

11. Échantillonnage de Riemann

Un autre point de vue, plus proche de l'esprit original de Bernhard Riemann dans sa thèse d'habilitation soutenue en 1854 à l'Université de Göttingen en Allemagne, mérite d'être présenté.

Comme précédemment, soit $f\colon A \longrightarrow \mathbb{R}$ une fonction bornée définie sur un sous-ensemble quarrable du plan $A \subset \mathbb{R}^2$. Pour toute μ-partition $(A_i)_{1 \leqslant i \leqslant n}$ de A, nous posons à nouveau :

$$m_i := \inf_{(x,y) \in A_i} f(x,y) \qquad \text{et} \qquad M_i := \sup_{(x,y) \in A_i} f(x,y).$$

Par définition, le *diamètre* de chaque A_i est le nombre réel positif :

$$\delta_i := \sup_{\substack{(x,y) \in A_i \\ (x',y') \in A_i}} \big\|(x,y) - (x',y')\big\|.$$

Définition 11.1. Le *pas* de la μ-partition $p = (A_i)_{1 \leqslant i \leqslant n}$ de A est :

$$\mathsf{pas}(p) := \delta := \max_{1 \leqslant i \leqslant n} \delta_i.$$

Maintenant, pour tout indice fixé $1 \leqslant i \leqslant n$, choisissons arbitrairement un réel $\theta_i \in [m_i, M_i]$, par exemple :

$$\theta_i := f(x_i, y_i) \quad \text{avec} \ (x_i, y_i) \in A_i \ \text{quelconque},$$

et introduisons la *somme de Riemann :*

$$\begin{aligned}\mathscr{R}(p) &:= \theta_1\, \mu(A_1) + \theta_2\, \mu(A_2) + \cdots + \theta_n\, \mu(A_n) \\ &= \sum_{i=1}^{n} \theta_i\, \mu(A_i).\end{aligned}$$

Puisque $m_i \leqslant \theta_i \leqslant M_i$, il est clair que :

$$\sum_{i=1}^{n} m_i\, \mu(A_i) \leqslant \mathscr{R}(p) \leqslant \sum_{i=1}^{n} M_i\, \mu(A_i),$$

c'est-à-dire :

$$s(p) \leqslant \mathscr{R}(p) \leqslant S(p).$$

Comme dans la théorie des intégrales de Riemann en dimension 1, on établit le

Théorème 11.2. *Si $f\colon A \longrightarrow \mathbb{R}$ est intégrable sur $A \subset \mathbb{R}^2$ quarrable, alors :*

$$\lim_{\mathsf{pas}(p) \to 0} \mathscr{R}(p) = \iint_A f(x,y)\, dxdy. \qquad \square$$

L'intérêt, c'est qu'on approxime la valeur de l'intégrale par des sommes échantillonnées 'au hasard' : chaque morceau A_i 'pèse' :

$$f(x_i, y_i) \cdot \mathsf{aire}(A_i),$$

où $f(x_i, y_i)$ est une valeur de la fonction en un point (x_i, y_i) choisi arbitrairement dans A_i. Quels que soient les choix de valeurs, il y a convergence vers l'intégrale double : c'est remarquable !

Il y a aussi une réciproque à ce résultat, que nous admettrons aussi sans démonstration (austère, trop austère).

Théorème 11.3. *Toute fonction intégrable au sens de Riemann :*

$$\lim_{\mathsf{pas} \to 0} \mathscr{R}(p) \quad \textit{existe},$$

est aussi intégrable (au sens de Darboux), i.e. *appartient à* $\mathsf{Int}\,(A, \mathbb{R})$. $\square$

12. Ensembles cubables

Plaçons-nous dans l'espace $\mathbb{R}^3$ de la géométrie euclidienne à 3 dimensions, rapporté à un repère orthonormé $\{0, \mathbf{i}, \mathbf{j}, \mathbf{k}\}$. Appelons *pavé* (semi-ouvert) tout produit d'intervalles :

$$[a, b[\times[c, d[\times[e, f[,$$

avec $a < b$, $c < d$, $e < f$ des réels quelconques. Comme dans la Section **2**, pour tout $k \in \mathbb{N}$, on considère le *pavage de profondeur* k de l'espace $\mathbb{R}^3$ tout entier :

$$\boldsymbol{P}_k := \bigcup_{m\in\mathbb{Z}} \bigcup_{n\in\mathbb{Z}} \bigcup_{p\in\mathbb{Z}} \left[\frac{m}{2^k}, \frac{m+1}{2^k}\right[\times \left[\frac{n}{2^k}, \frac{n+1}{2^k}\right[\times \left[\frac{p}{2^k}, \frac{p+1}{2^k}\right[.$$

Nous voulons donner un sens mathématique précis à la notion intuitive de « *volume* », ou « *étendue* », d'une partie bornée $B \subset \mathbb{R}^2$. Pour $k \in \mathbb{N}$ quelconque, introduisons à cet effet :

$$\mathrm{N}_k^- := \mathsf{Card}\,\big\{\text{pavés de } \boldsymbol{P}_k \text{ contenus dans } B\big\},$$
$$\mathrm{N}_k^+ := \mathsf{Card}\,\big\{\text{pavés de } \boldsymbol{P}_k \text{ intersectant } B\big\}.$$

Un raisonnement analogue à celui vu en dimension 2 dans la Section **3** donne, pour tout $k \geqslant 1$:

$$\frac{1}{2^{3(k-1)}}\,\mathrm{N}_{k-1}^- \leqslant \frac{1}{2^{3k}}\,\mathrm{N}_k^- \leqslant \frac{1}{2^{3k}}\,\mathrm{N}_k^+ \leqslant \frac{1}{2^{3(k-1)}}\,\mathrm{N}_{k-1}^+.$$

De la sorte, nous obtenons deux suites :

$$\left(\frac{\mathrm{N}_k^-}{2^{3k}}\right)_{k\in\mathbb{N}} \quad \text{croissante et majorée,}$$
$$\left(\frac{\mathrm{N}_k^+}{2^{3k}}\right)_{k\in\mathbb{N}} \quad \text{décroissante et minorée.}$$

Ces deux suites sont donc convergentes, et nous noterons :

$$\mathsf{volume}^-(B) := \lim_{k\to\infty} \frac{\mathrm{N}_k^-}{2^{3k}},$$
$$\mathsf{volume}^+(B) := \lim_{k\to\infty} \frac{\mathrm{N}_k^+}{2^{3k}},$$

quantités que nous nommerons *volume inférieur* de B et *volume supérieur* de B.

Définition 12.1. On dira qu'une partie bornée $B \subset \mathbb{R}^3$ est *cubable* si :

$$\mathsf{volume}^-(B) = \mathsf{volume}^+(B).$$

Cette valeur commune s'appellera *volume* de B, et sera abrégée :

$$\mathsf{volume}(B) \equiv \mu(B).$$

Enfin, l'ensemble des parties cubables de $\mathbb{R}^3$ sera noté :

$$\mathscr{C} := \big\{B \subset \mathbb{R}^3 : B \text{ est cubable}\big\}.$$

Maintenant, comme pour les parties quarrables du plan, on démontre :

$$A \in \mathscr{C} \quad \text{et} \quad B \in \mathscr{C} \qquad \Longrightarrow \qquad A \cap B \in \mathscr{C} \quad \text{et} \quad A \cup B \in \mathscr{C},$$

et aussi naturellement, que :

$$\mu\big(A \cup B\big) = \mu(A) + \mu(B) - \mu\big(A \cap B\big).$$

De plus, on vérifie que :

$$\begin{aligned}\big(b-a\big)\,\big(d-c\big)\,\big(f-e\big) &= \mu\Big([a, b[\times[c, d[\times[e, f[\Big) \\ &= \mu\Big([a, b] \times [c, d] \times [e, f]\Big).\end{aligned}$$

13. Signification de l'intégrale double

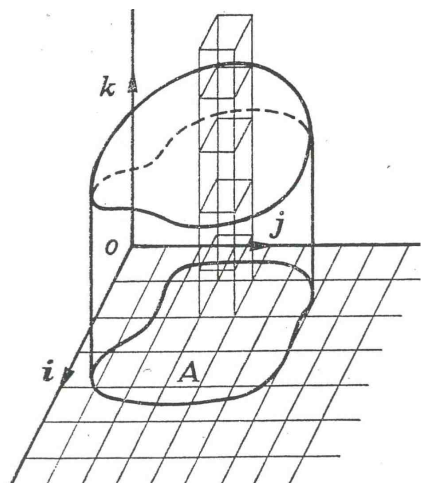
Sur une partie quarrable $A \subset \mathbb{R}^2$, soit une fonction à valeurs réelles positives :

$$A \ni (x, y) \longmapsto f(x, y) \in \mathbb{R}_+.$$

Introduisons l'*hypographe* de f :

$$B := \Big\{(x, y, z) \in \mathbb{R}^3 \colon\ (x, y) \in A,\ 0 \leqslant z \leqslant f(x, y)\Big\},$$

et étudions les volumes inférieur et supérieur de cet ensemble B, au moyen des pavages $\boldsymbol{P}_k$ de l'espace $\mathbb{R}^3$.



Tout pavé inclus dans B se projette orthogonalement sur $\mathbb{R}^2$ en un pavé-plan inclus dans A. Tout pavé qui rencontre B se projette en un pavé-plan qui rencontre A.

D'autre part, les intersections avec A des pavés-plans de profondeur k constituent une μ-partition de A, que l'on désignera par :

$$p_k \in \mathscr{P}_\mu(A).$$

À partir de ces remarques, il est aisé de se convaincre que :

$$\frac{\mathrm{N}_k^-}{2^{3k}} \leqslant s(p_k) \leqslant S(p_k) \leqslant \frac{\mathrm{N}_k^+}{2^{3k}},$$

d'où il résulte, en laissant $k \longrightarrow \infty$, que :

$$\mathsf{volume}^-(B) \leqslant \sigma_f(A) \leqslant \Sigma_f(A) \leqslant \mathsf{volume}^+(B).$$

De plus, comme la frontière de A a une aire nulle, la réunion des pavés-plans de profondeur k qui rencontrent cette frontière possède une aire qui tend vers 0 quand $k \longrightarrow \infty$. On peut alors démontrer rigoureusement que :

$$\mathsf{volume}^-(B) = \sigma_f(A) \qquad \text{et} \qquad \Sigma_f(A) = \mathsf{volume}^+(B).$$

Il en découle l'équivalence suivante.

Théorème 13.1. *Soit $A \subset \mathbb{R}^2$ une partie quarrable, et soit $f \colon A \longrightarrow \mathbb{R}_+$ une fonction positive bornée. Pour que f soit intégrable sur A, il faut et il suffit que son hypographe :*

$$B := \Big\{(x, y, z) \in \mathbb{R}^3 \colon\ (x, y) \in A,\ 0 \leqslant z \leqslant f(x, y)\Big\},$$

soit cubable. Dans ce cas :

$$\mathsf{volume}(B) = \mu(B) = \iint_A f(x, y)\, dxdy. \qquad \square$$

14. Calculs d'intégrales doubles sur un pavé compact

Soit un pavé fermé (donc compact) :

$$A := [a,b] \times [c,d],$$

avec des réels :

$$-\infty < a < b < \infty \qquad \text{et} \qquad -\infty < c < d < \infty.$$

Nous nous proposons de ramener le calcul de l'intégrale double :

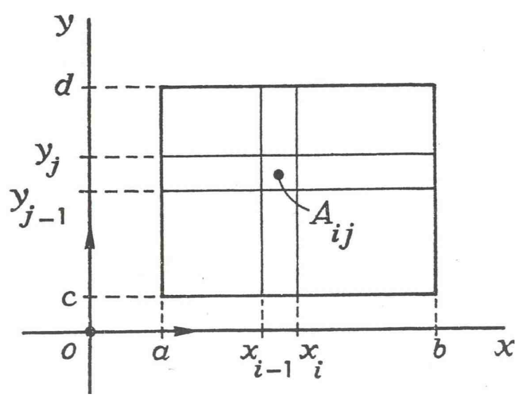
$$\iint_{[a,b]\times[c,d]} f(x,y)\,dxdy$$

d'une fonction continue $f\colon A \longrightarrow \mathbb{R}$ à *deux* calculs d'intégrales simples. À cette fin, effectuons une μ-partition $p = \big(A_{i,j}\big)_{\substack{1\leqslant i\leqslant n \\ 1\leqslant j\leqslant m}}$ de ce pavé A en les $n \cdot m$ sous-pavés :

$$A_{i,j} := \big[x_{i-1}, x_i\big] \times \big[y_{j-1}, y_j\big] \qquad\qquad (1\leqslant i\leqslant n,\, 1\leqslant j\leqslant m),$$

associés à deux subdivisions arbitraires des deux intervalles $[a,b]$ et $[c,d]$:

$$a = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n < x_n = b,$$
$$c = y_0 < y_1 < y_2 < \cdots < y_{m-1} < y_m < y_m = d.$$



Introduisons aussi les infima et les suprema :

$$m_{i,j} := \inf_{(x,y)\in A_{i,j}} f(x,y) \qquad \text{et} \qquad M_{i,j} := \sup_{(x,y)\in A_{i,j}} f(x,y).$$

Nous allons utiliser la définition de l'intégrale double qui passe par les sommes de Riemann. Soient des nombres arbitraires :

$$m_{i,j} \leqslant \theta_{i,j} \leqslant M_{i,j},$$

que nous choisirons plus tard, et envisageons, pour la μ-partition $p = \big(A_{i,j}\big)_{\substack{1\leqslant i\leqslant n \\ 1\leqslant j\leqslant m}}$ de $A = [a,b] \times [c,d]$, la somme de Riemann :

$$\mathscr{R}(p) := \sum_{i=1}^{n}\sum_{j=1}^{m} \big(x_i - x_{i-1}\big)\,\big(y_j - y_{j-1}\big)\cdot\theta_{i,j}.$$

Nous pouvons alors réorganiser cette sommation en deux moments successifs, dont le premier est :

$$\mathscr{R}(p) = \sum_{i=1}^{n} \big(x_i - x_{i-1}\big) \sum_{j=1}^{m} \big(y_j - y_{j-1}\big)\,\theta_{i,j}.$$

Quand on fixe l'indice i, pour tout j avec $1 \leqslant j \leqslant m$, nous décidons de choisir pour $\theta_{i,j}$ la *valeur moyenne* de la fonction continue $y \longmapsto f(x_i, y)$ dans l'intervalle $[y_{j-1}, y_j]$:

$$\theta_{i,j} := \frac{1}{y_j - y_{j-1}} \int_{y_{j-1}}^{y_j} f(x_i, y)\, dy.$$

Alors d'après la règle de Chasles pour les intégrales, on a :

$$\begin{aligned}\sum_{j=1}^{n} \big(y_j - y_{j-1}\big)\, \theta_{i,j} &= \sum_{j=1}^{m} \int_{y_{j-1}}^{y_j} f(x_i, y)\, dy \\ &= \int_c^d f(x_i, y)\, dy.\end{aligned}$$

Il en résulte que si on pose :

$$\psi(x) := \int_c^d f(x, y)\, dy,$$

on obtient :

$$\mathscr{R}(p) = \sum_{i=1}^{n} \big(x_i - x_{i-1}\big)\, \psi(x_i).$$

Comme on peut aisément démontrer que cette fonction $x \longmapsto \psi(x)$ est continue, donc intégrable, on reconnaît ici à droite une somme de Riemann pour le calcul de l'intégrale simple :

$$\int_a^b \psi(x)\, dx.$$

Par suite :

$$\iint_A f(x, y)\, dxdy = \int_a^b \psi(x)\, dx = \int_a^b dx \int_c^d f(x, y)\, dxdy.$$

En résumé, nous avons établi le

Théorème 14.1. *Sur un pavé compact $[a, b] \times [c, d] \subset \mathbb{R}^2$, l'intégrale d'une fonction réelle continue vaut :*

$$\begin{aligned}\iint_{[a,b]\times[c,d]} f(x, y)\, dxdy &= \int_a^b dx \int_c^d f(x, y)\, dy \\ &= \int_c^d dy \int_a^b f(x, y)\, dx.\end{aligned}$$

□

Évidemment, cette deuxième expression s'obtient simplement en échangeant $x \longleftrightarrow y$.

Un cas particulier mérite une mention, quand la fonction :

$$f(x, y) = g(x)\, h(y)$$

est produit de deux fonctions d'une variable. Dans ce cas, on voit que l'intégrale double est le produit de deux intégrales simples indépendantes :

$$\iint_{[a,b]\times[c,d]} f(x, y)\, dxdy = \bigg(\int_a^b g(x)\, dx\bigg) \bigg(\int_c^d h(y)\, dy\bigg).$$

Comme illustration de ce théorème fondamental qui permet de ramener le calcul d'une intégrale double à deux calculs d'intégrales simples, proposons-nous de calculer :

$$I := \iint_{[1,2]\times[0,1]} \frac{x\, y}{x^2 + y^2}\, dxdy.$$

D'après ce qui précède :

$$\begin{aligned}
I &= \int_1^2 x\,dx \int_0^1 \frac{y}{x^2+y^2}\,dy \\
&= \int_1^2 x\,dx \left[\tfrac{1}{2}\log\big(x^2+y^2\big)\right]_0^1 \\
&= \frac{1}{2}\int_1^2 x\log\frac{x^2+1}{x^2}\,dx \\
\text{[Poser } x^2 = t\text{]} \qquad &= \frac{1}{4}\int_1^4 \log\frac{t+1}{t}\,dt,
\end{aligned}$$

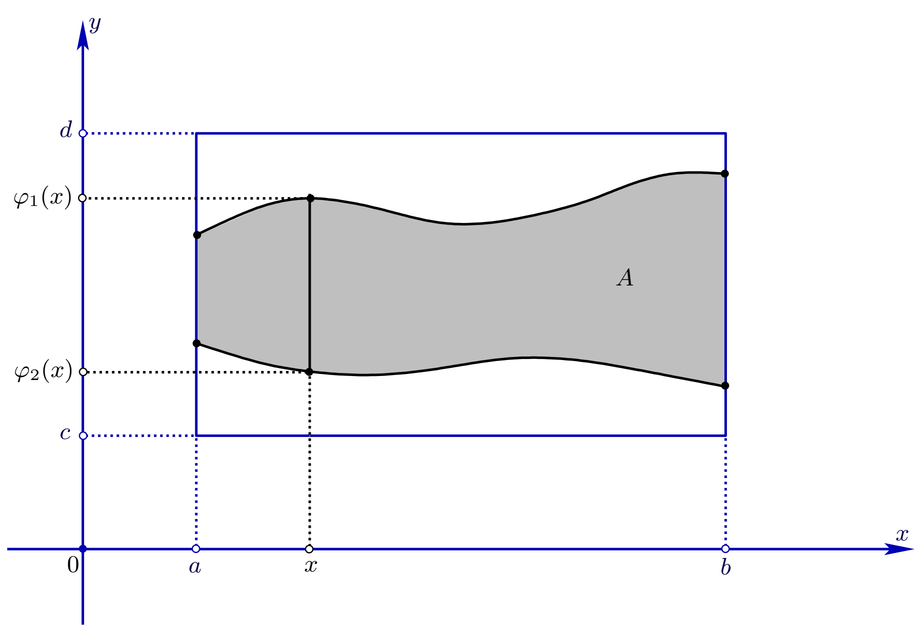
et en connaissant une primitive de $\log t$ qui s'obtient en observant que $\big(t\log t\big)' = \log t + 1$, on trouve :

$$\begin{aligned}
I &= \tfrac{1}{4}\Big[(t+1)\log t - t\log t\Big]_1^4 \\
&= \tfrac{1}{4}\Big(5\log 5 - 4\log 4 - 2\log 2 + \underline{1\log 1}\Big) \\
&= \tfrac{1}{4}\Big(5\log 5 - 5\log 4\Big) \\
&= \frac{5}{4}\log\frac{5}{4}.
\end{aligned}$$

$\square$

15. Intégrale double sur un compact simple

Au lieu d'un pavé, nous allons intégrer plus généralement sur un ensemble dont le toit et la cave sont éventuellement bosselés.



C'est-à-dire que nous allons intégrer sur un compact-type de la forme :

$$A := \Big\{(x,y)\in\mathbb{R}^2 :\ a \leqslant x \leqslant b,\ \varphi_1(x) \leqslant y \leqslant \varphi_2(x)\Big\},$$

où $a < b$ et où φ_1, φ_2 sont deux fonctions numériques *continues* sur l'intervalle $[a,b]$ satisfaisant $\varphi_1 \leqslant \varphi_2$.

Terminologie 15.1. On dit que A est un *compact simple.*

On peut démontrer — et nous l'admettrons — que ces compacts simples A sont quarrables.

Comme sur la figure, désignons alors par c un minorant de φ_1, et par d un majorant de φ_2 :

$$c \leqslant \varphi_1(x) \leqslant \varphi_2(x) \leqslant d \qquad\qquad (\forall\, x \in [a,b]).$$

Ensuite, sur le pavé rectangulaire quarrable :

$$B := [a,b] \times [c,d],$$

soit la nouvelle fonction :

$$g(x,y) := \begin{cases} f(x,y) & \text{lorsque } (x,y) \in A, \\ 0 & \text{lorsque } (x,y) \notin A. \end{cases}$$

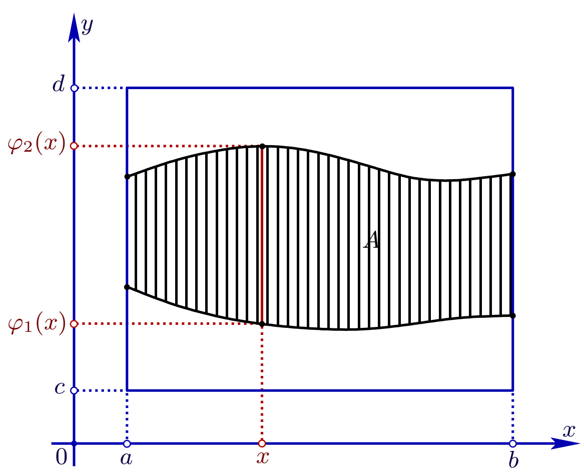
Cette fonction g est intégrable sur B, puisqu'elle l'est séparément sur A, et sur $B \backslash A$, et de plus, on a clairement :

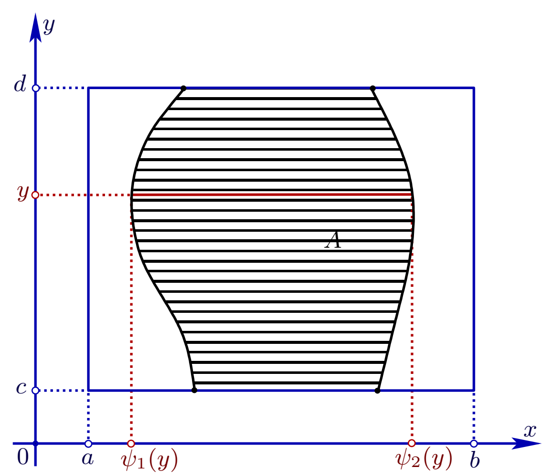
$$\iint_A f(x,y)\, dxdy = \iint_B g(x,y)\, dxdy.$$

Or maintenant, on peut calculer cette nouvelle intégrale en appliquant le Théorème 14.1 jusqu'à obtenir :

$$\begin{aligned} \iint_B g(x,y)\, dxdy &= \int_a^b dx \int_c^d g(x,y)\, dy \\ &= \int_a^b dx \int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\, dy, \end{aligned}$$

puisque g est identiquement nulle sur les deux intervalles $[c, \varphi_1(x)[$ et $]\varphi_2(x), d]$. En résumé, nous avons établi le





Théorème 15.2. *L'intégration d'une fonction f intégrable sur un compact simple :*

$$A := \Big\{ (x,y) \in \mathbb{R}^2 \colon\ a \leqslant x \leqslant b,\ \ \varphi_1(x) \leqslant y \leqslant \varphi_2(x) \Big\},$$

vaut :

$$\iint_A f(x,y)\, dxdy = \int_a^b dx \int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\, dy.$$

En échangeant $x \longleftrightarrow y$, avec un compact simple de la forme :

$$A := \Big\{ (x,y) \in \mathbb{R}^2 \colon\ \psi_1(y) \leqslant x \leqslant \psi_2(y),\ \ c \leqslant y \leqslant d \Big\},$$

l'intégrale vaut :

$$\iint_A f(x,y)\,dxdy \;=\; \int_c^d dy \int_{\psi_1(y)}^{\psi_2(y)} f(x,y)\,dx. \qquad \square$$

Parfois un compact simple peut s'écrire simultanément sous ces deux formes, à la fois « en piles » :

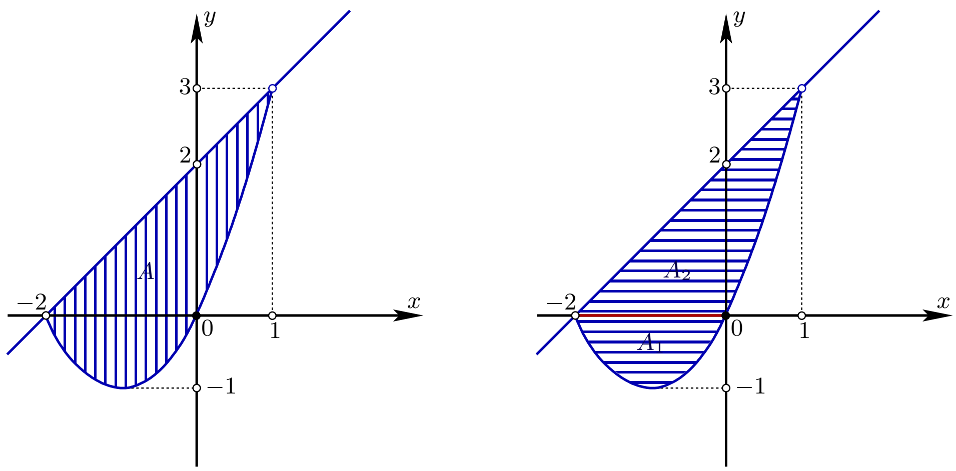
$$\big\{a \leqslant x \leqslant b,\;\; \varphi_1(x) \leqslant y \leqslant \varphi_2(x)\big\},$$

et aussi « en tranches » :

$$\big\{\psi_1(y) \leqslant x \leqslant \psi_2(y),\;\; c \leqslant y \leqslant d\big\}.$$

Voici un exemple déjà présenté :

$$A \;:=\; \overline{D}_3 \;:=\; \big\{(x,y)\in\mathbb{R}^2\colon\; x^2+2x \leqslant y \leqslant x+2\big\}.$$



En appliquant la formule du Théorème 15.2 qui découpe ce secteur parabolique en piles verticales, on calcule aisément :

$$\begin{aligned}
\mathsf{aire}(A) \;=\; \iint_A dxdy \;&=\; \int_{-2}^{1} dx \int_{x^2+2x}^{x+2} 1\cdot dy\\
&= \int_{-2}^{1} \big(x+2-(x^2+2\,x)\big)\,dx\\
&= \bigg[-\,\frac{x^2}{2}+2\,x-\frac{x^3}{3}\bigg]_{-2}^{1}\\
&= -\,\frac{1}{2}+2+2+4-\frac{1}{3}-\frac{8}{3}\\
&= \frac{9}{2}.
\end{aligned}$$

Toutefois, si on désire appliquer le découpage en tranches horizontales, il faut commencer — voir la figure — par décomposer $A = A_1\cup A_2$ en les deux domaines :

$$A_1 \;:=\; \big\{y\leqslant 0,\;\; x^2+2x-y\leqslant 0\big\} \qquad\Longleftrightarrow\qquad \Big\{-1\leqslant y\leqslant 0,\;\; \frac{-2-\sqrt{4+4y}}{2}\;\leqslant\; x\;\leqslant\;\frac{-2+\sqrt{4+4y}}{2}\Big\},$$

$$A_2 \;=\; \big\{y\geqslant 0,\;\; y-2\;\leqslant\; x\;\leqslant\; -1+\sqrt{1+y}\big\},$$

en observant la simplification dans la formule classique qui donne les deux racines de l'équation quadratique $x^2+2\,x-y=0$:

$$\frac{-2\pm\sqrt{4+4y}}{2} \;=\; -1\pm\sqrt{1+y}.$$

Grâce à cette décomposition adaptée, il vient :

$$\begin{aligned}\mathsf{aire}(A) &= \mathsf{aire}(A_1)+\mathsf{aire}(A_2)\\ &= \int_{-1}^{0}\Big(-1+\sqrt{1+y}-\big(-1-\sqrt{1+y}\big)\Big)\,dy+\int_{0}^{3}\Big(-1+\sqrt{1+y}-\big(y-2\big)\Big)\,dy\end{aligned}$$

mais le calcul devient quelque peu plus exigeant :

$$\begin{aligned}\mathsf{aire}(A) &= \Big[2\,\frac{2}{3}\,(1+y)^{\frac{3}{2}}\Big]_{-1}^{0}+\Big[y+\frac{2}{3}\,(1+y)^{\frac{3}{2}}-\frac{y^2}{2}\Big]_{0}^{3}\\ &= \frac{4}{3}-0+3-0+\frac{2}{3}\,4^{\frac{3}{2}}-\frac{2}{3}-\frac{9}{2}+0\\ &= \frac{9}{2}.\end{aligned}$$

Moralité : Avec les intégrales doubles, il vaut mieux réfléchir avant d'emprunter un chemin de calcul !

16. Simplifications en présence de symétries

Très souvent (*e.g.* en physique), le domaine quarrable A sur lequel on intègre possède une symétrie visible à l'œil nu, c'est-à-dire qu'il existe une application S qui est *involutive* au sens où $S\circ S=\mathsf{Id}$:

$$\begin{aligned}S\colon\qquad A &\longrightarrow A\\ (x,y) &\longmapsto S(x,y),\end{aligned}$$

telle que :

$$f(x,y) = f\big(S(x,y)\big)\qquad\qquad (\forall\,(x,y)\in A).$$

Il en découle que $A=A_1\cup A_2$ se décompose en deux ensembles quarrables :

$$A_2 = S(A_1)\qquad\qquad\text{avec}\quad A_1\cap A_2 = \emptyset_\mu,$$

et par additivité des intégrales doubles, on obtient :

$$\iint_A f(x,y)\,dxdy = 2\iint_{A_1} f(x,y)\,dxdy.$$

Si, à l'opposé :

$$-f(x,y) = f\big(S(x,y)\big),$$

alors tout disparaît :

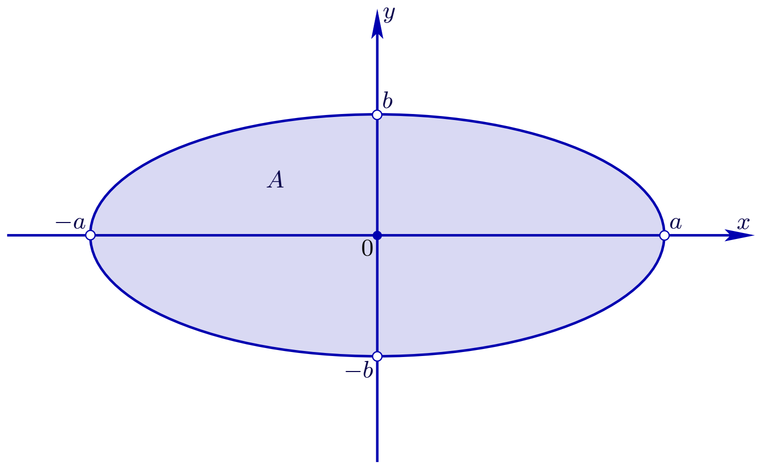
$$\iint_A f(x,y)\,dxdy = 0.$$

Exemple 16.1. Proposons-nous de calculer l'intégrale double :

$$I := \iint_A \big(x^2-y^2\big)\,dxdy,$$

sur le disque elliptique :

$$A := \Big\{\frac{x^2}{a^2}+\frac{y^2}{b^2}-1\leqslant 0\Big\}.$$

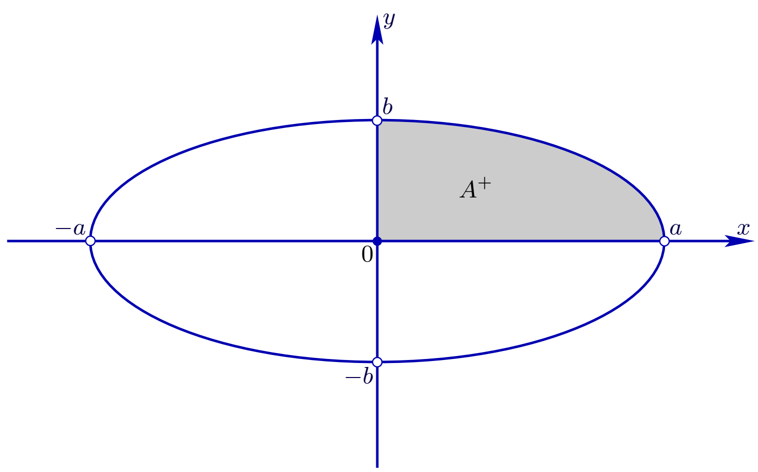


Observons que les pliages autour de $0x$ et autour de $0y$:

$$(x, y) \longmapsto (x, -y) \qquad\qquad \text{et} \qquad\qquad (x, y) \longmapsto (-x, y)$$

laissent A invariant, et ne changent pas non plus les valeurs de la fonction à intégrer :

$$x^2 - (-y)^2 \;=\; (-x)^2 - y^2 \;=\; x^2 - y^2.$$



Par conséquent, en introduisant le quart supérieur droit :

$$A^+ \;:=\; \Big\{ x \geqslant 0,\;\; y \geqslant 0,\;\; \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 \leqslant 0 \Big\},$$

il est clair que :

$$I \;=\; 4 \iint_A \big(x^2 - y^2\big)\, dxdy.$$

Ce quadrant elliptique est aussi représenté par les inégalités :

$$0 \;\leqslant\; x \;\leqslant\; a \qquad\qquad \text{et} \qquad\qquad 0 \;\leqslant\; y \;\leqslant\; b\sqrt{1 - \frac{x^2}{a^2}},$$

donc en application du Théorème 15.2, il vient :

$$
\begin{aligned}
I &= 4 \int_0^a \left(\int_0^{b\sqrt{1-\frac{x^2}{a^2}}} \left(x^2 - y^2\right) dy \right) dx \\
&= 4 \int_0^a \left[x^2\, y - \frac{1}{3}\, y^3 \right]_0^{b\sqrt{1-\frac{x^2}{a^2}}} dx \\
&= 4 \int_0^a \left\{ x^2\, b \sqrt{1-\frac{x^2}{a^2}} - \frac{1}{3}\, b^3 \sqrt{1-\frac{x^2}{a^2}} \left(1 - \frac{x^2}{a^2}\right) \right\} dx.
\end{aligned}
$$

Afin de calculer commodément cette intégrale simple, effectuons le changement de variable :

$$
x = a \sin t \qquad \text{avec} \qquad t \in \left[0, \tfrac{\pi}{2}\right],
$$

d'où :

$$
dx = a \cos t\, dt \qquad \text{et aussi} \qquad \sqrt{1-\frac{x^2}{a^2}} = \cos t,
$$

ce qui transforme :

$$
\begin{aligned}
I &= 4 \int_0^{\frac{\pi}{2}} \left(a^2 \sin^2 t\ b \cos t - \frac{1}{3}\, b^3 \cos t\ \underbrace{\cos^2 t}_{1-\sin^2 t} \right) a \cos t\, dt \\
&= 4\, ab \int_0^{\frac{\pi}{2}} \cos^2 t \left[\left(a^2 + \frac{b^2}{3} \right) \sin^2 t - \frac{b^2}{3} \right] dt.
\end{aligned}
$$

À ce moment, nous constatons que nous devons intégrer la fonction $\cos^2 t \sin^2 t$, dont il n'est pas immédiat de trouver une primitive. Alors nous devons rappeler le

Théorème 16.2. **[Formule de de Moivre]** *Pour tout nombre réel $t \in \mathbb{R}$, et avec $i := \sqrt{-1}$, on a :*

$$
e^{it} = \cos t + i \sin t. \qquad \square
$$

En passant au conjugué complexe :

$$
e^{-it} = \cos t - i \sin t,
$$

et en additionnant comme il faut — exercice — on obtient le

Théorème 16.3. **[Fomules d'Euler]** *Pour tout $t \in \mathbb{R}$, on a :*

$$
\cos t = \frac{e^{it} + e^{-it}}{2} \qquad \textit{et} \qquad \sin t = \frac{e^{it} - e^{-it}}{2\, i}. \qquad \square
$$

Grâce à ces formules, nous pouvons énoncer et démontrer un lemme utile à la poursuite de nos calculs.

Assertion 16.4. *On a la formule de trigonométrie :*

$$
4 \cos^2 t\ \sin^2 t = \frac{1 - \cos 4t}{2}.
$$

Preuve. En passant aux exponentielles complexes grâces aux formules de de Moivre et d'Euler, le calcul devient assez aisé :

$$
\begin{aligned}
4 \left[\cos t\ \sin t\right]^2 &= 4 \left[\frac{\left(e^{it} + e^{-it}\right)\left(e^{it} - e^{-it}\right)}{2 \quad \cdot \quad 2i} \right]^2 = 4\, \frac{\left[e^{2it} - e^{-2it}\right]^2}{4 \cdot (-4)} \\
&= \frac{e^{4it} - 2 + e^{-4it}}{-4} = \frac{1 - \cos 4t}{2}. \qquad \square
\end{aligned}
$$

Grâce à cette formule, nous pouvons déterminer la valeur de :

$$\int_0^{\frac{\pi}{2}} 4\cos^2 t \sin^2 t \, dt \;=\; \int_0^{\frac{\pi}{2}} \frac{1-\cos 4t}{2}\, dt \;=\; \left[\frac{t}{2} - \frac{\sin 4t}{8}\right]_0^{\frac{\pi}{2}} \;=\; \frac{\pi}{4}.$$

Ensuite, grâce à une formule trigonométrique connue (ou que l'on re-démontre en raisonnant d'une manière analogue), il vient :

$$\int_0^{\frac{\pi}{2}} 4\cos^2 t \, dt \;=\; \int_0^{\frac{\pi}{2}} 2\left(1+\cos 2t\right) dt \;=\; \Big[2\,t + \sin 2t\Big]_0^{\frac{\pi}{2}} \;=\; \pi.$$

Grâce à ces deux calculs d'intégrales laborieux, nous pouvons enfin conclure, via $\frac{1}{4}\,\frac{1}{3} - \frac{1}{3} = -\frac{1}{4}$:

$$\begin{aligned} I \;&=\; a\,b\left[\frac{\pi}{4}\left(a^2 + \frac{b^2}{3}\right) - \pi\,\frac{b^2}{3}\right] \\ &=\; \frac{\pi}{4}\,a\,b\left(a^2 - b^2\right). \end{aligned}$$

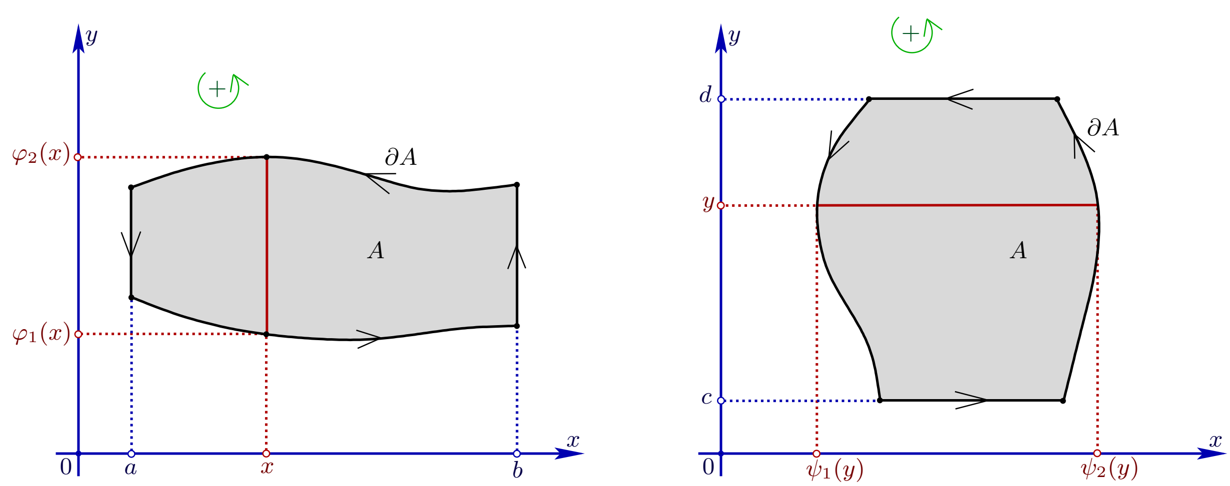
□

17. Formule de Green-Riemann

Comme dans le Théorème 15.2, considérons un compact simple $A \subset \mathbb{R}^2$ défini par :

$$a \;\leqslant\; x \;\leqslant\; b \qquad\qquad \text{et} \qquad\qquad \varphi_1(x) \;\leqslant\; y \;\leqslant\; \varphi_2(x),$$

où $\varphi_1 \colon [a,b] \longrightarrow \mathbb{R}$ et $\varphi_2 \colon [a,b] \longrightarrow \mathbb{R}$ sont deux fonctions *dérivables* et à dérivées *continues* sur $[a,b]$, avec $\varphi_1 \leqslant \varphi_2$.



La frontière ∂A de A est réunion de 4 courbes dont 2 segments verticaux, éventuellement réduits à un seul point — voir la figure gauche ci-dessus — :

$$\begin{aligned} \partial A \;=\; &\big\{a \leqslant x \leqslant b,\;\; y = \varphi_1(x)\big\} \\ &\bigcup\;\big\{x = b,\;\; \varphi_1(b) \leqslant y \leqslant \varphi_2(b)\big\} \\ &\bigcup\;\big\{b \geqslant x \geqslant a,\;\; y = \varphi_2(x)\big\} \\ &\bigcup\;\big\{x = a,\;\; \varphi_2(a) \geqslant y \geqslant \varphi_1(a)\big\}. \end{aligned}$$

Comme sur la figure, orientons la frontière ∂A de A de telle sorte qu'un point mobile la parcourant voie toujours A à sa gauche.

Soit alors $(x, y) \longmapsto P(x, y)$ une fonction continue sur ce compact simple A, admettant une dérivée partielle $\frac{\partial P}{\partial y}$ continue sur A. Le Théorème 15.2 permet de calculer :

$$\begin{aligned}\iint_A \frac{\partial P}{\partial y}\, dxdy \;&=\; \int_a^b dx \int_{\varphi_1(x)}^{\varphi_2(x)} \frac{\partial P}{\partial y}\, dy \\ &=\; \int_a^b \Big[P\big(x, \varphi_2(x)\big) - P\big(x, \varphi_1(x)\big)\Big]\, dx.\end{aligned}$$

On reconnaît ici l'*intégrale curviligne* de $-\, P\, dx$ le long du bord orienté ∂A :

$$\iint_A \frac{\partial P}{\partial y}\, dxdy \;=\; -\int_{\partial A} P(x, y)\, dx.$$

On observera que sur les deux segments verticaux, l'intégrale curviligne s'annule identiquement, car $x = \text{constante}$ implique $dx = 0$.

Considérons maintenant un autre compact simple A défini par des inégalités dans lesquelles les rôles de x et de y sont échangés :

$$\psi_1(y) \leqslant x \leqslant \psi_2(y) \qquad\qquad \text{et} \qquad\qquad c \leqslant y \leqslant d,$$

avec deux fonctions $\psi_1 \colon [c, d] \longrightarrow \mathbb{R}$ et $\psi_2 \colon [c, d] \longrightarrow \mathbb{R}$ dérivables et à dérivées continues sur $[c, d]$, satisfaisant $\psi_1 \leqslant \psi_2$.

La frontière ∂A de A est réunion de 4 courbes, dont 2 segments horizontaux éventuellement réduits à un seul point — voir la figure droite ci-dessus — :

$$\begin{aligned}\partial A \;=\; &\big\{\psi_1(c) \leqslant x \leqslant \psi_2(c),\;\; y = c\big\} \\ &\bigcup\, \big\{x = \psi_2(y),\;\; c \leqslant y \leqslant d\big\} \\ &\bigcup\, \big\{\psi_2(d) \geqslant x \geqslant \psi_1(d),\;\; y = d\big\} \\ &\bigcup\, \big\{x = \psi_1(y),\;\; d \geqslant y \geqslant c\big\}.\end{aligned}$$

Comme sur la figure, orientons la frontière ∂A de A de telle sorte qu'un point mobile la parcourant voie toujours A à sa gauche.

Soit alors $(x, y) \longmapsto Q(x, y)$ une fonction continue sur ce compact simple admettant une dérivée partielle $\frac{\partial Q}{\partial x}$ continue sur A. Le Théorème 15.2 permet de calculer :

$$\begin{aligned}\iint_A \frac{\partial Q}{\partial x}\, dxdy \;&=\; \int_c^d dy \int_{\psi_1(y)}^{\psi_2(y)} \frac{\partial Q}{\partial x}\, dx \\ &=\; \int_c^d \Big[Q\big(\psi_2(y), y\big) - Q\big(\psi_1(y), y\big)\Big]\, dy,\end{aligned}$$

ce qui s'exprime en termes d'intégrale curviligne comme :

$$\iint_A \frac{\partial Q}{\partial x}\, dxdy \;=\; \iint_A Q(x, y)\, dy.$$

Afin de pouvoir appliquer simultanément les deux formules précédentes, énonçons une

Définition 17.1. On appelle *compact élémentaire* tout compact défini indifféremment par :

$$a \leqslant x \leqslant b \qquad\qquad \text{et} \qquad\qquad \varphi_1(x) \leqslant y \leqslant \varphi_2(x),$$

les fonctions φ_1 et φ_2 étant de classe $\mathscr{C}^1$ sur $[a, b]$, ou bien par :

$$c \leqslant x \leqslant d \qquad\qquad \text{et} \qquad\qquad \psi_1(y) \leqslant x \leqslant \psi_2(y),$$

les fonctions ψ_1 et ψ_2 étant $\mathscr{C}^1$ sur $[c, d]$.

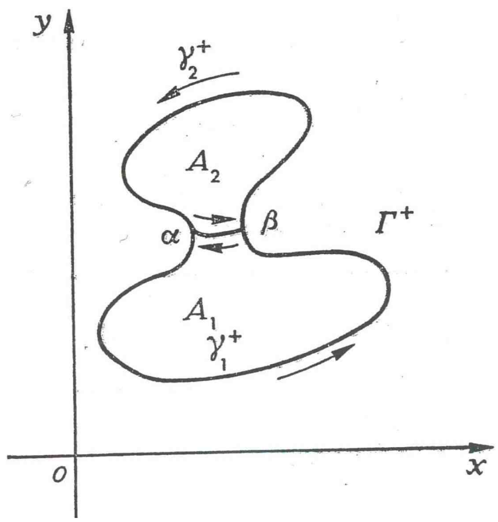
Par exemple, un disque circulaire fermé, un disque elliptique fermé, un pavé de côtés parallèles aux axes de coordonnées, sont des compacts élémentaires.

En retranchant membre à membre les deux formules qui précèdent, nous obtenons enfin le

Théorème 17.2. [de Riemann-Green] *Sur un compact élémentaire $A \subset \mathbb{R}^2$ de frontière orientée ∂A, si deux fonctions continues $P, Q\colon A \longrightarrow \mathbb{R}$ ont des dérivées partielles $\frac{\partial P}{\partial y}$ et $\frac{\partial Q}{\partial x}$ continues sur A, alors :*

$$\iint_A \Big(-\frac{\partial P}{\partial y} + \frac{\partial Q}{\partial x}\Big)\,dxdy \;=\; \int_{\partial A} \Big(P\,dx + Q\,dy\Big). \qquad \square$$

Plus généralement, ce théorème de Riemann-Green reste valable quand A est réunion d'un nombre fini de compacts élémentaires deux à deux μ-disjoints.



Il suffit de prouver cela pour la réunion de deux compacts élémentaires μ-disjoints A_1 et A_2 dont les frontières γ_1 et γ_2 ont une intersection non vide $\alpha\beta$.

Alors le Théorème 17.2 de Riemann-Green appliqué à A_1 et à A_2 séparément exige le calcul de deux intégrales curvilignes prises le long de γ_1 et de γ_2 ; l'arc $\alpha\beta$ est parcouru deux fois, mais en des sens contraires. Par suite, la somme des intégrales curvilignes précédentes est égale à l'intégrale curviligne prise le long de la frontière ∂A de la réunion $A_1 \cap A_2$. (Cette frontière est alors réunion de deux arcs à dérivée continue.)

18. Calculs d'aires planes en coordonnées cartésiennes et en coordonnées polaires

Considérons d'abord un compact simple A défini par :

$$a \leqslant x \leqslant b \qquad\qquad \text{et} \qquad\qquad \varphi_1(x) \leqslant y \leqslant \varphi_2(x),$$

avec $\varphi_1 \leqslant \varphi_2$. Pour évaluer l'aire de A, considérons la fonction :

$$P(x,y) := y, \qquad\qquad \text{d'où} \quad \frac{\partial P}{\partial y} \equiv A.$$

La formule de Riemann-Green donne alors :

$$\iint_A dxdy \;=\; -\int_{\partial A} y\,dx.$$

Or puisque la première intégrale est égale à l'aire de A, il vient :

$$\mu(A) \;=\; -\int_{\partial A} y\,dx.$$

De la même façon, si A est un compact simple défini par :

$$\psi_1(y) \leqslant x \leqslant \psi_2(y) \qquad \text{et} \qquad c \leqslant y \leqslant d,$$

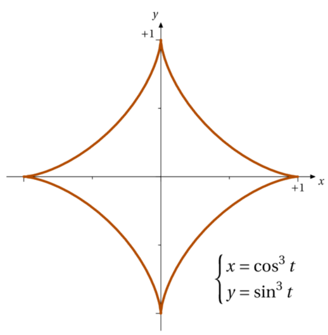
avec $\psi_1 \leqslant \psi_2$, alors en considérant la fonction $Q := x$, on obtient une formule alternative pour l'aire de A :

$$\mu(A) = \int_{\partial A} x\,dy.$$

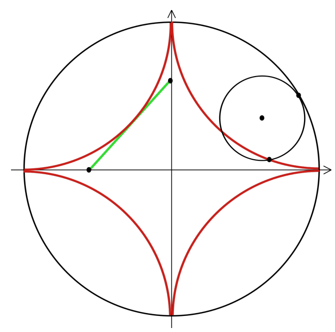
Par conséquent, nous obtenons le

Théorème 18.1. *L'aire d'un compact élémentaire $A \subset \mathbb{R}^2$, ou de la réunion A d'un nombre fini de compacts élémentaires μ-disjoints, vaut :*

$$\mu(A) = \frac{1}{2} \int_{\partial A} \big(x\,dy - y\,dx\big). \qquad \square$$



Astroïde



Exemple 18.2. [Aire de l'astroïde] Soit l'astroïde défini de manière paramétrique avec $t \in [0, 2\pi]$ par :

$$x(t) = a \cos^3 t \qquad \text{et} \qquad y(t) = a \sin^3 t.$$

Cette courbe est frontière d'un compact A, réunion de 4 sous-régions deux à deux symétriques par rapport aux axes de coordonnées. On a ainsi :

$$\mu(A) = \frac{4}{2} \int_{\gamma} \big(x\,dy - y\,dx\big),$$

γ étant la réunion de l'arc d'astroïde correspondant à $t \in \big[0, \frac{\pi}{2}\big]$ et de deux portions des axes $0x$ et $0y$.

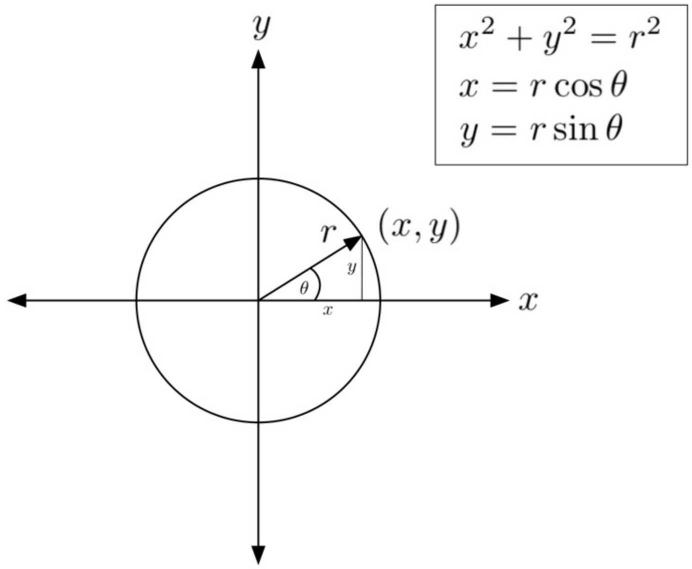
Sur l'astroïde, on a :

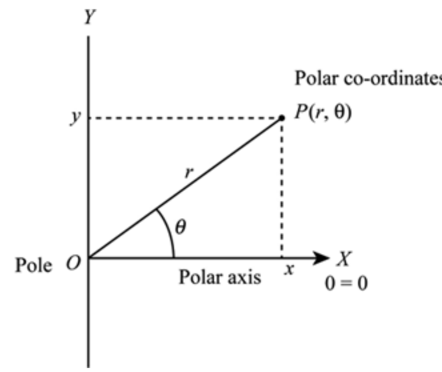
$$\begin{aligned} x\,dy - y\,dx &= a \cos^3 t \; 3\,a \cos t \; \sin^2 t - a \sin^3 t \; 3\,a \sin t \; \cos^2 t \\ &= 3\,a^2 \sin^2 t \; \cos^2 t \, \big(\cos^2 t + \sin^2 t\big)\,dt \\ &= 3\,a^2 \sin^2 t \; \cos^2 t \; dt. \end{aligned}$$

En se souvenant que $\sin^2 t \cos^2 t = \frac{1 - \cos 4t}{8}$, on obtient, par suite, pour l'aire demandée :

$$\mu(A) = 2 \int_0^{\frac{\pi}{2}} \frac{3}{8}\,a^2 \,\big(1 - \cos 4t\big)\,dt = \frac{6}{8}\,a^2 \left[t - \frac{\sin 4t}{4}\right]_0^{\frac{\pi}{2}} = \frac{3}{8}\,a^2\,\pi.$$

Passons maintenant à des calculs d'aire en coordonnées polaires.

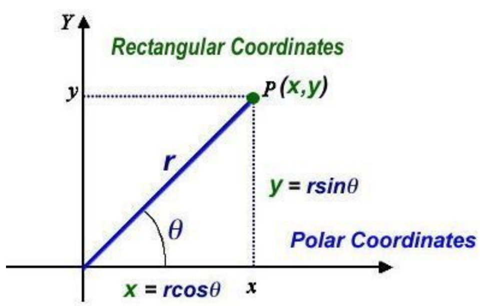


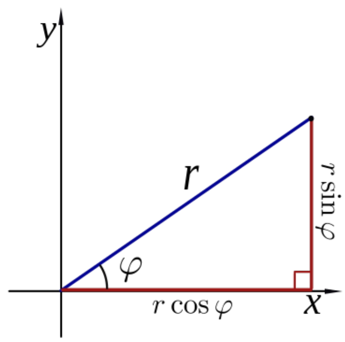


Commençons par un rappel sur les coordonnée polaires. On écrit :

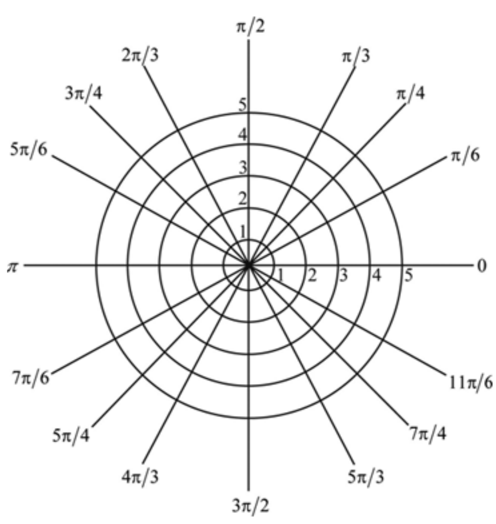
$$x \,=\, \rho \cos\theta, \qquad\qquad y \,=\, \rho \sin\theta,$$

avec un rayon $\rho \geqslant 0$ et un angle $0 \leqslant \theta < 2\pi$. Parfois, le rayon ρ se note r, et d'autre lettres que θ sont utilisées pour désigner l'angle.





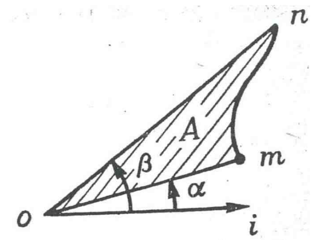
Les valeurs d'angles remarquables sont illustrées comme suit.



Maintenant, soit un arc mn donné par une équation :

$$\rho = f(\theta),$$

avec une certaine fonction f définie sur un intervalle $[\alpha, \beta]$, qui y possède une dérivée continue.

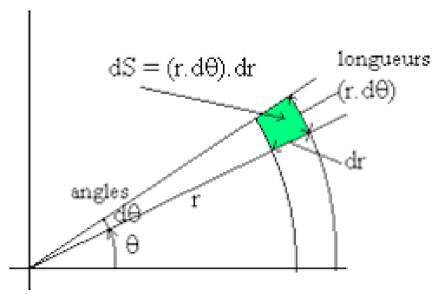


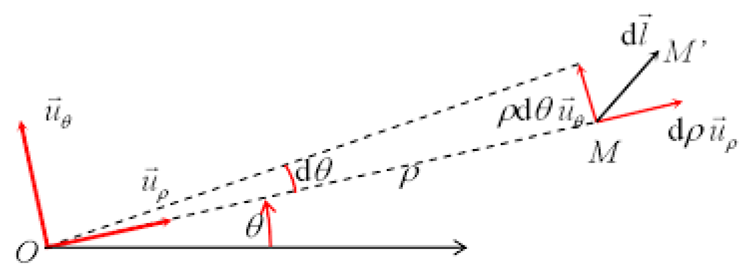
Nous supposerons de plus que la réunion de l'arc mn et des rayons $0m$ et $0n$ est la frontière d'un compact élémentaire. Une différentiation donne :

$$dx = d\rho \cos\theta - \rho \sin\theta \, d\theta, \qquad dy = d\rho \sin\theta + \rho \cos\theta \, d\theta,$$

et, par conséquent :

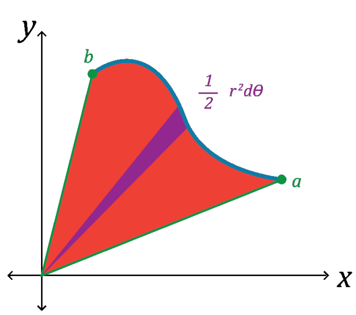
$$\begin{aligned} x\,dy - y\,dx &= \underline{\rho \cos\theta \sin\theta \; d\rho} + \rho^2 \cos^2\theta \; d\theta - \underline{\rho \sin\theta \cos\theta \; d\rho} + \rho^2 \sin^2\theta \; d\theta \\ &= \rho^2 \; d\theta. \end{aligned}$$



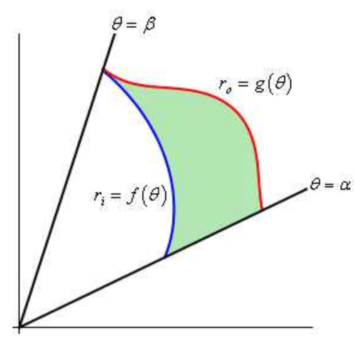


Par suite, pour l'aire du compact élémentaire A, la formule de Riemann-Green devient, puisque $\theta =$ constante entraîne $d\theta = 0$ sur les deux rayons $0m$ et $0n$:

$$\mu(A) = \frac{1}{2} \int_{mn} \rho^2 \, d\theta.$$



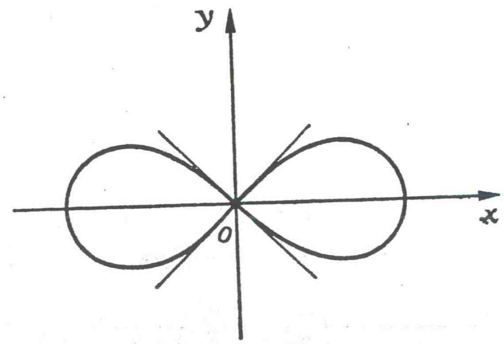
$$A = \frac{1}{2} \int_a^b r^2 d\theta$$



Exemple 18.3. [Aire de la leminiscate de Bernoulli] La leminiscate de Bernoulli est la courbe qui admet pour équation polaire :

$$\rho = a\sqrt{\cos 2\theta},$$

où $a > 0$ est une constante, et où la variable d'angle θ parcourt les deux intervalles $\left[-\frac{\pi}{4}, \frac{\pi}{4}\right]$ et $\left[\frac{3\pi}{4}, \frac{5\pi}{4}\right]$, ce qui donne les deux lobes bien visibles sur la figure.



Chacun de ces deux lobes est un secteur élémentaire. Pour l'aire totale de A, on a, par raison de symétrie :

$$\mu(A) \,=\, 4\int_0^{\frac{\pi}{4}} \frac{1}{2}\,\rho^2\,d\theta \,=\, 2\,a^2\int_0^{\frac{\pi}{4}} \cos 2\theta\,d\theta \,=\, a^2\,\Big[\sin 2\theta\Big]_0^{\frac{\pi}{4}} \,=\, a^2. \qquad \square$$

19. Formule de changement de variables dans les intégrales doubles

En dimension 1, à savoir sur la droite numérique $\mathbb{R}$, la formule de changement de variable dans une intégrale riemannienne s'exprime le plus souvent dans une circonstance différentiable bijective.

Théorème 19.1. *Soit un intervalle $[a,b] \subset \mathbb{R}$ avec $-\infty < a < b < \infty$, soit $\varphi\colon [a,b] \longrightarrow \mathbb{R}$ une application $\mathscr{C}^1$ avec $\varphi'(x) > 0$ pour tout $x \in [a,b]$, d'où le difféomorphisme :*

$$\varphi\big([a,b]\big) \,=\, \big[\varphi(a),\,\varphi(b)\big],$$

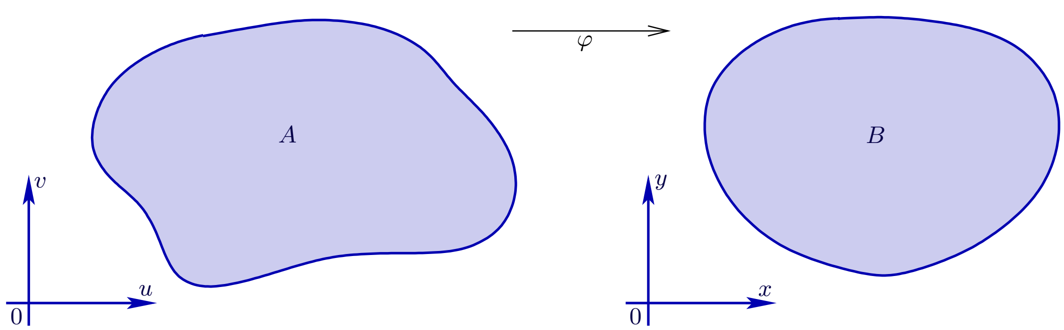
avec $-\infty < \varphi(a) < \varphi(b) < \infty$. Alors pour toute fonction Riemann-intégrable $f\colon [\varphi(a),\varphi(b)] \longrightarrow \mathbb{C}$, la composée $f\circ\varphi$ est aussi Riemann-intégrable et :

$$\int_{\varphi(a)}^{\varphi(b)} f(t)\,dt \,=\, \int_a^b f\big(\varphi(x)\big)\,\varphi'(x)\,dx.$$

Lorsque $\varphi' < 0$, cette formule est tout aussi satisfaite. $\square$

En prenant $f \equiv 1$, on retrouve la formule fondatrice du calcul intégral :

$$\varphi(b) - \varphi(a) \,=\, \int_a^b \varphi'(x)\,dx.$$



Notre objectif est de généraliser maintenant cette formule aux intégrales doubles. Soient donc deux plans euclidiens distincts $\mathbb{R}^2$ et $\mathbb{R}^2$ munis de deux systèmes de coordonnées (cartésiennes ou curvilignes) :

$$(u,v) \,\in\, \mathbb{R}^2 \qquad\qquad \text{et} \qquad\qquad (x,y) \,\in\, \mathbb{R}^2.$$

On suppose donnés deux compacts à bords quarrables :

$$A \,\subset\, \mathbb{R}^2 \qquad\qquad \text{et} \qquad\qquad B \,\subset\, \mathbb{R}^2,$$

ainsi qu'une application entre eux :

$$\begin{aligned} \varphi\colon \qquad A &\overset{\sim}{\longrightarrow} B \\ (u,v) &\longmapsto \big(f(u,v),\, g(u,v)\big) \;=:\; (x,y), \end{aligned}$$

qui satisfait toutes les hypothèses suivantes.

(1) φ est bijective $A \overset{\sim}{\longrightarrow} B$.

(2) En restriction aux bords, $\varphi\big|_{\partial A}\colon \partial A \overset{\sim}{\longrightarrow} \partial B$ est aussi bijective.

(3) φ est continue, et la bijection inverse $A \overset{\sim}{\longleftarrow} B : \varphi^{-1}$ est aussi continue.

(4) φ admet des dérivées partielles $\frac{\partial \varphi}{\partial u}$ et $\frac{\partial \varphi}{\partial v}$ qui sont continues sur A.

(5) φ conserve l'orientation.

Terminologie 19.2. On dit que φ est un $\mathscr{C}^1$-*difféomorphisme* entre A et B.

Nous supposerons de plus que A et B sont des compacts élémentaires, ou des réunions finies de compacts élémentaires μ-disjoints. Cette hypothèse générale déjà considérée précédemment, apparaîtra encore dans les énoncés des théorèmes.

$$\begin{tikzcd} A \arrow[r, "\varphi"] \arrow[dr, "G\circ\varphi"'] & B \arrow[d, "G"] \\ & \mathbb{R}. \end{tikzcd}$$

Problème 19.3. Exprimer l'intégrale d'une fonction intégrable $G(x,y)$ sur B :

$$\iint_B G(x,y)\,dxdy \;=\; \iint_A \text{quelque} \overset{?}{} \text{ chose},$$

en fonction d'une intégrale d'une certaine fonction sur A.

Nous allons commencer par le cas simple de la fonction $G \equiv 1$, c'est-à-dire :

$$\mathsf{aire}(B) \;=\; \iint_B 1\,dxdy \;=\; \iint_A \mathbf{?}.$$

Grâce au Théorème 18.1, on a :

$$\mathsf{aire}(B) \;=\; \frac{1}{2}\int_{\partial B} \big(x\,dy - y\,dx\big).$$

Ensuite, ∂B est une courbe 1-dimensionnelle, et l'application :

$$\begin{aligned} \varphi\colon \qquad \partial A &\overset{\sim}{\longrightarrow} \partial B \\ (u,v) &\longmapsto \big(f(u,v),\, g(u,v)\big) \end{aligned}$$

effectue un changement de variable 1-dimensionnel préservant l'orientation qui permet de ramener à A l'intégrale considérée :

$$\begin{aligned} \mathsf{aire}(B) &= \frac{1}{2}\int_{\partial A} \big(f\,dg - g\,df\big) \\ &= \frac{1}{2}\int_{\partial A} \bigg(f\,\Big[\frac{\partial g}{\partial u}\,du + \frac{\partial g}{\partial v}\,dv\Big] - g\,\Big[\frac{\partial f}{\partial u}\,du + \frac{\partial f}{\partial v}\,dv\Big]\bigg) \\ &= \frac{1}{2}\int_{\partial A} \bigg\{\underbrace{\bigg(f\,\frac{\partial g}{\partial u} - g\,\frac{\partial f}{\partial u}\bigg)}_{=:\,P(u,v)} du + \underbrace{\bigg(f\,\frac{\partial g}{\partial v} - g\,\frac{\partial f}{\partial v}\bigg)}_{=:\,Q(u,v)} dv\bigg\}. \end{aligned}$$

Alors le Théorème 17.2 de Riemann-Green s'applique :

$$\frac{1}{2}\int_{\partial A}\big(P\,du + Q\,dv\big) \;=\; \frac{1}{2}\iint_A\bigg(-\frac{\partial P}{\partial v}+\frac{\partial Q}{\partial u}\bigg)\,dudv,$$

donc on doit calculer :

$$\begin{aligned}
-\frac{\partial P}{\partial v} &= -\frac{\partial f}{\partial v}\,\frac{\partial g}{\partial u} - \underline{f\,\frac{\partial^2 g}{\partial v\partial u}}_{\circ} + \frac{\partial g}{\partial v}\,\frac{\partial f}{\partial u} + \underline{g\,\frac{\partial^2}{\partial v\partial u}}_{\circ\circ},\\
\frac{\partial Q}{\partial u} &= \frac{\partial f}{\partial u}\,\frac{\partial g}{\partial v} + \underline{f\,\frac{\partial^2 g}{\partial u\partial v}}_{\circ} - \frac{\partial g}{\partial u}\,\frac{\partial f}{\partial v} - \underline{g\,\frac{\partial^2 f}{\partial u\partial v}}_{\circ\circ}.
\end{aligned}$$

Quand on additionne, deux paires de termes (soulignées) s'annihilent, et par conséquent :

$$\begin{aligned}
\mathsf{aire}(B) \;=\; \iint_B 1\,dxdy \;&=\; \frac{1}{2}\iint_A\bigg(-\frac{\partial P}{\partial v}+\frac{\partial Q}{\partial u}\bigg)\,dudv\\
&= \iint_A\bigg(\frac{\partial f}{\partial u}\,\frac{\partial g}{\partial v}-\frac{\partial f}{\partial v}\,\frac{\partial g}{\partial u}\bigg)\,dudv.
\end{aligned}$$

On voit ainsi apparaître une quantité fondamentale.

Définition 19.4. La *matrice jacobienne* d'une application de classe $\mathscr{C}^1$:

$$\varphi\colon\qquad (u,v)\;\longmapsto\;\big(f(u,v),\,g(u,v)\big)$$

en un point (u,v) est la matrice 2×2 :

$$\mathsf{Jac}\,(\varphi)\,(u,v)\;:=\;\begin{pmatrix}\frac{\partial f}{\partial u} & \frac{\partial f}{\partial v}\\ \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v}\end{pmatrix}\,(u,v).$$

Le *déterminant jacobien* de φ est la fonction de (u,v) :

$$\begin{aligned}
\det\mathsf{Jac}(\varphi)\;&:=\;\det\begin{pmatrix}\frac{\partial f}{\partial u} & \frac{\partial f}{\partial v}\\ \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v}\end{pmatrix}\\
&= \frac{\partial f}{\partial u}\,\frac{\partial g}{\partial v}-\frac{\partial f}{\partial v}\,\frac{\partial g}{\partial u}.
\end{aligned}$$

Théorème 19.5. *Sous les hypothèses précédentes concernant le $\mathscr{C}^1$-difféomorphisme $\varphi\colon A\xrightarrow{\sim}B$ entre deux compacts élémentaires, ou entre deux réunions finies de compacts élémentaires μ-disjoints, on a :*

$$\mathsf{aire}(B)\;=\;\iint_B 1\,dxdy\;=\;\iint_A \det\mathsf{Jac}(\varphi)(u,v)\;dudv. \qquad\square$$

À présent, traitons le cas de l'intégration $\iint_B G$ d'une fonction continue quelconque, donc intégrable. Une astuce consiste à utiliser le Théorème 10.4 de la moyenne, qui stipule, pour une fonction F continue sur A, l'existence d'un point $(u_0,v_0)\in A$ avec :

$$\frac{1}{\mu(A)}\iint_A F(u,v)\,dudv\;=\;F(u_0,v_0).$$

Appliquée à $F:=\det\mathsf{Jac}(\varphi)$, cette formule donne le

Corollaire 19.6. *Il existe (au moins) un point $(u_0,v_0)\in A$ satisfaisant :*

$$\mu(B)\;=\;\det\mathsf{Jac}(\varphi)(u_0,v_0)\cdot\mu(A). \qquad\square$$

Rappelons que nous avons abrégé $\mathsf{aire}=\mu$. Nous pouvons enfin énoncer et démontrer le théorème principal de cette Section **19**.

Théorème 19.7. [Formule de changement de variables dans les intégrales doubles] *Soit un $\mathscr{C}^1$-difféomorphisme entre deux compacts élémentaires $A \subset \mathbb{R}^2_{u,v}$ et $B \subset \mathbb{R}^2_{x,y}$:*

$$\begin{aligned} \varphi\colon \qquad A &\overset{\sim}{\longrightarrow} B \\ (u,v) &\longmapsto \big(f(u,v),\, g(u,v)\big) =: (x,y), \end{aligned}$$

ou plus généralement, entre deux réunions finies de compacts élémentaires μ-disjoints. Alors pour toute fonction continue G sur B, on a :

$$\begin{aligned} \iint_B G(x,y)\,dxdy &= \iint_A G\circ\varphi \cdot \operatorname{det} \mathsf{Jac}(\varphi)(u,v)\; dudv \\ &= \iint_A G\big(f(u,v),\, g(u,v)\big)\cdot \operatorname{det}\mathsf{Jac}(\varphi)(u,v)\; dudv. \end{aligned}$$

Démonstration. Soit $\big(A_i\big)_{1\leqslant i\leqslant n}$ une μ-partition de A en (petits) compacts élémentaires au moyen de laquelle on approxime bien l'intégrale de $G\circ\varphi\cdot\operatorname{det}\mathsf{Jac}(\varphi)$. À travers φ, soit la μ-partition *image* :

$$\big(B_i\big)_{1\leqslant i\leqslant n} := \Big(\varphi\big(A_i\big)\Big)_{1\leqslant i\leqslant n},$$

qui est une μ-partition de B. Le Corollaire 19.6 donne, pour tous $i = 1, 2, \dots, n$, des points $(u_i, v_i) \in A_i$ tels que :

$$\mu\big(B_i\big) = \operatorname{det}\mathsf{Jac}(\varphi)\big(u_i, v_i\big)\cdot\mu\big(A_i\big) \qquad\qquad (1\leqslant i\leqslant n).$$

Soient alors les points-images :

$$\big(x_i, y_i\big) := \varphi\big(u_i, v_i\big) \qquad\qquad (1\leqslant i\leqslant n).$$

Les deux sommes de Riemann sont *égales* :

$$\sum_{i=1}^n G\big(x_i, y_i\big)\cdot\mu\big(B_i\big) = \sum_{i=1}^n G\circ\varphi\big(u_i, v_i\big)\cdot\operatorname{det}\mathsf{Jac}(\varphi)\big(u_i, v_i\big)\cdot\mu\big(A_i\big).$$

En raffinant la μ-partition, ces deux sommes de Riemann tendent vers les intégrales doubles respectives :

$$\begin{aligned} \iint_B G(x,y)\,dxdy \;\longleftarrow\; &\sum_{i=1}^n G\big(x_i, y_i\big)\,\mu\big(B_i\big) \\ = &\sum_{i=1}^n G\circ\varphi\big(u_i, v_i\big)\operatorname{det}\mathsf{Jac}(\varphi)\big(u_i, v_i\big)\,\mu\big(A_i\big) \\ \longrightarrow\; &\iint_A G\circ\varphi(u,v)\operatorname{det}\mathsf{Jac}(\varphi)(u,v)\,dudv. \end{aligned}$$ □

20. Intégrales doubles en coordonnées polaires

Dans le le plan euclidien $\mathbb{R}^2$ rapporté à un repère orthonormé $\big\{0, \mathbf{i}, \mathbf{j}\big\}$, considérons l'application φ de passage aux coordonnées polaires définie par :

$$\varphi\colon \qquad (r,\theta) \longmapsto \big(r\cos\theta,\; r\sin\theta\big) =: (x,y).$$

Nous savons que pour toute constante θ_0 fixée, si on se limite à faire varier $\theta \in \big[\theta_0, \theta_0 + 2\pi\big[$ dans un intervalle semi-ouvert de longueur 2π partant de θ_0, et si on se limite à des rayons $r > 0$ strictement positifs, alors φ établit une bijection :

$$\varphi\colon \qquad \mathbb{R}_+^* \times \big[\theta_0,\; \theta_0 + 2\pi\big[\;\overset{\sim}{\longrightarrow}\; \mathbb{R}^2\backslash\{0\}$$

sur le plan euclidien épointé de son origine. De plus, cette bijection est continûment différentiable avec :

$$\begin{aligned} dx &= \cos\theta\, dr - r\sin\theta\, d\theta, \\ dy &= \sin\theta\, dr + r\cos\theta\, d\theta, \end{aligned}$$

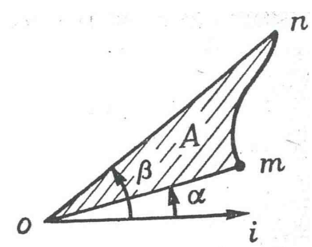
et par suite, le déterminant jacobien de φ en un point (r, θ) vaut :

$$\begin{aligned} \det \mathrm{Jac}(\varphi) &= \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} \\ &= \cos\theta\, r\cos\theta - \big(-r\sin\theta\big)\sin\theta \\ &= r. \end{aligned}$$

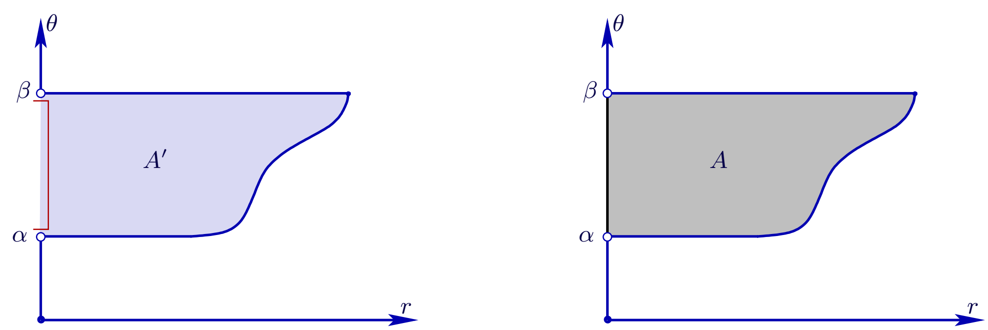
En application du Théorème 19.7, pour une réunion finie A de compacts élémentaires inclus dans $\mathbb{R}_+^* \times \big[\theta_0, \theta_0 + 2\pi\big[$, pour $B := \varphi(A)$, et pour toute application continue G sur B, nous obtenons la formule très importante de calculs d'intégrales en coordonnées polaires :

$$\int\!\!\int_B G(x,y)\, dxdy \;=\; \int\!\!\int_A G\big(r\cos\theta,\; r\sin\theta\big)\, r\, dr\, d\theta$$

Un cas spécial intéressant se produit quand A n'est pas forcément contenu dans $\mathbb{R}_+^* \times \big[\theta_0, \theta_0 + 2\pi\big[$, à savoir quand $A \ni 0$ contient l'origine, comme nous l'avons déjà vu.



Supposons en effet que B soit un secteur angulaire élémentaire, limité par deux rayons $0m$ et $0n$, ainsi que par un arc mn, rencontré en un seul point par une droite passant par θ, d'équation polaire $r = f(\theta)$.



Désignons par A' la partie du pavé $\mathbb{R}_+^* \times \big[\theta_0, \theta_0 + 2\pi\big[$ définie par :

$$0 < r \leqslant f(\theta) \qquad\qquad \text{et} \qquad\qquad \alpha \leqslant \theta \leqslant \beta.$$

Alors A' est en bijection, *via* φ, avec $B \backslash \{0\}$:

$$\varphi(A') = B \backslash \{0\}.$$

Désignons par A le compact déduit de A' par adjonction du segment :

$$r = 0, \qquad\qquad \alpha \leqslant \theta \leqslant \beta.$$

Alors A et A' ont même aire et pour toute fonction continue G sur B, la composée $G \circ \varphi$ est continue sur A et on a l'égalité :

$$\iint_{A'} G \circ \varphi \; r \, drd\theta \;=\; \iint_{A} G \circ \varphi \; r \, drd\theta.$$

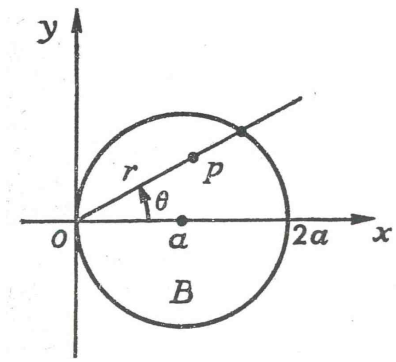
En définitive, *la formule de changement de variables peut encore s'appliquer au cas où B est réunion finie de secteurs angulaires élémentaires, deux à deux μ-disjoints*.

Exemple 20.1. Proposons-nous de calculer l'intégrale double :

$$I \;:=\; \iint_B \big(x^2 + y^2\big) \, dxdy,$$

sur le disque circulaire B de rayon $a > 0$ centré en le point $(a, 0)$ défini par :

$$x^2 + y^2 - 2\,a\,x \;\leqslant\; 0.$$



Désignons par φ l'application :

$$x \;=\; r\cos\theta, \qquad\qquad y \;=\; r\sin\theta,$$

sur la partie A définie par :

$$-\frac{\pi}{2} \;\leqslant\; \theta \;\leqslant\; \frac{\pi}{2}, \qquad\qquad 0 \;\leqslant\; r \;\leqslant\; 2a\cos\theta.$$

On a :

$$\varphi(A) \;=\; B,$$

et de plus, si on désigne par A' la partie définie par :

$$-\frac{\pi}{2} \;<\; \theta \;<\; \frac{\pi}{2}, \qquad\qquad 0 \;<\; r \;\leqslant\; 2a\cos\theta,$$

alors φ est une bijection de A' sur $B\backslash\{0\}$.

On peut donc appliquer la formule du changement de variables en coordonnées polaires :

$$I \;=\; \iint_A r^2\, r\, drd\theta \;=\; 2\int_0^{\frac{\pi}{2}} d\theta \int_0^{2a\cos\theta} r^3\, dr \;=\; 2\int_0^{\frac{\pi}{2}} \frac{1}{4}\big(2a\cos\theta\big)^4 d\theta \;=\; 8\,a^4 \int_0^{\frac{\pi}{2}} \cos^4\theta\, d\theta.$$

Enfin, en remarquant que :

$$\cos^4\theta \;=\; \frac{\big(1+\cos 2\theta\big)^2}{4} \;=\; \frac{1}{4} + \frac{1}{2}\cos 2\theta + \frac{1}{4}\,\frac{\big(1+\cos 4\theta\big)}{2},$$

on trouve aisément :

$$I \;=\; \frac{3\,\pi\,a^4}{2}. \qquad \square$$

Exemple 20.2. Revenons à l'Exemple 16.1 :

$$I := \iint_B \left(x^2 - y^2\right) dxdy,$$

où B est le disque elliptique :

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 \leqslant 0.$$

Nous avons vu que $I = \frac{\pi}{4}\, ab\left(a^2 - b^2\right)$, et nous souhaitons montrer qu'un changement de variables approprié permet d'arriver plus rapidement au résultat.

Considérons en effet l'application φ :

$$x = a\,u\cos v, \qquad\qquad y = b\,u\sin v,$$

sur le pavé A_1 défini par :

$$0 < u \leqslant 1, \qquad\qquad 0 \leqslant v \leqslant \frac{\pi}{2}.$$

Ici, φ est une bijection de A_1 sur $B_1 \backslash \{0\}$, où B_1 est le secteur elliptique :

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 \leqslant 0, \qquad\qquad x \geqslant 0, \quad y \geqslant 0.$$

Remarquons que φ est un $\mathscr{C}^1$-difféomorphisme de A_1 sur B_1, dont le jacobien vaut :

$$\begin{vmatrix} a\cos v & -a\,u\sin v \\ b\sin v & b\,u\cos v \end{vmatrix} = a\,b\,u.$$

Le Théorème 19.7 de changement de variables donne alors :

$$I = 4\iint_A \left(x^2 - y^2\right) dxdy = 4\iint_{A_1} \left(a^2\cos^2 v - b^2\sin^2 v\right)^2 u^2\, a\,b\,u \; dudv,$$

d'où :

$$I = 4\,a\,b\int_0^1 u^3\, du \cdot \int_0^{\frac{\pi}{2}} \left(a^2\cos^2 v - b^2\sin^2 v\right) dv,$$

et nous sommes ramenés au calcul de deux intégrales simples indépendantes. En utilisant (exercice) :

$$1 = \int_0^{\frac{\pi}{2}} \cos^2 v\, dv = \int_0^{\frac{\pi}{2}} \sin^2 v\, dv,$$

nous concluons que :

$$I = 4\,a\,b\,\frac{1}{4}\,\frac{\pi}{4}\left(a^2 - b^2\right) = \frac{\pi}{4}\, ab\left(a^2 - b^2\right).$$

□

21. Exercices

Exercice 1. EE

Intégrales triples

François DE MARÇAY
Département de Mathématiques d'Orsay
Université Paris-Sud, France

1. Introduction

2. Définition et propriétés de l'intégrale triple

La définition et les propriétés des intégrales doubles peuvent se généraliser, en utilisant des μ-partitions d'ensembles cubables.

Au chapitre précédent, nous avons défini les parties *cubables* de l'espace $\mathbb{R}^3$ de la géométrie euclidienne à 3 dimensions, comme celles auxquelles on peut attribuer un *volume*. Sur l'ensemble $\mathscr{C}$ des parties cubables de $\mathbb{R}^3$, nous avons défini une application $A \longmapsto \mu(A)$ qui associe à tout $A \in \mathscr{C}$ un nombre réel positif, son *volume* $\mu(A)$. Ce dernier satisfait :

$$A \cap B = \emptyset \qquad \Longrightarrow \qquad \mu\big(A \cup B\big) = \mu(A) + \mu(B).$$

Cette implication est encore exacte lorsque A et B sont *μ-disjoints*, c'est-à-dire ont une intersection qui, sans nécessairement être vide, a un volume *nul*, propriété que nous avons notée :

$$A \cap B = \emptyset_\mu.$$

Définition 2.1. On appelle *μ-partition* d'un ensemble cubable $A \in \mathscr{C}$ toute famille finie $\big(A_i\big)_{1 \leqslant i \leqslant n}$ constituée de parties cubables deux à deux μ-disjointes, dont la réunion est égale à :

$$A = A_1 \cup A_2 \cup \cdots \cup A_n.$$

Pour construire la notion d'intégrale triple, considérons une une fonction bornée $f \colon A \longrightarrow \mathbb{R}$ définie sur un ensemble cubable $A \subset \mathbb{R}^3$. Pour toute μ-partition de A :

$$p = \big(A_i\big)_{1 \leqslant i \leqslant n},$$

posons, pour tout $1 \leqslant i \leqslant n$:

$$m_i := \inf_{(x,y,z) \in A_i} f(x,y,z) \qquad \text{et} \qquad M_i := \sup_{(x,y,z) \in A_i} f(x,y,z).$$

Ensuite, formons les deux sommes :

$$\begin{aligned} s(p) &:= m_1\,\mu\big(A_1\big) + m_2\,\mu\big(A_2\big) + \cdots + m_n\,\mu\big(A_n\big), \\ S(p) &= M_1\,\mu\big(A_1\big) + M_2\,\mu\big(A_2\big) + \cdots + M_n\,\mu\big(A_n\big). \end{aligned}$$

Lorsque la μ-partition p parcourt l'ensemble $\mathscr{P}_\mu(A)$ de toutes les μ-partitions de A, on prouve, comme pour les intégrales doubles, que l'ensemble des nombres $s(p)$ est majoré par toute somme $S(q)$ où $q \in \mathscr{P}_\mu(A)$ est fixée, et par conséquent, le supremum suivant existe :

$$\sup_{p \in \mathscr{P}_\mu(A)} s(p) =: \sigma_f(A).$$

De même, l'ensemble des nombres $S(p)$ est minoré par toute somme $s(q)$, et par suite, existe l'infimum :

$$\inf_{p\in\mathscr{P}_\mu(A)} S(p) =: \Sigma_f(A).$$

On a évidemment :

$$\sigma_f(A) \leqslant \Sigma_f(A).$$

Définition 2.2. On dit qu'une fonction bornée $f\colon A \longrightarrow \mathbb{R}$ sur un ensemble cubable $A \subset \mathbb{R}^3$ est *intégrable* (au sens de Darboux) si :

$$\sigma_f(A) = \Sigma_f(A).$$

On note alors cette valeur commune :

$$\iiint_A f(x,y,z)\,dxdydz,$$

et on la nomme *intégrale triple* de f sur la partie cubable A.

L'ensemble des fonctions intégrables sur A sera noté :

$$\mathsf{Int}\big(A,\mathbb{R}\big).$$

Sans attendre, venons-en aux propriétés de l'intégrale triple, que nous admettrons, puisqu'elles se démontrent comme celles de l'intégrale double.

Théorème 2.3. *Soient A et B deux parties cubables de l'espace $\mathbb{R}^3$ qui sont μ-disjointes :*

$$0 = \mu\big(A\cap B\big).$$

Pour toute fonction bornée $f\colon A\cup B \longrightarrow \mathbb{R}$, on a l'équivalence :

$$f \text{ est intégrable sur } A \text{ et sur } B \qquad\Longleftrightarrow\qquad f \text{ est intégrable sur } A\cup B.$$

Dans ce cas :

$$\iiint_{A\cup B} f(x,y,z)\,dxdydz = \iiint_A f(x,y,z)\,dxdydz + \iiint_B f(x,y,z)\,dxdydz. \qquad \square$$

Ensuite, le théorème suivant montre que l'espace $\mathsf{Int}\,(A,\mathbb{R})$ est un $\mathbb{R}$-espace vectoriel.

Théorème 2.4. *Si f et g sont deux fonctions bornées intégrables sur une partie cubable $A \subset \mathbb{R}^3$, alors pour tous réels $\lambda, \mu \in \mathbb{R}$, la combinaison linéaire :*

$$\lambda\, f + \mu\, g \in \mathsf{Int}\,(A,\mathbb{R})$$

est aussi intégrable sur A, avec :

$$\iiint_A \Big(\lambda\, f(x,y,z)+\mu\, g(x,y,z)\Big)\,dxdy = \lambda\iiint_A f(x,y,z)\,dxdydz+\mu\iiint_A g(x,y,z)\,dxdydz.$$

$\square$

Exprimons aussi un critère d'intégrabilité tout à fait analogue, lui aussi, à ce qui a déjà été vu en dimension 2.

Lemme 2.5. **[Critère d'intégrabilité]** *Pour que f soit intégrable (au sens de Darboux) sur A, il faut et il suffit que pour tout $\varepsilon > 0$, il existe une μ-partition p de A telle que :*

$$(0 \leqslant) \qquad S(p) - s(p) \leqslant \varepsilon. \qquad \square$$

Le théorème probablement le plus utile et le plus important est le suivant.

Théorème 2.6. *Toute fonction continue sur un compact cubable y est intégrable.* $\square$

De plus, il existe une relation d'ordre sur les intégrales triples.

Théorème 2.7. *Si $f\colon A \longrightarrow \mathbb{R}$ et $g\colon A \longrightarrow \mathbb{R}$ sont bornées intégrables sur une partie cubable $A \subset \mathbb{R}^3$, alors :*

$$f \leqslant g \qquad \Longrightarrow \qquad \iiint_A f(x,y,z)\,dxdydz \;\leqslant\; \iiint_A g(x,y,z)\,dxdydz.$$

De plus :

$$\left| \iiint_A f(x,y,z)\,dxdydz \right| \;\leqslant\; \iiint_A \big|f(x,y,z)\big|\,dxdydz. \qquad \square$$

Il existe aussi des inégalités élémentaires.

Théorème 2.8. *Soit $f \in \mathsf{Int}\,(A,\mathbb{R})$ une fonction intégrable sur $A \subset \mathbb{R}^3$ cubable. Alors en posant :*

$$m := \inf_A f \qquad \textit{et} \qquad M := \inf_A f,$$

on a :

$$m \cdot \mu(A) \;\leqslant\; \iiint_A f(x,y,z)\,dxdydz \;\leqslant\; M \cdot \mu(A). \qquad \square$$

Le nombre :

$$\lambda := \frac{1}{\mu(A)} \iiint_A f(x,y,z)\,dxdydz$$

se nomme alors *valeur moyenne* de f sur A.

Théorème 2.9. *Si $A \subset \mathbb{R}^2$ est cubable et si $f\colon A \longrightarrow \mathbb{R}$ est continue, il existe un $(x_0, y_0, z_0) \in A$ tel que :*

$$\frac{1}{\mu(A)} \iiint_A f(x,y,z)\,dxdydz \;=\; f(x_0,y_0,z_0). \qquad \square$$

3. Intégrale triple de Riemann

Soit f une fonction bornée sur une partie cubable $A \subset \mathbb{R}^3$. Pour toute μ-partition $\big(A_i\big)_{1 \leqslant i \leqslant n}$ de A, nous avons posé plus haut, pour tout indice $1 \leqslant i \leqslant n$:

$$m_i := \inf_{(x,y,z)\in A_i} f(x,y,z) \qquad \text{et} \qquad M_i := \sup_{(x,y,z)\in A_i} f(x,y,z).$$

Dans cet espace euclidien $\mathbb{R}^3$, rappelons que la distance euclidienne entre deux points (x,y,z) et (x',y',z') est :

$$\big\|(x,y,z)-(x',y',z')\big\| := \sqrt{\big(x-x'\big)^2+\big(y-y'\big)^2+\big(z-z'\big)^2}.$$

Par définition, le *diamètre* de chaque A_i est le nombre réel positif :

$$\delta_i := \sup_{\substack{(x,y,z)\in A_i \\ (x',y',z')\in A_i}} \big\|(x,y,z)-(x',y',z')\big\|.$$

Par définition aussi, le *pas* de la μ-partition $\big(A_i\big)_{1 \leqslant i \leqslant n}$ est le nombre :

$$\mathsf{pas}(\delta) := \max_{1 \leqslant i \leqslant n} \delta_i.$$

Pour chaque indice $1 \leqslant i \leqslant n$, choisissons *arbitrairement* un réel $\theta_i \in [m_i, M_i]$, par exemple une valeur $\theta_i = f(x_i, y_i, z_i)$ de la fonction en un point arbitraire $(x_i, y_i, z_i) \in A_i$. On appelle *somme de Riemann* le nombre réel :

$$\mathscr{R}(p) := \theta_1\,\mu\big(A_1\big) + \theta_2\,\mu\big(A_2\big) + \cdots + \theta_n\,\mu\big(A_n\big).$$

On a évidemment :

$$s(p) \;\leqslant\; \mathscr{R}(p) \;\leqslant\; S(p).$$

Comme dans la théorie des intégrales doubles de Riemann, on établit le

Théorème 3.1. *Si $f\colon A \longrightarrow \mathbb{R}$ est intégrable sur $A \subset \mathbb{R}^2$ quarrable, alors :*

$$\lim_{\mathrm{pas}(p)\to 0} \mathscr{R}(p) \,=\, \iiint_A f(x,y,z)\,dxdydz. \qquad \square$$

L'intérêt, c'est qu'on approxime la valeur de l'intégrale par des sommes échantillonnées 'au hasard' : chaque morceau A_i 'pèse' :

$$f(x_i,y_i,z_i)\cdot \mathrm{volume}(A_i),$$

où $f(x_i,y_i,z_i)$ est une valeur de la fonction en un point (x_i,y_i,z_i) choisi arbitrairement dans A_i.

Il y a aussi une réciproque à ce résultat.

Théorème 3.2. *Toute fonction intégrable au sens de Riemann :*

$$\lim_{\mathrm{pas}\to 0} \mathscr{R}(p) \quad \textit{existe},$$

est aussi intégrable (au sens de Darboux), i.e. *appartient à* $\mathrm{Int}\,(A,\mathbb{R})$. $\square$

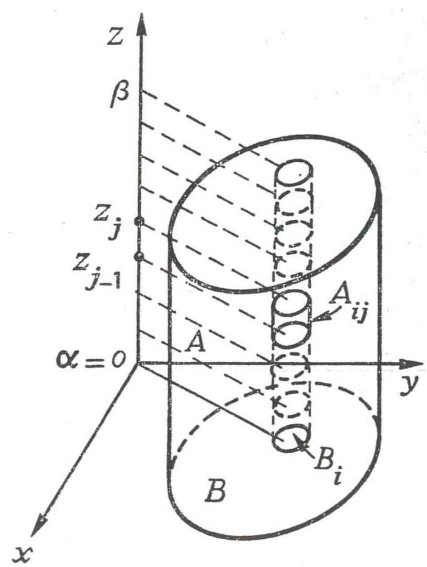
4. Calcul d'une intégrale triple sur un compact cylindrique

Rapportons l'espace euclidien $\mathbb{R}^3$ à un repère orthonormé $\{0,\mathbf{i},\mathbf{j},\mathbf{k}\}$. Soit $[a,b]$ un intervalle fermé, et considérons deux fonctions numériques φ_1 et φ_2 continues sur $[a,b]$. Soit le compact simple du plan horizontal :

$$B \,:=\, \big\{(x,y)\in\mathbb{R}^2\colon\ a\leqslant x\leqslant b,\ \ \varphi_1(x)\leqslant y\leqslant \varphi_2(x)\big\}.$$

Soient enfin deux constantes $\alpha<\beta$, et soit le *compact cylindrique* de l'espace tridimensionnel :

$$A \,:=\, \big\{(x,y,z)\in\mathbb{R}^3\colon\ (x,y)\in B,\ \ \alpha\leqslant z\leqslant\beta\big\}$$



Sur la figure, le nombre α vaut 0. On a vu que tout compact simple est cubable. Ici, le volume de A vaut :

$$\mu_v(A) \,=\, \mu_a(B)\,\big(\beta-\alpha\big),$$

puisque l'aire de la base B vaut $\mu_a(B)$, et le cylindre est de hauteur $(\beta-\alpha)$. On remarquera la différence notationnelle :

$\mu_a(\bullet)$ pour l'aire 2-dimensionnelle $\qquad\neq\qquad$ $\mu_v(\bullet)$ pour le volume 3-dimensionnel.

Proposons-nous alors de calculer l'intégrale triple :

$$\iiint_A f(x,y,z)\,dxdydz,$$

où $f\colon A \longrightarrow \mathbb{R}$ est une fonction continue (donc bornée). À cette fin, considérons une μ-partition quelconque de B :

$$\big(B_i\big)_{1\leqslant i\leqslant n},$$

ainsi qu'une subdivision de l'intervalle vertical $[\alpha,\beta]$:

$$\alpha\,=\,z_0\,<\,z_1\,<\,\cdots\,<\,z_{q-1}\,<\,z_q\,=\,\beta.$$

Comme l'illustre la figure, au-dessus de chaque élément-plan B_i de la μ-partition du plan de base, nous avons empilé q compacts cylindriques $A_{i,1},\dots,A_{i,q}$ définis par :

$$A_{i,j}\,=\,B_i\times\big[z_{j-1},z_j\big] \qquad\qquad (1\leqslant i\leqslant n,\ 1\leqslant j\leqslant q).$$

Nous avons ainsi constitué une μ-partition du compact 3-dimensionnel A :

$$p\,=\,\big(A_{i,j}\big)_{\substack{1\leqslant i\leqslant n\\1\leqslant j\leqslant q}},$$

en compacts cylindriques chacun de volume égale à :

$$\mu_v\big(A_{i,j}\big)\,=\,\big(z_j-z_{j-1}\big)\,\mu\big(B_i\big).$$

Ensuite, avec une telle μ-partition p, choisissons des nombres $\theta_{i,j}$ arbitrairement parmis les valeurs de f sur $A_{i,j}$ et associons-leur la somme de Riemann :

$$\begin{aligned}\mathscr{R}(p)\,:=\,\sum_{i=1}^n\sum_{j=1}^q\theta_{i,j}\,\mu_v\big(A_{i,j}\big)\;&=\;\sum_{i=1}^n\sum_{j=1}^q\theta_{i,j}\,\big(z_j-z_{j-1}\big)\,\mu_a\big(B_i\big)\\&=\;\sum_{i=1}^n\mu_a\big(B_i\big)\,\sum_{j=1}^q\big(z_j-z_{j-1}\big)\,\theta_{i,j},\end{aligned}$$

— en fait, nous allons effectuer astucieusement le choix de ces $\theta_{i,j}$.

En effet, pour tout $1\leqslant i\leqslant n$, prenons tout d'abord arbitrairement un point $(x_i,y_i)\in B_i$, mais ensuite, pour tout $1\leqslant j\leqslant q$, prenons la valeur moyenne de la fonction partielle $z\longmapsto f\big(x_i,y_i,z\big)$ sur l'intervalle $[z_{j-1},z_j]$:

$$\theta_{i,j}\,:=\,\frac{1}{z_j-z_{j-1}}\int_{z_{j-1}}^{z_j}f\big(x_i,y_i,z\big)\,dz.$$

Alors grâce à la règle de Chasles pour les intégrales :

$$\sum_{j=1}^q\big(z_j-z_{j-1}\big)\,\theta_{i,j}\;=\;\sum_{j=1}^q\int_{z_{j-1}}^{z_j}f\big(x_i,y_i,z\big)\;=\;\int_\alpha^\beta f\big(x_i,y_i,z\big)\,dz.$$

On peut vérifier rigoureusement que la fonction :

$$\psi(x,y)\,=\,\int_\alpha^\beta f\big(x,y,z\big)\,dz$$

est continue dans le compact simple $B=\big\{a\leqslant x\leqslant b,\,\varphi_1(x)\leqslant y\leqslant\varphi_2(x)\big\}$. Par conséquent, la somme de Riemann considérée devient :

$$\mathscr{R}(p)\,=\,\sum_{i=1}^q\mu_a(B_i)\,\psi\big(x_i,y_i\big),$$

de telle sorte que nous reconnaissons ici une somme de Riemann pour l'intégrale double :

$$\iint_B \psi(x,y)\,dxdy.$$

Notons de plus que le pas de la μ-partition $\big(B_i\big)_{1\leqslant i\leqslant n}$ est au plus égal au pas de la μ-partition $\big(A_{i,j}\big)_{\substack{1\leqslant i\leqslant n\\ 1\leqslant j\leqslant q}}$. Lorsque le second pas tend vers zéro, *a fortiori*, le premier pas tend vers zéro. Nous obtenons donc le

Théorème 4.1. *Sur un compact cylindrique $A \subset \mathbb{R}^3$ délimité par des fonctions continues :*

$$a \leqslant x \leqslant b, \qquad \varphi_1(x) \leqslant y \leqslant \varphi_2(x), \qquad \alpha \leqslant z \leqslant \beta,$$

de base planaire $B \subset \mathbb{R}^2$ définie par :

$$a \leqslant x \leqslant b, \qquad \varphi_1(x) \leqslant y \leqslant \varphi_2(x),$$

on a pour toute fonction continue $f\colon A \longrightarrow \mathbb{R}$:

$$\iiint_A f(x,y,z)\,dxdydz \;=\; \iint_B dxdy\left(\int_\alpha^\beta f(x,y,z)\,dz\right). \qquad \square$$

Le calcul d'une intégrale triple et ainsi ramené à celui d'une intégrale simple suivi du calcul d'une intégrale double.

Exemple 4.2. Proposons-nous de calculer l'intégrale triple :

$$I := \iiint_A z^{x^2+y^2}\,dxdydz,$$

sur le compact (vraiment) cylindrique A défini par :

$$x^2+y^2 \leqslant a, \qquad 0 \leqslant z \leqslant 1.$$

Désignons par B le disque circulaire :

$$x^2+y^2 \leqslant a,$$

projection orthogonale de A sur le plan horizontal $0xy$. D'après le Théorème 4.1, on a :

$$I = \iint_B dxdy \int_0^1 z^{x^2+y^2}\,dz \;=\; \iint_B dxdy\left[\frac{z^{x^2+y^2+1}}{x^2+y^2+1}\right]_0^1 \;=\; \iint_B dxdy\,\frac{1}{x^2+y^2+1}.$$

Le passage aux coordonnées polaires :

$$x = r\cos\theta, \qquad y = r\sin\theta,$$

avec, comme nous nous en souvenons :

$$dxdy = r\,drd\theta$$

est évidemment tout indiqué, et il nous ramène au produit de deux intégrales simples indépendantes :

$$I = \int_0^{2\pi} d\theta \int_0^a \frac{r\,dr}{r^2+1} = 2\pi\left[\frac{1}{2}\log\big(r^2+1\big)\right]_0^a = \pi\log\big(a^2+1\big).$$

5. Calcul d'une intégrale triple sur un compact simple

Proposons-nous maintenant de calculer l'intégrale triple :

$$\iiint_A f(x,y,z)\,dxdyz,$$

f étant une fonction continue sur le compact simple général de $\mathbb{R}^3$ défini par :

$$a \leqslant x \leqslant b, \qquad \varphi_1(x) \leqslant y \leqslant \varphi_2(x), \qquad \psi_1(x,y) \leqslant z \leqslant \psi_2(x,y),$$

où φ_1 et φ_2 sont, comme précédemment, deux fonctions continues sur $[a,b]$, et où ψ_1 et ψ_2 sont deux fonctions continues sur le compact simple B de $\mathbb{R}^2$ défini par :

$$a \leqslant x \leqslant b, \qquad \varphi_1(x) \leqslant y \leqslant \varphi_2(x).$$

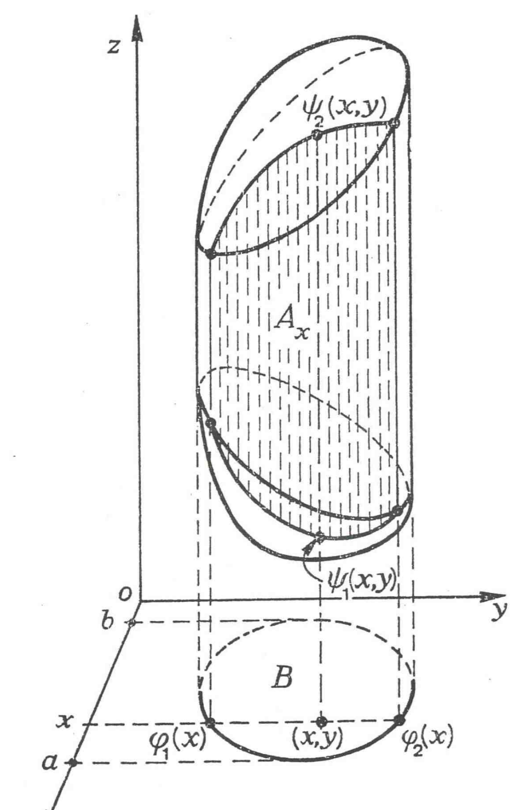


FIG. 2

Soit α un minorant de ψ_1 sur B et soit aussi β un majorant de ψ_2 sur B :

$$\alpha \leqslant \psi_1(x,y) \leqslant \psi_2(x,y) \leqslant \beta \qquad (\forall\,(x,y) \in B).$$

Désignons par C le compact cylindrique associé, ensemble des points $(x,y,z) \in \mathbb{R}^3$ satisfaisant :

$$(x,y) \in B \qquad \alpha \leqslant z \leqslant \beta.$$

Sur cette partie cubable, définissons la fonction auxiliaire :

$$g(x,y,z) := \begin{cases} f(x,y,z) & \text{lorsque } (x,y,z) \in A, \\ 0 & \text{lorsque } (x,y,z) \in C\backslash A. \end{cases}$$

Cette fonction g est intégrable sur C puisqu'elle l'est séparément sur A et sur $C \backslash A$. Comme elle s'annule identiquement sur $C \backslash A$, on a clairement :

$$\iiint_A f(x,y,z)\,dxdydz \;=\; \iiint_C g(x,y,z)\,dxdydz.$$

En application du Théorème 4.1, nous pouvons calculer la deuxième intégrale double jusqu'à obtenir :

$$F(x,y) \;:=\; \int_\alpha^\beta g(x,y,z)\,dz \;=\; \int_{\psi_1(x,y)}^{\psi_2(x,y)} f(x,y,z)\,dz,$$

grâce au fait que la fonction g est nulle sur les deux intervalles :

$$\big[\alpha,\,\psi_1(x,y)\big[\qquad\qquad \text{et} \qquad\qquad \big]\psi_2(x,y),\,\beta\big].$$

La fonction $(x,y) \longmapsto F(x,y)$ est continue sur le compact B, donc intégrable sur B. On obtient par conséquent :

$$\iiint_A f(x,y,z)\,dxdydz \;=\; \iint_B F(x,y)\,dxdy.$$

En calculant finalement cette intégrale double comme dans le chapitre précédent, nous concluons par le

Théorème 5.1. *Sur un compact simple $A \subset \mathbb{R}^3$ délimité par des fonctions continues :*

$$a \leqslant x \leqslant b, \qquad\qquad \varphi_1(x) \leqslant y \leqslant \varphi_2(x), \qquad\qquad \psi_1(x,y) \leqslant z \leqslant \psi_2(x,y),$$

on a pour toute fonction continue $f\colon A \longrightarrow \mathbb{R}$:

$$\iiint_A f(x,y,z)\,dxdydz \;=\; \int_a^b dx \int_{\varphi_1(x)}^{\varphi_2(x)} dy \int_{\psi_1(x,y)}^{\psi_2(x,y)} f(x,y,z)\,dz. \qquad \square$$

Exemple 5.2. Soit à calculer l'intégrale :

$$I \;:=\; \iiint_A z^2\,dxdydz,$$

sur la boule fermée :

$$x^2+y^2+z^2 \;\leqslant\; a^2.$$

Remarquons que A admet les symétries d'axes $0x$, $0y$, $0z$ qui laissent invariante la fonction à intégrer. On a donc :

$$I \;=\; 8 \iiint_{A^+} z^2\,dxdydz,$$

où A^+ est le $\frac{1}{8}$ de boule défini par :

$$A^+\colon \qquad x \geqslant 0, \qquad y \geqslant 0, \qquad z \geqslant 0, \qquad\qquad x^2+y^2+z^2 \;\leqslant\; a^2,$$

qui se projette orthogonalement sur le $\frac{1}{4}$ de cercle :

$$B^+\colon \qquad x \geqslant 0, \qquad y \geqslant 0, \qquad\qquad x^2+y^2-a^2 \;\leqslant\; 0.$$

Il en résulte que :

$$\begin{aligned} I \;=\; 8 \iint_B dxdy \int_0^{\sqrt{a^2-x^2-y^2}} z^2\,dz \;&=\; \iint_B dxdy \left[\frac{1}{3}\,z^3\right]_0^{\sqrt{a^2-x^2-y^2}} \\ &=\; \frac{8}{3} \iint_B dxdy \left(a^2-x^2-y^2\right)^{\frac{3}{2}} dxdy. \end{aligned}$$

Finalement, en passant aux coordonnées polaires, on trouve :

$$\begin{aligned} I &= \frac{8}{3} \int_0^{\frac{\pi}{2}} d\theta \int_0^a \left(a^2 - r^2\right)^{\frac{3}{2}} d\,dr \\ &= \frac{8}{3}\,\frac{\pi}{2} \Big[-\, \frac{1}{5} \left(a^2 - r^2\right)^{\frac{5}{2}} \Big]_0^a \\ &= \frac{4\,\pi}{15}\, a^5. \end{aligned}$$

□

6. Volume par intégration de sections planes

Comme dans le Théorème 5.1, soit un compact général $A \subset \mathbb{R}^3$ défini par :

$$a \leqslant x \leqslant b, \qquad \varphi_1(x) \leqslant y \leqslant \varphi_2(x), \qquad \psi_1(x,y) \leqslant z \leqslant \psi_2(x,y).$$

Désignons par A_x la partie de A constituée des points d'abscisse fixée $x \in [a, b]$:

$$A_x\colon \quad \varphi_1(x) \leqslant y \leqslant \varphi_2(x), \qquad \psi_1(x,y) \leqslant z \leqslant \psi_2(x,y).$$

D'un point de vue géométrique, A_x est l'intersection de A avec le plan parallèle au plan $y0z$ passant par x.

Grâce au calcul d'une intégrale double vu au chapitre précédent, nous pouvons écrire :

$$\int_{\varphi_1(x)}^{\varphi_2(x)} dy \int_{\psi_1(x,y)}^{\psi_2(x,y)} f(x,y,z)\, dz \;=\; \iint_{A_x} f(x,y,z)\, dydz,$$

et par suite :

$$\iiint_A f(x,y,z)\, dxdydz \;=\; \int_a^b dx \left(\iint_{A_x} f(x,y,z)\, dydz \right).$$

En particulier, si on considère la fonction :

$$f(x,y,z) := 1,$$

l'intégrale triple considérée représente évidemment le volume $\mu_v(A)$ du compact cubable A :

$$\mu_v(A) \;=\; \iiint_A dxdydz.$$

Quant à l'intégrale double, elle vaut l'aire $\mu_a(A_x)$ de la partie quarrable A_x :

$$\mu_a\big(A_x\big) \;=\; \iint_{A_x} dydz.$$

La formule précédente exprime alors que le calcul d'un volume se ramène à une intégrale simple de calculs d'aires.

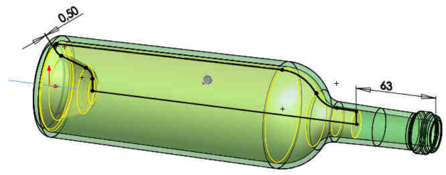
Théorème 6.1. *Soit $A \subset \mathbb{R}^3$ un compact simple cubable, et soit $\big(A_x\big)_{a \leqslant x \leqslant b}$ la famille de ses sections planes par des plans perpendiculaires à une droite fixée $0x$. Alors le volume $\mu_v(A)$ est donné par l'intégrale simple :*

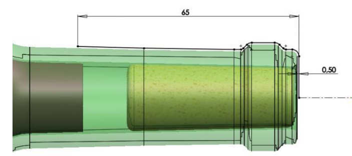
$$\mu_v(A) \;=\; \int_a^b \mu_a\big(A_x\big)\, dx,$$

des aires $\mu_a\big(A_x\big)$ des ses sections A_x.

Ce résultat est encore valable pour A cubable sans être compact.

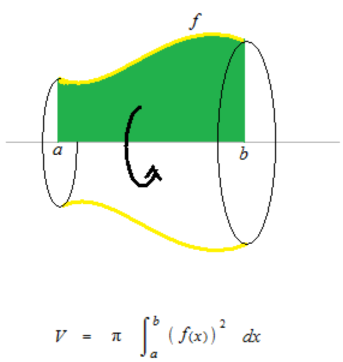
7. Volume d'un compact de révolution

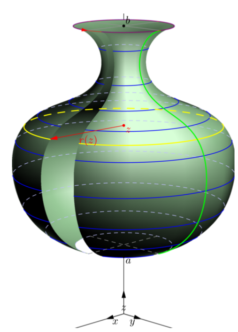


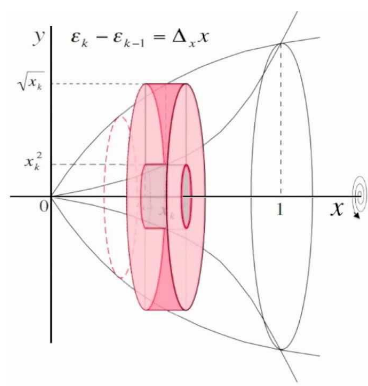


Considérons le cas particulier où le compact $A \subset \mathbb{R}^3$ est de révolution autour d'un axe $0x$, de telle sorte que toute section plane A_x, pour $a \leqslant x \leqslant b$, par un plan perpendiculaire à $0x$ soit un disque de rayon $r(x)$ centré sur l'axe. On a alors pour le volume :

$$\mu_v(A) = \pi \int_a^b \big(r(x)\big)^2 \, dx.$$







Cette formule peut s'étendre au cas où le compact est de révolution de telle sorte que la section A_x ne soit pas un disque complet, mais une couronne circulaire de rayons respectifs $r_1(x) \leqslant r_2(x)$. On a alors, pour le volume, la formule :

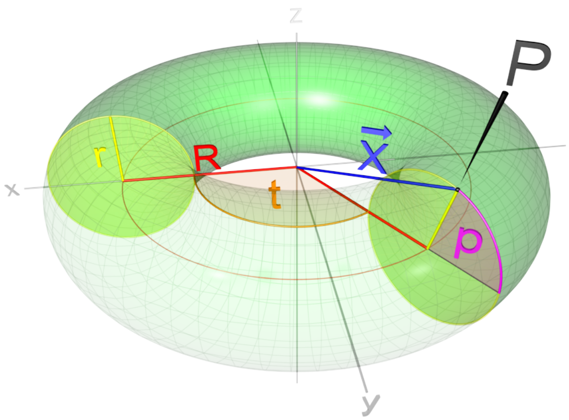
$$\mu_v(A) = \pi \int_a^b \Big[\big(r_2(x)\big)^2 - \big(r_1(x)\big)^2\Big] \, dx.$$

Ensuite, cherchons à transformer cette formule. Introduisons la *méridienne* de A, c'est-à-dire la section de A par tout plan contenant l'axe $0x$, cette figure ne dépendant pas de l'angle que fait le plan autour de cet axe, puisque A est de révolution. Notons Γ le bord de cette méridienne.

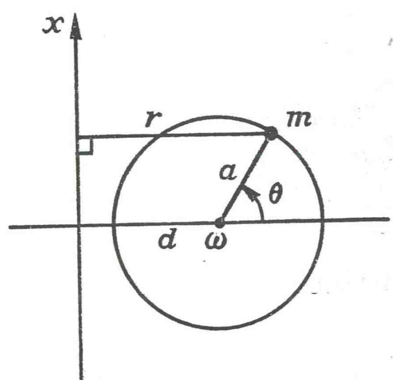
Alors la formule qui précède pour le volume de A peut se récrire sous forme d'une intégrale curviligne :

$$\mu_v(A) = \pi \int_\Gamma r^2 \, dx,$$

et cette formule est vraie pour tout solide A de révolution autour de $0x$. Par exemple, soit un solide comme l'intérieur d'un pneu, ou d'une chambre à air dans l'espace.



Exemple 7.1. [Volume d'un tore] Le terme mathématique est moins imagé. Soit donc un *tore* balayé par le disque de centre ω, de rayon a, en tournant autour de l'axe $0x$, ne rencontrant pas le disque, comme sur la figure suivante.



La distance de ω à $0x$ étant désignée par d, et θ étant l'angle repérant le point m du cercle indiqué sur la figure, on a :

$$r = d + a\cos\theta, \qquad\qquad x = a\sin\theta,$$

et donc pour le volume du tore, nous avons grâce à la formule qui précède :

$$\begin{aligned}\mu_v(A) &= \pi\, a \int_0^{2\pi} \big(d + a\cos\theta\big)^2 \cos\theta\, d\theta \\ &= \pi\, a \int_0^{2\pi} \Big(d^2\cos\theta + 2\, ad\cos^2\theta + a^2\cos^3\theta\Big)\, d\theta.\end{aligned}$$

En tenant compte des formules élémentaires (exercice : utiliser des formules de trigonométrie) :

$$\begin{aligned}0 &= \int_0^{2\pi} \cos\theta\, d\theta = \int_0^{2\pi} \cos^3\theta\, d\theta, \\ \pi &= \int_0^{2\pi} \cos^2\theta\, d\theta,\end{aligned}$$

on trouve aisément :

$$\mu_v(A) = 2\,\pi^2\, a^2\, d. \qquad \square$$

8. Formule de changement de variables dans les intégrales triples

Soient deux espaces euclidiens distincts $\mathbb{R}^3$ et $\mathbb{R}^3$ munis de deux systèmes de coordonnées (cartésiennes ou curvilignes) :

$$(u,v,w) \in \mathbb{R}^2 \qquad \text{et} \qquad (x,y,z) \in \mathbb{R}^2.$$

On suppose donnés deux compacts à bords quarrables :

$$A \subset \mathbb{R}^3 \qquad \text{et} \qquad B \subset \mathbb{R}^3,$$

ainsi qu'une application entre eux :

$$\begin{aligned} \varphi\colon \qquad A &\xrightarrow{\sim} B \\ (u,v,w) &\longmapsto \big(f(u,v,w),\, g(u,v,w),\, h(u,v,w)\big) =: (x,y,z), \end{aligned}$$

qui satisfait toutes les hypothèses suivantes.

(1) φ est bijective $A \xrightarrow{\sim} B$.

(2) En restriction aux bords, $\varphi\big|_{\partial A}\colon \partial A \xrightarrow{\sim} \partial B$ est aussi bijective.

(3) φ est continue, et la bijection inverse $A \xleftarrow{\sim} B : \varphi^{-1}$ est aussi continue.

(4) φ admet des dérivées partielles $\frac{\partial \varphi}{\partial u}$, $\frac{\partial \varphi}{\partial v}$, $\frac{\partial \varphi}{\partial w}$ qui sont continues sur A.

(5) φ conserve l'orientation.

Terminologie 8.1. On dit que φ est un $\mathscr{C}^1$-*difféomorphisme* entre A et B.

Définition 8.2. La *matrice jacobienne* d'une application de classe $\mathscr{C}^1$:

$$\varphi\colon \quad (u,v,w) \longmapsto \big(f(u,v,w),\, g(u,v,w),\, h(u,v,w)\big)$$

en un point (u,v,w) est la matrice 3×3 :

$$\mathsf{Jac}\,(\varphi)\,(u,v,w) := \begin{pmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} & \frac{\partial f}{\partial w} \\ \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} & \frac{\partial g}{\partial w} \\ \frac{\partial h}{\partial u} & \frac{\partial h}{\partial v} & \frac{\partial h}{\partial w} \end{pmatrix} (u,v,w).$$

Le *déterminant jacobien* de φ est la fonction de (u,v,w) :

$$\begin{aligned} \det \mathsf{Jac}(\varphi) &:= \det \begin{pmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} & \frac{\partial f}{\partial w} \\ \frac{\partial g}{\partial u} & \frac{\partial g}{\partial v} & \frac{\partial g}{\partial w} \\ \frac{\partial h}{\partial u} & \frac{\partial h}{\partial v} & \frac{\partial h}{\partial w} \end{pmatrix} \\ &= \frac{\partial f}{\partial u}\,\frac{\partial g}{\partial v}\,\frac{\partial h}{\partial w} + \frac{\partial f}{\partial w}\,\frac{\partial g}{\partial u}\,\frac{\partial h}{\partial v} + \frac{\partial f}{\partial v}\,\frac{\partial g}{\partial w}\,\frac{\partial h}{\partial u} - \\ &\quad - \frac{\partial f}{\partial w}\,\frac{\partial g}{\partial v}\,\frac{\partial h}{\partial u} - \frac{\partial f}{\partial u}\,\frac{\partial g}{\partial w}\,\frac{\partial h}{\partial v} - \frac{\partial f}{\partial v}\,\frac{\partial g}{\partial u}\,\frac{\partial h}{\partial w}. \end{aligned}$$

Théorème 8.3. *Sous les hypothèses précédentes concernant le $\mathscr{C}^1$-difféomorphisme $\varphi\colon A \xrightarrow{\sim} B$ entre deux compacts élémentaires, ou entre deux réunions finies de compacts élémentaires μ-disjoints, on a :*

$$\mathsf{volume}(B) = \iiint_B 1\,dxdydz = \iiint_A \det \mathsf{Jac}(\varphi)(u,v,w)\; dudvdw. \qquad \square$$

Théorème 8.4. [Formule de changement de variables dans les intégrales triples] *Soit un $\mathscr{C}^1$-difféomorphisme entre deux compacts élémentaires $A \subset \mathbb{R}^3_{u,v,w}$ et $B \subset \mathbb{R}^3_{x,y,z}$:*

$$\begin{aligned} \varphi\colon \qquad A &\xrightarrow{\sim} B \\ (u,v,w) &\longmapsto \big(f(u,v,w),\, g(u,v,w),\, h(u,v,w)\big) =: (x,y,z), \end{aligned}$$

ou plus généralement, entre deux réunions finies de compacts élémentaires μ-disjoints. Alors pour toute fonction continue G sur B, on a :

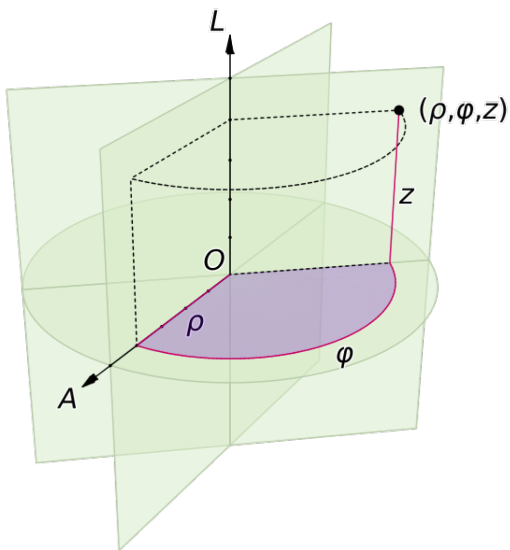
$$\iiint_B G(x,y,z)\,dxdydz \;=\; \iiint_A G\circ\varphi\cdot\det\mathsf{Jac}(\varphi)(u,v,w)\;dudvdw$$

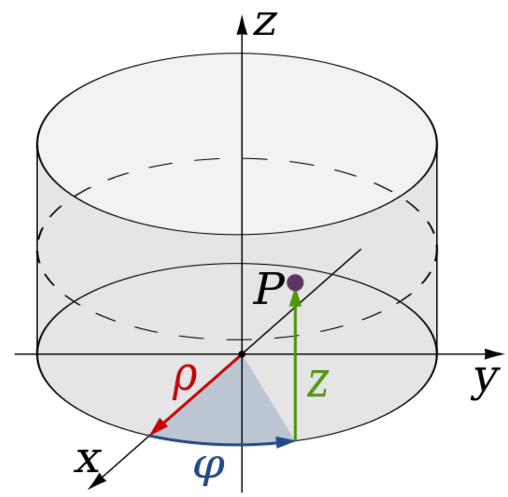
$$= \iiint_A G\big(f(u,v,w),\; g(u,v,w),\; h(u,v,w)\big)\cdot\det\mathsf{Jac}(\varphi)(u,v,w)\;dudvdw. \qquad\square$$

9. Intégrales triples en coordonnées cylindriques

Dans l'espace $\mathbb{R}^3$ muni des coordonnées (x,y,z) et du repère orthonormé $\{0,\mathbf{i},\mathbf{j},\mathbf{k}\}$, on peut repérer la projection $p' := (x,y)$ d'un point $p=(x,y,z)$ par des coordonnées polaires, sans modifier z :

$$\begin{cases} x \;=\; r\cos\theta, \\ y \;=\; r\sin\theta, \\ z \;=\; z. \end{cases}$$





Comme nous le savons, d'un cours à l'autre, les notations peuvent varier, et donc, il faut lire les figures en tenant compte de ce flottement des lettres. Introduisons l'espace euclidien $\mathbb{R}^3$ privé de l'axe vertical $0z$:

$$\mathbb{R}^3_{*z} \;:=\; \big\{(x,y,z)\in\mathbb{R}^3\colon\;(x,y)\neq(0,0)\big\}.$$

En fixant une constante quelconque $\theta_0\in\mathbb{Z}$, on a, comme pour les coordonnées polaires, qui 'oublient' la coordonnée z, l'énoncé suivant pour les coordonnées cylindriques.

Lemme 9.1. *L'application :*

$$\varphi\colon\quad (r,\theta,z)\;\longmapsto\;\big(r\cos\theta,\;r\sin\theta,\;z\big)$$

est une bijection continûment différentiable de $\mathbb{R}^*_+\times[\theta_0,\theta_0+2\pi[\times\mathbb{R}$ sur $\mathbb{R}^3_{*z}$. $\square$

La bijection inverse de φ est donnée en résolvant r et θ en fonction de (x,y) :

$$x^2+y^2 \;=\; \big(r\cos\theta\big)^2+\big(r\sin\theta\big)^2 \;=\; r^2\,\big(\cos^2\theta+\sin^2\theta\big) \;=\; r^2,$$

ce qui donne :

$$r \;=\; \sqrt{x^2+y^2},$$

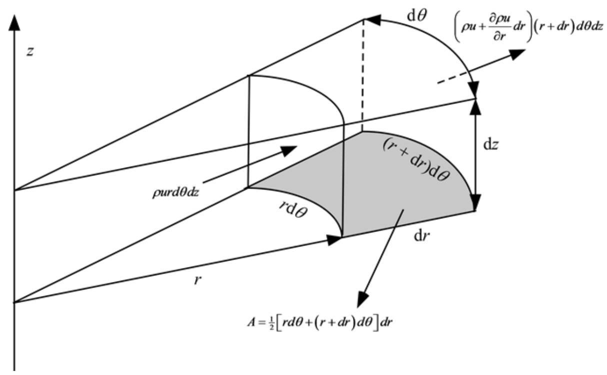
et aussi :

$$\frac{y}{x} \;=\; \frac{r\sin\theta}{r\cos\theta} \;=\; \frac{\sin\theta}{\cos\theta} \;=\; \tan\theta,$$

ce qui donne :

$$\theta \;=\; \operatorname{arctan} \frac{y}{x}.$$

Une dernière figure illustre l'élément de volume infinitésimal en coordonnées cylindriques, qui n'est autre que le produit de dz par l'élément infinitésimal $r\,drd\theta$ des coordonnées polaires.



Soit $f\colon A \longrightarrow \mathbb{R}$ une fonction continue sur un compact quarrable $A \subset \mathbb{R}^3$. Proposons-nous de calculer l'intégrale triple :

$$I \;:=\; \iiint_A f(x,y,z)\,dxdydz.$$

En d'autres termes, il s'agit de calculer I en passant aux coordonnées cylindriques.

En nous référant au Théorème 6.1, considérons la famille :

$$\big(A_z\big)_{\alpha\leqslant z\leqslant\beta}$$

des sections de A par les plans horizontaux perpendiculaires à $0z$. On obtient :

$$I \;=\; \int_\alpha^\beta dz \iint_{A_z} f(x,y,z)\,dxdydz.$$

Calculons alors l'intégrale double ainsi mise en évidence par le changement de variables :

$$\varphi\colon \qquad \begin{cases} x \;=\; r\cos\theta, \\ y \;=\; r\sin\theta, \\ z \;=\; z. \end{cases}$$

Puisque φ laisse z invariant, on a :

$$\varphi^{-1}\big(A_z\big) \;=\; \varphi^{-1}(A)_z.$$

En passant aux coordonnées polaires dans $\mathbb{R}^2_{x,y}$:

$$dxdy \;=\; r\,drd\theta,$$

nous obtenons :

$$\iint_{A_z} f(x,y,z)\,dxdydz \;=\; \iint_{\varphi^{-1}(A_z)} f\big(r\cos\theta,\; r\sin\theta,\; z\big)\; r\,drd\theta,$$

et par suite pour l'intégrale triple :

$$I \;=\; \int_\alpha^\beta dz \iint_{\varphi^{-1}(A)} f\big(r\cos\theta,\; r\sin\theta,\; z\big)\; r\,drd\theta.$$

Théorème 9.2. *Sur un compact cubable $A \subset \mathbb{R}^3$, soit une fonction continue $f\colon A \longrightarrow \mathbb{R}$, et soit le passage aux coordonnées cylindriques :*

$$\varphi\colon \qquad \begin{cases} x \,=\, r \cos\theta, \\ y \,=\, r \sin\theta, \\ z \,=\, z. \end{cases}$$

Alors :

$$\iiint_A f(x,y,z)\,dxdydz \,=\, \iiint_{\varphi^{-1}(A)} f\big(r\cos\theta,\ r\sin\theta,\ z\big)\ r\,drd\theta dz. \qquad \square$$

Nous avons déjà donné plus haut des exemples de calcul d'intégrales triples en coordonnées cylindriques.

Il est utile de constater, eu égard au Théorème 8.4 de changement de variables dans les intégrales triples, que le déterminant jacobien de l'application φ vaut :

$$\det \mathrm{Jac}(\varphi) \,=\, \det \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial z} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial z} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial z} \end{pmatrix} \,=\, \det \begin{pmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \,=\, r,$$

ce qui explique :

$$dxdydz \,=\, r\,drd\theta dz.$$

10. Intégrales triples en coordonnées sphériques

Dans l'espace $\mathbb{R}^3$, soit un repère orthonormé direct $\big\{0, \mathbf{i}, \mathbf{j}, \mathbf{k}\big\}$, où le vecteur unitaire $\mathbf{i} = (1,0,0)$ dirige l'axe $0x$, puis $\mathbf{j} = (0,1,0)$ dirige l'axe $0y$, et enfin $\mathbf{k} = (0,0,1)$ dirige l'axe $0z$.

Comme pour les coordonnées cylindriques — mais il va y avoir des différences importantes —, considérons l'espace privé de l'axe vertical $0z$:

$$\mathbb{R}^3_{*z} \,:=\, \big\{(x,y,z) \in \mathbb{R}^3\colon\ (x,y) \neq (0,0)\big\}.$$

Alors à tout point $\mathbf{m} = (x,y,z) \in \mathbb{R}^3_{*z}$ correspondent trois quantités réelles.

(1) Un nombre positif $r := \sqrt{x^2+y^2+z^2}$, et un seul, $r \in \mathbb{R}^*_+$.

(2) Un *demi*-plan unique de bord $0\mathbf{k}$ contenant $\mathbf{m} = (x,y,z)$ qui coupe le plan horizontal $\mathbb{R}^2_{x,y}$ suivant une droite D_0 d'origine 0 unique, et par conséquent, il existe un unique angle $\theta \in \mathbb{R} \bmod 2\pi\mathbb{Z}$ tel que :

$$\mathrm{Angle}\,\big(\mathbf{i},\ D_0\big) \,\equiv\, \theta \bmod 2\pi.$$

(3) Un angle non orienté unique ψ représentant la position du point $\mathbf{m}$ dans ce demi-plan :

$$\mathrm{Angle}\,\big(0\mathbf{k},\ 0\mathbf{m}\big) \,\equiv\, \psi\ \in]0,\pi[.$$

En résumé, à tout point $\mathbf{m} \in \mathbb{R}^3_{*z}$ correspond un unique triplet $\big(r,\theta,\psi\big)$ appartenant à l'ensemble :

$$\mathbb{R}^*_+ \times \big(\mathbb{R}/2\pi\mathbb{Z}\big) \times]0,\pi[.$$

Réciproquement, nous affirmons qu'à tout triplet $\big(r,\theta,\psi\big)$ de cet ensemble produit correspond un point $\mathbf{p} \in \mathbb{R}^3_{*z}$ et un seul.

En effet, θ définit un unique demi-plan de bord $0\mathbf{k}$: celui qui contient la demi-droite D_0 du plan $\mathbb{R}^2_{x,y}$ telle que $\mathrm{Angle}\big(\mathbf{i}, D_0\big) \,\equiv\, \theta \bmod 2\pi$. Ensuite, le nombre ψ définit une demi-droite Δ_0 unique — voir la figure —, incluse dans ce demi-plan (ouvert) et telle que :

$$\mathrm{Angle}\,\big(0\mathbf{k},\ \Delta_0\big) \,=\, \psi.$$

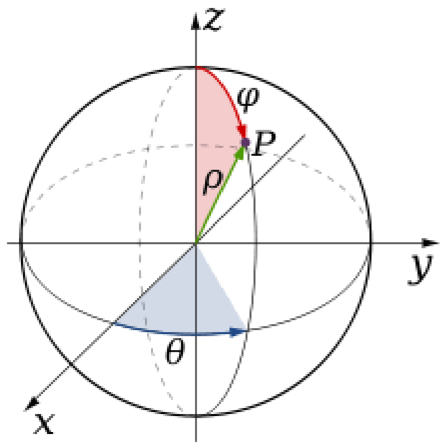
Enfin, il existe un point $\mathbf{m} \in \mathbb{R}^3_{*z}$ et un seul tel que :

$$\mathbf{m} \,\in\, \Delta_0 \qquad\qquad \text{et} \qquad\qquad \|0\mathbf{m}\| \,=\, r.$$

Nous avons donc démontré le

Lemme 10.1. *L'application* $\mathbf{m} \longmapsto \big(r, \theta, \psi\big)$ *établit une bijection de l'espace* $\mathbb{R}^3_{*z}$ *privé de l'axe vertical* $0z$ *sur l'ensemble :*

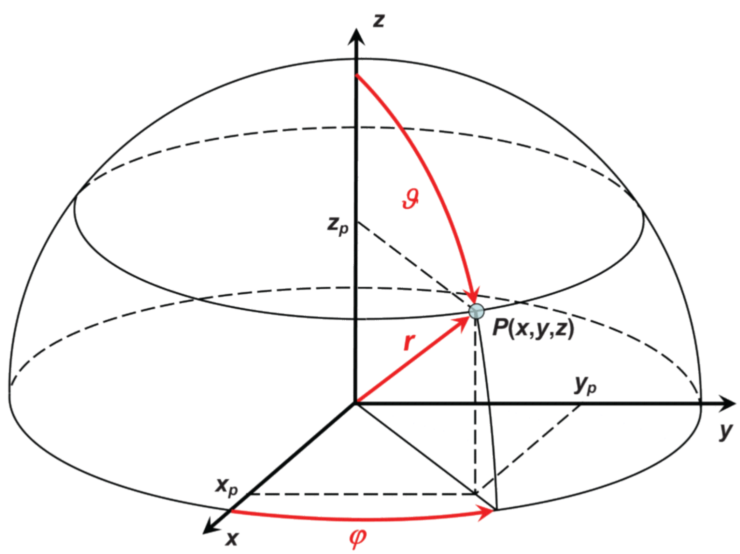
$$\mathbb{R}_+^* \times \big(\mathbb{R}/2\pi\mathbb{Z}\big) \times]0, \pi[. \qquad \square$$



Terminologie 10.2. Le triplet $\big(r, \theta, \psi\big)$ sera appelé *coordonnées sphériques* du point $\mathbf{p} = (x, y, z) \in \mathbb{R}^3$.

Un examen de la figure convainc que les coordonnées cartésiennes s'expriment en fonction des coordonnées sphériques par les formules fondamentales :

$$\begin{cases} x = r \sin\psi \, \cos\theta, \\ y = r \sin\psi \, \sin\theta, \\ z = r \cos\psi. \end{cases}$$



Inversement, si (x, y, z) est donné avec $(x, y) \neq (0, 0)$, on trouve ses coordonnées sphériques $\big(r, \theta, \psi\big)$ comme suit. Tout d'abord :

$$r = \sqrt{x^2 + y^2 + z^2} \in \mathbb{R}_+^*,$$

ensuite :

$$\cos\psi = \frac{z}{r} \qquad \Longleftrightarrow \qquad \psi = \arccos\frac{z}{r},$$

ce qui donne un unique $\psi \in]0, \pi[$, puisque :

$$x^2 + y^2 \neq 0 \qquad\qquad \Longrightarrow \qquad\qquad \frac{z}{r} < 1,$$

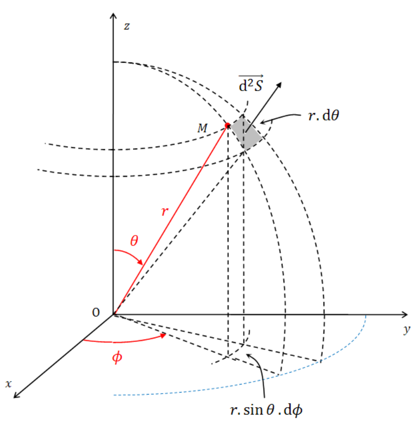
et enfin, les représentations de $x = r \sin\psi \cos\theta$ et de $y = r \sin\psi \sin\theta$ donnent un unique angle $\theta \in \mathbb{R}\big/2\pi\mathbb{Z}$ défini *via :*

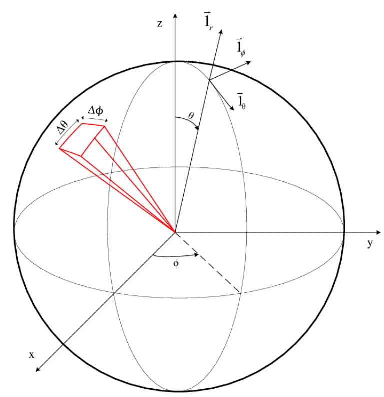
$$\cos\theta = \frac{x}{r \sin\psi} \qquad\qquad \text{et} \qquad\qquad \sin\theta = \frac{y}{r \sin\psi},$$

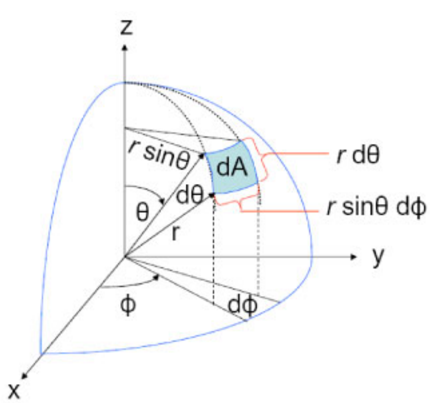
c'est-à-dire :

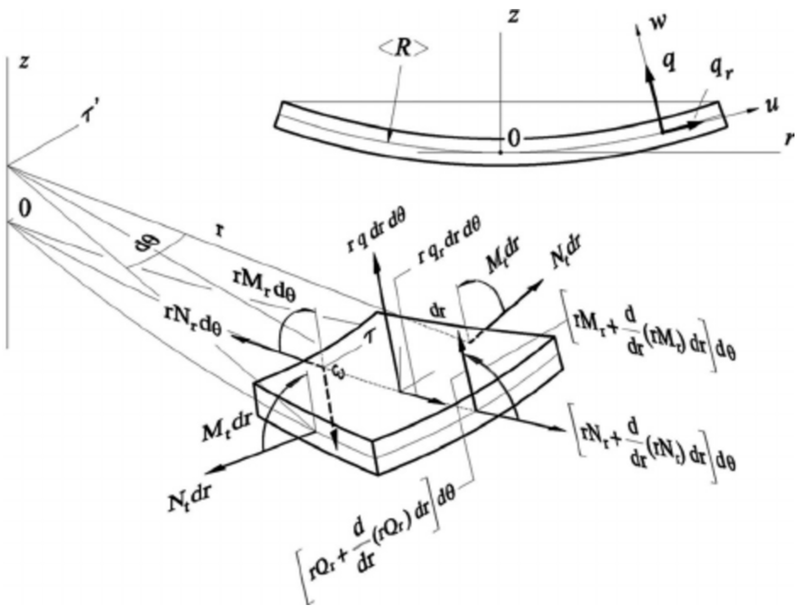
$$\theta = \arctan\frac{y}{x}.$$

Trois figures illustrent l'élément d'aire infinitésimale sur une sphère, suivie d'une figure plus artistique et technique qu'on est invité à contempler.









Soit maintenant f une fonction continue sur un compact cubable $A \subset \mathbb{R}^3$. Proposons-nous de calculer l'intégrale triple :

$$I := \iiint_A f(x,y,z)\, dxdydz,$$

en coordonnées sphériques, *via* le changement de variables :

$$\varphi\colon \qquad \begin{cases} x = \rho \sin\psi \cos\theta, \\ x = \rho \sin\psi \sin\theta, \\ z = \rho \cos\psi. \end{cases}$$

Une astuce très intersidérale consiste à remarquer que φ est la composée de *deux* transformations cylindriques :

$$\varphi_1\colon \quad \begin{cases} x = r\cos\theta, \\ y = r\sin\theta, \\ z = z, \end{cases} \qquad \text{et} \qquad \varphi_2\colon \quad \begin{cases} r = \rho\sin\psi, \\ z = \rho\cos\psi, \\ \theta = \theta, \end{cases}$$

d'où :

$$dxdydz = r\, drd\theta dz \qquad \text{et} \qquad drdzd\theta = \rho\, d\rho d\psi d\theta,$$

selon le schéma :

$$\big(\rho,\theta,\psi\big) \xrightarrow{\varphi_2} \big(r,\theta,z\big) \xrightarrow{\varphi_1} \big(x,y,z\big),$$
$$\varphi = \varphi_1 \circ \varphi_2.$$

D'après le Théorème 9.2, on a :

$$I = \iiint_{\varphi_1^{-1}(A)} f\big(r\cos\theta,\ r\sin\theta,\ z\big)\ r\, drd\theta dz.$$

En appliquant à nouveau à cette intégrale la même formule du Théorème 9.2, et en remarquant que :

$$\varphi_2^{-1} \circ \varphi_1^{-1} = \big(\varphi_1 \circ \varphi_2\big)^{-1} = \varphi^{-1},$$

nous obtenons le

Théorème 10.3. *Sur un compact cubable $A \subset \mathbb{R}^3$, soit une fonction continue $f\colon A \longrightarrow \mathbb{R}$, et soit le passage aux coordonnées sphériques :*

$$\varphi\colon \qquad \begin{cases} x \ = \ \rho \sin\psi \cos\theta, \\ x \ = \ \rho \sin\psi \sin\theta, \\ z \ = \ \rho\cos\psi. \end{cases}$$

Alors :

$$\iiint_A f(x,y,z)\,dxdydz \ = \ \iiint_{\varphi^{-1}(A)} f\big(r\sin\psi\cos\theta,\ r\sin\psi\sin\theta,\ \rho\cos\theta\big)\ \rho^2\sin\psi\,d\rho d\theta d\psi.$$

□

C'est la formule de changement de variables dans les intégrales triples, en coordonnées sphériques.

Il est utile de constater, eu égard au Théorème 8.4 de changement de variables dans les intégrales triples, que le déterminant jacobien de l'application φ vaut :

$$\begin{aligned}
\det \mathrm{Jac}(\varphi) \ &= \ \det \begin{pmatrix} \frac{\partial x}{\partial\rho} & \frac{\partial x}{\partial\theta} & \frac{\partial x}{\partial\psi} \\ \frac{\partial y}{\partial\rho} & \frac{\partial y}{\partial\theta} & \frac{\partial y}{\partial\psi} \\ \frac{\partial z}{\partial\rho} & \frac{\partial z}{\partial\theta} & \frac{\partial z}{\partial\psi} \end{pmatrix} \\
&= \ \det \begin{pmatrix} \sin\psi\cos\theta & -\rho\sin\psi\sin\theta & \rho\cos\psi\cos\theta \\ \sin\psi\sin\theta & \rho\sin\psi\cos\theta & \rho\cos\psi\sin\theta \\ \cos\psi & 0 & -\rho\sin\psi \end{pmatrix} \\
&= \ \cos\psi \begin{vmatrix} -\rho\sin\psi\sin\theta & \rho\cos\psi\cos\theta \\ \rho\sin\psi\cos\theta & \rho\cos\psi\sin\theta \end{vmatrix} - \rho\sin\psi \begin{vmatrix} \sin\psi\cos\theta & -\rho\sin\psi\sin\theta \\ \sin\psi\sin\theta & \rho\sin\psi\cos\theta \end{vmatrix} \\
&= \ \cos\psi\,\rho^2\sin\psi\cos\psi\big(-\sin^2\theta - \cos^2\theta\big) - \rho\sin\psi\,\rho\sin^2\psi\,\big(\cos^2\theta + \sin^2\theta\big) \\
&= \ -\rho^2\cos^2\psi\sin\psi - \rho^2\sin\psi\sin^2\psi \\
&= \ -\rho^2\sin\psi,
\end{aligned}$$

ce qui, en tenant compte du renversement d'orientation, explique :

$$dxdydz \ = \ \rho^2\sin\psi\,d\rho d\theta d\psi.$$

Exemple 10.4. Reprenons le calcul de l'intégrale déjà vue :

$$I \ := \ \iiint_A z^2\,dxdydz,$$

sur la boule fermée :

$$A\colon \qquad x^2+y^2+z^2-a^2 \ \leqslant \ 0.$$

Utilisons à cette fin les coordonnées sphériques. Remarquons que si A' est déduit de A par suppression de l'intersection de A et de $0z$, alors $\varphi^{-1}(A')$ est le pavé défini par :

$$0 \ < \ \rho \ \leqslant \ a, \qquad 0 \ \leqslant \ \theta \ \leqslant \ 2\pi, \qquad 0 \ < \ \psi \ < \ \pi.$$

On est donc ramené au calcul de :

$$I \ = \ \int_0^{2\pi} d\theta \int_0^a \rho^4\,d\rho \int_0^\pi \cos^2\psi\sin\psi\,d\psi,$$

c'est-à-dire au produit de 3 intégrales indépendantes :

$$I \ = \ 2\pi\,\frac{a^5}{5}\,\frac{2}{3}\,\frac{4\,\pi\,a^5}{15}.$$

L'utilisation de coordonnées sphériques a donc permis d'arriver plus rapidement au résultat.

11. Exercices

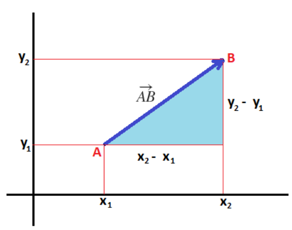
Exercice 1. EE

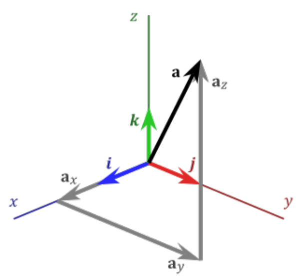
Espaces vectoriels

François DE MARÇAY
Département de Mathématiques d'Orsay
Université Paris-Sud, France

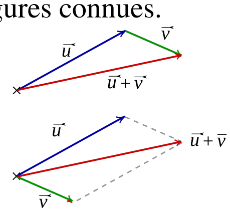
1. Introduction

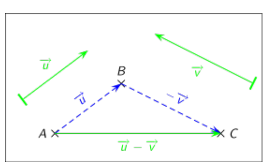
Dans l'enseignement secondaire, on parle de vecteurs dans le plan euclidien $\mathbb{R}^2$ muni de coordonnées (x, y), et aussi dans l'espace euclidien $\mathbb{R}^3$, muni de coordonnées (x, y, z).

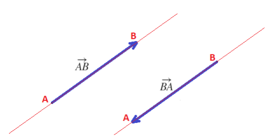


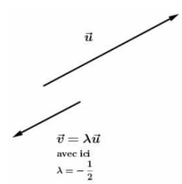


On peut aussi parler abstraitement des vecteurs, sans référence à aucun système de coordonnées. Sommes, soustractions, multiplications par un scalaire peuvent alors être représentées par des figures connues.



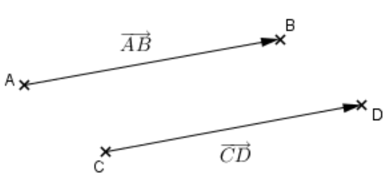


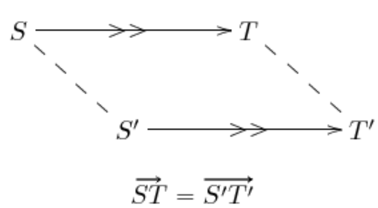


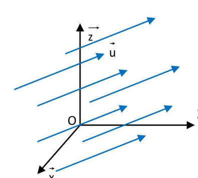


Une notion fondamentale est alors la relation d'*équipollence* entre paires de vecteurs, qui s'illustre d'une manière claire dans le cadre de la géométrie plane euclidienne. En premier lieu, on introduit la notion de *vecteur lié* ou *segment orienté*, caractérisé par :

- une longueur ou norme ;
- une direction ou support,
- un sens ou orientation ;
- une origine ou point d'application.







En comparant des vecteurs liés entre eux, on constate que certains d'entre eux sont « les mêmes », à l'origine près : ils sont parallèles, de même longueur et de même direction. On dit alors que deux vecteurs liés sont *équipollents* s'ils ont même direction, même sens, et même longueur.

L'équipollence est ainsi une *relation d'équivalence* dont les classes peuvent être représentées comme des vecteurs dont l'origine n'est pas fixe : ce sont donc des vecteurs libres, par opposition aux vecteurs liés, dont l'origine est fixe.

La notion de vecteur lié est une notion de géométrie euclidienne, qui n'a de sens que dans un espace affine euclidien. Les vecteurs libres obtenus appartiennent ainsi à un espace vectoriel euclidien.

Plus précisément, dans le plan euclidien $\mathbb{R}^2$, la classe d'un bipoint (A, B) est un vecteur libre dit vecteur généralisé et noté :

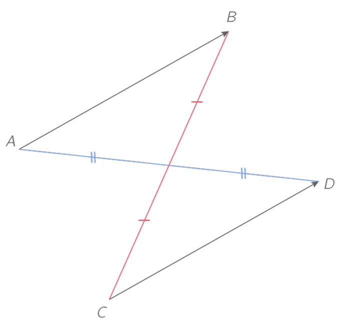
$$\overrightarrow{AB}$$

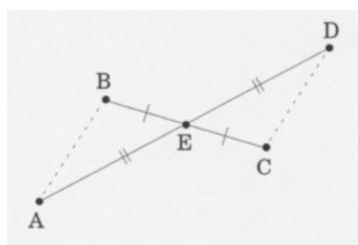
On sait alors démontrer que si deux vecteurs sont équipollents :

$$\overrightarrow{AB} = \overrightarrow{CD},$$

le quadrilatère $ABCD$ est un *parallélogramme*, et on démontre ensuite que cela implique l'équipollence de l'autre paire de côtés opposés :

$$\overrightarrow{AC} = \overrightarrow{BD}.$$





Il est possible de généraliser les propriétés « évidentes » de l'espace euclidien en une définition qui s'applique à un ensemble quelconque non vide, grâce une définition axiomatique abstraite.

Soit donc E un ensemble non vide. On considère dans $E \times E$ une relation d'équivalence notée $\sim$, qui doit satisfaire deux conditions postulées.

Axiome 1. Pour tous points A, B, C dans E, il existe un unique point D dans E tel que le vecteur issu de (C, D) soit égal à celui issu de (A, B) :

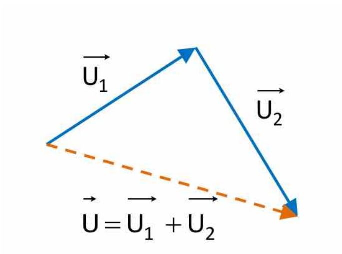
$$\overrightarrow{AB} = \overrightarrow{CD}.$$

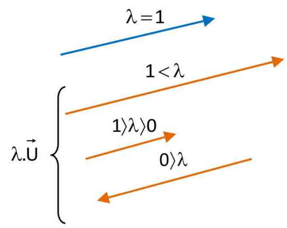
Axiome 1. Pour tous A, B, C dans E, si le vecteur issu de (A, B) est égal à celui issu de (C, D), alors les vecteurs issus des bipoints obtenus en échangeant les extrémités finales sont encore égaux entre eux :

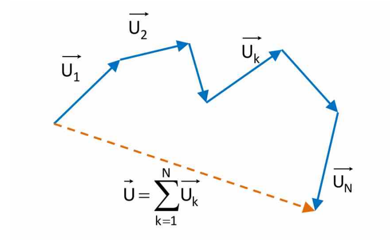
$$\overrightarrow{AB} = \overrightarrow{CD} \qquad \Longrightarrow \qquad \overrightarrow{AC} = \overrightarrow{BD}.$$

Cette dernière formule est parfois appelée *formule du croisement des équipollences*.

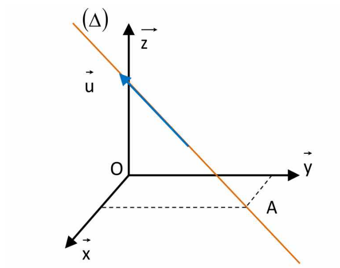
Ensuite, on démontre que les vecteurs sont satisfont les règles usuelles d'addition et de multiplication par un scalaire. L'espace des vecteurs de $\mathbb{R}^3$, ainsi que l'espace des vecteurs-bipoints de E, constitue un *espace de vecteurs*, un type d'espace que nous allons définir axiomatiquement dans ce chapitre en introduisant le concept mathématique d'*espace vectoriel*.







En physique, on introduit la notion de *vecteur glissant*. Un vecteur glissant est défini par sa droite d'action (support), son sens et sa valeur, son point d'application pouvant être quelconque sur la droite d'action.



Par exemple, une force appliquée à un solide indéformable peut glisser sur sa droite d'action sans modifier l'effet qu'elle produit. On dira que la force est représentée par un vecteur glissant.

En contraste, un vecteur lié est défini par sa droite d'action son sens, sa valeur *et* son point d'application.

Par exemle, le poids d'un corps $P \subset \mathbb{R}^3$ dans l'espace est un vecteur lié. C'est un vecteur qui a un point d'application bien défini qui est le barycentre ou le centre de gravité du corps.

On peut remplacer un vecteur libre par un vecteur équipollent quelconque, dans tout l'espace.

On peut glisser un vecteur glissant le long de son support.

Mais un vecteur lié ne peut pas être remplacé par un autre vecteur.

En résumé, on distingue les vecteurs liés, glissants et libres comme suit :

- Un vecteur lié a un point d'application bien défini.
- Un vecteur glissant a un point d'application variant sur une droite qui est son support.
- Un vecteur libre a un point d'application quelconque.

L'histoire de l'algèbre linéaire commence avec Al-Khawarizmi qui a traduit des textes de mathématiques indiens, réinterprété les travaux de l'école grecque et qui est la source du développement conscient de l'algèbre qui s'étendra pendant des siècles après lui. Elle a été reprise par René Descartes qui pose des problèmes de géométrie, comme la détermination de l'intersection de deux droites, en termes d'équation linéaire, établissant dès lors un pont entre deux branches mathématiques jusqu'alors séparées : l'algèbre et la géométrie. S'il ne définit pas la notion de base de l'algèbre linéaire qu'est celle d'espace vectoriel, il l'utilise déjà avec succès, et cette utilisation naturelle des aspects linéaires des équations manipulées demeurera utilisée de manière *ad hoc*, fondée essentiellement sur les idées géométriques sous-jacentes. Après cette découverte, les progrès en algèbre linéaire vont se limiter à des études ponctuelles comme la définition et l'analyse des premières propriétés des déterminants par Jean d'Alembert.

Ce n'est qu'au XIX$^{\text{ième}}$ siècle que l'algèbre linéaire devient une branche des mathématiques à part entière. Carl Friedrich Gauss trouve une méthode générique pour la résolution des systèmes d'équations linéaires et Camille Jordan résout définitivement le problème de la réduction des endomorphismes. En 1843, William Rowan Hamilton (inventeur du terme '*vector*') découvre les quaternions (extension de degré 4 du corps des nombres réels). En 1844, Hermann Grassmann publie son traité *Die lineale Ausdehnungslehre*, *La théorie de l'extension linéaire*, qui est la première tentative de formalisation générale de la notion d'espace vectoriel. Si son œuvre resta grandement inaperçue, elle contenait l'essentiel des idées modernes de l'algèbre linéaire, et cette étape fondamentale dans le développement de l'algèbre linéaire est reconnue comme telle tant par Hamilton que par Giuseppe Peano, qui axiomatise entièrement la théorie en 1888. Les espaces vectoriels deviennent alors une structure générale omniprésente dans presque tous les domaines mathématiques, notamment en analyse (espaces de fonctions).

Sous leur forme la plus simple, les applications linéaires dans les espaces vectoriels représentent intuitivement les déplacements dans les espaces géométriques élémentaires comme la droite, le plan ou notre espace physique. Les bases de cette théorie remplacent maintenant la représentation construite par Euclide au III$^{\text{ième}}$ siècle av. J.-C. La construction moderne permet de généraliser la notion d'espace à des dimensions quelconques.

L'algèbre linéaire permet de résoudre tout un ensemble d'équations dites linéaires utilisées non seulement en mathématiques ou en mécanique, mais aussi dans de nombreuses autres branches comme les sciences naturelles ou les sciences sociales.

Les espaces vectoriels forment aussi un outil fondamental pour les sciences de l'ingénieur et servent de base à de nombreux domaines dans la recherche opérationnelle.

Enfin, c'est un outil utilisé en mathématiques dans des domaines aussi divers que la théorie des groupes, des anneaux ou des corps, l'analyse fonctionnelle, la géométrie différentielle ou la théorie des nombres.

L'algèbre linéaire moderne s'intéresse beaucoup aux espaces de dimension arbitraire, éventuellement infinie. La plupart des résultats obtenus en dimension 2 ou 3 peuvent être étendus aux dimensions finies supérieures. Les vecteurs étant des listes ordonnées à n composantes, on peut manipuler ces données efficacement dans cet environnement. Par exemple en économie, on peut créer et utiliser des vecteurs à huit dimensions pour représenter le produit national brut de huit pays.

Les espaces vectoriels forment le support et le fondement de l'algèbre linéaire. Ils sont aussi présents dans de nombreux domaines distincts. S'il n'est pas possible d'indiquer ici tous les cas d'utilisation, on peut tout de même citer pour les principales structures objet de théories, des exemples significatifs. Leurs rôles dans de vastes théories ne traitant pas d'une structure particulière, comme celles des nombres algébriques ou de Galois peuvent aussi être évoqués.

Les espaces vectoriels utilisés sont d'une grande diversité. On y trouve les classiques espaces vectoriels de dimension 2 ou 3 sur les nombres réels, cependant la dimension peut être quelconque, même infinie. Les nombres complexes sont aussi très utilisés, ainsi que les rationnels. Il n'est pas rare qu'une partie des nombres réels ou complexes soit considérés comme un espace vectoriel rationnel. Le corps de base peut aussi contenir un nombre fini d'éléments, définissant parfois un espace vectoriel fini.

Les propriétés géométriques de la structure permettent la démonstration de nombreux théorèmes. Elles ne se limitent pas aux cas où l'espace est réel, même dans le cas de corps plus insolites comme les corps finis ou les extensions finies des rationnels, les propriétés géométriques s'avèrent parfois essentielles.

En algèbre, on rencontre des êtres mathématiques qui semblent très éloignés des vecteurs, par exemple l'espace $\mathbb{R}[x]$ des polynômes :

$$P(x) \,=\, a_0\, x^d + a_1\, x^{d-1} + \cdots + a_{d-1}\, x + a_d \qquad\qquad (a_0 \neq 0),$$

à coefficients réels $a_i \in \mathbb{R}$. Toutefois, cet espace $\mathbb{R}[x]$ jouit des mêmes propriétés de stabilité par addition et multiplication par des scalaires quelconques $\lambda, \mu \in \mathbb{R}$:

$$\Big(P(x) \,\in\, \mathbb{R}[x] \quad \text{et} \quad Q(x) \,\in\, \mathbb{R}[x]\Big) \qquad \Longrightarrow \qquad \lambda\, P(x) + \mu\, \mathbb{Q}(x) \,\in\, \mathbb{R}[x].$$

Dans ce premier chapitre du Cours d'Algèbre Linéaire, nous nous proposons donc d'étudier d'une manière systématique tout espace abstrait vérifiant les axiomes d'*espace vectoriel*, en un sens que nous allons préciser.

2. Structure d'espace vectoriel sur $\mathbb{R}$ et premières propriétés

Nous travaillerons la plupart du temps avec les nombres réels $\mathbb{R}$. De temps à autre, nous considérerons plus généralement un *corps commutatif* quelconque $\mathbb{K}$, mais nous garderons en tête que seulement deux corps concrets nous intéressent réellement :

$$\mathbb{K} = \mathbb{R} \qquad \text{ou} \qquad \mathbb{K} = \mathbb{C}.$$

En mathématiques, la notion de *corps commutatif* est une des structures algébriques les plus fondamentales de l'algèbre générale. Il s'agit d'un ensemble muni de deux opérations binaires rendant possibles les additions, soustractions, multiplications et divisions. Plus précisément, un corps commutatif est un anneau commutatif dans lequel l'ensemble des éléments non nuls est un groupe commutatif pour la multiplication. Si cela est nécessaire, rappelons la formulation mathématique rigoureuse.

Définition 2.1. Un *corps commutatif* est un ensemble $\mathbb{K}$ muni de deux lois internes, notées en général $+$ et $\times$, vérifiant les conditions suivantes.

(1) $(\mathbb{K}, +)$ forme un groupe abélien — on dit aussi groupe commutatif —, dont l'élément neutre est noté 0.

(2) $(\mathbb{K}\backslash 0, \times)$ forme un groupe abélien multiplicatif, dont l'élément neutre est noté 1.

(3) La multiplication est *distributive*, à gauche comme à droite, pour l'addition c'est-à-dire que :

$$\forall\,(\lambda, \mu, \nu) \in \mathbb{K}^3 \qquad \lambda\times(\mu+\nu) = \lambda\times\mu+\lambda\times\nu \qquad \text{et} \qquad (\lambda+\mu)\times\nu = \lambda\times\nu+\mu\times\nu.$$

On parle alors du *corps commutatif* $(\mathbb{K}, +, \times)$.

En fait, la multiplication $\lambda \cdot \nu$ est souvent notée avec un simple $\cdot$, voire même $\lambda\,\nu$ sans aucun signe.

Les trois exemples canoniques de corps commutatifs élémentaires sont :

- l'ensemble des nombres rationnels $(\mathbb{Q}, +, \times)$;
- l'ensemble des nombres réels $(\mathbb{R}, +, \times)$;
- l'ensemble des complexes $(\mathbb{C}, +, \times)$.

L'Appendice **8** à la fin de ce chapitre expose des rappels de base sur les nombres complexes.

Définition 2.2. Un *sous-corps* d'un corps commutatif $\mathbb{K}$ est une partie $\mathbb{L} \subset K$, stable par $+$ et $\times$, telle que $\mathbb{L}$ munie des lois induites soit un corps.

Évidemment, on a les inclusions de sous-corps :

$$\mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}.$$

Nous pouvons enfin entrer dans le vif du sujet.

Définition 2.3. On dit qu'un ensemble E, dont les éléments seront appelés des *vecteurs*, est un *espace vectoriel* sur $\mathbb{R}$, ou un $\mathbb{R}$-*espace vectoriel*, s'il satisfait les deux axiomes suivants.

Axiome I. E est un groupe commutatif pour une loi additive notée $+$, c'est-à-dire que pour tous vecteurs $\vec{x}, \vec{y}, \vec{z} \in E$, on a :

$$\vec{x} + (\vec{y} + \vec{z}) = (\vec{x} + \vec{y}) + \vec{z}, \qquad \vec{x} + \vec{y} = \vec{y} + \vec{x},$$
$$\vec{x} + (-\,\vec{x}) = \vec{0}, \qquad \vec{x} + \vec{0} = \vec{0} + \vec{x} = \vec{x},$$

où $-\,\vec{x}$ est l'opposé de $\vec{x}$ pour l'addition, et où $\vec{0} \in E$ est le vecteur nul.

Axiome II. E est muni d'une loi de composition externe sur $\mathbb{R}$, noté :

$$\begin{aligned} \mathbb{R} \times E &\longrightarrow E \\ (\lambda, \vec{x}) &\longmapsto \lambda\,\vec{x}, \end{aligned}$$

satisfaisant les propriétés suivantes, pour tous vecteurs $\vec{x}, \vec{y} \in E$, et tous nombres réels $\lambda, \mu \in \mathbb{R}$:

(1) $\lambda\,\big(\vec{x}+\vec{y}\big) = \lambda\,\vec{x} + \lambda\,\vec{y}$;

(2) $\big(\lambda+\mu\big)\,\vec{x} = \lambda\,\vec{x} + \mu\,\vec{y}$;

(3) $\lambda\,\big(\mu\,\vec{x}\big) = (\lambda\,\mu)\,\vec{x}$;

(4) $1\,\vec{x} = \vec{x}$.

Il importe d'exprimer que la caractéristique fondamentale d'un espace vectoriel est la stabilité de ses éléments par combinaisons linéaires quelconques :

$$\forall\,\vec{x}\in E,\ \ \forall\,\vec{y}\in E,\ \ \forall\,\lambda\in\mathbb{R},\ \ \forall\,\mu\in\mathbb{R} \qquad \lambda\,\vec{x}+\mu\,\vec{y}\,\in\,E.$$

Pour les différencier des *vecteurs* $\vec{x} \in E$, les éléments $\lambda \in \mathbb{R}$ seront nommés des *scalaires*. Cette définition se généralise en remplaçant $\mathbb{R}$ par n'importe quel corps commutatif $\mathbb{K}$.

En se basant sur ces axiomes, on développe une théorie que l'on nomme *Algèbre linéaire*, qui est une généralisation, à des espaces abstraits vérifiant ces axiomes, des propriétés que l'on rencontre en géométrie euclidienne.

Assertion 2.4. *Pour tout vecteur $\vec{x} \in E$ et tout scalaire $\lambda \in \mathbb{R}$, les quatre propriétés suivantes sont satisfaites.*

(1) $0\cdot\vec{x} = \vec{0}$.

(2) $\lambda\,\vec{0} = \vec{0}$.

(3) $\lambda\,\vec{x} = \vec{0} \Longleftrightarrow \big(\lambda = 0 \text{ ou } \vec{x} = \vec{0}\big)$.

(4) $(-\lambda)\,\vec{x} = -(\lambda\,x) = \lambda\,(-\vec{x})$.

Preuve. **(1)** En effet :

$$\lambda\,\vec{x} \,=\, \big(\lambda+0\big)\,\vec{x} \,=\, \lambda\,\vec{x}+0\,\vec{x},$$

d'où par soustraction $\vec{0} = 0\,\vec{x}$.

(2) En effet :

$$\lambda\,\vec{x} \,=\, \lambda\,\big(\vec{x}+\vec{0}\big) \,=\, \lambda\,\vec{x}+\lambda\,\vec{0},$$

d'où par soustraction $\vec{0} = \lambda\,\vec{0}$.

(3) Il reste à voir l'implication '$\Longrightarrow$'. En partant de $\vec{0} = \lambda\,\vec{x}$, où on peut supposer $\lambda \neq 0$ grâce à **(1)**, et en multipliant par λ^{-1}, on obtient bien :

$$\vec{0} \,=\, \lambda^{-1}\,\vec{0} \,=\, \lambda^{-1}\big(\lambda\,\vec{x}\big) \,=\, \big(\lambda^{-1}\lambda\big)\,\vec{x} \,=\, \vec{x}.$$

(4) En effet, d'après **(1)** :

$$0 \,=\, \lambda+(-\lambda) \qquad\Longrightarrow\qquad \begin{aligned}\vec{0} &= \big[\lambda+(-\lambda)\big]\,\vec{x}\\ &= \lambda\,\vec{x}+(-\lambda)\,\vec{x},\end{aligned}$$

ce qui donne la première égalité $-\lambda\,\vec{x} = (-\lambda)\,\vec{x}$. Avec $\lambda = 1$, on en déduit en particulier :

$$(-1)\,\vec{x} \,=\, -\,\big(1\,\vec{x}\big) \,=\, -\,\vec{x}.$$

Alors la déuxième égalité va s'en déduire *via :*

$$\lambda\,\big(-\vec{x}\big) \,=\, \lambda\,\big[(-1)\,\vec{x}\big] \,=\, \big(\lambda\,(-1)\big)\,\vec{x} \,=\, (-\lambda)\,\vec{x}. \qquad\square$$

Exemple 2.5. Soit $\mathbb{K}$ un corps commutatif quelconque (on peut s'imaginer qu'il s'agit de $\mathbb{R}$). Munissons l'ensemble $\mathbb{K}\times\mathbb{K}$ des lois suivantes :

$$\begin{aligned}(a,b)+(a',b') \,&=\, \big(a+a',\,b+b'\big),\\ \lambda\,(a,b) \,&=\, \big(\lambda\,a,\,\lambda\,b\big).\end{aligned}$$

La première loi définit sur $\mathbb{K} \times \mathbb{K}$ une structure de groupe commutatif : l'associativité et la commutativité résultent de l'associativité et de la commutativité de l'addition dans $\mathbb{K}$; l'élément neutre est $(0,0)$; l'opposé de (a,b) est $(-a,-b)$.

La seconde loi vérifie les 4 Axiomes II de la Définition 2.3.

$$
\begin{aligned}
&\textbf{(1)} \qquad \lambda\big[(a,b)+(a',b')\big] &&= \lambda\big(a+a',\, b+b'\big) = \big(\lambda\, a + \lambda\, a',\ \lambda\, b + \lambda\, b'\big)\\
& &&= \big(\lambda\, a,\, \lambda\, b\big) + \big(\lambda\, a',\, \lambda\, b'\big) = \lambda\,(a,b) + \lambda\,(a',b'),\\
&\textbf{(2)} \qquad \big(\lambda+\mu\big)\,(a,b) &&= \big[(\lambda+\mu)\,a,\, (\lambda+\mu)\,b\big] = \big(\lambda\, a + \mu\, a,\, \lambda\, b + \mu\, b\big)\\
& &&= \big(\lambda\, a,\, \lambda\, b\big) + \big(\mu\, a,\, \mu\, b\big) = \lambda\,(a,b) + \mu\,(a,b),\\
&\textbf{(3)} \qquad \lambda\big[\mu\,(a,b)\big] &&= \lambda\big(\mu\, a,\, \mu\, b\big) = \big(\lambda\mu\, a,\, \lambda\mu\, b\big) = (\lambda\mu)\,(a,b),\\
&\textbf{(4)} \qquad 1\,(a,b) &&= \big(1\, a,\, 1\, b\big) = (a,b).
\end{aligned}
$$

Par conséquent, $\mathbb{K} \times \mathbb{K}$ constitue un $\mathbb{K}$-espace vectoriel.

Exemple 2.6. Plus généralement, considérons l'ensemble $\mathbb{K}^n$ des n-uplets ordonnés :

$$\vec{a} := \big(a_1, a_2, \dots, a_n\big),$$

de nombres $a_i \in \mathbb{K}$. Si on munit $\mathbb{K}^n$ des lois suivantes :

$$
\begin{aligned}
\big(a_1, a_2, \dots, a_n\big) + \big(b_1, b_2, \dots, b_n\big) &= \big(a_1 + b_1, a_2 + b_2, \dots, a_n + b_n\big),\\
\lambda\big(a_1, a_2, \dots, a_n\big) &= \big(\lambda\, a_1,\, \lambda\, a_2,\, \dots, \lambda\, a_n\big),
\end{aligned}
$$

on vérifie, comme précédemment, que les Axiomes I et II de la Définition 2.3 sont satisfaits. De la sorte, $\mathbb{K}^n$ est un espace vectoriel sur le corps $\mathbb{K}$.

Observons que dans le cas $n = 1$, *tout corps commutatif* $\mathbb{K}$ *est un espace vectoriel sur lui-même*. De plus, pour $n = 1$, les scalaires sont *en même temps* des vecteurs.

Ensuite, si $\mathbb{K}' \subset \mathbb{K}$ est un sous-corps de $\mathbb{K}$, alors tout espace vectoriel E sur $\mathbb{K}$ est aussi un espace vectoriel sur $\mathbb{K}'$, lorsqu'on prend pour loi de composition externe la restriction à $\mathbb{K}' \times \mathbb{E}$ de la loi définie sur $\mathbb{K} \times E$.

C'est ainsi qu'un espace vectoriel sur le corps $\mathbb{C}$ des nombres complexes est un espace vectoriel sur le corps $\mathbb{R}$ des nombres réels, lequel est à son tour un espace vectoriel sur les corps $\mathbb{Q}$ des nombres rationnels, au vu des inclusions de corps :

$$\mathbb{Q} \subset \mathbb{R} \subset \mathbb{R}.$$

3. Sous-espaces vectoriels

Dans l'espace vectoriel $\mathbb{R}^3$, constitué des vecteurs (a,b,c) comme ci-dessus, l'ensemble des vecteurs $(a,b,0)$ dont la troisième coordonnées est nulle est encore stable par les Axiomes I et II. Plus généralement, énonçons une

Définition 3.1. On appelle *sous-espace vectoriel*, ou simplement *sous-espace*, d'un $\mathbb{R}$-espace vectoriel E, tout sous-ensemble $F \subset E$ qui est à lui tout seul un $\mathbb{R}$-espace vectoriel.

En d'autres termes, une partie F d'un $\mathbb{R}$-espace vectoriel E est dite sous-espace lorsque les deux conditions suivantes sont réalisées :

(1) F est un sous-groupe de E pour l'addition ;

(2) la restriction à $\mathbb{R} \times F$ de la loi externe, définie sur $\mathbb{R} \times E$, est une application à valeurs dans F.

Lemme 3.2. *Pour que F soit un sous-espace vectoriel de E, il faut et il suffit que :*

$$\forall\, \lambda,\, \mu \in \mathbb{R} \qquad \big(\vec{x} \in F \quad \text{et} \quad \vec{y} \in F\big) \qquad \Longrightarrow \qquad \lambda\, \vec{x} + \mu\, \vec{y} \in F.$$

Preuve. La condition est évidemment nécessaire. Prouvons qu'elle suffit, c'est-à-dire qu'une telle partie $F \subset E$ est un sous-espace vectoriel. En prenant d'abord $\lambda = 1$ et $\mu = -1$, il vient :

$$\big(\vec{x} \in F \quad \text{et} \quad \vec{y} \in F\big) \qquad \Longrightarrow \qquad \vec{x} - \vec{y} \,\in\, F,$$

ce qui montre que F est un sous-groupe, pour l'addition, de E.

Ensuite, en prenant $\mu = 0$, il vient :

$$\big(\lambda \in \mathbb{R} \quad \text{et} \quad \vec{x} \in F\big) \qquad \Longrightarrow \qquad \lambda\,\vec{x} \,\in\, F.$$

Par conséquent, la loi externe $(\lambda, \vec{x}) \longmapsto \lambda\,\vec{x}$ applique $\mathbb{R} \times F$ dans F, ce qui conclut. □

Exemples 3.3. **(1)** Le singleton $\{\vec{0}\}$ contenant comme unique élément l'élément neutre de l'addition des vecteurs de E est un sous-espace vectoriel de E.

(2) Soit $\vec{x} \in E$. L'ensemble F des vecteurs $\lambda\,\vec{x}$ quand λ parcourt $\mathbb{R}$ est un sous-espace vectoriel de E. Il se réduit à $\{\vec{0}\}$ quand $\vec{x} = \vec{0}$.

(3) Soient deux vecteurs $\vec{x}, \vec{y} \in E$. L'ensemble des vecteurs $\lambda\,\vec{x} + \mu\,\vec{y}$ quand λ et μ parcourent $\mathbb{R}$ est un sous-espace vectoriel de E. Dans un instant, cet exemple va être généralisé aux combinaisons linéaires d'une famille quelconque finie de vecteurs.

(4) Dans l'espace vectoriel V des vecteurs de la géométrie euclidienne dans $\mathbb{R}^3$, toute partie V_D consistant en les vecteurs de même direction qu'une droite D donnée est un sous-espace vectoriel de V. Toute partie V_P consistant en les vecteurs parallèles à un plan donné est un sous-espace vectoriel de V.

Lemme 3.4. *L'intersection $F_1 \cap F_2$ de deux sous-espaces $F_1 \subset E$ et $F_2 \subset E$ est encore un sous-espace de E.*

Preuve. En effet, si $\vec{x}$ et $\vec{y}$ appartiennent à $F_1 \cap F_2$, alors ils appartiennent à la fois à F_1 et à F_2. Par suite, quels que soient les scalaires λ et μ, l'élément $\lambda\vec{x} + \mu\vec{y}$ appartient à la fois à F_1 et à F_2, donc à $F_1 \cap F_2$. □

En géométrie euclidienne, si P et Q sont deux plans sécants le long d'une droite D, alors l'intersection des sous-espaces V_P et V_Q est le sous-espace V_D.

4. Combinaisons linéaires de vecteurs

Une partie finie d'un espace vectoriel E se nomme *famille finie de vecteurs*.

Définition 4.1. Pour toute famille finie $\big\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p\big\}$ d'un nombre $p \geqslant 1$ de vecteurs de E, on appelle *combinaison linéaire* de ces vecteurs tout vecteur $\vec{x} \in E$ qui s'écrit sous la forme :

$$\vec{x} \,=\, \lambda_1\,\vec{x}_1 + \lambda_2\,\vec{x}_2 + \cdots + \lambda_p\,\vec{x}_p,$$

avec des scalaires $\lambda_1, \lambda_2, \dots, \lambda_p \in \mathbb{R}$ quelconques. Ces scalaires sont nommés *coefficients* de la combinaison linéaire.

Exemples 4.2. **(1)** Le vecteur nul $\vec{0}$ est combinaison linéaire de toute famille finie de vecteurs, les coefficients étant tous nuls.

(2) Tout vecteur $\vec{x}$ est combinaison linéaire de toute famille contenant $\vec{x}$, le coefficient de $\vec{x}$ étant égal à 1, tous les autres égaux à 0.

(3) Dans l'espace vectoriel $\mathbb{R}^3$, soient :

$$\begin{aligned}
\vec{e}_1 &:= \big(1, 0, 0\big), \\
\vec{e}_2 &:= \big(0, 1, 0\big), \\
\vec{e}_3 &:= \big(0, 0, 1\big).
\end{aligned}$$

Tout vecteur (a, b, c) de $\mathbb{R}^3$ est combinaison linéaire de la famille $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$, car :

$$\begin{aligned}(a, b, c) &= (a, 0, 0) + (0, b, 0) + (0, 0, c) \\ &= a\,\vec{e}_1 + b\,\vec{e}_2 + c\,\vec{e}_3.\end{aligned}$$

Venons-en maintenant au concept de sous-espace engendré par une famille de vecteurs.

Théorème 4.3. *Soit $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p\}$ une famille de $p \geqslant 1$ vecteurs d'un $\mathbb{R}$-espace vectoriel E. L'ensemble F des combinaisons linéaires de cette famille est un sous-espace vectoriel sur $\mathbb{R}$ de E.*

De plus c'est le plus petit sous-espace contenant la famille donnée.

Démonstration. Partons de deux éléments quelconques de F :

$$\begin{aligned}\vec{x} &= \lambda_1\,\vec{x}_1 + \cdots + \lambda_p\,\vec{x}_p, \\ \vec{y} &= \mu_1\,\vec{x}_1 + \cdots + \mu_p\,\vec{x}_p.\end{aligned}$$

Quels que soient les scalaires γ et δ, on a :

$$\gamma\,\vec{x} + \delta\,\vec{y} = (\gamma\,\lambda_1 + \delta\,\mu_1)\,\vec{x}_1 + \cdots + (\gamma\,\lambda_p + \delta\,\mu_p)\,\vec{x}_p.$$

On obtient une combinaison linéaire de la famille proposée, donc un élément de F, qui est par conséquent un sous-espace vectoriel de E.

Ensuite, F contient clairement chacun des vecteurs $\vec{x}_i$ de la famille : pour chaque rang i, prendre $a_i = 1$, et tous les autres coefficients nuls.

D'autre part, tout sous-espace contenant la famille $\{\vec{x}_1, \dots, \vec{x}_p\}$ doit contenir $\lambda_1\vec{x}_1 + \cdots + \lambda_p\vec{x}_p$ quels que soient les coefficients, et donc doit contenir aussi la somme $\lambda_1\vec{x}_1 + \cdots + \lambda_p\vec{x}_p$. Un tel sous-espace coïncide donc avec F qui est, par conséquent, le plus petit sous-espace contenant $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p\}$. □

Terminologie 4.4. Le sous-espace vectoriel $F \subset E$ est dit *engendré* par la famille $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p\}$, et on note :

$$F = \mathrm{Vect}\,(\vec{x}_1, \dots, \vec{x}_p).$$

Inversement, $\{\vec{x}_1, \dots, \vec{x}_p\}$ est dite une *famille de générateurs* de F.

5. Familles libres et familles liées

Voici un concept important qui signifie qu'on ne peut pas supprimer des vecteurs.

Définition 5.1. On dit qu'une famille $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p\}$ de vecteurs d'un espace vectoriel E est *libre*, ou que les vecteurs $\vec{x}_i$ sont *linéairement indépendants*, si la relation :

$$a_1\,\vec{x}_1 + a_2\,\vec{x}_2 + \cdots + a_p\,\vec{x}_p = \vec{0}$$

entraîne :

$$a_1 = a_2 = \cdots = a_p = 0.$$

Exemples 5.2. **(1)** Dans l'espace vectoriel $\mathbb{R}^4$, les trois vecteurs :

$$\begin{aligned}\vec{x}_1 &:= (1, 0, 0, 0), \\ \vec{x}_2 &:= (0, 1, 0, 0), \\ \vec{x}_3 &:= (0, 0, 1, 0),\end{aligned}$$

constituent une famille libre. En effet :

$$a_1\,\vec{x}_1 + a_2\,\vec{x}_2 + a_3\,\vec{x}_3 = (a_1, a_2, a_3, 0),$$

et :

$$(a_1, a_2, a_3) = \vec{0} \qquad \Longrightarrow \qquad a_1 = a_2 = a_3 = 0.$$

(2) Toujours dans $\mathbb{R}^4$, les trois vecteurs :

$$\begin{aligned}\vec{x}_1 &:= \big(1,0,0,0\big),\\ \vec{x}_2 &:= \big(0,1,0,0\big),\\ \vec{x}_3 &:= \big(1,1,0,0\big),\end{aligned}$$

ne sont *pas* linéairement indépendants. En effet :

$$a_1\,\vec{x}_1 + a_2\,\vec{x}_2 + a_3\,\vec{x}_3 \,=\, \big(a_1+a_3,\; a_2+a_3,\; 0,0\big),$$

et il existe des scalaires non nuls tels que cette combinaison linéaire soit nulle, car il suffit de prendre par exemple :

$$a_1 \,=\, a_2 \,=\, -\,a_3 \,\neq\, 0.$$

Exprimons cela comme suit.

Définition 5.3. Si une famille $\big\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p\big\}$ n'est pas libre, on dit qu'elle est *liée*, ou que les vecteurs $\vec{x}_i$ *linéairement dépendants*, ou encore qu'ils ne sont *pas linéairement indépendants*. Il existe alors des scalaires $a_i \in \mathbb{R}$ non tous nuls avec :

$$\vec{0} \,=\, a_1\,\vec{x}_1 + a_2\,\vec{x}_2 + \cdots + a_p\,\vec{x}_p.$$

Si dans la famille $\big\{\vec{x}_1, \dots, \vec{x}_p\big\}$, il existe un indice k avec $1 \leqslant k \leqslant p$ tel que $\vec{x}_k = \vec{0}$, alors cette famille est liée. En effet, on satisfait à :

$$a_1\,\vec{x}_1 + \cdots + a_k\,\vec{x}_k + \cdots + a_p\,\vec{x}_p \,=\, \vec{0},$$

en prenant tous les $a_i = 0$ nuls, sauf a_k. En résumé, on a une

Observation 5.4. *Une famille est liée dès lors qu'elle contient un vecteur nul.* □

Nous pouvons maintenant prouver un premier résultat technique déterminant pour enclencher ce cours d'algèbre linéaire.

Théorème 5.5. *Soit $F \subset E$ un sous-espace vectoriel d'un $\mathbb{R}$-espace vectoriel E qui est engendré par $p \geqslant 1$ vecteurs. Alors toute famille d'au moins $p+1$ vecteurs de E est toujours nécessairement liée.*

Il importe que ces $p+1$ vecteurs soient tous dans F.

Démonstration. Par hypothèse :

$$F \,=\, \mathrm{Vect}\big(\vec{x}_1, \dots, \vec{x}_p\big),$$

est engendré par une certaine famille $\big\{\vec{x}_1, \dots, \vec{x}_p\big\}$ de vecteurs $\vec{x}_i \in E$.

Il s'agit de prouver que toute famille de $p+1$ vecteurs de F :

$$\big\{\vec{y}_1,\, \vec{y}_2,\, \dots,\, \vec{y}_p,\, \vec{y}_{p+1}\big\} \,\subset\, F$$

est liée. Toute famille de F, dont le cardinal est supérieur à $p+1$, sera alors liée *a fortiori*.

Raisonnons par récurrence sur l'entier p.

• Pour $p = 1$, l'ensemble F est engendré par un unique vecteur $\{\vec{x}_1\}$, et alors :

$$\begin{aligned}\vec{y}_1 \,\in\, F \qquad &\Longrightarrow \qquad \exists\, a_1 \,\in\, \mathbb{R}, \quad \vec{y}_1 \,=\, a_1\,\vec{x}_1,\\ \vec{y}_2 \,\in\, F \qquad &\Longrightarrow \qquad \exists\, a_2 \,\in\, \mathbb{R}, \quad \vec{y}_2 \,=\, a_2\,\vec{x}_2.\end{aligned}$$

On *élimine* $\vec{x}_1$, ce qui donne :

$$a_2\,\vec{y}_1 - a_1\,\vec{y}_2 \,=\, \vec{0}.$$

Si les scalaires a_1 et a_2 ne sont pas simultanément nuls, alors la famille $\{\vec{y}_1, \vec{y}_2\}$ est liée par définition.

Si les deux scalaires sont nuls, c'est que $\vec{y}_1 = \vec{y}_2 = \vec{0}$, et la famille $\{\vec{y}_1, \vec{y}_2\}$ est encore liée.

Par conséquent, le théorème est prouvé pour $p = 1$.

• Ensuite, supposons le théorème vrai pour $p-1$ vecteurs, avec $p \geqslant 2$, et démontrons-le pour p vecteurs. Alors, si on désigne par F' le sous-espace engendré par la famille des $p-1$ premiers vecteurs :

$$F' := \mathsf{Vect}\big\{\vec{x}_1, \ldots, \vec{x}_{p-1}\big\},$$

l'hypothèse de récurrence s'applique à F', et donc, toute famille de p vecteurs de F' est liée.

Maintenant, tout vecteur $\vec{y}$ de F, c'est-à-dire toute combinaison linéaire des vecteurs $\vec{x}_1, \ldots, \vec{x}_{p-1}, \vec{x}_p$, peut se mettre sous la forme :

$$\vec{y} = \vec{z} + a\,\vec{x}_p \qquad \text{avec} \qquad \vec{z} \in F' \quad \text{et} \quad a \in \mathbb{R}.$$

En conséquence, pour tout j avec $1 \leqslant j \leqslant p+1$, il existe :

$$\vec{z}_j \in F' \qquad \text{et} \qquad a_j \in \mathbb{R},$$

tels que, pour tout $1 \leqslant j \leqslant p+1$:

$$\vec{y}_j = \vec{z}_j + a_j\,\vec{x}_p. \tag{5.6}$$

On a ainsi $p+1$ relations. Envisageons alors deux cas.

Premier cas : Tous les $a_j = 0$ sont nuls. Alors :

$$\vec{y}_j = \vec{z}_j \in F' \qquad (\forall\, 1 \leqslant j \leqslant p+1).$$

L'hypothèse de récurrence s'applique à F', et dit que p vecteurs quelconques — *a fortiori* les $p+1$ vecteurs $\vec{y}_j$ — constituent toujours une famille liée. Le théorème est donc prouvé sans aucun effort dans ce cas.

Deuxième cas : Il existe des $a_j \neq 0$ non nuls. En changeant au besoin l'ordre de la numérotation des $\vec{y}_j$, on peut supposer par exemple que $a_{p+1} \neq 0$. Alors, de la dernière des relations (5.6), *i.e.* celle pour $j = p+1$, on tire, puisque a_{p+1} possède un inverse $\frac{1}{a_{p+1}}$ dans $\mathbb{R}$:

$$\vec{x}_p = \tfrac{1}{a_{p+1}}\big(\vec{y}_{p+1} - \vec{z}_{p+1}\big).$$

En portant cela dans les p premières relations (5.6), on *élimine* $\vec{x}_p$:

$$\vec{y}_j = \vec{z}_j + \tfrac{a_j}{a_{p+1}}\big(\vec{y}_{p+1} - \vec{z}_{p+1}\big) \qquad (\forall\, 1 \leqslant j \leqslant p+1),$$

d'où l'on tire :

$$\vec{y}_j - \tfrac{a_j}{a_{p+1}}\,\vec{y}_{p+1} = \vec{z}_j - \tfrac{a_j}{a_{p+1}}\,\vec{z}_{p+1},$$

ce qui prouve que les p vecteurs $\vec{y}_j - \frac{a_j}{a_{p+1}}\,\vec{y}_{p+1}$ des premiers membres appartiennent, comme ceux $\vec{z}_j - \frac{a_j}{a_{p+1}}\,\vec{z}_{p+1}$ des seconds membres, à F'. Par hypothèse de récurrence, ces p vecteurs constituent une famille liée. Il existe par conséquent des scalaires $b_1, \ldots, b_p$ *non tous nuls* tels que :

$$b_1\big(\vec{y}_1 - \tfrac{a_1}{a_{p+1}}\,\vec{y}_{p+1}\big) + \cdots + b_p\big(\vec{y}_p - \tfrac{a_p}{a_{p+1}}\,\vec{y}_{p+1}\big) = \vec{0}.$$

Si l'on pose alors :

$$b_{p+1} := -\tfrac{1}{a_{p+1}}\big(a_1\,b_1 + \cdots + a_p\,b_p\big),$$

on obtient :

$$b_1\,\vec{y}_1 + \cdots + b_p\,\vec{y}_p + b_{p+1}\,\vec{y}_{p+1} = \vec{0},$$

les scalaires $b_1, \ldots, b_p, b_{p+1}$ n'étant pas tous nuls. La famille $\big\{\vec{y}_1, \ldots, \vec{y}_p, \vec{y}_{p+1}\big\}$ est donc liée. Le Théorème 5.5 est démontré. □

Exemple 5.7. Soient, dans un $\mathbb{R}$-espace vectoriel E, une famille $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$. Prenons :

$$\begin{aligned}
\vec{y}_1 &= \vec{x}_1 + \vec{x}_2 + \vec{x}_3,\\
\vec{y}_2 &= \vec{x}_1 - \vec{x}_2 + \vec{x}_3,\\
\vec{y}_3 &= \vec{x}_1 + \vec{x}_2 - \vec{x}_3,\\
\vec{y}_4 &= -\vec{x}_1 + \vec{x}_2 + \vec{x}_3.
\end{aligned}$$

Nous affirmons que quels que soient $\vec{x}_1$, $\vec{x}_2$, $\vec{x}_3$, la famille $\{\vec{y}_1, \vec{y}_2, \vec{y}_3, \vec{y}_4\}$ est liée. Une relation de dépendance s'obtient facilement en ajoutant membre à membre :

$$\vec{y}_1 + \vec{y}_2 + \vec{y}_3 + \vec{y}_3 \;=\; 2\,\vec{x}_1 + 2\,\vec{x}_2 + 2\,\vec{x}_3 \;=\; 2\,\vec{y}_1,$$

d'où :

$$-\,\vec{y}_1 + \vec{y}_2 + \vec{y}_3 + \vec{y}_4 \;=\; \vec{0}.$$

La démonstration du Théorème 5.5 est basée sur l'élimination. Recommençons la démonstration sur cet exemple, en tirant $\vec{x}_3$ de la dernière relation :

$$\vec{x}_3 \;=\; \vec{y}_4 + \vec{x}_1 - \vec{x}_2,$$

et remplaçons dans les trois autres, ce qui donne :

$$\begin{aligned} \vec{y}_1 &= 2\,\vec{x}_1 + \vec{y}_4, \\ \vec{y}_2 &= 2\,\vec{x}_1 - 2\,\vec{x}_2 + \vec{y}_4, \\ \vec{y}_3 &= 2\,\vec{x}_2 - \vec{y}_4. \end{aligned}$$

Si une relation ne contient plus les $\vec{x}_i$ — ce n'est pas le cas ici —, elle offre une relation de dépendance entre les $\vec{y}_j$, et on s'arrête. Sinon, on réitère le procédé : on tire $2\,\vec{x}_2$ de la dernière :

$$2\,\vec{x}_2 \;=\; \vec{y}_3 + \vec{y}_4,$$

et on porte cela dans les précédentes, ce qui donne :

$$\begin{aligned} \vec{y}_1 &= 2\,\vec{x}_1 + \vec{y}_4, \\ \vec{y}_2 &= 2\,\vec{x}_1 - \vec{y}_3. \end{aligned}$$

En réitérant encore une fois le procédé, on élimine $\vec{x}_1$, et on obtient finalement la relation de dépendance cherchée :

$$\vec{y}_1 \;=\; \vec{y}_2 + \vec{y}_3 + \vec{y}_4.$$

Corollaire 5.8. *Pour qu'une famille $\{\vec{x}_1, \dots, \vec{x}_p\}$ soit liée, il faut et il suffit qu'il existe un indice i avec $1 \leqslant i \leqslant p$ tel que $\vec{x}_i$ appartienne au sous-espace engendré par les $\vec{x}_j$ d'indices $j \neq i$.*

Preuve. En effet, s'il existe un tel indice i, on applique le Théorème 5.5, et la famille $\{\vec{x}_1, \dots, \vec{x}_p\}$ est liée.

Réciproquement, supposons que $\{\vec{x}_1, \dots, \vec{x}_p\}$ soit liée. Alors il existe des scalaires $a_1, \dots, a_p \in \mathbb{R}$ non tous nuls tels que :

$$a_1\,\vec{x}_1 + a_2\,\vec{x}_2 + \cdots + a_p\,\vec{x}_p \;=\; \vec{0}.$$

En changeant au besoin l'ordre de la numérotation, on peut supposer que $a_1 \neq 0$. On en déduit alors :

$$\vec{x}_1 \;=\; -\,\tfrac{a_2}{a_1}\,\vec{x}_2 - \cdots - \tfrac{a_p}{a_1}\,\vec{x}_p.$$

Cette relation prouve que $\vec{x}_1$ appartient au sous-espace vectoriel engendré par $\{\vec{x}_2, \dots, \vec{x}_p\}$. □

Par contraposition, on a immédiatement un énoncé équivalent à celui du corollaire qui précède.

Corollaire 5.9. *Pour que la famille $\{\vec{x}_1, \dots, \vec{x}_p\}$ soit libre, il faut et il suffit que, pour tout indice $1 \leqslant i \leqslant p$, le vecteur $\vec{x}_i$ n'appartienne* pas *au sous-espace engendré par tous les $\vec{x}_j$ d'indices $j \neq i$.* □

On a aussi le théorème suivant, qui établit l'*unicité* de la représentation.

Théorème 5.10. *Pour qu'une famille $\{\vec{x}_1, \dots, \vec{x}_p\}$ soit libre, il faut et il suffit qu'à tout vecteur $\vec{x}$ du sous-espace qu'elle engendre corresponde une* unique *famille $\{a_1, \dots, a_p\}$ de scalaires tels que :*

$$\vec{x} \;=\; a_1\,\vec{x}_1 + \cdots + a_p\,\vec{x}_p.$$

Démonstration. Donnons un nom au sous-espace engendré par cette famille $\{\vec{x}_1, \dots, \vec{x}_p\}$:

$$F := \mathsf{Vect}\,(\vec{x}_1, \dots, \vec{x}_p).$$

Supposons qu'à un vecteur $\vec{x} \in F$ correspondent deux familles $\{a_1, \dots, a_p\}$ et $\{b_1, \dots, b_p\}$ de scalaires tels que :

$$\vec{x} = a_1\,\vec{x}_1 + \cdots + a_p\,\vec{x}_p = b_1\,\vec{x}_1 + \cdots + b_p\,\vec{x}_p.$$

Par soustraction, on déduit :

$$(a_1 - b_1)\,\vec{x}_1 + \cdots + (a_p - b_p)\,\vec{x}_p = \vec{0}. \tag{5.11}$$

Démontrons l'implication '$\Longrightarrow$'. Si $\{\vec{x}_1, \dots, \vec{x}_p\}$ est une famille libre, alors cette relation (5.11) entraîne :

$$a_1 - b_1 = \cdots = a_p - b_p = 0,$$

c'est-à-dire :

$$a_1 = b_1, \;\; \dots\dots, \;\; a_p = b_p.$$

Ainsi, à tout $\vec{x}$ du sous-espace F engendré par la famille libre correspond une *unique* famille de scalaires tels que $\vec{x} = a_1\,\vec{x}_1 + \cdots + a_p\,\vec{x}_p$.

Démontrons l'implication inverse '$\Longleftarrow$'. Si $\{\vec{x}_1, \dots, \vec{x}_p\}$ n'est pas libre, la relation (5.11) est vérifiée par certains scalaires $a_i - b_i$ non tous nuls. Au vecteur $\vec{x}$ de F correspondent alors deux familles *distinctes* de scalaires $\{a_1, \dots, a_p\}$ et $\{b_1, \dots, b_p\}$ vérifiant :

$$\vec{x} = a_1\,\vec{x}_1 + \cdots + a_p\,\vec{x}_p = b_1\,\vec{x}_1 + \cdots + b_p\,\vec{x}_p.$$

Le Théorème 5.10 est démontré. □

Exemples 5.12. [en Géométrie] Voici trois illustrations dans l'espace $V = \mathbb{R}^3$ de la géométrie euclidienne spatiale.

(1) Une famille $\{\vec{x}_1, \vec{x}_2\}$ de deux vecteurs est libre si l'un deux, par exemple $\vec{x}_1$, n'appartient pas au sous-espace engendré par l'autre, c'est-à-dire si $\vec{x}_1$ et $\vec{x}_2$ ne sont *pas colinéaires*.

(2) Une famille $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ de trois vecteurs est libre si l'un quelconque d'entre eux n'appartient pas au sous-espace engendré par les deux autres : cette propriété exige que les trois vecteurs ne soient *pas coplanaires*.

(3) Nous affirmons que pour $p \geqslant 4$, une famille $\{\vec{x}_1, \vec{x}_2, \vec{x}_3, \dots, \vec{x}_p\}$ est toujours liée. En effet, supposons au contraire cette famille libre, toujours avec $p \geqslant 4$, et prouvons qu'on aboutit à une contradiction.

L'hypothèse entraîne que $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ est libre, aussi. Or nous savons, en prenant ces trois vecteurs pour directions d'axes de coordonnées dans l'espace, que tout vecteur quelconque $\vec{x} \in V$ est alors combinaison linéaire de ces trois-là. En particulier, $\vec{x}_4$ serait combinaison linéaire de $\vec{x}_1$, $\vec{x}_2$, $\vec{x}_3$, et l'hypothèse serait contredite. Donc quand $p \geqslant 4$, la famille est nécessairement liée.

Dans cet exemple de l'espace vectoriel $V = \mathbb{R}^3$ de la géométrie euclidienne, on voit que toute famille libre contient au plus trois vecteurs. C'est pour cela que $V = \mathbb{R}^3$ est dit en physique *espace vectoriel à 3 dimensions*, la famille libre $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ constituant une *base* de cet espace. Nous allons maintenant définir mathématiquement ces notions pour un espace vectoriel quelconque.

6. Bases, dimensions, coordonnées

Commençons par deux définitions très importantes.

Définition 6.1. On dit qu'un espace vectoriel E est *de dimension finie* s'il existe une famille *finie* de générateurs de E.

Définition 6.2. On appelle *base* d'un espace vectoriel E toute famille libre $\{\vec{e}_1, \dots, \vec{e}_n\}$ de générateurs de E.

Dans la notion de base, les deux conditions sont importantes et indépendantes :

- liberté ;
- générativité.

Si la famille finie $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ engendre l'espace E tout entier, alors, pour tout vecteur $\vec{x} \in E$, il existe au moins une collection $\{a_1, a_2, \ldots, a_n\}$ de scalaires $a_i \in \mathbb{R}$ tels que :

$$\vec{x} = a_1 \vec{e}_1 + a_2 \vec{e}_2 + \cdots + a_n \vec{e}_n.$$

Si, de plus, la famille $\{\vec{e}_1, \ldots, \vec{e}_n\}$ est libre, c'est-à-dire si elle constitue une base de E, alors, pour tout vecteur $\vec{x} \in E$, la famille de scalaires $\{a_1, \ldots, a_n\}$ est de plus *unique*, d'après le Théorème 5.10.

Terminologie 6.3. Les scalaires $a_1, a_2, \ldots, a_n$ sont alors appelés *coordonnées de $\vec{x}$ dans la base* $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$.

Exemples 6.4. Dans l'espace vectoriel $\mathbb{R}^3$, considérons les trois vecteurs :

$$\begin{aligned} \vec{e}_1 &= (1,0,0), \\ \vec{e}_2 &= (0,1,0), \\ \vec{e}_3 &= (0,0,1). \end{aligned}$$

Quel que soit le vecteur $(a,b,c) \in \mathbb{R}^3$, on a :

$$(a,b,c) = a\,\vec{e}_1 + b\,\vec{e}_2 + c\,\vec{e}_3.$$

Par conséquent, $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ est une famille de générateurs de $\mathbb{R}^3$, qui est donc de dimension finie.

De plus, cette famille est libre, car :

$$a\,\vec{e}_1 + b\,\vec{e}_2 + c\,\vec{e}_3 = \vec{0}$$

entraîne :

$$(a,b,c) = \vec{0},$$

ce qui exige :

$$a = b = c = 0.$$

Donc $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ est une base de $\mathbb{R}^3$.

(2) L'espace vectoriel $\mathbb{R}[x]$ des polynômes réels à une indéterminée sur le corps $\mathbb{R}$ n'est *pas* de dimension finie. Montrons en effet que toute famille finie incluse dans $\mathbb{R}[x]$ ne peut pas être famille de générateurs de $\mathbb{R}[x]$. En effet, l'ensemble des degrés des polynômes est une partie finie de $\mathbb{N}$; elle admet donc un plus grand élément m.

Tout polynôme de degré strictement supérieur à m, et il en existe dans $\mathbb{R}[x]$, ne peut s'obtenir comme combinaison linéaire de polynômes de degré au plus égal à m, puisque le degré d'une combinaison linéaire est toujours inférieur ou égal au maximum des degrés de ses composants. Donc $\mathbb{R}[x]$ n'est pas de dimension finie.

(3) Par contre, donnons-nous un nombre naturel $n \geqslant 0$, et considérons la partie de $\mathbb{R}[x]$, notée $\mathbb{R}_n[x]$, constituée des polynômes de degrés au plus égaux à n :

$$\mathbb{R}_n[x] := \{p \in \mathbb{R}[x] \colon \operatorname{deg} p \leqslant n\}.$$

Prouvons d'abord que $\mathbb{R}_n[x]$ est un sous-espace vectoriel de $\mathbb{R}[x]$. En effet, quels que soient les scalaires λ et μ, et les polynômes p et q, on a :

$$\operatorname{deg}(\lambda\, p + \mu\, q) \leqslant \max\{\operatorname{deg} p, \operatorname{deg} q\}.$$

Par suite :

$$\Big(\operatorname{deg} p \leqslant n \qquad \text{et} \qquad \operatorname{deg} q \leqslant n\Big) \qquad \Longrightarrow \qquad \operatorname{deg}(\lambda\, p + \mu\, q) \leqslant n.$$

Donc $\mathbb{R}_n[x]$ est bien un sous-espace vectoriel de $\mathbb{R}[x]$.

Montrons maintenant que la famille :

$$B := \{1, x^1, x^2, x^3, x^4, x^5, \dots, x^n\}$$

des puissances successives $0, 1, 2, 3, 4, 5, \dots, n$ de x constitue une *base* de $\mathbb{R}_n[x]$.

Assertion 6.5. *B est une famille de générateurs de $\mathbb{R}_n[x]$.*

Preuve. En effet, pour tout polynôme $p \in \mathbb{R}_n[x]$, il existe une famille de scalaires $a_0, a_1, \dots, a_n$ tels que :

$$p = a_0 + a_1 x + \cdots + a_n x^n.$$

Donc tout polynôme $p \in \mathbb{R}_n[x]$ est une combinaison linéaire des « vecteurs » de B, qui est par conséquent une famille de générateurs de $\mathbb{R}_n[x]$. Donc $\mathbb{R}_n[x]$ est de dimension finie. □

Assertion 6.6. *B est une famille libre.*

Preuve. En effet, la relation :

$$a_0 + a_1 x + \cdots + a_n x^n = 0$$

veut dire que le polynôme du premier membre est le polynôme nul, donc tous ses coefficients sont nuls :

$$a_0 = a_1 = \cdots = a_n = 0.$$

Par conséquent, B est bien une famille libre, donc constitue une base de $\mathbb{R}_n[x]$. □

Les coefficients $a_0, a_1, \dots, a_n$ du polynôme p sont alors les *coordonnées* de p dans cette base B.

Nous allons maintenant prouver l'existence de bases dans tout espace vectoriel de dimension finie. Remarquons d'abord que, dans tout espace vectoriel $E \neq \{\vec{0}\}$ non réduit au vecteur nul, il existe des familles *libres* et *finies* — par exemple, pour tout $\vec{x} \neq \vec{0}$, la partie $\{\vec{x}\}$ est libre et finie.

Théorème 6.7. [de la base incomplète] *Pour toute famille libre et finie L d'un espace vectoriel $E \neq \{\vec{0}\}$ de dimension finie, il existe une base B finie de E telle que $B \supset L$.*

Démonstration. Puisque E est un espace de dimension finie, il existe une famille G finie de générateurs de E. Si on opère la réunion de G et d'une famille quelconque de générateurs de E, on obtient évidemment encore une famille de générateurs de E.

Soit donc L une famille libre et finie de E. La réunion :

$$F := G \cup L$$

est encore une famille de générateurs de E. De plus, F est finie et :

$$L \subset F.$$

Rappelons que l'on note $\mathscr{P}(F)$ l'ensemble des *parties* de F, c'est-à-dire des *sous-ensembles* de F :

$$\mathscr{P}(F) := \{F' \subset F\}.$$

Rappelons aussi que l'on démontre que :

$$\mathsf{Card}\, \mathscr{P}(F) = 2^{\mathsf{Card}\, F}.$$

Ainsi, puisque $\mathsf{Card}\, F < \infty$, on a aussi la finitude de :

$$\mathsf{Card}\, \mathscr{P}(F) < \infty.$$

Maintenant, introduisons l'ensemble $\mathscr{L}$ de toutes les parties de F qui contiennent L :

$$\mathscr{L} := \{F' \in \mathscr{P}(F) : F' \text{ libre et } F' \supset L\}.$$

Cet ensemble $\mathscr{L}$ n'est pas vide, puisque $L \in \mathscr{L}$. De plus, $\mathsf{Card}\, \mathscr{L} < \infty$ est fini, puisque $\mathsf{Card}\, \mathscr{P}(F) < \infty$, et puisque $\mathscr{L} \subset \mathscr{P}(F)$.

L'ensemble des cardinaux des éléments F' de $\mathscr{L}$ est alors une partie non vide de $\mathbb{N}$ qui admet par conséquent un plus grand élément n.

Soit alors B un élément de $\mathscr{L}$ ayant ce nombre maximal n d'éléments :

$$L \subset B \subset F.$$

Observons que :

$$\mathsf{Card}\, L \leqslant \mathsf{Card}\, B = n \leqslant \mathsf{Card}\, F.$$

À présent, envisageons deux cas.

Cas 1. $n = \mathsf{Card}\, F$ Alors, évidemment, $B = F$. La famille F de générateurs est libre. Elle constitue par conséquent une base de E. Le théorème est prouvé sans travail dans ce cas.

Cas 2. $n \leqslant \mathsf{Card}\, F - 1$. Alors, par définition de l'élément maximum n, pour tout $\vec{x}$ appartenant à $F\backslash B$, la famille $B \cup \{\vec{x}\}$ n'est *pas* libre, puisque c'est une partie de F à $n+1$ éléments. Si on écrit :

$$B = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\},$$

il existe alors des scalaires non tous nuls $a_1, \dots, a_n, b$ tels que :

$$a_1\, \vec{e}_1 + \cdots + a_n\, \vec{e}_n + b\, \vec{x} = \vec{0}.$$

On a nécessairement $b \neq 0$, sinon les vecteurs de B vérifieraient une combinaison linéaire nulle à coefficients non tous nuls, et B ne serait pas libre. Donc, dans le corps $\mathbb{R}$ des scalaires, l'inverse $\frac{1}{b}$ existe, et :

$$\vec{x} = -\tfrac{1}{b}\,\big(a_1\, \vec{e}_1 + \cdots + a_n\, \vec{e}_n\big).$$

Tout vecteur $\vec{x}$ de la famille F des générateurs de E est une combinaison linéaire des vecteurs de B. Donc B est elle-même une famille de générateurs de E, et par conséquent, une base de E.

Le Théorème 6.7 est démontré. □

Ce Théorème 6.7 exprime deux propriétés importantes.

Corollaire 6.8. [Existence de bases] *Dans tout espace vectoriel de dimension finie, il existe des bases.* □

Corollaire 6.9. [Complétion d'une famille libre pour obtenir une base] *Dans tout espace vectoriel de dimension finie, toute famille libre peut être complétée en une base.* □

Ceci veut dire que L peut être complétée par les vecteurs de $B\backslash L$ pour obtenir la base B de E.

Théorème 6.10. *Si, dans un espace vectoriel E, une base B possède n éléments, alors toute famille L libre dans E vérifie :*

$$\mathsf{Card}\, L \leqslant n.$$

Preuve. Appliquons le Théorème 5.5 : puisque B engendre E, une famille de $n+1$ (ou plus) vecteurs quelconques de E est liée. Il ne peut donc pas exister dans E de famille *libre* L vérifiant $\mathsf{Card}\, L \geqslant n+1$. □

Théorème 6.11. *Dans un espace vectoriel de dimension finie, toutes les bases ont le même cardinal.*

Preuve. En effet, si B et B' sont deux bases de E, alors en appliquant le Théorème 6.10 à la base B et à la famille libre B', on obtient :

$$\mathsf{Card}\, B' \leqslant \mathsf{Card}\, B.$$

En échangeant B et B' qui jouent le même rôle, il vient :

$$\mathsf{Card}\, B \leqslant \mathsf{Card}\, B'.$$

Par conséquent :

$$\mathsf{Card}\, B = \mathsf{Card}\, B'.$$ □

Maintenant, recommençons à travailler avec un corps commutatif $\mathbb{K}$ quelconque, au lieu de $\mathbb{R}$. Les nombres de $\mathbb{K}$ se comportent exactement de la même manière que les nombres réels, vis-à-vis de l'addition et de la multiplication. Le lecteur qui éprouverait des réticences à penser en cette généralité est invité à remplacer mentalement $\mathbb{K}$ par $\mathbb{R}$ jusqu'à la fin de ce chapitre. Il va de soi que tous les résultats qui précèdent sont valables, avec la même démonstration, pour un corps $\mathbb{K}$ quelconque.

Puisque toutes les bases de E ont même cardinal, la définition suivante est justifiée.

Définition 6.12. [de la dimension] Soit E un $\mathbb{K}$-espace vectoriel de dimension finie. On appelle *dimension* de E le cardinal de toute base de E.

La dimension de E sur $\mathbb{K}$ sera notée :

$$\dim_{\mathbb{K}} E.$$

Lorsqu'aucune confusion n'est à craindre, notamment lorsque $\mathbb{K} = \mathbb{R}$, on notera simplement :

$$\dim E.$$

On convient que :

$$\dim \{\vec{0}\} \,=\, 0.$$

Exemple 6.13. $\mathbb{R}^n$ et $\mathbb{K}^n$ sont de dimension n sur le corps $\mathbb{R}$ et sur le corps $\mathbb{K}$:

$$\dim_{\mathbb{R}} \mathbb{R}^n \,=\, n \qquad\qquad \text{et} \qquad\qquad \dim_{\mathbb{K}} \mathbb{K}^n \,=\, n.$$

Soit E un espace vectoriel sur le corps $\mathbb{C}$ des nombres complexes, de dimension $n \geqslant 1$, et soit :

$$B \,=\, \big\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\big\}$$

une base de E.

Puisque $\mathbb{R} \subset \mathbb{C}$, on peut voir E aussi comme espace vectoriel sur le corps $\mathbb{R}$ des nombres réels. Rappelons que l'on note :

$$i \,:=\, \sqrt{-1}.$$

Assertion 6.14. *L'ensemble :*

$$B' \,:=\, \big\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n,\, i\,\vec{e}_1,\, i\,\vec{e}_2,\, \dots,\, i\,\vec{e}_n\big\}$$

constitue une base de E sur $\mathbb{R}$. De plus :

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

Démonstration. En effet, pour tout vecteur $\vec{x}$ de E, il existe une famille de scalaires *complexes* $\alpha_1, \dots, \alpha_n$ tels que :

$$\vec{x} \,=\, \alpha_1\,\vec{e}_1 + \cdots + \alpha_n\,\vec{e}_n. \tag{6.15}$$

Ces α_i sont les *coordonnées* de $\vec{x}$ dans la base B.

Pour tout indice k avec $1 \leqslant k \leqslant n$, faisons intervenir les parties réelle et imaginaire de α_k :

$$\alpha_k \,=\, a_k + i\,b_k \qquad\qquad (a_k \in \mathbb{R},\, b_k \in \mathbb{R}).$$

La relation (6.15) s'écrit alors :

$$\vec{x} \,=\, a_1\,\vec{e}_1 + \cdots + a_n\,\vec{e}_n + b_1\,\big(i\,\vec{e}_1\big) + \cdots + b_n\,\big(i\,\vec{e}_n\big). \tag{6.16}$$

Ainsi, pour tout $\vec{x} \in E$, il existe une famille de scalaires réels $a_1, \dots, a_n$, $b_1, \dots, b_n$ vérifiant la relation (6.16). Par conséquent, B' est une famille de générateurs de E sur $\mathbb{R}$.

De plus, puisque B est une base de E sur $\mathbb{C}$, à tout $\vec{x}$ de E correspond une *unique* famille $\alpha_1, \dots, \alpha_n$ de scalaires de $\mathbb{C}$ vérifiant (6.15). Comme à tout $\alpha \in \mathbb{C}$ correspond un couple unique $(a, b) \in \mathbb{R} \times \mathbb{R}$ tel que $\alpha = a + i\,b$, alors la famille $a_1, \dots, a_n$, $b_1, \dots, b_n$ de scalaires de $\mathbb{R}$ correspondant à $\vec{x}$ est unique aussi. Par conséquent, d'après le Théorème 5.10, B' est une famille libre de E sur $\mathbb{R}$, donc une base de ce espace.

Enfin, comme $\mathsf{Card}\, B' = 2\, \mathsf{Card}\, B$, on a bien :

$$\mathsf{dim}_{\mathbb{R}}\, E \,=\, 2\, \mathsf{dim}_{\mathbb{C}}\, E. \qquad \square$$

Théorème 6.17. *Tout sous-espace F d'un $\mathbb{K}$-espace vectoriel E de dimension finie est lui-même de dimension finie, avec :*

$$\mathsf{dim}_{\mathbb{K}}\, F \,\leqslant\, \mathsf{dim}_{\mathbb{K}}\, E.$$

Démonstration. Soit donc E un $\mathbb{K}$-espace vectoriel, et soit $F \subset E$ un $\mathbb{K}$-sous-espace vectoriel de E.

Le théorème est évident lorsque $F = \{\vec{0}\}$.

Supposons donc que $F \neq \{\vec{0}\}$, et introduisons l'ensemble $\mathscr{A}$ des cardinaux de toutes les familles *libres* dans F.

Cet ensemble $\mathscr{A}$ n'est pas vide, car pour tout vecteur de F non nul $\vec{x} \in F \backslash \{\vec{0}\}$, la famille $\{\vec{x}\}$ est libre, donc $1 \in \mathscr{A}$.

D'autre part, toute famille libre incluse dans F est une famille libre incluse dans E, puisque $F \subset E$. Alors d'après le Théorème 6.10, $\mathscr{A}$ est majoré par $n = \mathsf{dim}\, E$.

Comme $\mathscr{A}$ est un ensemble non vide de $\mathbb{N}$ et majoré, il admet un élément maximum :

$$p \,:=\, \mathsf{max}\, A \qquad\qquad \text{avec} \qquad\qquad p \,\leqslant\, n.$$

Cela veut dire qu'une famille quelconque d'au moins $p+1$ vecteurs de F est liée.

Soit B' une famille libre incluse dans F et ayant ce nombre maximum p d'éléments. Désignons par F' le sous-espace engendré par B'. Comme B' est finie, F' est de dimension finie, et comme B' est libre, alors $\mathsf{dim}\, F' = \mathsf{Card}\, B' = p$, ce qui entraîne :

$$\mathsf{dim}\, F' \,\leqslant\, \mathsf{dim}\, E.$$

Pour finir, nous allons prouver que $F' = F$, ce qui établira complètement le Théorème 6.17.

Comme F' est le plus petit sous-espace contenant B' d'après le Théorème 4.3, on a :

$$F' \,\subset\, \mathsf{Vect}\, B' \,\subset\, F.$$

D'autre part, pour tout $\vec{x} \in F \backslash B'$, la famille $B' \cup \{\vec{x}\}$, de cardinal $p+1$, est nécessairement liée, donc le Corollaire 5.8 assure que $\vec{x}$ appartient au sous-espace F' engendré par B'. Par suite :

$$F \,\subset\, F',$$

et en définitive :

$$F \,=\, F'. \qquad \square$$

7. Somme de deux sous-espaces vectoriels

Soient F et G deux sous-espaces vectoriels d'un espace vectoriel E. Étudions l'ensemble H des $\vec{z} \in E$ tels qu'il existe un $\vec{x} \in F$ et un $\vec{y} \in G$ de façon que :

$$\vec{z} \,=\, \vec{x} + \vec{y}.$$

En d'autres termes, étudions :

$$H \,:=\, \big\{\vec{z} \in E \colon\, \exists\, \vec{x} \in F,\; \exists\, \vec{y} \in G,\;\; \vec{z} = \vec{x} + \vec{y}\big\}.$$

Assertion 7.1. *H est un sous-espace vectoriel de E, c'est-à-dire que, quels que soient les scalaires a et a' on a :*

$$\big(\vec{z} \in H \qquad \textit{et} \qquad \vec{z}\,' \in H\big) \qquad\qquad \Longrightarrow \qquad\qquad a\, \vec{z} + a'\, \vec{z}\,' \,\in\, H$$

Preuve. Par hypothèse, il existe $\vec{x}$ et $\vec{x}\,'$ dans F et il existe $\vec{y}$ et $\vec{y}\,'$ dans G tels que :

$$\vec{z} \,=\, \vec{x} + \vec{y} \qquad \text{et} \qquad \vec{z}\,' \,=\, \vec{x}\,' + \vec{y}\,'.$$

Or :

$$a\,\vec{z} + a'\,\vec{z}\,' \,=\, a\,\vec{x} + a'\,\vec{x}\,' + b\,\vec{y} + b'\,\vec{y}\,'.$$

Comme F et G sont deux sous-espaces de E :

$$\begin{aligned} \big(\vec{x}\in F \quad \text{et} \quad \vec{x}\,'\in F\big) \quad &\Longrightarrow \quad a\,\vec{x} + a'\,\vec{x}\,' \,\in\, F,\\ \big(\vec{y}\in G \quad \text{et} \quad \vec{y}\,'\in G\big) \quad &\Longrightarrow \quad a\,\vec{y} + a'\,\vec{y}\,' \,\in\, G. \end{aligned}$$

Par conséquent, par définition de H :

$$\big(a\,\vec{x} + a'\,\vec{x}\,'\big) + \big(a\,\vec{y} + a'\vec{y}\,'\big) \,\in\, H,$$

ce qui montre bien que H est un sous-espace vectoriel de E. □

Assertion 7.2. *H est le plus petit sous-espace vectoriel contenant $F \cup G$.*

Preuve. En prenant $\vec{y} = \vec{0}$, on voit que $H \supset F$. En prenant $\vec{x} = \vec{0}$, on voit que $H \supset G$. Par conséquent :

$$H \,\supset\, F \cup G.$$

D'autre part, tout sous-espace contenant $F \cup G$ doit contenir tout $\vec{x}$ de F, tout $\vec{y}$ de G, puis la somme $\vec{x} + \vec{y}$. Donc tout sous-espace contenant $F \cup G$ contient aussi H.

En conclusion, H est le plus petit sous-espace contenant $F \cup G$. □

Définition 7.3. Le plus petit sous-espace contenant la réunion $F \cup G$ de deux sous-espaces F et G d'un espace vectoriel E est appelé *somme* de F et de G, et se note :

$$F + G \,:=\, \big\{\vec{z}\in E\colon\, \exists\,\vec{x}\in F,\ \exists\,\vec{y}\in G,\ \vec{z} = \vec{x} + \vec{y}\big\}.$$

On démontre aisément (exercice) que cette addition entre sous-espaces vectoriels est associative, commutative, et admet comme élément neutre l'espace vectoriel nul $\{\vec{0}\}$. De plus :

$$F \,\subset\, G \qquad \Longleftrightarrow \qquad F + G \,=\, G.$$

Introduisons maintenant le conept de *somme directe* de deux sous-espaces vectoriels.

Considérons un élément quelconque $\vec{z}$ de la somme $F + G$ définie à l'instant. Alors il existe $\vec{x} \in F$ et $\vec{y} \in G$ tels que :

$$\vec{z} \,=\, \vec{x} + \vec{y}.$$

En d'autres termes, $\vec{z}$ est l'image du couple $(\vec{x}, \vec{y})$ par l'application :

$$\begin{aligned} f\colon \quad F \times G \,&\longrightarrow\, E\\ \big(\vec{x}, \vec{y}\big) \,&\longmapsto\, \vec{x} + \vec{y}. \end{aligned}$$

L'image de $F \times G$ par f vient précisément d'être notée $F + G$.

Cette application f n'est pas nécessairement injective, ce qui veut dire qu'un vecteur $\vec{z}$ de $F+G$ peut être l'image de deux couples distincts $(\vec{x}, \vec{y})$ et $\big(\vec{x}\,', \vec{y}\,'\big)$. Or de :

$$\vec{z} \,=\, \vec{x} + \vec{y} \,=\, \vec{x}\,' + \vec{y}\,',$$

on tire :

$$\vec{x} - \vec{x}\,' \,=\, \vec{y}\,' - \vec{y}, \tag{7.4}$$

avec :

$$\vec{x} - \vec{x}\,' \,\in\, F \qquad \text{et} \qquad \vec{y}\,' - \vec{y} \,\in\, G.$$

La coïncidence de ces deux éléments exige leur appartenance à $F \cap G$. Remarquons que deux sous-espaces ont toujours $\vec{0}$ en commun, et envisageons deux cas.

Premier cas : $F \cap G \neq \{\vec{0}\}$. Alors si $\vec{u}$ est un élément de cette intersection, on a :

$$\begin{aligned} \vec{x} \in F \qquad &\Longrightarrow \qquad \vec{x} + \vec{u} \in F, \\ \vec{y} \in G \qquad &\Longrightarrow \qquad \vec{y} - \vec{u} \in G, \end{aligned}$$

et :

$$\vec{z} = \vec{x} + \vec{y} = \big(\vec{x} + \vec{u}\big) + \big(\vec{y} - \vec{u}\big).$$

Ainsi, $\vec{z}$ est l'image par f de deux éléments distincts $(\vec{x}, \vec{y})$ et $\big(\vec{x} + \vec{u},\, \vec{y} - \vec{u}\big)$ de $F \times G$.

L'appplication f de $F \times G$ sur $F + G$ n'est donc *pas* injective.

Deuxième cas : $F \cap G = \{\vec{0}\}$. Alors l'égalité (7.4) ne peut être réalisée que si :

$$\vec{x} - \vec{x}\,' = \vec{0} \qquad\qquad \text{et} \qquad\qquad \vec{y}\,' - \vec{y} = \vec{0},$$

d'où :

$$\big(\vec{x}, \vec{y}\big) = \big(\vec{x}\,', \, \vec{y}\,'\big).$$

L'application est donc *injective :* tout $\vec{z}$ de $F + G$ admet alors une *décomposition unique* $\vec{x} + \vec{y}$ avec $\vec{x} \in F$ et $\vec{y} \in G$.

Ces considérations motivent les définitions suivantes.

Définition 7.5. Deux sous-espaces vectoriels F et G d'un espace vectoriel E sont dits *linéairement indépendants* lorsque :

$$F \cap G = \{\vec{0}\}.$$

Définition 7.6. La somme $F+G$ de deux sous-espaces F et G linéairement indépendants se nomme *somme directe* de F et de G, et se note :

$$F \oplus G.$$

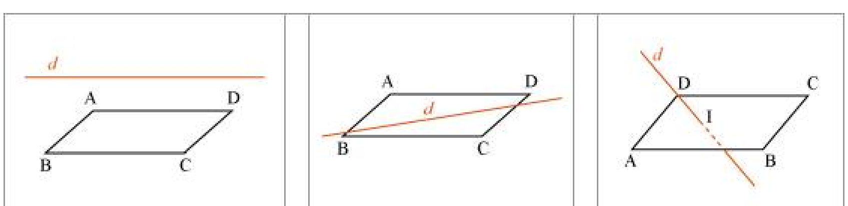
Les raisonnements que nous venons de conduire donnent alors :

$$F \cap G = \{\vec{0}\} \qquad \Longrightarrow \qquad F + G = F \oplus G.$$

Pour terminer ce chapitre, présentons la notion de *sous-espaces supplémentaires.*

Définition 7.7. Si F et G sont linéairement indépendants, et si $F \oplus G = H$, alors F et G sont dits *supplémentaires* dans H.

Par exemple, considérons, en géométrie euclidienne dans l'espace $V = \mathbb{R}^3$, une droite D et un plan P. Soit V_D l'espace des vecteurs de $\mathbb{R}^3$ de même direction que la droite D. Soit aussi V_P l'espace des vecteurs parallèles au plan P. Deux cas peuvent se produire.



La droite d est strictement parallèle au plan (ABC).	La droite d est contenue dans le plan (ABC).	Le plan (ABC) et la droite d sont sécants en I.

Premier cas : La droite $D \| P$ est parallèle au plan. Alors $V_D \subset V_P$, et par conséquent :

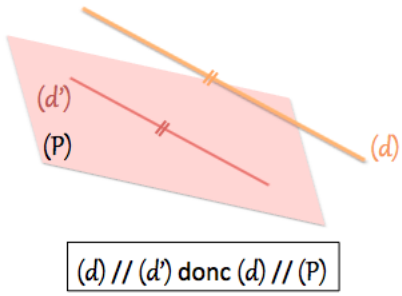
$$V_D + V_P = V_P.$$

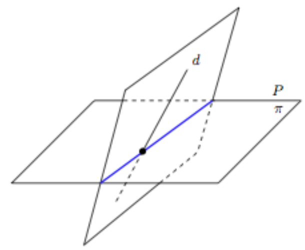
Les espaces V_D et V_P ne sont *pas* linéairement indépendants, et on ne peut donc *pas* parler ici de somme directe $V_D \oplus V_P$.

D'ailleurs, on peut voir directement que l'application $(\vec{x}, \vec{y}) \longmapsto \vec{x} + \vec{y}$ de $V_D \times V_P$ dans E n'est pas injective. En effet, pour tout $\vec{x}\,' \in V_D$, on a :

$$\vec{z} \,=\, \vec{x} + \vec{y} \,=\, \big(\vec{x} + \vec{x}\,'\big) + \big(\vec{y} - \vec{x}\,'\big).$$

On obtient là deux décompositions distinctes de $\vec{z}$.





Deuxième cas : La droite $D \not\parallel P$ n'est pas parallèle au plan. Alors :

$$V_D \cap V_P \,=\, \{\vec{0}\}.$$

Les deux sous-espaces V_D et V_P sont linéairement indépendants. Leur somme se nomme somme directe :

$$V_D + V_P \,=\, V_D \oplus V_P.$$

Le résultat de cette somme est ici l'espace vectoriel $V = \mathbb{R}^3$ tout entier : pour tout vecteur $\vec{z} \in V$, il existe un unique $\vec{x} \in V_D$ et un unique $\vec{y} \in V_P$ tels que $\vec{z} = \vec{x} + \vec{y}$.

L'application $(\vec{x}, \vec{y}) \longmapsto \vec{x} + \vec{y}$ de $V_D \times V_P$ dans V est ici *bijective*. Ainsi, V_D et V_P sont supplémentaires dans V.

D'une façon générale, de l'étude qui précède, on déduit la propriété suivante.

Théorème 7.8. *Si, dans un espace vectoriel E, deux sous-espaces F et G sont supplémentaires, alors à tout $\vec{z} \in E$ correspond un unique couple $(\vec{x}, \vec{y})$ de $F \times G$ tel que :*

$$\vec{z} \,=\, \vec{x} + \vec{y}. \qquad \square$$

En d'autres termes :

$$F \oplus G \,=\, E$$

veut dire :

$$\forall\, \vec{z} \in E \qquad \exists\,!\, x \in F \qquad \exists\,!\, y \in G$$

tels que :

$$\vec{z} \,=\, \vec{x} + \vec{y}.$$

Le point d'exclamation $\exists\,!$ derrière le symbole $\exists$ signifie « *il existe un unique* » ceci ou cela.

8. Appendice : Nombres complexes et similitudes complexes

Un *nombre complexe* prend la forme :

$$z = x + i\,y,$$

où x et y sont deux nombres réels et où i est un nombre (abstrait) satisfaisant :

$$i^2 = -1,$$

que les mathématiciens asiatiques préfèrent noter systématiquement :

$$\sqrt{-1},$$

mais il s'agit d'un autre continent, où l'on ne craint aucun caractère. Classiquement, on note :

$$\boxed{\mathbb{C}}$$

l'ensemble de ces nombres.

On appelle x la *partie réelle* de z et y la *partie imaginaire* de z :

$$x = \mathrm{Re}\, z,$$
$$y = \mathrm{Im}\, z.$$

Les nombres réels sont donc précisément les nombres complexes dont la partie imaginaire est nulle, et les nombres *purement imaginaires :*

$$\{i\,y\colon\ y \in \mathbb{R}\},$$

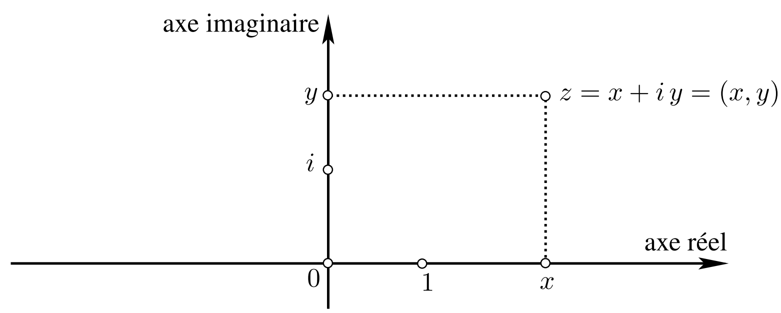
ceux dont la partie réelle est nulle.

Grâce à Argand et à Gauss, on peut visualiser les nombres complexes dans le plan euclidien usuel $\mathbb{R}^2$ en identifiant :

$$\mathbb{C} \ni x + i\,y \longleftrightarrow (x, y) \in \mathbb{R}^2.$$

Par exemple, $0 \in \mathbb{C}$ correspond à l'origine $(0, 0) \in \mathbb{R}^2$, et le nombre $i = \sqrt{-1}$ correspond au point $(0, 1) \in \mathbb{R}^2$.

Naturellement, l'axe des x est appelé l'*axe réel*, tandis que l'axe des y est appelé l'*axe imaginaire*.



Les règles pour additionner et pour soustraire les nombres complexes sont naturelles : il suffit de conserver en mémoire que $i^2 = -1$.

Par exemple, étant donné deux nombres complexes :

$$z_1 = x_1 + i\,y_1 \qquad \text{et} \qquad z_2 = x_2 + i\,y_2,$$

leur somme vaut :

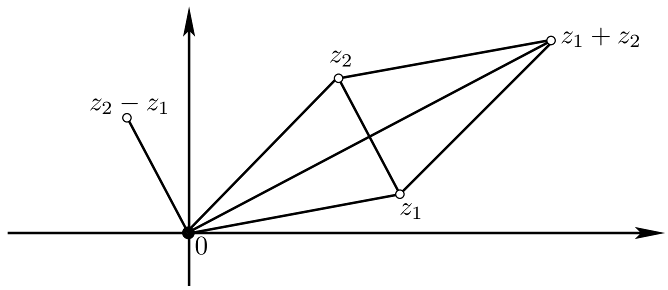
$$z_1 + z_2 = (x_1 + x_2) + i\,(y_1 + y_2),$$

et leur produit vaut :

$$\begin{aligned} z_1 z_2 &= (x_1 + i\,y_1)\,(x_2 + i\,y_2) \\ &= x_1x_2 + i\,x_1y_2 + i\,y_1\,x_2 + i^2\,y_1y_2 \\ &= (x_1x_2 - y_1y_2) + i\,(x_1y_2 + y_1x_2). \end{aligned}$$

Si l'on prend ces deux expressions pour définitions de l'addition et de la multiplication, il est facile de vérifier les trois propriétés suivantes.

- Commutativité : $z_1 + z_2 = z_2 + z_1$ et $z_1z_2 = z_2z_1$ pour tous $z_1, z_2 \in \mathbb{C}$.
- Associativité : $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$ et $(z_1z_2)z_3 = z_1(z_2z_3)$ pour tous $z_1, z_2, z_3 \in \mathbb{C}$.
- Distributivité : $z_1(z_2 + z_3) = z_1z_2 + z_1z_3$ pour tous $z_1, z_2, z_3 \in \mathbb{C}$.



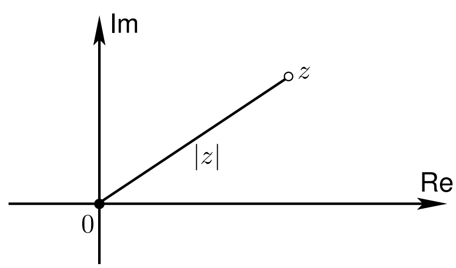
Géométriquement, l'addition des nombres complexes correspond à l'addition des vecteurs dans le plan $\mathbb{R}^2$. La multiplication quant à elle consiste en une rotation suivie d'une dilatation, un fait qui devient intuitivement transparent une fois qu'on a introduit la forme polaire d'un nombre complexe (voir ci-dessous). À ce stade, observons au moins que *la multiplication par i correspond à une rotation d'angle $\frac{\pi}{2}$.*

La notion de valeur absolue d'un nombre complexe est identique à la norme euclidienne des vecteurs dans $\mathbb{R}^2$.

Définition 8.1. La *valeur absolue* ou le *module* d'un nombre complexe $z = x + i\,y$ est la quantité positive :

$$|z| := \big(x^2 + y^2\big)^{\frac{1}{2}},$$

de telle sorte que $|z|$ est précisément la *distance euclidienne* entre l'origine $(0,0)$ et le point $(x,y) \in \mathbb{R}^2$.



Bien entendu, $|z| = 0$ si et seulement si $z = 0$, et l'*inégalité du triangle* est tout aussi valable :

$$|z + w| \leqslant |z| + |w| \qquad\qquad (z, w \in \mathbb{C}).$$

D'autres inégalités seront aussi utiles. Pour tout $z \in \mathbb{C}$, on a :

$$\big|\mathrm{Re}\, z\big| \leqslant |z|,$$
$$\big|\mathrm{Im}\, z\big| \leqslant |z|,$$

et pour tous $z, w \in \mathbb{C}$, on a :

$$\big||z| - |w|\big| \leqslant |z - w|,$$

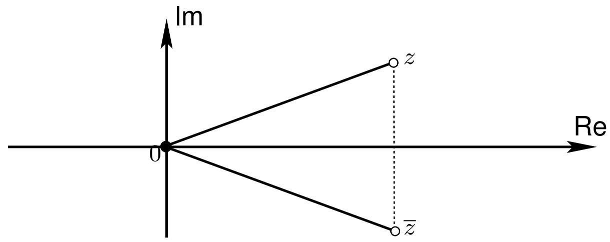
ce qui découle par soustraction de l'inégalité du triangle, puisque :

$$|z| \leqslant |z - w| + |w| \qquad \text{et} \qquad |w| \leqslant |z - w| + |z|.$$

Définition 8.2. Le *conjugué* d'un nombre complexe $z = x + i\,y$ est le nombre :

$$\overline{z} = x - i\,y,$$

que l'on obtient géométriquement en appliquant la symétrie le long de l'axe réel du plan complexe.



Évidemment, un nombre complexe z est réel si et seulement si :

$$z = \overline{z},$$

et il est imaginaire pur si et seulement si :

$$z = -\,\overline{z}.$$

On vérifie aussi sans difficulté que :

$$\begin{aligned}\mathsf{Re}\, z &= \frac{z+\overline{z}}{2},\\ \mathsf{Im}\, z &= \frac{z-\overline{z}}{2\,i}.\end{aligned}$$

De plus :

$$|z|^2 = z\,\overline{z},$$

et aussi :

$$\frac{1}{z} = \frac{\overline{z}}{|z|^2} \qquad\qquad (z\neq 0).$$

Définition 8.3. Un nombre complexe quelconque $z \in \mathbb{C}\backslash\{0\}$ non nul peut toujours s'écrire sous forme *polaire :*

$$z = r\,e^{i\,\theta},$$

avec $r > 0$ réel égal au module $|z|$ et $\theta \in \mathbb{R}$ appelé l'*argument* de z, qui est défini à un multiple entier $\in \mathbb{Z}$ de 2π près, traditionellement noté :

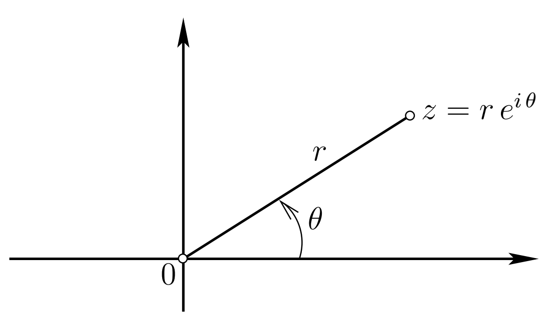
$$\theta = \arg z,$$

sachant que :

$$\begin{aligned}e^{i\,\theta} &= \cos\theta + i\sin\theta,\\ e^{i(\theta+2k\pi)} &= \cos\big(\theta+2k\pi\big) + i\sin\big(\theta+2k\pi\big)\\ &= \cos\theta + i\sin\theta,\end{aligned}$$

pour tout entier $k \in \mathbb{Z}$ en effet.

Puisque $|e^{i\,\theta}| = 1$, le nombre θ est l'*angle* que fait la demi-droite $0z$ avec l'axe des réels positifs.



Enfin, notons que la multiplication entre deux nombres complexes $z = r\,e^{i\,\theta}$ et $w = s\,e^{i\,\varphi}$ donne :

$$z\,w = rs\,e^{i(\theta+\varphi)},$$

de telle sorte que *la multiplication complexe consiste toujours, géométriquement, en une homothétie composée (commutativement) avec une rotation*.

9. Exercices

Exercice 1. EE

Applications linéaires

François DE MARÇAY
Département de Mathématiques d'Orsay
Université Paris-Sud, France

1. Introduction

2. Homomorphismes linéaires entre espaces vectoriels

Les concepts et définitions de ce chapitre sont valables sur n'importe quel corps commutatif $\mathbb{K}$, mais nous travaillerons le plus souvent avec le corps $\mathbb{K} = \mathbb{R}$ des nombres réels.

Définition 2.1. Soient E et F deux $\mathbb{R}$-espaces vectoriels. Une application $f\colon E \longrightarrow F$ est dite *linéaire* si les deux conditions suivantes sont satisfaites :

$$\begin{aligned} &\forall\,\vec{x}\in E \quad \forall\,\vec{y}\in E && f\big(\vec{x}+\vec{y}\big) \,=\, f(\vec{x})+f(\vec{y}), \\ &\forall\,\vec{x}\in E \quad \forall\,a\in\mathbb{R} && f\big(a\,\vec{x}\big) \,=\, a\,f(\vec{x}). \end{aligned}$$

La première condition exprime que f est un *homomorphisme* de E dans F pour l'addition. C'est pourquoi, au lieu d'application linéaire, on parle parfois d'*homomorphisme linéaire* entre espaces vectoriels.

Terminologie 2.2. Dans le cas où l'espace-arrivée $F = E$ coïncide avec l'espace-source, on dit que f est un *endomorphisme linéaire* de E.

Exemples 2.3. **(1)** La notion d'application linéaire dans le plan euclidien $V = \mathbb{R}^2$ ou dans l'espace euclidien $V = \mathbb{R}^3$ est connue dans les cours de géométrie de l'enseignement secondaire, puisqu'on y parle de *projections*, sur une droite, ou sur un plan. Si f est un projecteur quelconque, alors la relation :

$$f\big(\vec{x}+\vec{y}\big) \,=\, f(\vec{x})+f(\vec{y})$$

est nommée *théorème des projections.*

La relation :

$$f\big(a\,\vec{x}\big) \,=\, a\,f(\vec{x}) \qquad\qquad (\forall\,a\in\mathbb{R}),$$

est le fameux *Théorème de Thalès.*

Tout projecteur dans $V = \mathbb{R}^2$ ou $V = \mathbb{R}^3$ est un endomorphisme linéaire — nous y reviendrons au cours de ce chapitre.

(2) Dans un $\mathbb{R}$-espace vectoriel E, fixons un scalaire a, et définissons une application h de E dans E comme suit :

$$\forall\,\vec{x}\in E \qquad\qquad h(\vec{x}) \,:=\, a\,\vec{x}.$$

On a évidemment :

$$h\big(\vec{x}+\vec{y}\big) \,=\, a\,\big(\vec{x}+\vec{y}\big) \,=\, a\,\vec{x}+a\,\vec{y} \,=\, h(\vec{x})+h(\vec{y}),$$

et aussi, puisque $\mathbb{R}$ est commutatif :

$$\forall\,b\in\mathbb{R} \qquad\qquad h\big(b\,\vec{x}\big) \,=\, a\,\big(b\,\vec{x}\big) \,=\, b\,\big(a\,\vec{x}\big) \,=\, b\,h(\vec{x}).$$

Par conséquent, l'application h est un endomorphisme linéaire de E. On la nomme *homothétie de rapport* a, et on la note parfois h_a.

Assertion 2.4. *Pour $a \neq 0$ fixée, l'homothétie h_a est une bijection de E.*

Preuve. Pour vérifier qu'on a en effet :

$$\forall \vec{y} \in E \qquad \exists ! \, \vec{x} \in E \qquad \vec{y} \,=\, h(\vec{x}),$$

partons de :

$$\vec{y} \,=\, a \, \vec{x}.$$

Comme $a \neq 0$, son inverse a^{-1} pour la multiplication dans $\mathbb{R}$ existe, et donc l'existence de $\vec{x}$ est assurée par :

$$a^{-1} \, \vec{y} \,=\, \big(a^{-1} \, a\big) \, \vec{x} \,=\, \vec{x}.$$

Enfin, l'unicité de $\vec{x}$ résulte de :

$$a \, \vec{x} \,=\, a \, \vec{x}\,' \qquad \Longrightarrow \qquad a \, \big(\vec{x} - \vec{x}\,'\big) \,=\, \vec{0} \qquad \overset{a \neq 0}{\Longrightarrow} \qquad \vec{x} - \vec{x}\,' \,=\, \vec{0}. \qquad \square$$

Terminologie 2.5. Une application linéaire $h \colon E \longrightarrow E$ d'un $\mathbb{R}$-espace vectoriel E dans lui même qui est *bijective* sera nommée *automorphisme linéaire* de E.

D'une façon plus générale, il se peut que la bijection linéaire s'exerce entre deux espaces distincts.

Terminologie 2.6. Une application linéaire bijective $h \colon E \longrightarrow F$ entre deux $\mathbb{R}$-espaces vectoriels E et F qui est *bijective* sera nommée *isomorphisme linéaire* entre E et F.

On dit aussi que E et F sont *isomorphes*. Quand $E = F$, on parle donc d'un *automorphisme*.

Proposition 2.7. [Critère de linéarité] *Pour qu'une application $f \colon E \longrightarrow F$ entre deux $\mathbb{R}$-espaces vectoriels soit linéaire, il faut et il suffit que :*

$$\forall \, a, b \in \mathbb{R} \qquad \forall \, \vec{x}, \vec{y} \in E \qquad\qquad f\big(a \, \vec{x} + b \, \vec{y}\big) \,=\, a \, f(\vec{x}) + b \, f(\vec{y}).$$

Démonstration. Prouvons '$\Longrightarrow$'. Si l'application f est linéaire, alors on a :

$$f\big(a \, \vec{x}\big) \,=\, a \, f(\vec{x}), \qquad\qquad f\big(b \, \vec{x}\big) \,=\, b \, f(\vec{x}), \qquad\qquad f\big(a \, \vec{x} + b \, \vec{y}\big) \,=\, f\big(a \, \vec{x}\big) + f\big(b \, \vec{y}\big),$$

d'où :

$$f\big(a \, \vec{x} + b \, \vec{y}\big) \,=\, a \, f(\vec{x}) + b \, f(\vec{y}).$$

Inversement, prouvons '$\Longleftarrow$'. Partons donc de cette dernière relation. Puisqu'elle est vraie pour tous $a, b \in \mathbb{R}$ et tous vecteurs $\vec{x}, \vec{y} \in E$, nous pouvons prendre $a = b = 1$ pour obtenir la première condition de la Définition 2.1 ; ensuite, prenons $b = 0$ pour obtenir la deuxième condition. En définitive, f est bien linéaire. $\square$

Observation 2.8. *Toute application linéaire $f \colon E \longrightarrow F$ entre $\mathbb{R}$-espaces vectoriels satisfait :*

$$f(\vec{0}_E) \,=\, \vec{0}_F.$$

Preuve. Il suffit de prendre $a = 0$ dans la deuxième condition de la Définition 2.1. $\square$

3. Image et noyau d'une application linéaire

Commençons par étudier l'image d'une application linéaire.

Théorème 3.1. *Soit $f \colon E \longrightarrow F$ une application linéaire entre deux $\mathbb{R}$-espaces vectoriels E et F. Alors pour tout sous-espace vectoriel $E' \subset E$, l'image :*

$$f(E') \,:=\, \big\{\vec{y} \in F \colon \exists \, \vec{x} \in E', \;\; \vec{y} = f(\vec{x})\big\}$$

est aussi un sous-espace vectoriel de F.

Démonstration. Prouvons que $f(E')$ est un sous-espace vectoriel de F, c'est-à-dire :

$$\Big(a,\, a' \in \mathbb{R} \quad \text{et} \quad \vec{y},\, \vec{y}\,' \in f(E')\Big) \qquad \Longrightarrow \qquad a\,\vec{y} + a'\,\vec{y}\,' \in f(E').$$

En effet, si $\vec{y}$ et $\vec{y}\,'$ appartiennent à $f(E')$, c'est qu'il existe deux vecteurs $\vec{x}$ et $\vec{x}\,'$ dans E' tels que :

$$\vec{y} = f(\vec{x}) \qquad \text{et} \qquad \vec{y}\,' = f(\vec{x}\,').$$

Comme E' est un sous-espace de E, alors, quels que soient les scalaires a et a', on a :

$$\big(\vec{x} \in E' \quad \text{et} \quad \vec{x}\,' \in E'\big) \qquad \Longrightarrow \qquad a\,\vec{x} + a'\,\vec{x}\,' \in E' \qquad \Longrightarrow \qquad f\big(a\,\vec{x} + a'\,\vec{x}'\big) \in f(E').$$

Mais comme f est linéaire :

$$f\big(a\,\vec{x} + a'\,\vec{x}\,'\big) = a\,f(\vec{x}) + a'\,f(\vec{x}'\,).$$

Par conséquent :

$$a\,\vec{y} + a'\vec{y}\,' \in f(E').$$

Le théorème est démontré. □

Définition 3.2. Quand le sous-espace $E' = E$ coïncide avec l'espace vectoriel ambiant tout entier, alors $f(E)$ est un sous-espace vectoriel de F que l'on nomme *image de* f, et que l'on note $\mathsf{Im}(f)$.

Ce sous-espace vectoriel $\mathsf{Im}(f)$ s'écrit :

$$\mathsf{Im}\, f := \big\{\vec{y} \in F \colon \exists\, \vec{x} \in E',\ \ \vec{y} = f(\vec{x})\big\}.$$

Nous pouvons maintenant introduire le concept de *noyau* d'une application linéaire.

Définition 3.3. Soit $f \colon E \longrightarrow F$ une application linéaire entre deux $\mathbb{R}$-espaces vectoriels. L'ensemble des vecteurs de E qui ont pour image le vecteur $\vec{0} \in F$ se nomme *noyau* de f et se note :

$$\mathsf{Ker}(f) := \big\{\vec{x} \in E \colon f(\vec{x}) = \vec{0}\big\}.$$

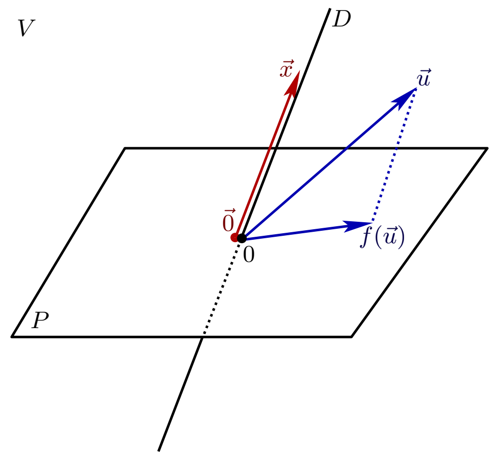
Noyau se dit *kernel* en anglais, et *Kern* en allemand. C'est pourquoi on rencontre très souvent la notation $\mathsf{Ker}(f)$, que nous adopterons.

Observons que $\mathsf{Ker}(f)$ n'est jamais vide, car, quelle que soit l'application linéaire f, on a $f(\vec{0}) = \vec{0}$, donc :

$$\vec{0} \in \mathsf{Ker}(f) \qquad \text{(toujours)}.$$

Exemple 3.4. [en géométrie euclidienne] Dans l'espace euclidien $V = \mathbb{R}^3$, soit un plan $P \ni 0$ passant par l'origine, et soit une droite $D \ni 0$ passant par l'origine, avec $D \not\parallel P$ non parallèle au plan.

Rappelons que les vecteurs $\vec{y} \in P$ sont les vecteurs liés ayant comme point-origine $0 \in V$, dont la flèche-extrémité appartient à P. De même, les vecteurs $\vec{x} \in D$ ont pour point-origine $0 \in D$, et une flèche-extrémité appartenant à D. On peut donc voir P et V comme deux sous-espaces vectoriels de V, de dimensions 2 et 1.



Désignons par f la projection de V sur P parallèle à D. On sait que f est une application linéaire de V dans lui-même, avec $f(V) = P$.

Cherchons le noyau de f.

Pour qu'un vecteur $\vec{x} \in V$ vérifie $f(\vec{v}) = \vec{0}$, c'est-à-dire ait une projection nulle dans V, il faut et il suffit que $\vec{x}$ appartienne à D. Ainsi :

$$\mathsf{Ker}(f) \,=\, D.$$

Théorème 3.5. *Pour toute application linéaire $f \colon E \longrightarrow F$ entre $\mathbb{R}$-espaces vectoriels, le noyau $\mathsf{Ker}(f)$ est un sous-espace vectoriel de E.*

Preuve. En effet, partons de :

$$\vec{x} \,\in\, \mathsf{Ker}(f) \qquad\qquad \text{et} \qquad\qquad \vec{x}\,' \,\in\, \mathsf{Ker}(f),$$

ce qui équivaut à :

$$f(\vec{x}) \,=\, \vec{0} \qquad\qquad \text{et} \qquad\qquad f(\vec{x}\,') \,=\, \vec{0}.$$

Comme f est linéaire, alors, quels que soient les scalaires a et a' :

$$f\big(a\,\vec{x} + a'\,\vec{x}\,'\big) \,=\, a\,f(\vec{x}) + a'\,f(\vec{x}\,') \,=\, \vec{0} + \vec{0} \,=\, \vec{0}.$$

Par conséquent :

$$a\,\vec{x} + a'\,\vec{x}\,' \,\in\, \mathsf{Ker}(f).$$

Le théorème est démontré. □

Voici maintenant une caractérisation des applications linéaires injectives.

Théorème 3.6. *Pour une application linéaire $f \colon E \longrightarrow F$ entre $\mathbb{R}$-espaces vectoriels, on a l'équivalence :*

$$f \text{ est injective} \qquad\qquad \Longleftrightarrow \qquad\qquad \mathsf{Ker}(f) \,=\, \{\vec{0}\}.$$

Démonstration. Prouvons '$\Longrightarrow$'. Supposons donc que l'application linéaire f est injective. Alors $\vec{x} \in \mathsf{Ker}(f)$ signifie :

$$f(\vec{x}) \,=\, \vec{0} \,=\, f(\vec{0}).$$

Comme f est injective, on en déduit que $\vec{x} = \vec{0}$, ce qui conclut :

$$\mathsf{Ker}(f) \,=\, \{\vec{0}\}.$$

Inversement, prouvons '$\Longleftarrow$'. Supposons $\mathsf{Ker}(f) = \{\vec{0}\}$. Alors, quels que soient $\vec{x}$ et $\vec{x}\,'$ dans E :

$$\begin{aligned} f(\vec{x}) \,=\, f(\vec{x}\,') \quad &\Longleftrightarrow \quad f(\vec{x}) - f(\vec{x}\,') \,=\, \vec{0} \quad &&\Longleftrightarrow \quad f\big(\vec{x} - \vec{x}\,'\big) \,=\, \vec{0} \\ &\Longleftrightarrow \quad \vec{x} - \vec{x}\,' \,\in\, \mathsf{Ker}(f) \quad &&\Longleftrightarrow \quad \vec{x} - \vec{x}\,' \,=\, \vec{0} \\ & &&\Longleftrightarrow \quad \vec{x} \,=\, \vec{x}\,'. \end{aligned}$$

Ceci signifie que f est injective, conclut l'implication inverse, et termine la démonstration du théorème. □

Parlons maintenant de l'image d'une famille de générateurs.

Théorème 3.7. **(1)** *Une application linéaire f de E dans F applique toute famille de générateurs de tout sous-espace vectoriel $E' \subset E$ sur une famille de générateurs de $f(E')$.*

(2) *Si f est de plus injective, elle applique toute famille libre de E sur une famille libre de F.*

Démonstration. Soit donc une famille de vecteurs de E :

$$L := \{\vec{x}_1, \dots, \vec{x}_p\}.$$

Par une application linéaire $f \colon E \longrightarrow F$, elle devient la famille de vecteurs de F :

$$f(F) = \{f(\vec{x}_1), \dots, f(\vec{x}_p)\}.$$

(1) Désignons par $E' := \mathsf{Vect}(L)$ le sous-espace de E engendré par L. Prouvons que $f(L)$ engendre le sous-espace $f(E')$. En effet :

$$\vec{y} \in f(E') \qquad \text{signifie} \qquad \exists\, \vec{x} \in E' \quad \vec{y} = f(\vec{x}).$$

Puisque L est une famille de générateurs de E', à un vecteur $\vec{x} \in E'$ correspond une famille $a_1, \dots, a_p \in \mathbb{R}$ de scalaires tels que :

$$\vec{x} = a_1\,\vec{x}_1 + \cdots + a_p\,\vec{x}_p.$$

Par linéarité de f, on obtient :

$$f(\vec{x}) = a_1\,f(\vec{x}_1) + \cdots + a_p\,f(\vec{x}_p).$$

Tout vecteur $\vec{y}$ de $f(E')$ est donc une combinaison linéaire de vecteurs de $f(L)$. La première partie du théorème est prouvée.

(2) Supposons de plus la famille L *libre*, et l'application f *injective* ; prouvons que $f(L)$ est libre. En effet :

$$a_1\,f(\vec{x}_1) + \cdots + a_p\,f(\vec{x}_p) = \vec{0} \qquad \Longrightarrow \qquad f(\vec{x}) = \vec{0} \qquad \Longleftrightarrow \qquad \vec{x} \in \mathsf{Ker}(f).$$

Comme f est injective, cela force $\vec{x} = \vec{0}$, grâce au Théorème 3.6. Comme $L = \{\vec{x}_1, \dots, \vec{x}_p\}$ est libre, il vient :

$$\vec{0} = \vec{x} = a_1\,\vec{x}_1 + \cdots + a_p\,\vec{x}_p \qquad \Longrightarrow \qquad a_1 = \cdots = a_p = 0.$$

En définitive, la famille $f(L)$ est bien libre, ce qui termine la démonstration de la seconde partie du théorème. □

Corollaire 3.8. *Si E est un $\mathbb{R}$-espace vectoriel de dimension finie et si F est un $\mathbb{R}$-espace vectoriel,* pas forcément de dimension finie*, alors pour toute application linéaire $f \colon E \longrightarrow F$, l'image $f(E) \subset F$ est un sous-espace vectoriel de dimension finie $\leqslant \mathsf{dim}_{\mathbb{R}}\, E$.*

Preuve. En effet, l'hypothèse que E est de dimension finie veut dire qu'il existe une famille finie L de générateurs de E. Alors d'après le Théorème 3.7 qui précède, $f(L)$ est une famille *finie* de générateurs de $f(E)$. Si L est de plus une base de E, on déduit bien que :

$$\mathsf{dim}_{\mathbb{R}}\, f(E) \leqslant \mathsf{Card}\, L = \mathsf{dim}_{\mathbb{R}}\, E < \infty. \qquad \square$$

Voici un résultat central qui exprime qu'une application linéaire est complètement déterminée lorsqu'on connaît seulement les images d'un nombre fini de vecteurs qui forment une *base* de l'espace vectoriel source.

Théorème 3.9. *Soit $B = \{\vec{e}_1, \dots, \vec{e}_n\}$ une* base *d'un $\mathbb{R}$-espace vectoriel E de dimension $n \geqslant 1$, et soit $C = \{\vec{c}_1, \dots, \vec{c}_n\}$ une famille quelconque de n vecteurs d'un autre $\mathbb{R}$-espace vectoriel F.*

Alors il existe une application linéaire, et une seule, $f\colon E \longrightarrow F$ telle que :

$$f(\vec{e}_1) = \vec{c}_1, \quad f(\vec{e}_2) = \vec{c}_2, \quad \dots\dots, \quad f(\vec{e}_n) = \vec{c}_n.$$

De plus, si la famille C est libre, alors f est injective.

Démonstration. Commençons par l'existence de f. Définissons une application f de E dans F de la manière suivante : à tout vecteur $\vec{x} \in E$, associons d'abord l'unique famille de scalaires $a_1, \dots, a_n$ qui sont les coordonnées de $\vec{x}$ dans la base B :

$$\vec{x} = a_1\,\vec{e}_1 + \cdots + a_n\,\vec{e}_n.$$

À cet $\vec{x}$, faisons alors correspondre le vecteur $\vec{y} \in F$ défini par :

$$\vec{y} = f(\vec{x}) = a_1\,\vec{c}_1 + \cdots + a_n\,\vec{c}_n. \tag{3.10}$$

Le domaine de définition de f est bien sûr E tout entier, et l'image $f(E)$ est bien incluse dans F.

Prouvons que f est linéaire. Partons de deux vecteurs $\vec{x}, \vec{x}\,' \in E$, que nous écrirons sous forme condensée :

$$\vec{x} = \sum_i a_i\,\vec{e}_i \qquad \text{et} \qquad \vec{x}\,' = \sum_i a'_i\,\vec{e}_i.$$

Pour deux scalaires quelconques b et b', on a :

$$b\,\vec{x} + b'\,\vec{x}\,' = \sum_i b\,a_i\,\vec{e}_i + \sum_i b'\,a'_i\,\vec{e}_i = \sum_i \big(b\,a_i + b'\,a'_i\big)\,\vec{e}_i.$$

Par définition de l'application linéaire f, on a alors :

$$f(\vec{x}) = \sum_i a_i\,\vec{c}_i, \qquad f(\vec{x}\,') = \sum_i a'_i\,\vec{c}_i\,',$$
$$f\big(b\,\vec{x} + b'\,\vec{x}\,'\big) = \sum_i \big(b\,a_i + b'\,a'_i\big)\,\vec{c}_i.$$

Dans l'espace vectoriel F, on voit immédiatement que :

$$\sum_i \big(b\,a_i + b'\,a'_i\big)\,\vec{c}_i = b\sum_i a_i\,\vec{c}_i + b'\sum_i a'_i\,\vec{c}_i.$$

c'est-à-dire :

$$f\big(b\,\vec{x} + b'\,\vec{x}\,'\big) = b\,f(\vec{x}) + b'\,f(\vec{x}\,'),$$

ce qui montre que f est bien linéaire.

Passons à l'unicité de f. Par définition, toute autre application linéaire g de E dans F envoie :

$$\vec{x} = a_1\,\vec{e}_1 + \cdots + a_n\,\vec{e}_n$$

sur :

$$g(\vec{x}) = a_1\,g(\vec{e}_1) + \cdots + a_n\,g(\vec{e}_n).$$

Par conséquent, si on fixe dans F comme pour f :

$$g(\vec{e}_1) := \vec{c}_1, \quad \dots\dots, \quad g(\vec{e}_n) := \vec{c}_n,$$

on obtient une application g particulière qui vérifie :

$$g(\vec{x}) = f(\vec{x}) \qquad (\forall\,\vec{x} \in E),$$

d'où :

$$g = f,$$

ce qui montre l'unicité de l'application f.

Terminons l'argumentation en montrant que f sera de plus injective, quand on suppose que $C = \{\vec{c}_1, \dots, \vec{c}_n\}$ est libre.

Soit donc $\vec{x} \in E$ avec $f(\vec{x}) = \vec{0}$. Comme $\{\vec{e}_1, \dots, \vec{e}_n\}$ est une base, on peut décomposer selon cette base $\vec{x} = a_1\,\vec{e}_1 + \cdots + a_n\,\vec{e}_n$, d'où :

$$\vec{0} \,=\, f(\vec{x}) \,=\, a_1\,\vec{c}_1 + \cdots + a_n\,\vec{c}_n \qquad\qquad \Longrightarrow \qquad\qquad a_1 \,=\, \cdots \,=\, a_n \,=\, 0,$$

puisque $\{\vec{c}_1, \dots, \vec{c}_n\}$ est libre. Ainsi, nous avons bien démontré que :

$$\Big(f(\vec{x}) \,=\, \vec{0} \qquad \Longrightarrow \qquad \vec{x} \,=\, \vec{0}\Big) \qquad\qquad \Longleftrightarrow \qquad\qquad \mathsf{Ker}(f) \,=\, \{\vec{0}\}. \qquad \square$$

Des deux Théorèmes 3.7 et 3.9, on déduit immédiatement le

Corollaire 3.11. *Pour qu'une application $f\colon E \longrightarrow F$ entre $\mathbb{R}$-espaces vectoriels soit injective, il faut et il suffit que l'image d'une base de E soit une base de $f(E)$.* $\square$

4. Applications linéaires et dimension

Nous savons que l'ensemble $\mathbb{R}^n$ des n-uplets $(a_1, \dots, a_n)$ constitués de n réels $a_i \in \mathbb{R}$ est un $\mathbb{R}$-espace vectoriel.

Définition 4.1. On appelle *base canonique* de $\mathbb{R}^n$ la base $\big\{\vec{u}_1, \dots, \vec{u}_n\big\}$ suivante :

$$\begin{aligned}
\vec{u}_1 \,&:=\, \big(1, 0, \dots, 0\big),\\
\vec{u}_2 \,&:=\, \big(0, 1, \dots, 0\big),\\
&\cdots\cdots\cdots\cdots\cdots\cdots\\
\vec{u}_n \,&:=\, \big(0, 0, \dots, 1\big).
\end{aligned}$$

Pour tout indice $1 \leqslant i \leqslant n$, le vecteur $\vec{u}_i$ est le n-uplet dont tous les termes sont nuls, sauf celui de rang i qui vaut 1. Dans tout ce qui suit, nous supposerons toujours que $\mathbb{R}^n$ est rapporté à la base canonique :

$$\big(a_1,\, a_2,\, \dots,\, a_n\big) \,=\, a_1\,\vec{u}_1 + a_2\,\vec{u}_2 + \cdots + a_n\,\vec{u}_n.$$

Voici un théorème qui exprime une propriété fondamentale d'isomorphisme.

Théorème 4.2. *Pour qu'un $\mathbb{R}$-espace vectoriel E soit de dimension n, il faut et il suffit qu'il soit* isomorphe *à $\mathbb{R}^n$.*

L'isomorphisme est alors déterminé de façon unique dès qu'on fixe une base de E comme image de la base canonique de $\mathbb{R}^n$.

Démonstration. Prouvons l'implication '$\Longrightarrow$'. Supposons donc que $\mathsf{dim}_{\mathbb{R}}\, E = n$, et soit une base de E :

$$B \,:=\, \big\{\vec{e}_1, \dots, \vec{e}_n\big\}.$$

Comme ci-dessus, soit $C = \{\vec{u}_1, \dots, \vec{u}_n\}$ la base canonique de $\mathbb{R}^n$.

D'après le Théorème 3.9, il existe une application linéaire f et une seule telle que :

$$f(\vec{e}_1) \,=\, \vec{u}_1, \ \ \dots\dots, \ \ f(\vec{e}_n) \,=\, \vec{u}_n,$$

et de plus, f est injective. Mais elle est aussi surjective, car C engendre à la fois $f(E)$ et $\mathbb{R}^n$, donc $f(E) = \mathbb{R}^n$.

Puisque f est bijective, elle établit bien un *isomorphisme* de E sur $\mathbb{R}^n$.

Observons que f est d'ailleurs l'unique application linéaire de E dans F qui envoie B sur C.

Prouvons maintenant l'implication inverse '$\Longleftarrow$'. Si donc il existe un isomorphisme f de E sur $\mathbb{R}^n$, posons :

$$\vec{e}_1 \,:=\, f^{-1}(\vec{u}_1), \ \ \dots\dots, \ \ \vec{e}_n \,:=\, f^{-1}(\vec{u}_n).$$

D'après le Théorème 3.7, l'application linéaire f^{-1} envoie la famille libre $C = \{\vec{u}_1, \dots, \vec{u}_n\}$ de générateurs de $\mathbb{R}^n$ sur la famille libre $B = \{\vec{e}_1, \dots, \vec{e}_n\}$ de générateurs de $f^{-1}(\mathbb{R}^n) = E$. Donc B est bien une base de E, et $\mathsf{dim}_{\mathbb{R}}\, E = n$, ce qui conclut. $\square$

Voici maintenant un théorème très important qui établit une relation entre les dimensions du noyau et de l'image.

Théorème 4.3. *Pour toute application linéaire $f\colon E \longrightarrow F$ entre deux espaces vectoriels, on a :*

$$\dim \mathsf{Ker}(f) + \dim \mathsf{Im}(f) \;=\; \dim E.$$

Démonstration. Soient :

$$\begin{aligned} B \;&:=\; \big\{\vec{e}_1, \dots, \vec{e}_p\big\} \quad \text{une base de } \mathsf{Ker}(f) \,\subset\, E, \\ C \;&:=\; \big\{\vec{\ell}_1, \dots, \vec{\ell}_q\big\} \quad \text{une base de } \mathsf{Im}(f) \,\subset\, F, \end{aligned}$$

où on a posé :

$$\begin{aligned} p \;&:=\; \dim \mathsf{Ker}(f), \\ q \;&:=\; \dim \mathsf{Im}(f). \end{aligned}$$

On sait que $\vec{y} \in \mathsf{Im}(f)$ signifie qu'il existe un vecteur $\vec{x} \in E$ avec $\vec{y} = f(\vec{x})$. Donc, pour tout indice $1 \leqslant i \leqslant q$, il existe un vecteur de E, que nous désignerons par $\vec{e}_{p+i}$, tel que :

$$f\big(\vec{e}_{p+i}\big) \;=\; \vec{\ell}_i.$$

On obtient ainsi une famille B' de q vecteurs de E :

$$B' \;:=\; \big\{\vec{e}_{p+1}, \dots, \vec{e}_{p+q}\big\}.$$

Nous affirmons que $B \cup B'$ est alors une *base* de E, ce que nous démontrons en deux temps.

Assertion 4.4. *$B \cup B'$ engendre E.*

Preuve. Montrons que tout $\vec{x} \in E$ est combinaison linéaire de vecteurs appartenant à $B \cup B'$. En effet :

$$\vec{x} \,\in\, E \qquad\qquad \Longrightarrow \qquad\qquad f(\vec{x}) \,\in\, f(E) \,=\, \mathsf{Im}\, f.$$

Il existe donc des scalaires $a_1, \dots, a_q \in \mathbb{R}$, coordonnées de $f(\vec{x})$ dans la base de $\mathsf{Im}\, f$, tels que :

$$\begin{aligned} f(\vec{x}) \;&=\; a_1\, \vec{\ell}_1 + \cdots + a_q\, \vec{\ell}_q \\ &=\; a_1\, f\big(\vec{e}_{p+1}\big) + \cdots + a_q\, f\big(\vec{e}_{p+q}\big) \\ &=\; f\, \big(a_1\, \vec{e}_{p+1} + \cdots + a_q\, \vec{e}_{p+q}\big). \end{aligned}$$

Si on pose :

$$\vec{x}\,' \;:=\; a_1\, \vec{e}_{p+1} + \cdots + a_q\, \vec{e}_{p+q}, \tag{4.5}$$

alors :

$$f(\vec{x}) \;=\; f\big(\vec{x}\,'\big) \qquad\qquad \Longrightarrow \qquad\qquad \vec{x} - \vec{x}\,' \,\in\, \mathsf{Ker}(f).$$

Il existe donc des scalaires $a_1', \dots, a_p'$, coordonnées de $\vec{x} - \vec{x}\,'$ dans B, tels que :

$$\vec{x} - \vec{x}\,' \;=\; a_1'\, \vec{e}_1 + \cdots + a_p'\, \vec{e}_p,$$

d'où, en remplaçant $\vec{x}\,'$ par sa valeur dans (4.5) :

$$\vec{x} \;=\; a_1'\, \vec{e}_1 + \cdots + a_p'\, \vec{e}_p + a_1\, \vec{e}_{p+1} + \cdots + a_q\, \vec{e}_{p+q}.$$

Donc $B \cup B'$ engendre E. □

Assertion 4.6. *$B \cup B'$ est une famille libre dans E.*

Preuve. Partons de :

$$a'_1\,\vec{e}_1+\cdots+a'_p\,\vec{e}_p+a_1\,\vec{e}_{p+1}+\cdots+a_q\,\vec{e}_{p+q}\;=\;\vec{0}, \tag{4.7}$$

et montrons que tous les coefficients de cette combinaison linéaire sont nuls. Prenons l'image par f des deux membres de (4.7) ; comme, pour tout indice $1\leqslant i\leqslant p$, on a :

$$\vec{e}_i\;\in\;\mathsf{Ker}(f)\qquad\qquad\Longrightarrow\qquad\qquad f(\vec{e}_i)\;=\;\vec{0},$$

on obtient :

$$a_1\,f\big(\vec{e}_{p+1}\big)+\cdots+a_q\,f\big(\vec{e}_{p+q}\big)\;=\;\vec{0},$$

c'est-à-dire :

$$a_1\,\vec{\ell}_1+\cdots+a_q\,\vec{\ell}_q\;=\;\vec{0}.$$

Comme C est une famille libre, cette relation entraîne :

$$a_1\;=\;\cdots\;=\;a_q\;=\;0.$$

La relation (4.7) s'écrit alors :

$$a'_1\,\vec{e}_1+\cdots+a'_p\,\vec{e}_p\;=\;\vec{0}.$$

Comme B est une famille libre, cette relation entraîne à son tour :

$$a'_1\;=\;\cdots\;=\;a'_p\;=\;0.$$

Ainsi, $B\cup B'$ est bien une famille libre. □

Grâce à la conjonction de ces deux assertions, le théorème est démontré. □

Le cas où $\mathsf{dim}\,E=\mathsf{dim}\,F$ mérite d'être mis en valeur.

Corollaire 4.8. *Soit $f\colon E\longrightarrow F$ une application linéaire entre deux $\mathbb{R}$-espaces vectoriels de même dimension. Alors les quatre propriétés suivantes sont équivalentes :*

(i) *f est bijective ;*

(ii) *f est surjective ;*

(iii) *f est injective ;*

(iv) $\mathsf{Ker}(f)=\{\vec{0}\}$.

Preuve. On sait déjà que les propriétés **(iii)** et **(iv)** sont équivalentes.

D'autre part, **(i)** est équivalente à la conjonction des propriétés **(ii)** et **(iii)**.

Il suffit donc d'établir l'équivalence entre **(ii)** et **(iv)**. Or d'après le Théorème 4.3, on a :

$$\mathsf{Ker}(f)\;=\;\{\vec{0}\}\qquad\qquad\Longleftrightarrow\qquad\qquad\mathsf{dim}\,\mathsf{Im}(f)\;=\;\mathsf{dim}\,E.$$

Par conséquent, **(ii)** et **(iv)** sont bien équivalentes, ce qui conclut. □

5. Espace vectoriel des applications linéaires de E dans F

Soient E et F deux $\mathbb{R}$-espaces vectoriels. L'ensemble des applications linéaires de E dans F sera noté :

$$\mathscr{L}(E,F)\;:=\;\big\{f\colon\;E\longrightarrow F\;\;\text{linéaires}\big\}.$$

Nous allons munir cet ensemble $\mathscr{L}(E,F)$ d'une structure de $\mathbb{R}$-espace vectoriel.

Commençons par observer que la somme de deux applications linéaires et encore une application linéaire.

Lemme 5.1. *Si f et g sont deux applications linéaires appartenant à $\mathscr{L}(E,F)$, alors l'application $h\colon E\longrightarrow F$ définie par :*

$$h(\vec{x})\;:=\;f(\vec{x})+g(\vec{x})\qquad\qquad(\forall\,\vec{x}\in E)$$

est aussi linéaire.

Preuve. En effet, pour tous vecteurs $\vec{x}, \vec{y} \in E$ et tous scalaires $a, b \in \mathbb{R}$, on a :

$$h\big(a\,\vec{x} + b\,\vec{y}\big) \,=\, f\big(a\,\vec{x} + b\,\vec{y}\big) + g\big(a\,\vec{x} + b\,\vec{y}\big).$$

Comme f et g sont linéaires, et comme l'addition est commutative, on déduit :

$$\begin{aligned} h\big(a\,\vec{x} + b\,\vec{y}\big) \,&=\, a\,f(\vec{x}) + a\,g(\vec{x}) + b\,f(\vec{y}) + b\,g(\vec{y}) \\ &=\, a\,h(\vec{x}) + b\,g(\vec{y}). \end{aligned}$$

Ceci démontre le lemme. □

La définition suivante est alors justifiée.

Définition 5.2. À tout couple ordonné f, g de deux éléments de $\mathscr{L}(E, F)$, on associe un élément de $\mathscr{L}(E, F)$, nommé *somme* de f et de g, et noté $f + g$, qui est défini par :

$$\big(f + g\big)(\vec{x}) \,:=\, f(\vec{x}) + g(\vec{x}) \qquad (\forall\,\vec{x} \in E).$$

Cette addition est donc bien une loi interne dans $\mathscr{L}(E, F)$.

Théorème 5.3. *$\mathscr{L}(E, F)$ possède une structure naturelle de $\mathbb{R}$-espace vectoriel.*

Démonstration. Tout d'abord, l'addition dans $\mathscr{L}(E, F)$ est commutative et associative, puisqu'il en va ainsi de l'addition dans $\mathbb{R}$.

Ensuite, cette addition admet comme élément neutre la *fonction nulle*, notée $f = 0$, définie par :

$$0(\vec{x}) \,:=\, \vec{0} \qquad (\forall\,\vec{x} \in E).$$

Ici, remarquons au passage que $\mathrm{Ker}(0) = E$, et que la fonction nulle est la seule dont le noyau coïncide avec E tout entier.

Ensuite, toute application linéaire $f \in \mathscr{L}(E, F)$ a une opposée, notée $-f$, appartenant aussi à $\mathscr{L}(E, F)$, qui est définie par :

$$\big(-f\big)(\vec{x}) \,:=\, -f(\vec{x}) \qquad (\forall\,\vec{x} \in E).$$

Ensuite, parlons du produit d'une application linéaire par un scalaire. Pour $a \in \mathbb{R}$ quelconque, introduisons l'application h de E dans F définie par :

$$h(\vec{x}) \,:=\, a\,f(\vec{x}) \qquad (\forall\,\vec{x} \in E).$$

Alors pour tout couple de vecteurs $\vec{x}, \vec{y} \in E$ et tout couple de scalaires $b, c \in \mathbb{R}$, on a, puisque $\mathbb{R}$ est commutatif :

$$\begin{aligned} h\big(b\,\vec{x} + c\,\vec{y}\big) \,=\, a\,f\big(b\,\vec{x} + c\,\vec{y}\big) \,&=\, a\,b\,f(\vec{x}) + a\,c\,f(\vec{y}) \\ &=\, b\,a\,f(\vec{x}) + c\,a\,f(\vec{y}) \\ &=\, b\,h(\vec{x}) + c\,h(\vec{y}). \end{aligned}$$

Ce calcul justife la

Définition 5.4. À tout scalaire $a \in \mathbb{R}$ et à toute application linéaire $f \in \mathscr{L}(E, F)$, on associe le *produit de a par f*, qui est défini par :

$$\big(a\,f\big)(\vec{x}) \,:=\, a\,f(\vec{x}) \qquad (\forall\,\vec{x} \in E).$$

Ainsi, $a\,f \in \mathscr{L}(E, F)$. Pour terminer la démonstration du théorème, il est maintenant aisé de vérifier, pour toutes $f, g, \in \mathscr{L}(E, F)$ et tous $a, b \in \mathbb{R}$, que :

$$\begin{aligned} \big(a + b\big)\,f \,&=\, a\,f + b\,f, \\ a\,\big(f + g\big) \,&=\, a\,f + b\,g, \\ a\,\big(b\,f\big) \,&=\, (a\,b)\,f, \\ 1\,f \,&=\, f. \end{aligned}$$

Ces identités concluent la démonstration. □

6. Dual d'un espace vectoriel

Dans cette très courte section, introduisons le concept important de *dual* d'un espace vectoriel.

Pour cela, considérons dans la Section **5** qui précède le cas particulier où $F = \mathbb{R}$ est le corps des scalaires. Rappelons qu'on peut toujours regarder le corps $\mathbb{R}$ comme un espace vectoriel sur lui-même.

Terminologie 6.1. Toute application linéaire de E dans $\mathbb{R}$ se nomme *forme linéaire* sur E.

Notation 6.2. L'espace de ces formes linéaires sur E sera noté :

$$E^* = \mathscr{L}(E, \mathbb{R})$$

Le Théorème 5.3 montre que cet ensemble $E^* = \mathscr{L}(E, \mathbb{R})$ des formes linéaires de E dans $\mathbb{R}$ est lui-même un $\mathbb{R}$-espace vectoriel.

Terminologie 6.3. On nommera *dual de* E cet espace $E^* = \mathscr{L}(E, \mathbb{R})$ des formes linéaires sur E.

7. Anneau des endomorphismes linéaires d'un espace vectoriel E

Rappelons qu'une application linéaire de E dans E lui-même est appelée un *endomorphisme* de E. L'ensemble des endomorphismes de E sera noté :

$$\mathscr{L}(E) := \big\{ f \colon E \longrightarrow E \ \text{ linéaires} \big\},$$

au lieu de $\mathscr{L}(E, E)$. Le Théorème 5.3 s'applique : $\mathscr{L}(E)$ *est un espace vectoriel sur le corps* $\mathbb{R}$.

Nous allons considérer une deuxième loi de composition interne dans $\mathscr{L}(E)$ qui va le structurer en *anneau*. Commençons par un

Lemme 7.1. *Étant donné trois $\mathbb{R}$-espaces vectoriels E, F, G, pour deux applications linéaires quelconques :*

$$f \in \mathscr{L}(E, F) \qquad\qquad \textit{et} \qquad\qquad g \in \mathscr{L}(F, G),$$

l'application composée $g \circ f$ est encore linéaire, c'est-à-dire :

$$g \circ f \in \mathscr{L}(E, G).$$

$$\begin{tikzcd} E \arrow[r, "f"] \arrow[rr, bend left, "g\circ f"] & F \arrow[r, "g"] & G \end{tikzcd}$$

Preuve. Pour tout couple de scalaires $a, b \in \mathbb{R}$, puisque f et g sont linéaires, on a :

$$\begin{aligned} f\big(a\,\vec{x} + b\,\vec{x}\,'\big) &= a\,f(\vec{x}) + b\,f(\vec{x}\,'), && (\forall\,\vec{x} \in E,\ \forall\,\vec{x}\,' \in E) \\ g\big(a\,\vec{y} + b\,\vec{y}\,'\big) &= a\,g(\vec{y}) + b\,g(\vec{y}\,') && (\forall\,\vec{y} \in F,\ \forall\,\vec{y}\,' \in F). \end{aligned}$$

Prenons $\vec{y} := f(\vec{x})$ et $\vec{y}\,' := f(\vec{x}\,')$. Alors on a bien linéarité de $g \circ f$ grâce au calcul :

$$\begin{aligned} g \circ f\,\big(a\,\vec{x} + b\,\vec{x}\,'\big) &= g\big(a\,f(\vec{x}) + b\,f(\vec{x}\,')\big) \\ &= a\ g \circ f(\vec{x}) + b\ g \circ f(\vec{x}\,'). \end{aligned}$$

□

L'énoncé suivant établit la distributivité de la composition par rapport à l'addition, à droite et à gauche :

Théorème 7.2. **(1)** *Soient $f, g \in \mathscr{L}(E, F)$ et $h \in \mathscr{L}(F, G)$. Alors :*

$$h \circ \big(f + g\big) = h \circ f + h \circ g.$$

(2) *Soient $h \in \mathscr{L}(E, F)$ et $f, g \in \mathscr{L}(F, G)$. Alors :*

$$\big(f + g\big) \circ h = f \circ h + g \circ h.$$

Preuve. **(1)** Clairement :

$$\big(f+g\big)(\vec{x}) \,=\, f(\vec{x})+g(\vec{x}) \qquad\qquad (\forall\,\vec{x}\in E).$$

Par l'application linéaire h, on obtient :

$$h\circ\big(f+g\big)(\vec{x}) \,=\, h\big(f(\vec{x})+g(\vec{x})\big) \,=\, h\circ f(\vec{x})+h\circ g(\vec{x}),$$

ce qui est la première partie du théorème.

(2) Pour tout $\vec{x}\in E$, posons :

$$\vec{y} \,:=\, h(\vec{x}).$$

Or par définition de la somme de deux applications linéaires :

$$\big(f+g\big)(\vec{y}) \,=\, f(\vec{y})+g(\vec{y}) \qquad\qquad (\forall\,\vec{y}\in F),$$

ce qui entraîne :

$$\big(f+g\big)\circ h(\vec{x}) \,=\, f\circ h(\vec{x})+g\circ h(\vec{x}),$$

puis termine la deuxième partie, et conclut. □

Maintenant, la composition d'applications linéaires peut être vue comme une 'multiplication' $\times$ qui conduit à une structure d'anneau pour la collection $\mathscr{L}(E)$ des endomorphismes d'un $\mathbb{R}$-espace vectoriel E. Mais attention ! Il y a une grande différence entre $\circ$ et $\times$, car la composition n'est en général pas commutative :

$$f\circ g \,\neq\, g\circ f \qquad\qquad (f,g\in\mathscr{L}(E),\ \text{en général}),$$

nous y reviendrons, nous donnerons des exemples plus tard.

En tout cas, le Lemme 7.1 appliqué à $E=F=G$ montre que la loi de composition :

$$\big(f,g\big) \,\longmapsto\, g\circ f$$

est *interne* dans $\mathscr{L}(E)$, à savoir $g\circ f\in\mathscr{L}(E)$ lorsque $f,g\in\mathscr{L}(E)$. Le Théorème 7.2 montre que cette loi est distributive à droite et à gauche par rapport à l'addition. De plus, cette loi admet un élément neutre pour la composition, à savoir l'application identité, ou de coïncidence :

$$c(\vec{x}) \,:=\, \vec{x} \qquad\qquad (\forall\,\vec{x}\in E).$$

En conclusion, nous pouvons donc énoncer le

Théorème 7.3. *$\mathscr{L}(E)$, muni des opérations $(+,\circ)$, est un anneau à élément unité.* □

Terminologie 7.4. On appelle $\mathscr{L}(E)$ *anneau des endomorphismes* de E.

8. Groupe des automorphismes d'un $\mathbb{R}$-espace vectoriel E

Commençons par quelques rappels généraux de théorie des ensembles.

Terminologie 8.1. On appelle *permutation* toute application *bijective* $E \xrightarrow{\sim} E$ d'un ensemble E sur lui-même.

Les permutations peuvent être composées entre elles, et inversées, puisqu'elles sont bijectives. En fait, une structure plus riche existe, comme l'énonce le

Lemme 8.2. *La collection des permutations d'un ensemble E est toujours un* groupe *pour la composition.*

Preuve. La composition de deux bijections :

$$\begin{tikzcd} E \arrow[r, "f"] \arrow[rr, bend left, "g\circ f"] & E \arrow[r, "g"] & E \end{tikzcd}$$

est encore une bijection, et cette opération de composition admet comme élément neutre l'application identité. Ensuite, cette opération de composition est (trivialement) associative :

$$\begin{tikzcd} E \arrow[r, "f"] \arrow[rr, bend left, "g\circ f"] & E \arrow[r, "g"] & E \arrow[r, "h"] & E \end{tikzcd} \;=\; \begin{tikzcd} E \arrow[r, "f"] & E \arrow[r, "g"] \arrow[rr, bend left, "h\circ g"] & E \arrow[r, "h"] & E, \end{tikzcd}$$

à savoir :

$$h \circ \big(g \circ f\big) \,=\, \big(h \circ g\big) \circ f,$$

mais elle n'est la plupart du temps *pas* commutative, comme nous l'avons signalé plus haut. Enfin, l'inverse d'une bijection est encore une bijection :

$$\begin{tikzcd} E \arrow[r, bend left, "f"] & E. \arrow[l, bend left, "f^{-1}"] \end{tikzcd}$$ □

Quand E est un $\mathbb{R}$-espace vectoriel, on s'intéresse bien évidemment aux permutations qui sont, de plus, linéaires.

Il en existe ! Par exemple, l'application identité $E \longrightarrow E$ est linéaire ! Donnons un autre exemple. Si on munit E, supposé de dimension finie $n \geqslant 1$, d'une base :

$$B \,:=\, \big\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\big\},$$

toute permutation g de B qui envoie chaque vecteur $\vec{e}_i$ sur un certain autre vecteur $\vec{e}_j$ définit une autre base :

$$B' \,:=\, \big\{g(\vec{e}_1), g(\vec{e}_2), \dots, g(\vec{e}_n)\big\} \,=\, \big\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\big\} \,=\, B,$$

qui coïncide avec B à un changement d'ordre près, d'où :

$$\mathsf{Vect}_{\mathbb{R}}(B') \,=\, \mathsf{Vect}_{\mathbb{R}}(B) \,=\, E,$$

et alors, grâce au Théorème 3.9, on sait qu'il existe un endomorphisme $\mathbb{R}$-linéaire unique $f\colon E \longrightarrow E$ qui prolonge g au sens où pour tout indice $1 \leqslant i \leqslant n$, on a :

$$f(\vec{e}_i) \,=\, g(\vec{e}_i),$$

et on sait aussi que f est bijectif, *i.e.* est un *automorphisme linéaire* de E.

Notation 8.3. L'ensemble des automorphismes linéaires d'un $\mathbb{R}$-espace vectoriel E sera noté :

$$\mathscr{A}(E) \,:=\, \big\{f\colon E \xrightarrow{\sim} E \;\text{ linéaires et bijectives}\big\}.$$

Théorème 8.4. *L'ensemble $\mathscr{A}(E)$ des automorphismes linéaires d'un $\mathbb{R}$-espace vectoriel E est un sous-groupe de l'ensemble des permutations de E, envisagé comme ensemble.*

Preuve. Rappelons que, pour qu'une permutation de E appartienne à $\mathscr{A}(E)$, il faut et il suffit qu'elle soit linéaire. Le Lemme 7.1 a déjà fait voir que la composition d'applications linéaires préserve la linéarité :

$$\big(f \,\in\, \mathscr{A}(E) \quad \text{et} \quad g \,\in\, \mathscr{A}(E)\big) \qquad\Longrightarrow\qquad g \circ f \,\in\, \mathscr{A}(E).$$

Prouvons enfin que l'inversion préserve aussi la linéarité :

$$f \,\in\, \mathscr{A}(E) \qquad\Longrightarrow\qquad f^{-1} \,\in\, \mathscr{L}(E).$$

Il s'agit de démontrer que, quels que soient les scalaires $a, b \in \mathbb{R}$, et les vecteurs $\vec{x}, \vec{y} \in E$, on a :

$$f^{-1}\big(a\,\vec{x} + b\,\vec{y}\big) \,=\, a\,f^{-1}(\vec{x}) + b\,f^{-1}(\vec{y}). \tag{8.5}$$

Or puisque f est, entre autres, injective, il suffit de prouver que les deux membres ont même image par f. Pour le membre de droite, on a :

$$f \circ f^{-1}\big(a\,\vec{x} + b\,\vec{y}\big) \,=\, a\,\vec{x} + b\,\vec{y}.$$

Pour le membre de gauche, en utilisant la linéarité de f, on calcule :

$$\begin{aligned} f\Big(a\, f^{-1}(\vec{x}) + b\, f^{-1}(\vec{y})\Big) &= a\ f \circ f^{-1}(\vec{x}) + b\ f \circ f^{-1}(\vec{y}) \\ &= a\,\vec{x} + b\,\vec{y}. \end{aligned}$$

La relation (8.5) est donc vérifiée, et le théorème est ainsi démontré. □

Terminologie 8.6. Le groupe $\mathscr{A}(E)$ des automorphismes linéaires de E se nomme aussi *groupe linéaire*.

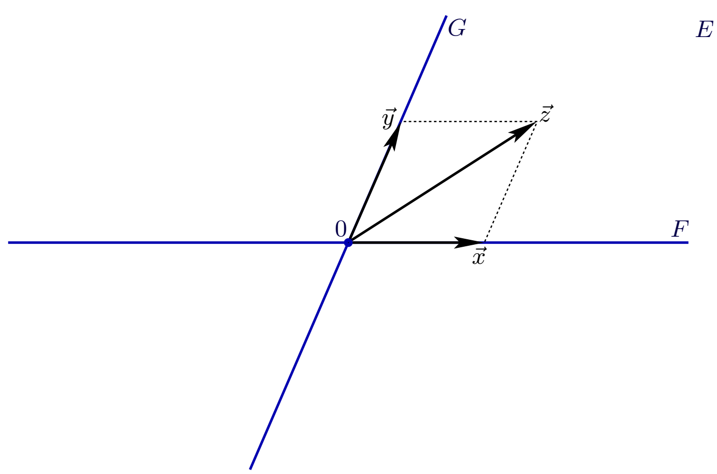
9. Projecteurs

À la fin du chapitre précédent, nous avons vu que deux sous-espaces $F \subset E$ et $G \subset E$ d'un $\mathbb{R}$-espace vectoriel E sont dits *supplémentaires* dans E s'ils sont d'intersection $\{\vec{0}\} = F \cap G$ nulle, et de somme $F + G = E$ qui engendre E, propriété que l'on a notée :

$$F \oplus G = E.$$

Alors pour tout vecteur $\vec{z} \in E$, il existe un unique vecteur $\vec{x} \in F$ et un unique vecteur $\vec{y} \in G$ tels que :

$$\vec{z} = \vec{x} + \vec{y}.$$

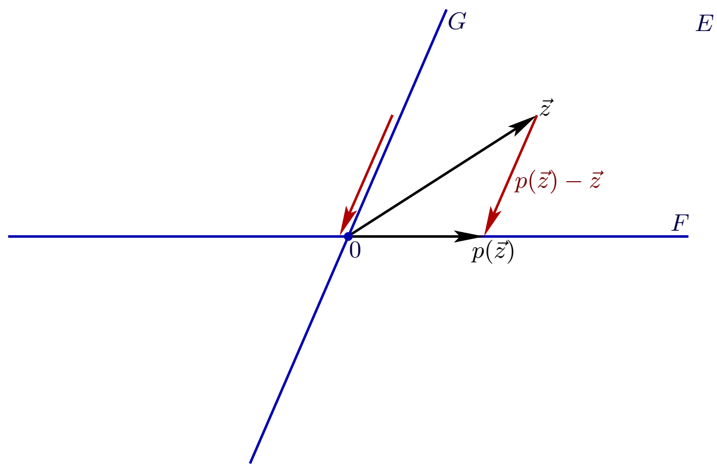


Par analogie avec la géométrie euclidienne, $\vec{x}$ se nomme *projection* de $\vec{z}$ sur F parallèlement à G, tandis que $\vec{y}$ se nomme *projection* de $\vec{z}$ sur G parallèlement à F.

Définition 9.1. F et G étant deux sous-espaces supplémentaires de $E = F \oplus G$, les deux applications $p\colon E \longrightarrow F$ et $q\colon E \longrightarrow G$ telles que :

$$p(\vec{z}) + q(\vec{z}) = \vec{z} \qquad (\forall\, \vec{z} \in E),$$

se nomment respectivement *projecteur de E sur F* parallèlement à G, et *projecteur de E sur G* parallèlement à F.



En d'autres termes, le projecteur p de E sur F associe à tout $\vec{z} \in E$ l'unique vecteur $p(\vec{z}) \in F$ vérifiant :

$$p(\vec{z}) - \vec{z} \in G.$$

De même, le projecteur q de E sur G associe à tout $\vec{z} \in E$ l'unique vecteur $q(\vec{z}) \in G$ tel que :

$$q(\vec{z}) - \vec{z} \in F.$$

Ces projecteurs satisfont quelques propriétés élémentaires.

Lemme 9.2. *Tout projecteur est une application linéaire.*

Preuve. Soit $E = F \oplus G$. Démontrons la linéarité du projecteur $p\colon E \longrightarrow F$.

Pour tout couple de scalaires $a, a' \in \mathbb{R}$ et tout couple de vecteurs $\vec{z}, \vec{z}\,' \in E$, on a :

$$p\big(a\,\vec{z} + a'\,\vec{z}\,'\big) - \big(a\,\vec{z} + a'\,\vec{z}\,'\big) \;\in\; G. \tag{9.3}$$

Or les relations :

$$p(\vec{z}) - \vec{z} \in G \qquad \text{et} \qquad p(\vec{z}\,') - \vec{z}\,' \in G$$

entraînent, puisque G est un sous-espace vectoriel de E :

$$a\,\big(p(\vec{z}) - \vec{z}\big) + a'\,\big(p(\vec{z}\,') - \vec{z}\,'\big) \;\in\; G,$$

soit :

$$a\,p(\vec{z}) + a'\,p(\vec{z}\,') - \big(a\,\vec{z} + a'\,\vec{z}\,'\big) \;\in\; G. \tag{9.4}$$

Comme, pour tout $\vec{y}$ de E, la projection $p(\vec{y})$ est l'unique vecteur de F tel que $p(\vec{y}) - \vec{y} \in G$, alors en comparant (9.3) avec (9.4), on obtient bien la linéarité :

$$p\big(a\,\vec{z} + a'\,\vec{z}\,'\big) \;=\; a\,p(\vec{z}) + a'\,p(\vec{z}\,').$$

Par symétrie, la démonstration de la linéarité de l'autre projecteur $q\colon E \longrightarrow G$ est en tout point analogue, et sera éludée. □

Lemme 9.5. *Si p et q sont les projecteurs de $E = F \oplus G$ sur les sous-espaces supplémentaires F et G, parallèlement à l'autre, alors :*

$$\mathsf{Ker}(p) \;=\; G \qquad\qquad \textit{et} \qquad\qquad \mathsf{Ker}(q) \;=\; F.$$

Démonstration. Puisque, pour tout vecteur $\vec{z} \in E$, on sait que $p(\vec{z})$ est l'unique vecteur tel que $p(\vec{z}) - \vec{z} \in G$, alors, la condition nécessaire et suffisante pour que $p(\vec{z}) = \vec{0}$ est que $\vec{0} - \vec{z} \in G$, soit $\vec{z} \in G$. Donc $\mathsf{Ker}(p) = G$.

La démonstration de $\mathsf{Ker}(q) = F$ est (très) analogue. □

Lemme 9.6. *Tout projecteur :*

$$p\colon\quad E\;=\;F\oplus G\;\longrightarrow\;F\;\subset\;E$$

est un endomorphisme de E vérifiant :

$$p\circ p\;=\;p.$$

Preuve. En effet, puisque $E = F\oplus G$ et puisque p est le projecteur de E sur F, on a :

$$p(\vec{z})-\vec{z}\;\in\;G\qquad\qquad\text{et}\qquad\qquad \mathsf{Ker}(p)\;=\;G,$$

ce qui entraîne :

$$p\big(p\,(\vec{z})-\vec{z}\big)\;=\;\vec{0}\qquad\qquad(\forall\,\vec{z}\in E).$$

Or comme p est linéaire, on en déduit :

$$p\circ p(\vec{z})-p(\vec{z})\;=\;\vec{0},$$

soit effectivement :

$$p\circ p\;=\;p.\qquad\square$$

Au sous-espace F de E ne correspond pas un unique supplémentaire G. On ne peut donc pas dire ‘le’ supplémentaire de F. Mais tous les supplémentaires de F sont isomorphes entre eux, comme le stipule le

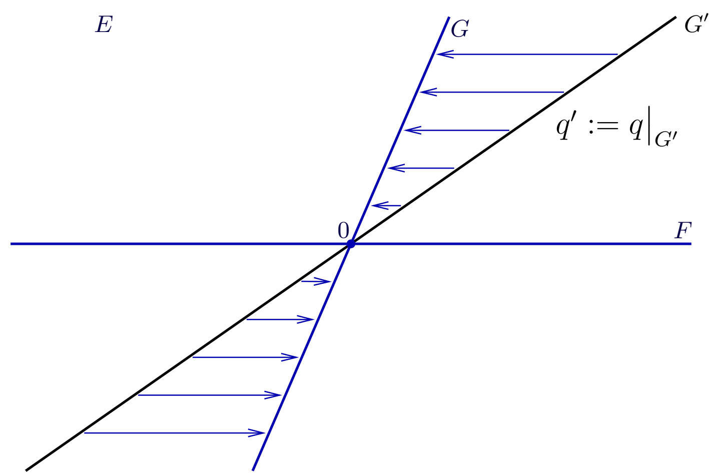
Théorème 9.7. *Soient G et G' deux sous-espaces supplémentaires d’un même sous-espace $F\subset E$ d’un espace vectoriel E :*

$$E\;=\;F\oplus G\;=\;F\oplus G'.$$

Soit q le projecteur de E sur G parallèlement à F. Alors la restriction q' de q à G' :

$$q'\;:=\;q\big|_{G'}\colon\quad G'\;\xrightarrow{\sim}\;G$$

est un isomorphisme linéaire de G' sur G.



Démonstration. Pour commencer, prouvons que cette restriction $q' = q\big|_{G'}$ est bijective.

En effet, puisqu’on a $\mathsf{Ker}(q) = F$, il est clair (exercice) que :

$$\mathsf{Ker}(q')\;=\;F\cap G'.$$

Or F et G' sont supplémentaires, donc $F\cap G' = \{\vec{0}\}$, d’où :

$$\mathsf{Ker}(q')\;=\;\{\vec{0}\},$$

ce qui montre que q' est injective, d’après le Théorème 3.6.

Ensuite, prouvons que q' est surjective, c’est-à-dire, prouvons que :

$$\forall\,\vec{y}\;\in\;G,\qquad\quad \exists\,\vec{x}\;\in\;G'\qquad\quad q'(\vec{x})\;=\;\vec{y}.$$

Or pour tout vecteur $\vec{y} \in E$, et en particulier pour tout $\vec{y} \in G$, il existe une décomposition unique :

$$\vec{y} \,=\, \vec{z} + \vec{x}, \tag{9.8}$$

avec :

$$\vec{z} \in F \qquad \text{et} \qquad \vec{x} \in G'.$$

Mais :

$$\begin{aligned} \vec{z} \in F \qquad &\Longrightarrow \qquad q(\vec{z}) \,=\, \vec{0}, \\ \vec{y} \in G \qquad &\Longrightarrow \qquad q(\vec{y}) \,=\, \vec{y}. \end{aligned}$$

Appliquons alors q aux deux membres de (9.8), ce qui donne :

$$q(\vec{y}) \,=\, q(\vec{z}) + q(\vec{x}),$$

soit :

$$\vec{y} \,=\, \vec{0} + q(\vec{x}) \,=\, q'(\vec{x}).$$

Ainsi, q' est bien surjective.

En définitive, q' est injective et surjective, donc bijective, et puisqu'elle est linéaire, c'est un isomorphisme linéaire de G' sur G. □

Le Théorème 9.7 vient d'être établi pour un espace vectoriel quelconque E, pas forcément de dimension finie, éventuellement de dimension infinie, mais spécifions maintenant le cas où E est de dimension finie.

Si E est de dimension finie $n \geqslant 1$, on peut remarquer que :

$$F \oplus G \,=\, E \qquad \Longrightarrow \qquad \dim F + \dim G \,=\, \dim E,$$

puisque, en prenant une base B de F et une base B' de G, alors $B \cup B'$ est une base de E.

Par conséquent, si $\dim F = p$, alors :

$$F \oplus G \,=\, F \oplus G' \,=\, E \qquad \Longrightarrow \qquad \dim G \,=\, \dim G' \,=\, n - p.$$

Puisque G et G' ont même dimension sur $\mathbb{R}$, le Théorème 4.2 prouve qu'ils sont isomorphes.

On obtient ainsi, dans le cas où E est de dimension finie, une autre démonstration de la principale des propriétés indiquées par le Théorème 9.7.

10. Formes linéaires et espace dual

Dans la Section **6** plus haut, nous avons déjà donné la définition de l'*espace dual* d'un $\mathbb{R}$-espace vectoriel E :

$$E^* \,:=\, \mathscr{L}(E, \mathbb{R}).$$

Cet espace E^* est lui-même un $\mathbb{R}$-espace vectoriel, d'après la même Section **6**. Les éléments de E^* sont les formes linéaires définies dans E, à valeurs dans $\mathbb{R}$.

D'après le Théorème 3.9, si $B = \{\vec{e}_1, \dots, \vec{e}_n\}$ est une base de l'espace E, supposé de dimension finie $n \geqslant 1$, toute forme linéaire $f \in E^*$ est définie de façon unique, dès que l'on se donne les scalaires $c_1, c_2, \dots, c_n$ tels que :

$$f(\vec{e}_1) \,=\, c_1, \;\; f(\vec{e}_2) \,=\, c_2, \;\; \dots\dots, \;\; f(\vec{e}_n) \,=\, c_n.$$

Nous allons maintenant définir n formes linéaires de cette manière, en posant :

$$\begin{aligned} &f_1(\vec{e}_1) \,:=\, 1, \;\; f_1(\vec{e}_2) \,:=\, 0, \;\; \dots\dots, \;\; f_1(\vec{e}_n) \,:=\, 0, \\ &f_2(\vec{e}_1) \,:=\, 0, \;\; f_2(\vec{e}_2) \,:=\, 1, \;\; \dots\dots, \;\; f_2(\vec{e}_n) \,:=\, 0, \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ &f_n(\vec{e}_1) \,:=\, 0, \;\; f_n(\vec{e}_2) \,:=\, 0, \;\; \dots\dots, \;\; f_n(\vec{e}_n) \,:=\, 1. \end{aligned}$$

Plus brièvement, si i et j désignent deux indices parcourant l'ensemble $\{1, 2, \ldots, n\}$, on a :

$$\begin{array}{lcl} i \neq j & \Longrightarrow & f_i(\vec{e}_j) \,:=\, 0, \\ i = j & \Longrightarrow & f_i(\vec{e}_j) \,:=\, 1. \end{array}$$

Théorème 10.1. *Soit E un $\mathbb{R}$-espace vectoriel de dimension $n \geqslant 1$, rapporté à une base $\{\vec{e}_1, \ldots, \vec{e}_n\}$. Les n formes linéaires $f_1, \ldots, f_n$ de l'espace dual E^* définies par :*

$$f_i(\vec{e}_j) \,:=\, \begin{cases} 0 & \textit{lorsque } \ i \neq j, \\ 1 & \textit{lorsque } \ i = j, \end{cases}$$

constituent une base de E^*, et par conséquent :

$$\dim E^* \,=\, \dim E.$$

Démonstration. Commençons par montrer que ces n formes linéaires engendrent l'espace E^*.

En effet, d'après le Théorème 3.9, toute forme linéaire $f \in E^*$ est déterminée de manière unique par la donné de ses valeurs :

$$f(\vec{e}_1) \,=\, c_1, \ \ldots\ldots, \ f(\vec{e}_n) \,=\, c_n.$$

Or pour tout indice $1 \leqslant j \leqslant n$, on peut écrire :

$$c_j \,=\, c_1\, f(\vec{e}_1) + \cdots + c_j\, f(\vec{e}_j) + \cdots + c_n\, f(\vec{e}_n),$$

simplement parce que :

$$\begin{array}{lcl} i \neq j & \Longrightarrow & f_i(\vec{e}_j) \,=\, 0, \\ i = j & \Longrightarrow & f_i(\vec{e}_j) \,=\, 1. \end{array}$$

Par conséquent, on a :

$$f(\vec{e}_j) \,=\, \sum_{i=1}^{n} c_i\, f_i(\vec{e}_j),$$

et d'après le Théorème 3.9, la forme f vérifie dans E^* :

$$f \,=\, c_1\, f_1 + \cdots + c_n\, f_n.$$

Par conséquent, toute forme linéaire $f \in E^*$ est une combinaison linéaire des f_i, les coefficients de la combinaison étant précisément les valeurs c_i qui déterminent f.

Montrons à présent que ces n formes linéaires constituent une famille libre dans E^*.

Partons de la combinaison linéaire nulle suivante :

$$a_1\, f_1 + \cdots + a_n\, f_n \qquad (a_i \in \mathbb{R}).$$

Cela veut dire que, pour tout vecteur $\vec{x} \in E$, on a :

$$a_1\, f_1(\vec{x}) + \cdots + a_n\, f_n(\vec{x}) \,=\, \vec{0}.$$

Prenons successivement :

$$\vec{x} \,:=\, \vec{e}_1, \ \ldots\ldots, \ \vec{x} \,:=\, \vec{e}_n.$$

Pour tout indice $1 \leqslant j \leqslant n$, en prenant $\vec{x} = \vec{e}_j$, et en appliquant la définition des f_i :

$$f_i(\vec{e}_j) \,=\, \begin{cases} 0 & \text{lorsque } \ i \neq j, \\ 1 & \text{lorsque } \ i = j, \end{cases}$$

on obtient :

$$a_j \,=\, 0 \qquad (\forall\, 1 \leqslant j \leqslant n).$$

En définitive :

$$a_1\, f_1 + \cdots + a_n\, f_n \,=\, 0 \qquad \Longrightarrow \qquad a_1 \,=\, \cdots \,=\, a_n \,=\, 0,$$

ce qui montre bien que la famille $\{f_1, \ldots, f_n\}$ est libre.

En conclusion, cette famille constitue une base de E^*, ce qui achève la démonstration. □

Notation 10.2. On nomme $\{f_1, \dots, f_n\}$ *base duale* de $\{\vec{e}_1, \dots, \vec{e}_n\}$, et on la note plutôt

$$\{\vec{e}_1^{\,*}, \dots, \vec{e}_n^{\,*}\}.$$

Terminons ce chapitre par une représentation utile.

Observation 10.3. *Pour tout vecteur $\vec{x} \in E$, on a :*

$$\begin{aligned}\vec{x} &= f_1(\vec{x})\,\vec{e}_1 + \cdots + f_n(\vec{x})\,\vec{e}_n \\ &= \vec{e}_1^{\,*}(\vec{x})\,\vec{e}_1 + \cdots + \vec{e}_n^{\,*}(\vec{x})\,\vec{e}_n.\end{aligned}$$

Preuve. En effet, $\vec{x} = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n$ entraîne, pour tout indice $1 \leqslant j \leqslant n$:

$$f_j(\vec{x}) = x_1\, f_j(\vec{e}_1) + \cdots + x_n\, f_j(\vec{e}_n) = x_j,$$

puisque $f_j(\vec{e}_i) = 0$ pour $i \neq j$ et $f_j(\vec{e}_j) = 1$. □

11. Exercices

Exercice 1. EE

Matrices

François DE MARÇAY
Département de Mathématiques d'Orsay
Université Paris-Sud, France

1. Introduction

2. Étude d'un cas particulier éclairant

Considérons d'abord deux espaces vectoriels F et E, de dimensions respectives 2 et 3, sur un même corps $\mathbb{K}$. On pourra penser, pour fixer les idées, que $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$.

Soient aussi :

$$\begin{aligned} &\big\{\vec{f}_1, \vec{f}_2\big\} && \text{une base de } F, \\ &\big\{\vec{e}_1, \vec{e}_2, \vec{e}_3\big\} && \text{une base de } E. \end{aligned}$$

Considérons enfin une application linéaire f de F dans E :

$$f \in \mathscr{L}\big(F, E\big),$$

c'est-à-dire :

$$F \xrightarrow{f} E.$$

D'après un théorème qui a déjà été vu dans un chapitre qui précède, f est déterminée de façon unique dès que l'on fixe les images $f(\vec{f}_1)$ et $f(\vec{f}_2)$. Donnons-nous alors ces images par leurs coordonnées dans la base $\big\{\vec{e}_1, \vec{e}_2, \vec{e}_3\big\}$ de E :

$$\begin{aligned} f(\vec{f}_1) &= a_{1,1}\,\vec{e}_1 + a_{2,1}\,\vec{e}_2 + a_{3,1}\,\vec{e}_3, \\ f(\vec{f}_2) &= a_{1,2}\,\vec{e}_1 + a_{2,2}\,\vec{e}_2 + a_{3,2}\,\vec{e}_3, \end{aligned}$$

où $a_{1,1}, \dots, a_{3,2}$ sont certains nombres bien déterminés appartenant au corps $\mathbb{K}$.

Pour tout vecteur $\vec{x} \in F$, considérons ses coordonnées dans la base $\big\{\vec{f}_1, \vec{}_2\big\}$:

$$\vec{x} = x_1\,\vec{f}_1 + x_2\,\vec{f}_2.$$

Son image $f(\vec{x}) =: \vec{y}$ est un vecteur $\vec{y} \in E$ dont nous désignons par y_1, y_2, y_3 les coordonnées dans la base $\big\{\vec{e}_1, \vec{e}_2, \vec{e}_3\big\}$:

$$\vec{y} = y_1\,\vec{e}_1 + y_2\,\vec{e}_2 + y_3\,\vec{e}_3.$$

Or on a aussi, puisque f est un homomorphisme linéaire :

$$\begin{aligned} \vec{y} &= f(\vec{x}) \\ &= x_1\,f(\vec{f}_1) + x_2\,f(\vec{f}_2). \end{aligned}$$

En remplaçant alors $f(\vec{f}_1)$ et $f(\vec{f}_2)$ par les valeurs précédemment choisies, nous obtenons :

$$\begin{aligned} \vec{y} = x_1\,\big(a_{1,1}\,\vec{e}_1 + a_{2,1}\,\vec{e}_2 + a_{3,1}\,\vec{e}_3\big) + \\ + x_2\,\big(a_{1,2}\,\vec{e}_1 + a_{2,2}\,\vec{e}_2 + a_{3,2}\,\vec{e}_3\big), \end{aligned}$$

et par conséquent, nous déduisons par identification que :

$$
\left\{
\begin{aligned}
y_1 &= a_{1,1}\,x_1 + a_{1,2}\,x_2, \\
y_2 &= a_{2,1}\,x_1 + a_{2,2}\,x_2, \\
y_3 &= a_{3,1}\,x_1 + a_{3,2}\,x_2.
\end{aligned}
\right.
\tag{2.1}
$$

Terminologie 2.2. Ces relations se nomment *équations de f par rapport à la base $\{\vec{f}_1, \vec{f}_2\}$ de F et à la base $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ de E.*

Ces équations sont caractérisées par les coefficients $a_{i,j}$ rangés sous forme d'un tableau :

$$
\begin{pmatrix}
a_{1,1} & a_{1,2} \\
a_{2,1} & a_{2,2} \\
a_{3,1} & a_{3,2}
\end{pmatrix}
=
\big(a_{i,j}\big)_{1\leqslant i\leqslant 3 \text{ lignes}}^{1\leqslant j\leqslant 2 \text{ colonnes}}.
$$

Terminologie 2.3. Un tel tableau se nomme *matrice à 3 lignes et à 2 colonnes*. Ses éléments $a_{i,j}$ se nomment *termes* ou *entrées* de la matrice.

Convention 2.4. Dans $\big(a_{i,j}\big)$, le premier indice i concernera toujours [1] les lignes, et le deuxième indice j concerne toujours les colonnes.

Mais les lettres peuvent changer ! On pourra aussi écrire parfois :

$$
\big(a_{j,i}\big)_{1\leqslant j\leqslant 3 \text{ lignes}}^{1\leqslant i\leqslant 2 \text{ colonnes}}.
$$

Observation 2.5. *Pour chaque $1 \leqslant j \leqslant 2$, la colonne de rang j de la matrice est constituée par les coordonnées, dans la base $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ de E, de l'image $f(\vec{f}_j)$ du vecteur $\vec{f}_j$ de rang j dans la base $\{\vec{f}_1, \vec{f}_2\}$ de F.*

Nous pouvons abréger la dénomination des deux bases :

$$
\begin{aligned}
\boldsymbol{B}_F &:= \{\vec{f}_1, \vec{f}_2\}, \\
\boldsymbol{B}_E &:= \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}.
\end{aligned}
$$

Terminologie 2.6. On dit aussi que *la matrice est associée à f par rapport à la donnée simultanée de la base $\boldsymbol{B}_F$ de F et de la base $\boldsymbol{B}_E$ de E*. On note alors cette matrice :

$$
\mathsf{Mat}_{\boldsymbol{B}_F \boldsymbol{B}_E}(f) \qquad \text{ou parfois simplement :} \qquad \mathsf{Mat}(f).
$$

Toutefois, la théorie va démontrer qu'*une unique application linéaire $f \in \mathscr{L}(F, E)$ peut donner lieu à des matrices très différentes quand on change les bases.*

Réciproquement, si on se donne une matrice à 3 lignes et à 2 colonnes, dont les termes sont des scalaires appartenant à $\mathbb{K}$, alors les relations (2.1) définissent une application f de F dans E, car tout vecteur $\vec{x} \in F$, de coordonnées (x_1, x_2), est envoyé par cette application sur un vecteur $\vec{y} \in F$, dont les coordonnées (y_1, y_2, y_3) sont justement données par (2.1).

Une vérification fondée sur calcul simple montre que cette application f est *linéaire* de F dans E, c'est-à-dire $f \in \mathscr{L}(F, E)$. Par conséquent, si on désigne par $\mathscr{M}_{3,2}(\mathbb{K})$ l'ensemble des matrices à 3 lignes et à 2 colonnes sur le corps $\mathbb{K}$, on constate que l'application :

$$
\begin{aligned}
\mathscr{L}(F, E) &\longrightarrow \mathscr{M}_{3,2}(\mathbb{K}) \\
f &\longmapsto M(f)
\end{aligned}
$$

est une *bijection*. Mais attention, cette bijection dépend du choix de deux bases $\boldsymbol{B}_F$ et $\boldsymbol{B}_E$!

1. Sauf dans la Section **12**.

3. Passage au cas général

Maintenant que les premières idées intuitives ont bien pénétré la douce cervelle de notre cerveau, nous pouvons généraliser ces notions à deux espaces vectoriels F et E de dimensions finies quelconque sur un corps $\mathbb{K}$ — penser que $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. Soient :

$$n := \dim_{\mathbb{K}} F \qquad\qquad \text{et} \qquad\qquad m := \dim_{\mathbb{K}} E.$$

Choisissons une base $\boldsymbol{B}_F$ de F :

$$\boldsymbol{B}_F = \big\{\vec{f}_1, \vec{f}_2, \dots, \vec{f}_n\big\},$$

ainsi qu'une base $\boldsymbol{B}_E$ de E :

$$\boldsymbol{B}_E = \big\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_m\big\}.$$

Exprimons tout vecteur $\vec{x} \in F$ et tout vecteur $\vec{y} \in E$ par leurs coordonnées dans $\boldsymbol{B}_F$ et dans $\boldsymbol{B}_E$:

$$\vec{x} = x_1\,\vec{f}_1 + x_2\,\vec{f}_2 + \cdots + x_n\,\vec{f}_n = \sum_{j=1}^{n} x_j\,\vec{f}_j, \tag{3.1}$$

$$\vec{y} = y_1\,\vec{e}_1 + y_2\,\vec{e}_2 + \cdots + y_m\,\vec{e}_m = \sum_{i=1}^{m} y_i\,\vec{e}_i. \tag{3.2}$$

Un diagramme résume la situation.

$$\boxed{\begin{array}{ccc}
\boldsymbol{B}_F = \big\{\vec{f}_1, \dots, \vec{f}_n\big\} & & \boldsymbol{B}_E = \big\{\vec{e}_1, \dots, \vec{e}_m\big\} \\[2em]
F & \xrightarrow{\qquad\qquad f \qquad\qquad} & E \\[2em]
\big(x_1, \dots, x_n\big) & & \big(y_1, \dots, y_m\big) \\[2em]
1 \leqslant j \leqslant n & & 1 \leqslant i \leqslant m \\[2em]
\vec{x} = \displaystyle\sum_{j=1}^{n} x_j\,\vec{f}_j & & \vec{y} = \displaystyle\sum_{i=1}^{m} y_i\,\vec{e}_i.
\end{array}}$$

Il faut bien faire la différence entre les vecteurs $\vec{f}_j$, $\vec{x}$, $\vec{e}_i$, $\vec{y}$ qui appartiennent à des espaces vectoriels et les coordonnées x_j, y_i qui sont des nombres scalaires appartenant au corps $\mathbb{K}$.

Ensuite, prenons une application linéaire de F dans E :

$$f \in \mathscr{L}(F, E).$$

Nous savons que f est déterminée de façon unique dès que l'on fixe les images :

$$f(\vec{f}_1),\ f(\vec{f}_2),\ \dots\dots,\ f(\vec{f}_n)$$

des vecteurs de la base $\boldsymbol{B}_F$ de F. Donnons-nous alors ces images par leurs coordonnées dans la base $\boldsymbol{B}_E$ de E :

$$\forall\, 1 \leqslant j \leqslant n: \qquad f(\vec{f}_j) = a_{1,j}\,\vec{e}_1 + a_{2,j}\,\vec{e}_2 + \cdots + a_{m,j}\,\vec{e}_m. \tag{3.3}$$

Il y a donc n relations (3.3), que l'on peut écrire, d'une façon plus condensée :

$$f(\vec{f}_j) = \sum_{i=1}^{m} a_{i,j}\,\vec{e}_i \qquad\qquad (1 \leqslant j \leqslant n).$$

Tout vecteur $\vec{x} \in E$, mis sous la forme (3.1), est envoyé, par l'application linéaire f, sur le vecteur de F suivant :

$$\begin{aligned}\vec{y} = f(\vec{x}) &= x_1\, f(\vec{f}_1) + \cdots + x_n\, f(\vec{f}_n)\\ &= \sum_{j=1}^{n} x_j\, f(\vec{f}_j).\end{aligned}$$

Cherchons les coordonnées de ce vecteur $\vec{y}$ dans la base $\boldsymbol{B}_E$ de E, c'est-à-dire mettons-le sous la forme (3.2) ; à cet effet, remplaçons les $f(\vec{f}_j)$ par leur valeur donnée par (3.3) :

$$\tag{3.4}\begin{aligned}\vec{y} &= \sum_{j=1}^{n} x_j\, f(\vec{f}_j)\\ &= \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{i,j}\,\vec{e}_i\\ &= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{i,j}\, x_j\right)\vec{e}_i,\end{aligned}$$

où nous avons utilisé les propriétés d'associativité et de commutativité de l'addition vectorielle ainsi que celles de la multiplication par des scalaires, et par conséquent, une identification donne :

$$y_i = \sum_{j=1}^{n} a_{i,j}\, x_j \qquad\qquad (1 \leqslant i \leqslant m).$$

Nous avons ainsi obtenu les coordonnées $(y_1, \dots, y_m)$ du vecteur $\vec{y} = f(\vec{x})$ dans la base $\boldsymbol{B}_E$ de E, ce qui s'écrit aussi de manière détaillée :

$$\tag{3.5}\left\{\begin{aligned} y_1 &= a_{1,1}\, x_1 + a_{1,2}\, x_2 + \cdots + a_{1,n}\, x_n,\\ y_2 &= a_{2,1}\, x_1 + a_{2,2}\, x_2 + \cdots + a_{2,n}\, x_n,\\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ y_m &= a_{m,1}\, x_1 + a_{m,2}\, x_2 + \cdots + a_{m,n}\, x_n.\end{aligned}\right.$$

Terminologie 3.6. Ces m équations à n variables $x_1, \dots, x_n$ se nomment *équations de f par rapport aux bases $\boldsymbol{B}_F$ de F et $\boldsymbol{B}_E$ de E.*

Elles peuvent aussi se représenter sous la forme :

$$\begin{pmatrix} y_1\\ y_2\\ \vdots\\ y_m\end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n}\\ a_{2,1} & a_{2,2} & \cdots & a_{2,n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{m,1} & a_{m,2} & \cdots & a_{m,n}\end{pmatrix}\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix},$$

et il est clair qu'elles sont caractérisées par le tableau de nombres scalaires :

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n}\\ a_{2,1} & a_{2,2} & \cdots & a_{2,n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{m,1} & a_{m,2} & \cdots & a_{m,n}\end{pmatrix}.$$

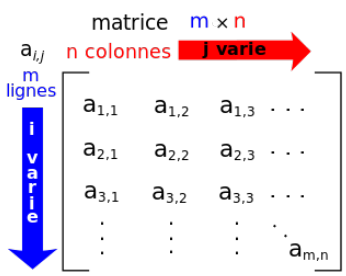
Terminologie 3.7. On nomme un tel tableau *matrice à m lignes et à n colonnes*. On dit aussi que cette matrice est *associée à f par rapport à la base B_F de F et à la base B_E de E*, et on la note :

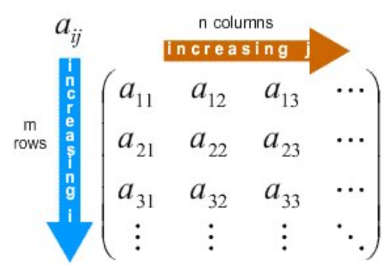
$$\mathsf{Mat}_{B_F B_E}(f) \;=\; \Big(a_{i,j}\Big)_{1\leqslant i\leqslant m \text{ lignes}}^{1\leqslant j\leqslant n \text{ colonnes}}.$$

Voici pour commencer une belle petite matrice à 3 lignes et à 3 colonnes :

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Ensuite, voici une représentation de la matrice générale à m lignes et à n colonnes, dans laquelle on fait bien voir comment le premier indice $a_{i,\bullet}$ augmente en descendant, tandis que le second indice $a_{\bullet,j}$ augmente pendant qu'on progresse vers la droite (sans gilet jaune — tenu correcte exigée !).





Et avec la même chose en version anglaise pour ceux qui aiment les cours de langue Polytech ! Et les bonnes couleurs flashy ! Welcome to London !

Observation 3.8. *Pour tout indice $1 \leqslant j \leqslant n$, la colonne de rang j de cette matrice est constituée par les coordonnées, dans la base B_E de E, de l'image $f(\vec{f}_j)$ du vecteur $\vec{f}_j$ de rang j de la base B_F de F :*

$$\big(a_{i,j}\big) \;=\; \begin{array}{c} \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \cdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \\ \begin{array}{cccc} f(\vec{f}_1) & f(\vec{f}_2) & \cdots & f(\vec{f}_n) \end{array} \end{array}. \qquad \square$$

Réciproquement, si on se donne une matrice à m lignes et à n colonnes, les formules (3.5) définissent une application f de F dans E ; en effet, à tout vecteur $\vec{x} \in F$ de coordonnées $x_1, \dots, x_n$ dans la base B_F, les formules 3.5 font correspondre un vecteur $\vec{y} \in E$ de coordonnées $y_1, \dots, y_m$ dans la base B_E. Ce vecteur $\vec{y}$ est donné par la relation (3.4).

À partir de cette relation (3.4), on peut prouver que cette application f est linéaire ; si on pose :

$$\vec{\ell}_j \;:=\; \sum_{i=1}^{m} a_{i,j}\, \vec{e}_j \qquad (1 \leqslant j \leqslant n),$$

il est facile de voir que, pour tout indice j, on a $\vec{\ell}_j = f(\vec{f}_j)$, et les relations (3.4) s'écrivent alors :

$$\vec{y} \;=\; f(\vec{x}) \;=\; \sum_{j=1}^{n} x_j\, \vec{\ell}_j,$$

et, pour tous scalaires a et a', et tous vecteurs $\vec{x}$ et $\vec{x}\,'$ de E, on a :

$$\begin{aligned} a\, f(\vec{x}) + a'\, f(\vec{x}\,') &= a \sum_{j=1}^{n} x_j\, \vec{\ell}_j + a' \sum_{j=1}^{n} x_j'\, \vec{\ell}_j \\ &= \sum_{j=1}^{n} \big(a\, x_j + a'\, x_j'\big)\, \vec{\ell}_j \;=\; f\big(a\, \vec{x} + a'\, \vec{x}\,'\big), \end{aligned}$$

donc f est bien linéaire.

Notation 3.9. On désignera par :

$$\mathscr{M}_{m,n}(\mathbb{K})$$

l'ensemble des matrices à m lignes et à n colonnes sur le corps $\mathbb{K}$.

L'étude qui précède permet alors d'énoncer le

Théorème 3.10. *Deux bases $\boldsymbol{B}_F$ et $\boldsymbol{B}_E$ étant choisies respectivement dans deux espaces vectoriels F et E, de dimensions n et m sur un corps $\mathbb{K}$, il existe, par rapport à $\boldsymbol{B}_F$ et $\boldsymbol{B}_E$, une bijection unique :*

$$\begin{aligned} \mathscr{L}(F,E) &\longrightarrow \mathscr{M}_{m,n}(\mathbb{K}) \\ f &\longmapsto \mathsf{Mat}_{\boldsymbol{B}_F \boldsymbol{B}_E}. \end{aligned}$$

□

4. Matrices de rotation en géométrie euclidienne planaire

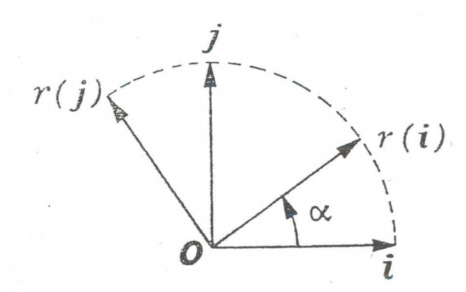
Considérons l'espace vectoriel $\mathbb{R}^2$ de la géométrie euclidienne plane. On sait que toute rotation $\boldsymbol{r}$ autour de l'origine est une application linéaire de $\mathbb{R}^2$ dans lui-même. On a donc $F = E = \mathbb{R}^2$.

Choisissons les bases confondues :

$$\boldsymbol{B}_F = \boldsymbol{B}_E = \{\vec{i}, \vec{j}\},$$

en termes des deux vecteurs de coordonnées :

$$\vec{i} := (1,0) \qquad \text{et} \qquad \vec{j} := (0,1).$$



Si α est l'angle de la rotation $\boldsymbol{r}$, on a :

$$\begin{aligned} \boldsymbol{r}(\vec{i}) &= \cos\alpha \cdot \vec{i} + \sin\alpha \cdot \vec{j}, \\ \boldsymbol{r}(\vec{j}) &= -\sin\alpha \cdot \vec{i} + \cos\alpha \cdot \vec{j}. \end{aligned}$$

La matrice $\mathsf{Mat}(\boldsymbol{r})$ associée à cette rotation dans la base $\{\vec{i}, \vec{j}\}$ est donc :

$$\mathsf{Mat}(\boldsymbol{r}) = \begin{array}{c} \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \\ \begin{array}{cc} \scriptstyle \boldsymbol{r}(\vec{i}) & \quad \scriptstyle \boldsymbol{r}(\vec{j}) \end{array} \end{array}.$$

Observons que la première colonne est constituée des coordonnées $\boldsymbol{r}(\vec{i})$, et la seconde colonne de celles de $\boldsymbol{r}(\vec{j})$.

Un vecteur général $\vec{x} = x_1\,\vec{i} + x_2\,\vec{j}$ devient le vecteur $\boldsymbol{r}(\vec{x}) =: y_1\,\vec{i} + y_2\,\vec{j}$ défini par :

$$\begin{aligned} y_1 &= \cos\alpha \cdot x_1 - \sin\alpha \cdot x_2, \\ y_2 &= \sin\alpha \cdot x_1 + \cos\alpha \cdot x_2. \end{aligned}$$

Enfin, remarquons que toute rotation $\boldsymbol{r}$ de $\mathbb{R}^2$ est caractérisée par son angle de rotation α, et ainsi, cette matrice $\mathsf{Mat}(\boldsymbol{r})$ est *indépendante* de la base *orthonormée* choisie.

5. Matrice ligne et matrice colonne

Soit toujours F un espace vectoriel de dimension $n \geqslant 1$ sur un corps $\mathbb{K}$. On sait que ce corps $\mathbb{K}$ lui-même est un espace vectoriel de dimension 1 sur le corps $\mathbb{K}$, et que la base canonique de $\mathbb{K}$ est constituée de l'unique vecteur $1 \in \mathbb{K}$. Prenons alors comme espace vectoriel d'arrivée :

$$E := \mathbb{K}.$$

Rappelons que toute application linéaire de E dans $\mathbb{K}$ s'appelle une *forme linéaire*. À tout vecteur $\vec{x} \in F$, une forme linéaire $f \in \mathscr{L}(F, \mathbb{K})$ associe un *scalaire* $f(\vec{x})$; et, pour tous scalaires a et a', et tous vecteurs $\vec{x}$ et $\vec{x}\,'$, on a :

$$f\big(a\,\vec{x} + a'\vec{x}\,'\big) = a\,f(\vec{x}) + a'\,f(\vec{x}\,').$$

Si on rapporte F à une base $\boldsymbol{B}_F = \big\{\vec{f}_1, \dots, \vec{f}_n\big\}$ et $\mathbb{K}$ à sa base canonique $\boldsymbol{B}_{\mathbb{K}} = \{1\}$, on obtient, pour image de $\vec{x} = x_1\,\vec{f}_1 + \cdots + x_n\,\vec{f}_n$ le scalaire :

$$f(\vec{x}) = x_1\,f(\vec{f}_1) + \cdots + x_n\,f(\vec{f}_n).$$

Pour caractériser f, il suffit donc de préciser les n scalaires :

$$a_1 := f(\vec{f}_1),\ \dots\dots\dots,\ a_n := f(\vec{f}_n),$$

qui se nomment *coefficients de la forme linéaire relativement à la base* $\boldsymbol{B}_F$.

La matrice $\mathsf{Mat}(f)$ associée à f dans ces bases $\boldsymbol{B}_F$ et $\boldsymbol{B}_{\mathbb{K}}$ comporte alors 1 ligne et n colonnes :

$$\mathsf{Mat}(f) = \big(a_1\ a_2\ \dots\ a_n\big)$$

c'est-à-dire :

$$y = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = a_1\,x_1 + a_2\,x_2 + \cdots + a_n\,x_n.$$

Ensuite, soit E un espace vectoriel de dimension $m \geqslant 1$ sur un corps $\mathbb{K}$, et soit $f\colon \mathbb{K} \longrightarrow E$ une appplication linéaire. Ainsi, on choisit $F := \mathbb{K}$. L'image d'un scalaire $x \in \mathbb{K}$ est un vecteur $f(x) \in E$. Rapportons $\mathbb{K}$ à la base canonique $\boldsymbol{B}_{\mathbb{K}} = \{1\}$, et E à une base $\boldsymbol{B}_E = \big\{\vec{e}_1, \dots, \vec{e}_m\big\}$. Alors :

$$x = x \cdot 1 \qquad\qquad \Longrightarrow \qquad\qquad f(x) = x\,f(1).$$

Pour caractériser f, il suffit par conséquent de préciser l'image $f(1)$ par ses coordonnées dans $\boldsymbol{B}_E$:

$$f(1) = a_1\,\vec{e}_1 + a_2\,\vec{e}_2 + \cdots + a_m\,\vec{e}_m.$$

La matrice associée à une telle application possède donc m lignes et 1 colonne :

$$\mathsf{Mat}(f) = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}.$$

Ainsi :

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix} \big(x\big) \qquad\qquad \Longleftrightarrow \qquad\qquad \begin{cases} y_1 = a_1\,x, \\ y_2 = a_2\,x, \\ \quad\vdots \\ y_m = a_m\,x. \end{cases}$$

6. Espace vectoriel $\mathscr{M}_{m,n}(\mathbb{K})$

Commençons par définir l'addition des matrices. Dans l'ensemble $\mathscr{M}_{m,n}(\mathbb{K})$ des matrices à m lignes et à n colonnes sur le corps $\mathbb{K}$, nous allons définir une addition de telle sorte que la bijection $f \longmapsto \mathsf{Mat}(f)$ de $\mathscr{L}(F,E)$ sur $\mathscr{M}_{m,n}(\mathbb{K})$ soit un *isomorphisme linéaire*.

Pour simplifier les notations, nous désignerons la matrice par son terme général simplement mis entre parenthèses :

$$\big(a_{i,j}\big) = \big(a_{i,j}\big)_{1\leqslant i\leqslant m \text{ lignes}}^{1\leqslant j\leqslant n \text{ colonnes}}.$$

Ici, rappelons que le premier indice i décrit les lignes, tandis que le second j décrit les colonnes.

Soient f et f' deux applications linéaires de F dans E, rapportées à deux bases $\boldsymbol{B}_F$ et $\boldsymbol{B}_E$:

$$\begin{aligned} \boldsymbol{B}_F &= \big\{\vec{f}_1, \dots, \vec{f}_n\big\}, \\ \boldsymbol{B}_E &= \big\{\vec{e}_1, \dots, \vec{e}_m\big\}. \end{aligned}$$

Soient aussi les deux matrices associées, dans ces bases, à f et à f' :

$$\begin{aligned} M(f) &= \mathsf{Mat}_{\boldsymbol{B}_F\boldsymbol{B}_E}(f) = \big(a_{i,j}\big), \\ M(f') &= \mathsf{Mat}_{\boldsymbol{B}_F\boldsymbol{B}_E}(f') = \big(a'_{i,j}\big). \end{aligned}$$

Elles appartiennent à $\mathscr{M}_{m,n}(\mathbb{K})$. Ainsi, on a :

$$\begin{aligned} f(\vec{f}_j) &= \sum_{i=1}^m a_{i,j}\,\vec{e}_i && (1\leqslant j\leqslant n), \\ f'(\vec{f}_j) &= \sum_{i=1}^m a'_{i,j}\,\vec{e}_i && (1\leqslant j\leqslant n), \end{aligned}$$

Par définition de la somme $f + f'$ vue au chapitre qui précède, on a :

$$\begin{aligned} \big(f+f'\big)(\vec{f}_j) &= \sum_{i=1}^m a_{i,j}\,\vec{e}_i + \sum_{i=1}^m a'_{i,j}\,\vec{e}_i \\ &= \sum_{i=1}^m \big(a_{i,j} + a'_{i,j}\big)\,\vec{e}_j && (1\leqslant j\leqslant n). \end{aligned}$$

Par conséquent, la matrice associée à $f + f'$ est :

$$\mathsf{Mat}(f+f') = \big(a_{i,j} + a'_{i,j}\big)_{1\leqslant i\leqslant m \text{ lignes}}^{1\leqslant j\leqslant n \text{ colonnes}}.$$

Ceci justifie la

Définition 6.1. À tout couple $\big(a_{i,j}\big)$ et $\big(a'_{i,j}\big)$ de deux matrices de $\mathscr{M}_{m,n}(\mathbb{K})$, on associe une matrice de $\mathscr{M}_{m,n}(\mathbb{K})$, nommée *somme* de ces matrices, noté et définie par :

$$\big(a_{i,j}\big) + \big(a'_{i,j}\big) := \big(a_{i,j} + a'_{i,j}\big).$$

$$
\begin{aligned}
\big(a_{i,j}\big)+\big(a'_{i,j}\big) &= \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} + \begin{pmatrix} a'_{1,1} & a'_{1,2} & \cdots & a'_{1,n} \\ a'_{2,1} & a'_{2,2} & \cdots & a'_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a'_{m,1} & a'_{m,2} & \cdots & a'_{m,n} \end{pmatrix} \\
&= \begin{pmatrix} a_{1,1}+a'_{1,1} & a_{1,2}+a'_{1,2} & \cdots & a_{1,n}+a'_{1,n} \\ a_{2,1}+a'_{2,1} & a_{2,2}+a'_{2,2} & \cdots & a_{2,n}+a'_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1}+a'_{m,1} & a_{m,2}+a'_{m,2} & \cdots & a_{m,n}+a'_{m,n} \end{pmatrix}
\end{aligned}
$$

Ainsi on a :

$$\mathsf{Mat}(f)+\mathsf{Mat}(f') \;=\; \mathsf{Mat}\big(f+f'\big).$$

La bijection $f \longmapsto \mathsf{Mat}(f)$ de $\mathscr{L}(F,E)$ sur $\mathscr{M}_{m,n}(\mathbb{K})$ est un isomorphisme pour l'addition. Par conséquent :

Lemme 6.2. *$\mathscr{M}_{m,n}(\mathbb{K})$ est un groupe commutatif pour l'addition.* □

L'élément neutre est la matrice nulle, *i.e.* celle dont tous les $m\,n$ éléments sont nuls, et on la notera 0. L'opposé d'une matrice est tout simplement :

$$-\big(a_{i,j}\big) \;:=\; \big(-\,a_{i,j}\big).$$

Ensuite, définissons la multiplication des matrices par un scalaire. Soient donc :

$$f \in \mathscr{L}(F,E) \qquad\qquad \text{et} \qquad\qquad \lambda \in \mathbb{K}.$$

Par définition de l'application linéaire $\lambda\, f$ vue au chapitre qui précède, on a :

$$f(\vec{f}_j) \;=\; \sum_{i=1}^{m} a_{i,j}\,\vec{e}_i \qquad\qquad \Longrightarrow \qquad\qquad (\lambda\, f)(\vec{f}_j) \;=\; \sum_{i=1}^{m} \big(\lambda\, a_{i,j}\big)\,\vec{e}_i.$$

- Calcul de la matrice λA

$$
\lambda \begin{pmatrix} a_{11} & \cdots & \cdots & \cdots & a_{1n} \\ a_{21} & \cdots & \cdots & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & a_{ij} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a_{p1} & \cdots & \cdots & \cdots & a_{pn} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \cdots & \cdots & \cdots & \lambda a_{1n} \\ \lambda a_{21} & \cdots & \cdots & \cdots & \lambda a_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \lambda a_{ij} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \lambda a_{p1} & \cdots & \cdots & \cdots & \lambda a_{pn} \end{pmatrix}
$$

Par conséquent, la matrice associée à $(\lambda\, f)$ est simplement la matrice dont tous les $m\,n$ termes $\big(\lambda\, a_{i,j}\big)$ sont multipliés par λ, ce qui justifie la

Définition 6.3. À toute matrice $\big(a_{i,j}\big)$ de $\mathscr{M}_{m,n}(\mathbb{K})$ et à tout scalaire $\lambda \in \mathbb{K}$, on associe une matrice de $\mathscr{M}_{m,n}(\mathbb{K})$, nommée *produit* de $\big(a_{i,j}\big)$ par λ, notée et définie par :

$$\lambda\,\big(a_{i,j}\big) \;:=\; \Big(\lambda\, a_{i,j}\Big)_{1\leqslant i\leqslant m \text{ lignes}}^{1\leqslant j\leqslant n \text{ colonnes}}.$$

Ainsi, nous avons :

$$\mathsf{Mat}\big(\lambda\, f\big) \,=\, \lambda\, \mathsf{Mat}(f).$$

Par conséquent, nous pouvons énoncer en résumé un

Théorème 6.4. *La bijection :*

$$\begin{aligned} \mathscr{L}(F,E) &\longrightarrow \mathscr{M}_{m,n}(\mathbb{K}) \\ f &\longmapsto M(f) \end{aligned}$$

est un isomorphisme de $\mathbb{K}$*-espaces vectoriels.* □

7. Multiplication des matrices

Après l'addition et la soustraction, que pourrait-être la *multiplication* des matrices ?

La seule réponse candidate à cette question s'inspire de la *composition* des applications linéaires. Il n'y a donc pas à proprement parler de « multiplication », notamment *il n'y a en aucun cas multiplication terme à terme des éléments de deux matrices données* — il faut s'interdire de faire cette bêtise ! Honte, fatale, à l'étudiant(e) idiot(e) de Polytech qui la commettrait !

Au contraire, le bon concept de multiplication va s'inspirer de ce qu'il se passe en présence de *trois* $\mathbb{K}$-espaces vectoriels.

En effet, soit un troisième $\mathbb{K}$-espace vectoriel G de dimension finie :

$$p \,:=\, \dim_{\mathbb{K}} G,$$

et soit un diagramme de composition de deux applications linéaires :

$$\begin{tikzcd} G \arrow[r, "g"] \arrow[rr, bend left, "f\circ g"] & F \arrow[r, "f"] & E. \end{tikzcd}$$

Dans le chapitre qui précède, nous avons vu qu'une telle composition préservait la linéarité :

$$\Big(g \,\in\, \mathscr{L}(G,F) \qquad \text{et} \qquad f \,\in\, \mathscr{L}(F,E)\Big) \qquad\qquad \Longrightarrow \qquad\qquad f\circ g \,\in\, \mathscr{L}(G,E).$$

Choisissons aussi trois bases quelconques de G, F, E :

$$\begin{aligned} \boldsymbol{B}_G &= \big\{\vec{g}_1,\dots,\vec{g}_p\big\}, \\ \boldsymbol{B}_F &= \big\{\vec{f}_1,\dots,\vec{f}_n\big\}, \\ \boldsymbol{B}_E &= \big\{\vec{e}_1,\dots,\vec{e}_m\big\}. \end{aligned}$$

Relativement aux deux paires de bases $\boldsymbol{B}_G$, $\boldsymbol{B}_F$ et $\boldsymbol{B}_F$, $\boldsymbol{B}_E$, soient aussi les deux matrices de ces deux applications linéaires :

$$\mathsf{Mat}(g) \,=\, \Big(b_{j,k}\Big)_{1\leqslant j\leqslant n \text{ lignes}}^{1\leqslant k\leqslant p \text{ colonnes}} \qquad\qquad \text{et} \qquad\qquad \mathsf{Mat}(f) \,=\, \Big(a_{i,j}\Big)_{1\leqslant i\leqslant m \text{ lignes}}^{1\leqslant j\leqslant n \text{ colonnes}}.$$

Observons, car cela sera important, que le nombre n de lignes de $\mathsf{Mat}(g)$ est égal au nombre n de colonnes de $\mathsf{Mat}(f)$, puisque la balle vectorielle F de dimension n est au centre !

Avant de poursuivre, élaborons un diagramme synthétique qui nous fera office de GPS dans cette forêt algébrique.

$$
\boxed{
\begin{array}{ccccc}
\boldsymbol{B}_G = \big\{\vec{g}_1, \dots, \vec{g}_p\big\} & & \boldsymbol{B}_F = \big\{\vec{f}_1, \dots, \vec{f}_n\big\} & & \boldsymbol{B}_E = \big\{\vec{e}_1, \dots, \vec{e}_m\big\} \\[2em]
G & \xrightarrow{\qquad\qquad g \qquad\qquad} & F & \xrightarrow{\qquad\qquad f \qquad\qquad} & E \\[2em]
\big(u_1, \dots, u_n\big) & & \big(x_1, \dots, x_n\big) & & \big(y_1, \dots, y_m\big) \\[2em]
1 \leqslant k \leqslant p & & 1 \leqslant j \leqslant n & & 1 \leqslant i \leqslant m \\[2em]
\vec{u} = \displaystyle\sum_{k=1}^{p} u_k\, \vec{g}_k & & \vec{x} = \displaystyle\sum_{j=1}^{n} x_j\, \vec{f}_j & & \vec{y} = \displaystyle\sum_{i=1}^{m} y_i\, \vec{e}_i.
\end{array}
}
$$

On a évidemment :

$$\mathsf{Mat}(f) \in \mathscr{M}_{m,n}(\mathbb{K}) \qquad\qquad \text{et} \qquad\qquad \mathsf{Mat}(g) \in \mathscr{M}_{n,p}(\mathbb{K}).$$

Dans la concaténation, l'entier n va disparaître :

$$(\bullet)_{m,n}(\bullet)_{n,p} = (\bullet)_{m,p}.$$

On a évidemment aussi :

$$\mathsf{Mat}\big(f \circ g\big) \in \mathscr{M}_{m,p}(\mathbb{K}).$$

Maintenant, écrivons les actions de f et de g dans ces bases :

$$
\begin{aligned}
f(\vec{f}_j) &= \sum_{i=1}^{m} a_{i,j}\, \vec{e}_i && (1 \leqslant j \leqslant n), \\
g(\vec{g}_k) &= \sum_{j=1}^{n} b_{j,k}\, \vec{f}_j && (1 \leqslant k \leqslant p),
\end{aligned}
$$

et pour déterminer l'action de l'application composée $f \circ g \colon G \longrightarrow E$, commençons un calcul très important qui utilise franchement la linéarité de f :

$$
\begin{aligned}
f \circ g\,(\vec{g}_k) = f\big(g(\vec{g}_k)\big) &= f\bigg(\sum_{j=1}^{n} b_{j,k}\, \vec{f}_j\bigg) \\
&= \sum_{j=1}^{n} b_{j,k}\, f(\vec{f}_j) \\
&= \sum_{j=1}^{n} b_{j,k} \sum_{i=1}^{m} a_{i,j}\, \vec{e}_i,
\end{aligned}
$$

et ensuite, utilisons la commutativité de la sommation/addition afin de réorganiser cette double somme en regroupant les termes multiples de chaque vecteur $\vec{e}_i$:

$$\begin{aligned} f \circ g\,(\vec{g}_k) &= \sum_{i=1}^{m} \bigg(\sum_{j=1}^{n} b_{j,k}\, a_{i,j} \bigg)\, \vec{e}_i \\ [\alpha \cdot \beta = \beta \cdot \alpha] \qquad &= \sum_{i=1}^{m} \underbrace{\bigg(\sum_{j=1}^{n} a_{i,j}\, b_{j,k} \bigg)}_{=:\, c_{i,k}}\, \vec{e}_i, \end{aligned}$$

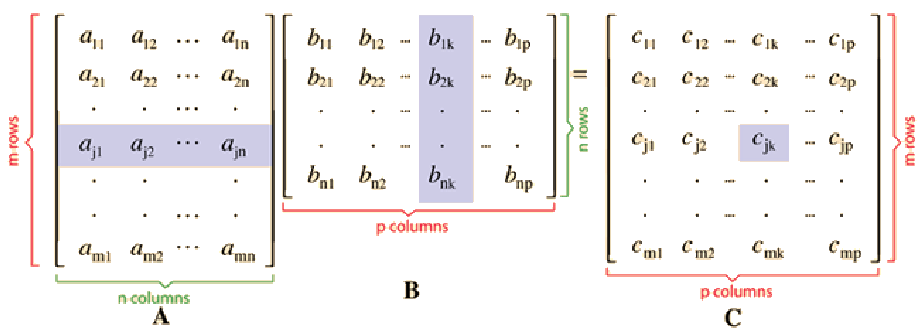
et par conséquent, si nous posons comme cela vient d'être souligné par en-dessous :

$$c_{i,k} := \sum_{j=1}^{n} a_{i,j}\, b_{j,k} \qquad\qquad (1 \leqslant i \leqslant m,\ 1 \leqslant k \leqslant p),$$

nous obtenons une représentation adéquate de l'action de $f \circ g$ sur les vecteurs de la base $\boldsymbol{B}_G$ de G :

$$f \circ g\,(\vec{g}_k) = \sum_{i=1}^{m} c_{i,k}\, \vec{e}_i \qquad\qquad (1 \leqslant k \leqslant p),$$

ce qui nous permet de conclure que nous venons effectivement de calculer la *matrice* de cette composée $f \circ g$ dans la paire de base(kets ?) $\boldsymbol{B}_G, \boldsymbol{B}_E$!



Tout ce beau travail de calculs utiles justifie alors la

Définition 7.1. À toute paire de matrices de tailles $m \times n$ et $n \times p$:

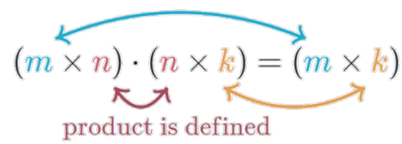
$$\big(a_{i,j}\big)_{1 \leqslant i \leqslant m}^{1 \leqslant j \leqslant n} \qquad \text{et} \qquad \big(b_{j,k}\big)_{1 \leqslant j \leqslant n}^{1 \leqslant k \leqslant p}$$

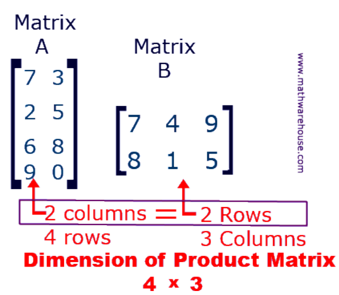
dont le nombre n de colonnes de la première *coïncide* avec le nombre n de lignes de la seconde, on associe une matrice :

$$\big(c_{i,k}\big)_{1 \leqslant i \leqslant m}^{1 \leqslant k \leqslant p} := \big(a_{i,j}\big)_{1 \leqslant i \leqslant m}^{\underline{1 \leqslant j \leqslant n}} \cdot \big(b_{j,k}\big)_{\underline{1 \leqslant j \leqslant n}}^{1 \leqslant k \leqslant p},$$

appelée *multiplication* de $\big(a_{i,j}\big)$ par $\big(b_{j,k}\big)$, dans laquelle a disparu le nombre n, et qui est définie par :

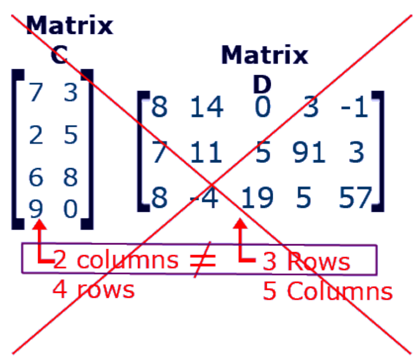
$$\boxed{\big(a_{i,j}\big)_{1 \leqslant i \leqslant m}^{\underline{1 \leqslant j \leqslant n}} \cdot \big(b_{j,k}\big)_{\underline{1 \leqslant j \leqslant n}}^{1 \leqslant k \leqslant p} := \bigg(\sum_{j=1}^{n} a_{i,j}\, b_{j,k} \bigg)_{1 \leqslant i \leqslant m}^{1 \leqslant k \leqslant p}.}$$





Malheureusement, plusieurs mises en garde doivent être formulées.

$\triangle$ Cela n'a aucun sens de composer $g \circ f$, puisque f arrive dans E, tandis que G part de $G \neq E$. C'est seulement $f \circ g$ qui a un sens, car nous avons fait l'hypothèse que g arrive dans l'espace F, et que f part du *même* espace F.



$\triangle$ Par conséquent, la multiplication de matrices dans l'autre sens $(b_{j,k}) \cdot (a_{i,j})$ n'a *pas de sens en général*, y compris parce qu'il est *absolument nécessaire* d'avoir coïncidence des nombres de colonnes et de lignes qui ont été soulignés dans l'encadré ci-dessus, sachant qu'on n'a pas en général $p = m$.

$\triangle$ Même lorsque $p = m = n$, cette multiplication de matrices n'est en général *pas* commutative, comme le montre l'exemple :

$$A := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{et} \qquad B := \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix},$$

qui oblige à constater que $A \cdot B \neq B \cdot A$:

$$A \cdot B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 \cdot 1 + 1 \cdot 0 & 0 \cdot 1 + 1 \cdot 0 \\ 1 \cdot 1 + 0 \cdot 0 & 1 \cdot 1 + 0 \cdot 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix},$$

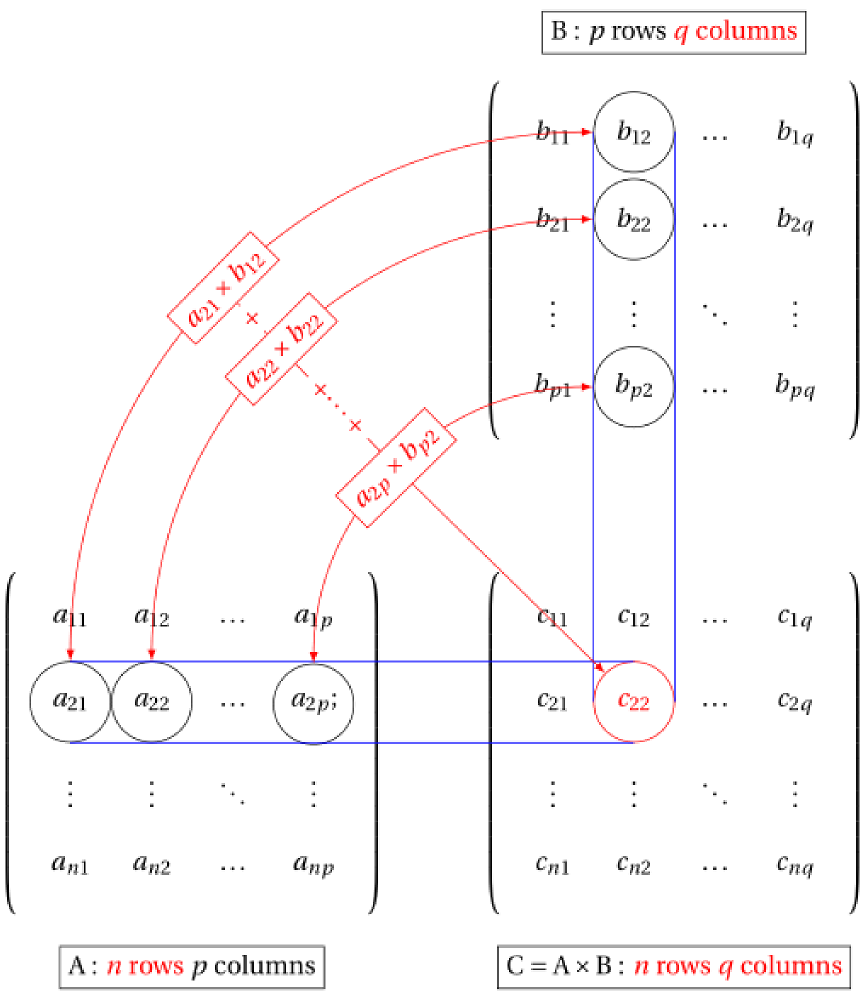
$$B \cdot A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 \cdot 0 + 1 \cdot 1 & 1 \cdot 1 + 1 \cdot 0 \\ 0 \cdot 0 + 0 \cdot 1 & 0 \cdot 1 + 0 \cdot 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

$\triangle$ Aussi, bien qu'on parle de *multiplication* de matrices, l'analogie avec la multiplication des nombres réels est plutôt trompeuse !

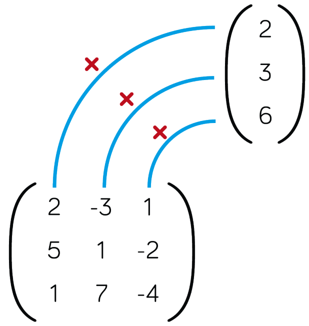
△ Il vaudrait mieux dire « *composition* » des matrices, puisqu'on *compose* des applications linéaires.

Maintenant, voici de nombreuses schématisations, dans des notations parfois différentes, de la manière dont les *colonnes* de la matrice à droite $\big(b_{j,k}\big)$ se *rabattent* sur les lignes de la matrice à gauche $\big(a_{i,j}\big)$, en tombant comme des sapins géants coupés par un impitoyable bûcheron-mathématicien (le prof, donc…).

Commençons par la meilleure de toutes ces illustrations. Elle a probablement dû demander plusieurs jours de travail au pauvre malheureux auquel elle a été *volée* sur internet !



Et comme elle est assez complète et complexe, il s'avère utile d'examiner d'autres illustrations plus simples qui permettront de se former les bonnes intuitions de la multiplication entre matrices. Que doit faire l'étudiant ? Il doit *étudier* ! Alors on lui montre plein de figures pour qu'il les *étudie*.



Oui, quelque chose se rabat de la verticale vers l'horizontale, afin de s'accoupler comme deux ADNs : c'est cette intuition dynamique qu'il faut conserver en mémoire, la plus riche d'entre toutes.

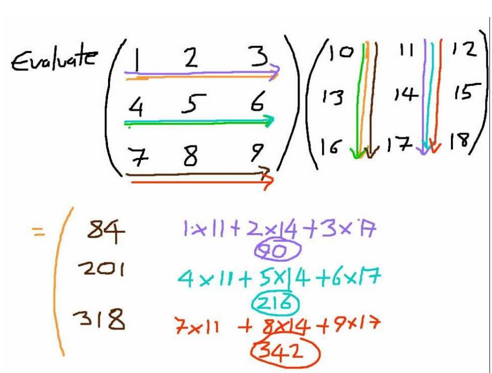
Passons maintenant à des illustrations moins dynamiques, bien que présentées en couleurs. Quand $n = 1$, un seul sapin à droite tombe au sol à gauche.

$$\begin{bmatrix} 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \\ 3 \end{bmatrix} = 1 \times 5 + 2 \times 2 + 0 \times 3 = 9 \qquad \begin{bmatrix} 1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} = 1 \times 1 + 2 \times 3 + 0 \times 4 = 7$$

$$\begin{bmatrix} 4 & 3 & -1 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \\ 3 \end{bmatrix} = 4 \times 5 + 3 \times 2 - 1 \times 3 = 23 \qquad \begin{bmatrix} 4 & 3 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} = 4 \times 1 + 3 \times 3 - 1 \times 4 = 9$$

$$\mathbf{x}^T \mathbf{y} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \sum_{i=1}^{n} x_i y_i$$

Voici encore de belles illustrations numériques. Exercice impératif : vérifier qu'il n'y a pas d'erreur de calcul ! Une première :



Une deuxième :

-3	3	0	-10		-8	-5		-98	84
7	-6	9	2		-4	-7		89	-83
-9	9	3	-8		11	-8		-19	30
					11	-9			

(-3)(-8)+3(-4)+0*11+(-10*11)=-98

-3	3	0	-10		-8	-5		-98	84
7	-6	9	2		-4	-7		89	-83
-9	9	3	-8		11	-8		-19	30
					11	-9			

7*(-8)+(-6)(-4)+9*11+2*11=89

Une troisième :

$$\begin{bmatrix} -1 & 4 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} 9 & -3 \\ 6 & 1 \end{bmatrix} = \begin{bmatrix} (-1)(9)+4(6) & (-1)(-3)+(4)(1) \\ 2(9)+3(6) & 2(-3)+3(1) \end{bmatrix}$$

$$= \boxed{\begin{bmatrix} 15 & 7 \\ 36 & -3 \end{bmatrix}}$$

Final answer

Sur le plan formel, c'est-à-dire quand les termes d'une matrice ne sont pas des nombres entiers explicites, mais des lettres, la multiplication de deux matrices 2×2 est la suivante.

$$\underset{A}{\begin{bmatrix} a & b \\ c & d \end{bmatrix}} \times \underset{B}{\begin{bmatrix} e & f \\ g & h \end{bmatrix}} = \underset{C}{\begin{bmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{bmatrix}}$$

Observons ici dans un cadre formel général que le produit dans un sens, puis dans un autre sens, de deux matrices 2×2 :

$$\begin{aligned} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix} &= \begin{pmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{pmatrix}, \\ \begin{pmatrix} e & f \\ g & h \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} &= \begin{pmatrix} ea+fc & eb+fd \\ ga+hc & gb+hd \end{pmatrix}, \end{aligned}$$

ne peut donner le même résultat que lorsque les 4 équations suivantes sont satisfaites :

$$\begin{aligned} ae+bg &= ea+fc, & af+bh &= eb+fd, \\ ce+dg &= ga+hc, & cf+dh &= gb+hd, \end{aligned}$$

et il est intuitivement clair que pour la plupart des choix possibles de nombres rationnels :

$$a,\, b,\, c,\, d,\, e,\, f,\, g,\, h \;\in\; \mathbb{Q},$$

ces quatre équations ne sont *pas* satisfaites, et ainsi, le produit entre les deux matrices correspondantes n'est *pas* commutatif. Évidemment, en dimension $n \geqslant 3$, ce sera encore pire ! Il y aura encore moins de chances que le produit entre deux matrices quelconques soit commutatif !

Or lorsqu'il s'agit de deux matrices 3×3, les choses se corsent. Voici ce qu'on écrivait lors du précédent millénaire, à Chicago, avec une vraie machine à écrire, ses touches en plomb, la cigarette entre les dents, et le *gun* à portée de main.

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \times \begin{bmatrix} j & k & l \\ m & n & o \\ p & q & r \end{bmatrix} =$$

$$\begin{bmatrix} aj + bm + cp & ak + bn + cq & al + bo + cr \\ dj + em + fp & dk + en + fq & dl + eo + fr \\ gj + hm + ip & gk + hn + iq & gl + ho + ir \end{bmatrix}$$

Voici le même produit dans un style plus contemporain, volé, lui aussi, sur internet, sans vergogne, et qui a des couleurs plutôt discrètes.

$$a_{11} \times b_{11} + a_{12} \times b_{21} + a_{13} \times b_{31} = c_{11}$$

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \times \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$$

Certaines fois, on peut multiplier des matrices dans les deux sens, par exemple une matrice 3×2 que multiplie une matrice 2×3, ainsi que dans l'autre sens :

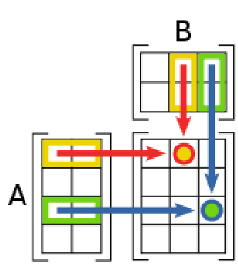
$$A := \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} \end{pmatrix}, \qquad B := \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \end{pmatrix},$$

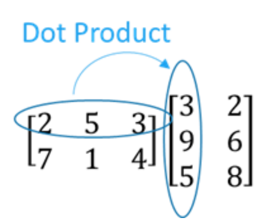
d'où :

$$A \cdot B = \begin{pmatrix} a_{1,1}b_{1,1} + a_{1,2}b_{2,1} & a_{1,1}b_{1,2} + a_{1,2}b_{2,2} & a_{1,1}b_{1,3} + a_{1,2}b_{2,3} \\ a_{2,1}b_{1,1} + a_{2,2}b_{2,1} & a_{2,1}b_{1,2} + a_{2,2}b_{2,2} & a_{2,1}b_{1,3} + a_{2,2}b_{2,3} \\ a_{3,1}b_{1,1} + a_{3,2}b_{2,1} & a_{3,1}b_{1,2} + a_{3,2}b_{2,2} & a_{3,1}b_{1,3} + a_{3,2}b_{2,3} \end{pmatrix},$$

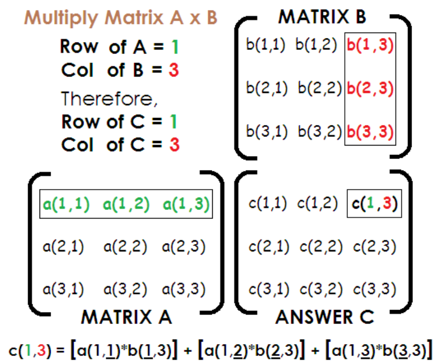
$$B \cdot A = \begin{pmatrix} b_{1,1}a_{1,1} + b_{1,2}a_{2,1} + b_{1,3}a_{3,1} & b_{1,1}a_{1,2} + b_{1,2}a_{2,2} + b_{1,3}a_{3,2} \\ b_{2,1}a_{1,1} + b_{2,2}a_{2,1} + b_{2,3}a_{3,1} & b_{2,1}a_{1,2} + b_{2,2}a_{2,2} + b_{2,3}a_{3,2} \end{pmatrix}$$

Voici encore deux autres illustrations intuitives en petites dimensions :

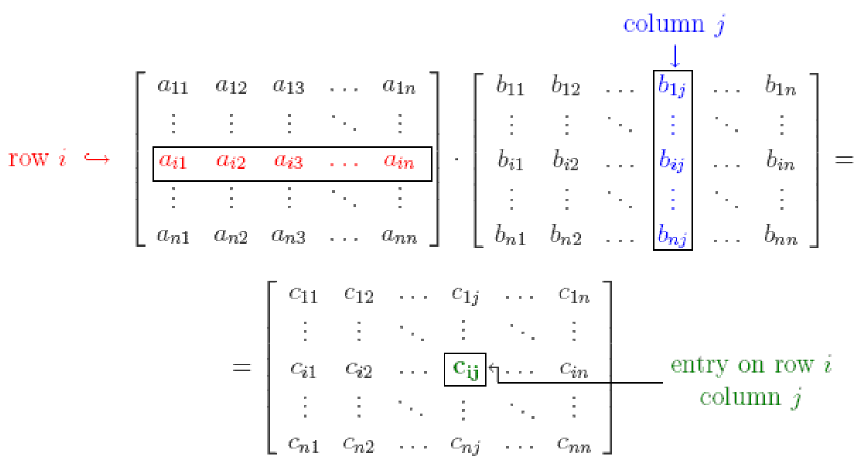




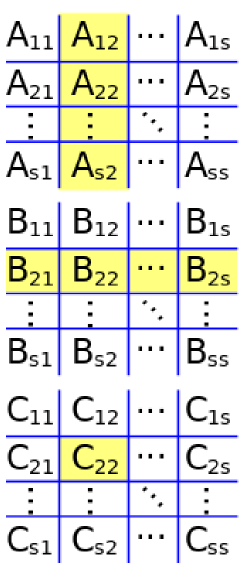
Et encore une :

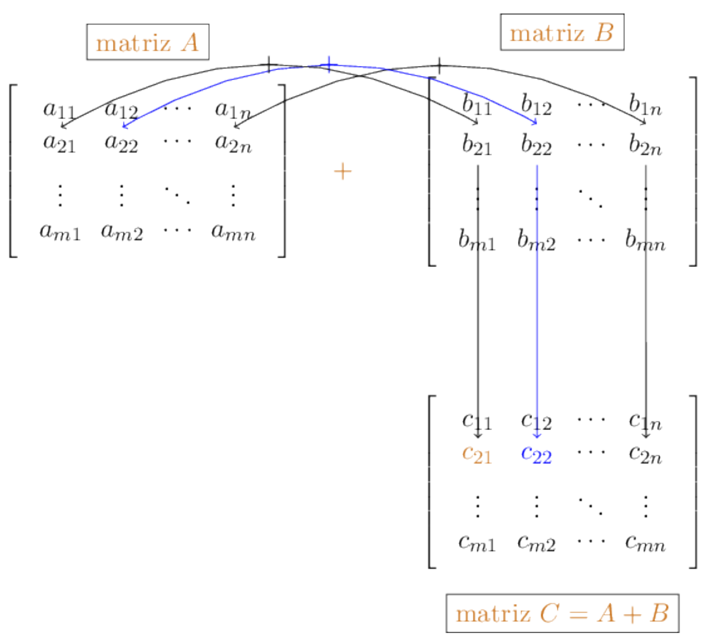


Enfin, redonnons un graphique en dimension quelconque :



et encore deux autres (on voit bien que le tout dernier a été volé ! « *matriz* » !) :





Voilà, une fois terminé cet interlude joyeux et fou, reprenons le chemin de la théorie austère et froide, le dos courbé, la cervelle en surchauffe, et le stylo droit comme un I.

Par définition de la multiplication matricielle, on a :

$$\mathsf{Mat}(f)\,\mathsf{Mat}(g) \;=\; \mathsf{Mat}\big(f \circ g\big).$$

Maintenant, si H est un quatrième $\mathbb{K}$-espace vectoriel de dimension finie $q \geqslant 1$, et si $h\colon H \longrightarrow G$ est une troisième application linéaire, de telle sorte qu'on a le diagramme :

$$H \xrightarrow{\quad h \quad} G \xrightarrow{\quad g \quad} F \xrightarrow{\quad f \quad} E,$$

le fait que la composition d'applications entre ensembles est associative :

$$\big[f \circ g\big] \circ h \;=\; f \circ \big[g \circ h\big],$$

implique immédiatement que la multiplication entre les matrices correspondantes est associative :

$$\big[\big(a_{i,j}\big) \cdot \big(b_{j,k}\big)\big] \cdot \big(c_{k,l}\big) \;\;=\;\; \big(a_{i,j}\big) \cdot \big[\big(b_{j,k}\big) \cdot \big(c_{k,\ell}\big)\big],$$

où, comme on l'aura deviné :

$$\Big(c_{k,\ell}\Big)_{1\leqslant k\leqslant p \text{ lignes}}^{1\leqslant \ell\leqslant q \text{ colonnes}} \;=\; \mathsf{Mat}(h).$$

Ainsi, avec quatre entiers $q, p, n, m \geqslant 1$, pour trois matrices quelconques :

$$M \,\in\, \mathscr{M}_{m,n}(\mathbb{K}), \qquad\qquad M' \,\in\, \mathscr{M}_{n,p}(\mathbb{K}), \qquad\qquad M'' \,\in\, \mathscr{M}_{p,q}(\mathbb{K}),$$

bien qu'on n'ait pas commutativité de la multiplication, on a au moins *associativité :*

$$\big[M \cdot M'\big] \cdot M'' \;=\; M \cdot \big[M' \cdot M''\big].$$

De plus, on déduit de propriétés déjà vues dans le chapitre qui précède que la multiplication matricielle est distributive, à gauche comme à droite, par rapport à l'addition matricielle. Plus précisément, étant donné deux paires d'applications représentées par le diagramme :

$$G \xrightarrow[\quad g_2 \quad]{\quad g_1 \quad} F \xrightarrow[\quad f_2 \quad]{\quad f_1 \quad} E,$$

on a toujours :

$$\begin{aligned} f_1 \circ \big(g_1 + g_2\big) \;&=\; f_1 \circ g_1 + f_1 \circ g_2, \\ \big(f_1 + f_2\big) \circ g_1 \;&=\; f_1 \circ g_1 + f_2 \circ g_1, \end{aligned}$$

et on en déduit que pour toutes matrices :

$$M_1,\, M_2 \,\in\, \mathscr{M}_{m,n} \qquad\qquad \text{et} \qquad\qquad M_1',\, M_2' \,\in\, \mathscr{M}_{n,p},$$

on a ladite distributivité :

$$\begin{aligned} M_1 \cdot \big(M_1' + M_2'\big) \;&=\; M_1 \cdot M_1' + M_1 \cdot M_2', \\ \big(M_1 + M_2\big) \cdot M_1' \;&=\; M_1 \cdot M_1' + M_2 \cdot M_1'. \end{aligned}$$

Enfin, pour terminer cette section, revenons à l'écriture matricielle (3.5) des applications linéaires :

$$y_i \;=\; \sum_{j=1}^{n} a_{i,j}\, x_j \qquad\qquad (1 \leqslant i \leqslant m).$$

Comme nous l'avons implicitement anticipé, ces formules peuvent s'interpréter comme multiplication matricielle. En effet, introduisons les deux matrices colonnes :

$$X \;:=\; \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad\qquad \text{et} \qquad\qquad Y \;:=\; \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix},$$

qui sont la matrice X des coordonnées du vecteur $\vec{x}$ dans la base $\boldsymbol{B}_F$ de F, et la matrice Y des coordonnées du vecteur $\vec{y}$ dans la base $\boldsymbol{B}_E$ de E.

D'après la définition de la multiplication d'une matrice de $\mathscr{M}_{m,n}(\mathbb{K})$ par une matrice de $\mathscr{M}_{n,1}(\mathbb{K})$, les formules en question expriment que pour tout indice $1 \leqslant i \leqslant m$, la coordonnée y_i est le terme de la ligne i de la matrice produit de $A := \big(a_{i,j}\big)$ par X, et par conséquent, on peut écrire ces m équations d'un seul bloc comme simple produit matriciel :

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{2,1} & \cdots & a_{n,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,m} & a_{2,m} & \cdots & a_{n,m} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

D'une façon condensée, nous pouvons aussi abréger le tout sous une forme :

$$Y = A \cdot X,$$

qui est la *traduction matricielle de la relation* $\vec{y} = f(\vec{x})$.

De même, l'application $g\colon G \longrightarrow F$, à savoir la relation $\vec{x} = g(\vec{u})$, s'écrit sous forme condensée :

$$X = B \cdot U,$$

et sous forme de polystyrène expansé comme :

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_{1,1} & b_{2,1} & \cdots & b_{p,1} \\ b_{1,2} & b_{2,2} & \cdots & b_{p,2} \\ \vdots & \vdots & \ddots & \vdots \\ b_{1,n} & b_{2,n} & \cdots & b_{p,n} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_p \end{pmatrix}.$$

Observons avant de changer de section que la composition s'écrit en abrégé :

$$Y = A \cdot X = \underbrace{A \cdot B}_{=:\, C} \cdot U,$$

c'est-à-dire avec tout le polystyrène qu'on aime et qu'on Dior-adore en se roulant dans le pop-corn :

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} = \underbrace{\begin{pmatrix} a_{1,1} & a_{2,1} & \cdots & a_{n,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,m} & a_{2,m} & \cdots & a_{n,m} \end{pmatrix} \begin{pmatrix} b_{1,1} & b_{2,1} & \cdots & b_{p,1} \\ b_{1,2} & b_{2,2} & \cdots & b_{p,2} \\ \vdots & \vdots & \ddots & \vdots \\ b_{1,n} & b_{2,n} & \cdots & b_{p,n} \end{pmatrix}}_{\text{multiplication matricielle}} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_p \end{pmatrix}.$$

8. Anneau des matrices carrées d'ordre n

Quand $m = n$, l'ensemble $\mathscr{M}_{m,n}(\mathbb{K}) = \mathscr{M}_{n,n}(\mathbb{K})$ jouit de propriétés particulières.

Terminologie 8.1. On appelle *matrice carrée* toute matrice dont le nombre de lignes égale celui des colonnes. Ce nombre s'appelle *ordre* de la matrice carrée. L'ensemble des matrices carrées sur le corps $\mathbb{K}$ sera noté :

$$\mathscr{M}_n(\mathbb{K}) \equiv \mathscr{M}_{n,n}(\mathbb{K}).$$

Ainsi, une matrice carrée d'ordre $n \geqslant 1$ est associée à une application linéaire $f \in \mathscr{L}(F, E)$ lorsque :

$$n = \dim F = \dim E.$$

Cela est en particulier le cas lorsque $F = E$, avec $n = \dim E$. Si, dans ce dernier cas, on choisit une base $\boldsymbol{B}_E$ de E, l'application $f \longmapsto \mathsf{Mat}(f)$ est une bijection de $\mathscr{L}(E)$ sur $\mathscr{M}_n(\mathbb{K})$.

Cette bijection est un isomorphisme pour l'addition :

$$\mathsf{Mat}\big(f+g\big) \;=\; \mathsf{Mat}(f)+\mathsf{Mat}(g),$$

et c'est un isomorphisme pour la multiplication, ou composition :

$$\mathsf{Mat}\big(f\circ g\big) \;=\; \mathsf{Mat}(f)\cdot\mathsf{Mat}(g).$$

Ici, les multiplications, comme les additions, sont des lois de composition *internes*, à savoir on a :

$$f,\, g \,\in\, \mathscr{M}_n(\mathbb{K}) \qquad\Longrightarrow\qquad \begin{cases} f+g \,\in\, \mathscr{M}_n(\mathbb{K}), \\ f\circ g \,\in\, \mathscr{M}_n(\mathbb{K}). \end{cases}$$

Or nous avons vu au chapitre précédent que $\mathscr{L}(E)$ est un anneau à élément unité. Par isomorphie, on peut donc énoncer le

Théorème 8.2. *L'ensemble $\mathscr{M}_n(\mathbb{K})$ des matrices carrées d'ordre n sur $\mathbb{K}$ est un anneau à élément unité.* □

Rappelons que nous avons, à deux reprises, insisté sur le fait que cet anneau n'est *pas* commutatif.

L'élément unité de cet anneau n'est autre que la matrice qui correspond à l'application identité :

$$\begin{aligned} \mathsf{Id}\colon\;\; E &\longrightarrow E \\ \vec{x} &\longmapsto \vec{x}, \end{aligned}$$

et dont la matrice est (exercice) :

$$\mathsf{Mat}(\mathsf{Id}) \;:=\; I_n \;:=\; \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

9. Matrices scalaires, diagonales et triangulaires

Commençons par introduire un objet mathématique qui permet souvent d'économiser et de contracter l'écriture des textes.

Notation 9.1. On appelle *symbole de Kronecker* les quantités $\delta_{i,j}$, définies pour des paires d'indices $1 \leqslant i,j \leqslant n$, par :

$$\delta_{i,j} \;:=\; \begin{cases} 1 & \text{lorsque}\;\; i=j, \\ 0 & \text{lorsque}\;\; i\neq j. \end{cases}$$

Il est symétrique par rapport à ses deux indices :

$$\delta_{j,i} \;=\; \delta_{i,j} \qquad\qquad {\scriptstyle (1\leqslant i,j\leqslant n)}.$$

En ces termes, nous pouvons alors écrire la matrice identité sous forme très condensée :

$$I_n \;=\; \big(\delta_{i,j}\big) \;=\; \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Observation 9.2. *Dans l'ensemble $\mathscr{M}_n(\mathbb{K})$ des matrices carrées de taille $n \times n$, la matrice identité est l'élément neutre pour la multiplication de matrices :*

$$A \cdot I_n = I_n \cdot A = A \qquad (\forall\, A \in \mathscr{M}_n(\mathbb{K})),$$

c'est-à-dire :

$$\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix},$$

ainsi que :

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \cdot \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}. \qquad \square$$

Introduisons maintenant les matrices scalaires. Comme précédemment, soit E un espace vectoriel de dimension $n \geqslant 1$ sur un corps $\mathbb{K}$. Soit $\lambda \in \mathbb{K}$ un scalaire fixé, et soit l'homothétie de rapport λ :

$$f\colon \quad \vec{x} \longmapsto \lambda\, \vec{x}.$$

Choisissons aussi une base de E :

$$\boldsymbol{B}_E = \big\{\vec{e}_1, \dots, \vec{e}_n\big\}.$$

On a donc pour tout $1 \leqslant i \leqslant n$:

$$f(\vec{e}_i) = \lambda\, \vec{e}_i.$$

Dans la matrice associée $\mathsf{Mat}(f)$, la colonne de rang i est constituée de zéros, sauf à la ligne i où l'élément vaut λ.

Par conséquent, en introduisant les symboles de Kronecker, nous avons :

$$\mathsf{Mat}(f) = \big(\lambda\, \delta_{i,j}\big) = \lambda\, I_n,$$

ce qui s'écrit sous forme détaillée :

$$\mathsf{Mat}(f) = \begin{pmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{pmatrix} = \lambda \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Terminologie 9.3. La matrice $\mathsf{Mat}(f) \in \mathscr{M}_n(\mathbb{K})$ d'une homothétie $f = \lambda\, \mathsf{Id}$ se nomme *matrice scalaire*.

On peut vérifier que la matrice scalaire, associée à une homothétie de rapport λ, est *indépendante* de la base $\boldsymbol{B}_E$ choisie pour E.

Passons maintenant à des matrices un peu plus générales, au sens où leurs éléments non nuls ne sont pas tous égaux à un même scalaire λ.

Définition 9.4. On appelle *matrice diagonale* toute matrice carrée dont tous les termes sont nuls en dehors de la diagonale principale. L'ensemble de ces matrices sera noté :

$$\mathscr{D}_n(\mathbb{K}).$$

Si on se donne n éléments scalaires :

$$a_1,\, a_2\, \dots,\, a_n \;\in\; \mathbb{K},$$

dont certains peuvent éventuellement être nuls, il leur correspond une matrice diagonale :

$$\mathsf{Diag}\big(a_1, a_2, \dots, a_n\big) \;=\; \big(a_i\,\delta_{i,j}\big) \;=\; \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}.$$

Les matrices scalaires sont alors des matrices diagonales (très) particulières dans lesquelles tous les éléments diagonaux $a_i = \lambda$ avec $1 \leqslant i \leqslant n$ sont égaux à un même et unique scalaire.

Observons qu'on peut tout aussi bien écrire en changeant $a_i \longmapsto a_j$ grâce à la symétrie du symbole de Kronecker :

$$\mathsf{Diag}\big(a_1, a_2, \dots, a_n\big) \;=\; \big(a_j\,\delta_{i,j}\big).$$

Proposition 9.5. *L'ensemble des matrices diagonales d'ordre n est un sous-anneau commutatif de $\mathscr{M}_n(\mathbb{K})$.*

Être un *sous-anneau* signifie être stable par l'addition et par la multiplication.

Démonstration. En effet, l'addition de deux matrices diagonales est clairement une matrice diagonale :

$$\big(a_i\,\delta_{i,j}\big) + \big(a_i'\,\delta_{i,j}\big) \;=\; \Big(\big(a_i + a_i'\big)\,\delta_{i,j}\Big).$$

Ensuite, nous affirmons que le produit de deux matrices diagonales est encore une matrice diagonale, ce que l'on peut représenter par l'équation complète :

$$\begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix} \begin{pmatrix} b_1 & 0 & \cdots & 0 \\ 0 & b_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & b_n \end{pmatrix} = \begin{pmatrix} a_1b_1 & 0 & \cdots & 0 \\ 0 & a_2b_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_nb_n \end{pmatrix}.$$

En effet, le calcul du produit de deux matrices diagonales quelconques :

$$\begin{aligned}
\mathsf{Diag}\big(a_1, \dots, a_n\big) \cdot \mathsf{Diag}\big(b_1, \dots, b_n\big) &= \big(a_i\,\delta_{i,j}\big) \cdot \big(b_j\,\delta_{j,k}\big) \\
[b_j \longleftrightarrow b_k] \qquad &= \big(a_i\,\delta_{i,j}\big) \cdot \big(b_k\,\delta_{j,k}\big) \\
[\text{Définition !}] \qquad &= \bigg(\sum_{j=1}^{n} a_i\,\delta_{i,j}\,b_k\,\delta_{j,k}\bigg) \\
&= \bigg(a_i\,b_k\,\sum_{j=1}^{n} \delta_{i,j}\,\delta_{j,k}\bigg),
\end{aligned}$$

fait apparaître une somme de produits de symboles de Kronecker, qu'il faut comprendre — mais cela est facile !

Parmi les n symboles $\delta_{i,1}, \dots, \delta_{i,i}, \dots, \delta_{i,n}$ dans cette somme, seul $\delta_{i,i} = 1$ est non nul. Par conséquent, dans la somme, il ne reste que le terme pour $j = i$, et en remplaçant alors $j := i$ dans le deuxième symbole de Kronecker, on obtient un réultat :

$$\begin{aligned}
\mathsf{Diag}\big(a_1, \dots, a_n\big) \cdot \mathsf{Diag}\big(b_1, \dots, b_n\big) &= \big(a_i\,b_k\,\delta_{i,k}\big)_{1 \leqslant i \leqslant n \text{ lignes}}^{1 \leqslant k \leqslant n \text{ colonnes}} \\
&= \mathsf{Diag}\big(a_1b_1, \dots, a_nb_n\big),
\end{aligned}$$

dans lequel un nouveau symbole de Kronecker fait bien voir une matrice diagonale.

De plus, la commutativité de la multiplication devient évidente avec cette formule, car la multiplication dans $\mathbb{K}$ est *commutative :*

$$\begin{aligned}\mathsf{Diag}\big(b_1,\dots,b_n\big)\cdot\mathsf{Diag}\big(a_1,\dots,a_n\big) &= \mathsf{Diag}\big(b_1a_1,\dots,b_na_n\big)\\ &= \mathsf{Diag}\big(a_1b_1,\dots,a_nb_n\big)\\ &= \mathsf{Diag}\big(a_1,\dots,a_n\big)\cdot\mathsf{Diag}\big(b_1,\dots,b_n\big).\end{aligned}$$ □

Maintenant, nous pouvons voir comment se comporte l'ensemble des matrices scalaires dans l'ensemble plus grand des matrices diagonales.

Proposition 9.6. *L'ensemble des matrices scalaires d'ordre n sur $\mathbb{K}$ est un corps, isomorphe à $\mathbb{K}$, et inclus dans l'anneau commutatif des matrices diagonales.*

Démonstration. Comme plus haut, soit I_n la matrice unité, avec des 1 sur la diagonale, et des 0 partout ailleurs. Introduisons l'application :

$$\begin{aligned}\Phi\colon\quad \mathbb{K} &\longrightarrow \mathscr{M}_n(\mathbb{K})\\ \lambda &\longmapsto \lambda\, I_n.\end{aligned}$$

Elle est visiblement injective (exercice mental). C'est donc une *bijection* de $\mathbb{K}$ sur l'ensemble $\Phi(\mathbb{K})$ des matrices scalaires.

Ensuite, on a :

$$\Phi(\lambda)+\Phi(\mu) = \lambda\, I_n+\mu\, I_n = \big(\lambda+\mu\big)\, I_n = \Phi(\lambda+\mu),$$

ainsi que :

$$\Phi(\lambda)\cdot\Phi(\mu) = \big(\lambda\, I_n\big)\cdot\big(\mu\, I_n\big) = \big(\lambda\,\mu\big)\, I_n = \Phi(\lambda\,\mu),$$

ce qui fait voir que Φ est bien un isomorphisme de corps. □

Évidemment, toutes ces petites matrices diagonales sont beaucoup trop simplettes pour nous, mathématiciens en herbe qui ambitionnent de travailler avec des matrices de taille 137 225 808 dans lesquelles il y aurait peu de zéros.

Définition 9.7. Une *matrice triangulaire supérieure* d'ordre n est une matrice carrée dans $\mathscr{M}_n(\mathbb{K})$ tel que tout terme situé en-dessous de la diagonale soit nul :

$$\begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n}\\ 0 & a_{2,2} & a_{2,3} & \cdots & a_{2,n}\\ 0 & 0 & a_{3,3} & \cdots & a_{3,n}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & 0 & \cdots & a_{n,n}\end{pmatrix}.$$

La notion de *matrice triangulaire inférieure* se définit d'une façon analogue et très évidente pour notre intuition instantanée.

Ainsi, une matrice est triangulaire supérieure si et seulement si :

$$i > j \qquad\qquad \Longrightarrow \qquad\qquad a_{i,j} = 0.$$

Notation 9.8. Le sous-ensemble de $\mathscr{M}_n(\mathbb{K})$ constituée des matrices triangulaires supérieures sera noté :

$$\mathscr{T}_n(\mathbb{K}).$$

Proposition 9.9. *$\mathscr{T}_n(\mathbb{K})$ est un sous-anneau de $\mathscr{M}_n(\mathbb{K})$.*

Rappelons que cela signigie que l'ensemble des matrices triangulaires supérieures est stable par addition et par multiplication.

Démonstration. Tout d'abord, étant donné deux matrices :

$$\big(a_{i,j}\big) \,\in\, \mathscr{T}_n(\mathbb{K}) \qquad\qquad \text{et} \qquad\qquad \big(a'_{i,j}\big) \,\in\, \mathscr{T}_n(\mathbb{K}),$$

ce qui s'exprime par :

$$i \,>\, j \qquad\qquad \Longrightarrow \qquad\qquad \Big(a_{i,j} \,=\, 0 \qquad \text{et} \qquad a'_{i,j} \,=\, 0\Big),$$

il est clair que :

$$i \,>\, j \qquad\qquad \Longrightarrow \qquad\qquad a_{i,j} + a'_{i,j} \,=\, 0 + 0 \,=\, 0,$$

ce qui établit la triangularité supérieure de la somme matricielle :

$$\big(a_{i,j} + a'_{i,j}\big) \,\in\, \mathscr{T}_n(\mathbb{K}).$$

Pour ce qui est de la triangularité du *produit* de deux matrices triangulaires supérieures $\big(a_{i,j}\big)$ et $\big(b_{j,k}\big)$, regardons :

$$\big(a_{i,j}\big) \cdot \big(b_{j,k}\big) \,=\, \bigg(\sum_{j=1}^{n} a_{i,j}\, b_{j,k}\bigg) \,=:\, \big(c_{i,k}\big).$$

Nous devons établir que tous ces termes $c_{i,k}$ sont nuls lorsque $i > k$.

Or si nous décomposons la somme $\sum_{j=1}^{n}$ qui les constitue en deux morceaux astucieux :

$$c_{i,k} \,=\, \sum_{j=1}^{i-1} \underline{a_{i,j}}_{\circ}\, b_{j,k} + \underbrace{\sum_{j=i}^{n}}_{j\geqslant i>k} a_{i,j}\, \underline{b_{j,k}}_{\circ},$$

dans le premier morceau, tous les termes s'annulent grâce à la première hypothèse que $a_{i,j} \,=\, 0$ lorsque $j < i$, et dans le second morceau, tous les termes s'annulent *aussi*, car pour eux, on voit que $j > k$, et l'autre hypothèse $b_{j,k} = 0$ lorsque $j > k$ s'applique ! □

Remarquons pour terminer que l'ensemble des matrices diagonales est un sous-anneau de $\mathscr{T}_n(\mathbb{K})$.

10. Matrices inversibles

Dans tout anneau commutatif A à élément unité noté $1 \in A$, on dit qu'un élément x est *inversible* lorsqu'il existe un élément $x' \in A$ satisfaisant :

$$x\, x' \,=\, x'\, x \,=\, 1.$$

On note alors :

$$x' \,=:\, x^{-1}.$$

Notation 10.1. L'ensemble des éléments inversible de l'anneau A sera noté :

$$\mathsf{Inv}_A.$$

Cet ensemble Inv_A n'est pas vide, tout bêtement parce que $1 \in \mathsf{Inv}_A$.

Lemme 10.2. Inv_A *est un groupe multiplicatif.*

Preuve. Il s'agit de vérifier que Inv_A est stable par multiplication et par inversion, et cela est aisé.

En effet, puisque pour $x \in \mathsf{Inv}_A$ et $y \in \mathsf{Inv}_A$, il existe des inverses $x^{-1} \in A$ et $y^{-1} \in A$, si nous calculons le produit :

$$\big(x\, y\big)\, \big(y^{-1}\, x^{-1}\big) \,=\, x\, \big(y\, y^{-1}\big)\, x^{-1} \,=\, x\, 1\, x^{-1} \,=\, 1,$$

nous constatons que $xy \in \mathsf{Inv}_A$, avec de plus l'information que $\big(xy\big)^{-1} = y^{-1}x^{-1}$.

Ensuite, il est clair (exercice mental) par définition que :

$$x \in \mathsf{Inv}_A \qquad\qquad \Longrightarrow \qquad\qquad x^{-1} \in \mathsf{Inv}_A.$$

Ceci démontre bien que Inv_A est un groupe multiplicatif. □

La multiplication entre matrices qui avait été notée jusqu'à présent avec un point $(\bullet) \cdot (\bullet)$ sera autorisée à être notée sans aucun signe mathématique, comme c'est le cas pour la multiplication entre scalaires $\lambda\, \mu$ dans le corps de référence $\mathbb{K}$.

Définition 10.3. On dit qu'une matrice $M \in \mathscr{M}_n(\mathbb{K})$ est *inversible*, ou *régulière*, lorsqu'il existe une matrice de $\mathscr{M}_n(\mathbb{K})$, notée M^{-1}, telle que :

$$M\, M^{-1} = M^{-1}\, M = I_n.$$

Rappelons que l'on note $\mathscr{L}(E) \equiv \mathscr{L}(E, E)$ l'anneau des endomorphismes d'un $\mathbb{K}$-espace vectoriel E de dimension $n \geqslant 1$ sur un corps $\mathbb{K}$. Choisissons une base $\boldsymbol{B}_E$ de E. Dans la bijection vue à de nombreuses reprises :

$$\begin{aligned} \mathscr{L}(E) &\longrightarrow \mathscr{M}_n(\mathbb{K}) \\ f &\longmapsto \mathsf{Mat}(f), \end{aligned}$$

entre f et sa matrice dans la base $\boldsymbol{B}_E$, l'application identité $\mathsf{Id} \in \mathscr{L}(E)$ a pour correspondante la matrice identité I_n. Ainsi, pour qu'une matrice M soit inversible, il faut et il suffit que son application correspondante $f = f_M \in \mathscr{L}(E)$ jouisse de la propriété analogue qu'il existe un endomorphisme $g \in \mathscr{L}(E)$, tel que :

$$f \circ g = g \circ f = \mathsf{Id},$$

et on note :

$$g =: f^{-1}.$$

Ainsi, pour que f possède un inverse dans $\mathscr{L}(E)$, il faut et il suffit que f soit *bijective*.

Réciproquement, si f est bijective, alors f^{-1} existe en tant qu'application, et :

$$\begin{aligned} \mathsf{Mat}(f) \cdot \mathsf{Mat}\big(f^{-1}\big) &= \mathsf{Mat}\big(f \circ f^{-1}\big) = \mathsf{Mat}\,(\mathsf{Id}) = I_n, \\ \mathsf{Mat}\big(f^{-1}\big) \cdot \mathsf{Mat}(f) &= \mathsf{Mat}\big(f^{-1} \circ f\big) = \mathsf{Mat}\,(\mathsf{Id}) = I_n. \end{aligned}$$

Par conséquent, $\mathsf{Mat}(f)$ est inversible, et on a :

$$\big[\mathsf{Mat}(f)\big]^{-1} = \mathsf{Mat}\big(f^{-1}\big).$$

Toutes ces considérations assez élémentaires peuvent maintenant être résumées sous la forme d'un

Théorème 10.4. *L'ensemble des matrices inversibles de $\mathscr{M}_n(\mathbb{K})$ est un groupe multiplicatif, isomorphe au groupe $\mathscr{A}(E) \subset \mathscr{L}(E)$ des automorphismes linéaires de E.* □

Notation 10.5. Dans l'ensemble $\mathscr{M}_n(\mathbb{K})$ des matrices carrées d'ordre $n \geqslant 1$, le sous-ensemble des matrices inversibles sera noté :

$$\mathscr{A}_n(\mathbb{K}).$$

Le groupe de ces matrices inversibles sera appelé *groupe linéaire à n variables* sur le corps $\mathbb{K}$, et il sera noté :

$$\mathsf{GL}_n(\mathbb{K}).$$

11. Changements de bases

À plusieurs reprises, nous avons signalé que la matrice $\mathsf{Mat}(f)$ d'une application linéaire dépend en général d'une base choisie dans les espaces vectoriels considérés — mais sans en dire plus. Cette section est destinée à étudier cette dépendance en la base. Nous travaillerons avec $\mathscr{L}(E)$, c'est-à-dire avec un unique espace vectoriel $E = F$.

Choisissons deux bases quelconques de E :

$$\begin{aligned} \boldsymbol{B}_E &= \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}, \\ \boldsymbol{B}'_E &= \{\vec{e}'_1, \vec{e}'_2, \dots, \vec{e}'_n\}. \end{aligned}$$

Problème 11.1. *Connaissant les coordonnées $(x_1, x_2, \dots, x_n)$ d'un vecteur $\vec{x} \in E$ dans la première base $\boldsymbol{B}_E$, déterminer ses nouvelles coordonnées $(x'_1, x'_2, \dots, x'_n)$ dans la deuxième base $\boldsymbol{B}'_E$.*

D'après un théorème vu dans le chapitre précédent, il existe un unique endomorphisme $f \in \mathscr{L}(E)$ qui envoie la base $\boldsymbol{B}_E$ sur la base $\boldsymbol{B}'_E$, c'est-à-dire tel que :

$$f(\vec{e}_1) = \vec{e}'_1, \; \dots\dots\dots, \; f(\vec{e}_n) = \vec{e}'_n.$$

De plus, cet endomorphisme est une bijection, et ainsi, c'est un *automorphisme* de E. Il en résulte que la matrice de f dans toute base de E est *inversible*.

Donnons-nous alors la matrice de f dans la première base $\boldsymbol{B}_E$:

$$\vec{e}'_i = f(\vec{e}_i) = \sum_{j=1}^{n} a_{j,i}\, \vec{e}_j \qquad\qquad (1 \leqslant i \leqslant n).$$

Terminologie 11.2. La matrice $\big(a_{i,j}\big)$ se nomme *matrice de passage* de la base $\boldsymbol{B}_E$ vers la base $\boldsymbol{B}'_E$, et sera notée :

$$P = \big(a_{i,j}\big).$$

Rappelons — au *passage !* — que les *colonnes* de cette matrice P sont constituée des coordonnées des vecteurs images $f(\vec{e}_1)$, $f(\vec{e}_2)$, ..., $f(\vec{e}_n)$ de la base $\boldsymbol{B}_E$, ce que nous pouvons représenter sous forme diagrammatique et intuitive comme :

$$P = \big(a_{i,j}\big) = \begin{array}{c} \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \\ \scriptstyle f(\vec{e}_1) \quad f(\vec{e}_2) \quad \cdots \quad f(\vec{e}_n) \end{array},$$

en ajoutant la représentation explicite de ces images de vecteurs :

$$f(\vec{e}_1) = \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{n,1} \end{pmatrix}, \quad f(\vec{e}_2) = \begin{pmatrix} a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{n,2} \end{pmatrix}, \quad \dots\dots\dots, \quad f(\vec{e}_n) = \begin{pmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{n,n} \end{pmatrix}.$$

Soit alors un vecteur quelconque $\vec{x} \in E$, de coordonnées $(x_1, x_2, \dots, x_n)$ dans la première base $\boldsymbol{B}_E$:

$$\vec{x} = x_1\, \vec{e}_1 + x_2\, \vec{e}_2 + \cdots + x_n\, \vec{e}_n = \sum_{i=1}^{n} x_i\, \vec{e}_i.$$

Cherchons les coordonnées $(x'_1, x'_2, \dots, x'_n)$ de ce vecteur $\vec{x}$ dans la nouvelle base $\boldsymbol{B}'_E$:

$$\vec{x} = x'_1\, \vec{e}'_1 + x'_2\, \vec{e}'_2 + \cdots + x'_n\, \vec{e}'_n = \sum_{i=1}^{n} x'_i\, \vec{e}'_i.$$

À cette fin, remplaçons et calculons :

$$
\begin{aligned}
\vec{x} = \sum_{i=1}^{n} x_i\,\vec{e}_i = \sum_{i=1}^{n} x_i'\,\vec{e}_i' &= \sum_{i=1}^{n} x_i' \bigg(\sum_{j=1}^{n} a_{j,i}\,\vec{e}_j\bigg) \\
&= \sum_{j=1}^{n} \bigg(\sum_{i=1}^{n} a_{j,i}\,x_i'\bigg)\,\vec{e}_j \\
[i \longleftrightarrow j] \qquad &= \sum_{i=1}^{n} \bigg(\sum_{j=1}^{n} a_{i,j}\,x_j'\bigg)\,\vec{e}_i,
\end{aligned}
$$

où, pour des raisons notationnelles, nous avons *permuté rigoureusement* les lettres i et j à la dernière ligne, afin d'obtenir par identification des coefficients des n vecteurs $\vec{e}_i$ entre les deux membres de l'équation obtenue les relations :

$$
x_i = \sum_{j=1}^{n} a_{i,j}\,x_j' \qquad\qquad (1 \leqslant i \leqslant n),
$$

qui expriment les coordonnées $x_\bullet$ en fonction des coordonnées $x'_\bullet$.

Ce n'est pas encore tout à fait la solution à notre problème, qui demandait à l'inverse d'exprimer les coordonnées $x'_\bullet$ en fonction des coordonnées $x_\bullet$, mais nous pouvons toutefois exprimer ce que nous avons obtenu avant de progresser plus avant.

En effet, au moyen des deux matrices colonnes :

$$
X := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad \text{et} \qquad X' := \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix},
$$

ces n équations scalaires qui reviennent à une équation incorporant un produit matriciel :

$$
\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \cdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \cdot \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix},
$$

peuvent alors s'exprimer de manière beaucoup plus condensée sous la forme agréable :

$$
\boxed{X = P\,X'.}
$$

Mêzalorre — Mais alors pour recevoir X' en fonction de X, il suffit de faire passer P à gauche — comme l'Assemblée Nationale en 1997 — :

$$
\boxed{P^{-1}\,X = X',}
$$

et notre problème est résolu !

Pas tout à fait, car cette équation sous-entend que nous sachions calculer l'inverse :

$$
P^{-1} = \big(a_{i,j}\big)^{-1}
$$

d'une matrice inversible $P = \big(a_{i,j}\big)$, et ce problème de calcul s'avère très difficile, y compris pour les ordinateurs quand la taille n d'une matrice devient un peu grande.

En tout cas, il est absolument interdit de s'imaginer que l'inverse d'une matrice soit la matrice idiote $\big(\frac{1}{a_{i,j}}\big)$ dont les termes sont les inverses, car nous avons insisté sur le fait que la multiplication entre matrice n'était *pas* la multiplication $\big(a_{i,j}\,b_{i,j}\big)$

Ainsi, calculer l'inverse d'une matrice sera un travail difficile, et nous en parlerons plus tard.

Pour le fun, et pour créer du suspens, importons trois exemples amusants d'inverses de matrices $(\bullet)^{-1}$ concoctés sur ordinateur.

$$\begin{pmatrix} 3 & 1 \\ 2 & 4 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{2}{5} & -\frac{1}{10} \\ -\frac{1}{5} & \frac{3}{10} \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 & -1 \\ 3 & 5 & -2 \\ -2 & 3 & -5 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{3}{38} & -\frac{5}{38} \\ -\frac{1}{2} & \frac{7}{38} & \frac{1}{38} \\ -\frac{1}{2} & \frac{3}{38} & -\frac{5}{38} \end{pmatrix},$$

$$\begin{pmatrix} 4 & -4 & 8 & 7 \\ 2 & 4 & -9 & -9 \\ 2 & -3 & -9 & 1 \\ 9 & 1 & -3 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{222}{2141} & \frac{182}{2141} & -\frac{19}{2141} & \frac{103}{2141} \\ -\frac{750}{2141} & -\frac{557}{2141} & -\frac{283}{2141} & \frac{520}{214} \\ \frac{241}{2141} & \frac{159}{2141} & -\frac{146}{2141} & -\frac{110}{214} \\ -\frac{525}{2141} & -\frac{604}{2141} & \frac{16}{2141} & \frac{364}{214} \end{pmatrix} \text{—}$$

— oh, allez, encore un gros dernier pour la route, l'inverse de la matrice pas très simple de taille 5×5 :

$$A := \begin{bmatrix} -46 & 80 & -74 & 51 & -17 \\ -29 & 20 & 13 & 20 & -25 \\ 9 & 39 & 32 & -46 & 78 \\ 81 & -35 & 48 & 35 & 23 \\ 35 & 26 & -60 & -54 & -67 \end{bmatrix}$$

est la matrice pas vraiment plus simple :

$$\begin{bmatrix} \frac{1621261}{788355480} & -\frac{4434691}{1970888700} & \frac{2489327}{1313925800} & \frac{39156173}{3941777400} & \frac{11694193}{1970888700} \\ \frac{1297599}{262785160} & \frac{9265871}{656962900} & \frac{11598439}{1313925800} & \frac{6004687}{1313925800} & \frac{3501467}{656962900} \\ -\frac{2358731}{394177740} & \frac{20284211}{985444350} & \frac{2983383}{656962900} & \frac{6599117}{1970888700} & \frac{269947}{985444350} \\ \frac{4857913}{788355480} & \frac{1555577}{1970888700} & -\frac{5626469}{1313925800} & \frac{24765569}{3941777400} & -\frac{9236471}{1970888700} \\ \frac{533365}{157671096} & -\frac{1164527}{78835548} & \frac{199707}{52557032} & -\frac{173375}{157671096} & -\frac{490807}{78835548} \end{bmatrix}$$

On voit bien que le problème d'inverser une matrice ne va pas être simple, n'est-il pas, *Milady* ? Alors là, pendant les examens, on va en voir ce qu'on va en voir, des inversions de matrices !

Maintenant que nous savons comment se transforment les coordonnées d'un vecteur quand on change de base, une deuxième question se présente à nous.

Problème 11.3. *Connaissant la matrice* $M = \mathrm{Mat}(f)$ *dans une base* $\boldsymbol{B}_E$ *d'un homomorphisme linéaire* $f \in \mathscr{L}(E)$, *trouver la matrice* M' *de cet homomorphisme dans une autre base* $\boldsymbol{B}'_E$ *de* E.

Soit donc $f \in \mathscr{L}(E)$, et soit $M := \mathrm{Mat}(f)$ sa matrice dans la base $\boldsymbol{B}_E$ de E. Si X et Y sont les matrices colonnes des coordonnées d'un vecteur $\vec{x} \in E$ et de son image $\vec{y} := f(\vec{x})$ dans la base $\boldsymbol{B}_E$, nous avons vu plus haut que :

$$Y = M\,X.$$

Soit aussi P la matrice de passage de la base $\boldsymbol{B}_E$ à la base $\boldsymbol{B}'_E$. Si X' et Y' désignent les matrices colonnes des coordonnées de ces mêmes vecteurs $\vec{x}$ et $\vec{y} = f(\vec{x})$ dans la base $\boldsymbol{B}'_E$, nous venons de voir que :

$$X = P\,X' \qquad \text{et} \qquad Y = P\,Y'.$$

La dernière équivaut à :

$$Y' = P^{-1}\,Y.$$

Remplaçons alors, dans cette relation, Y par $M\,X = M\,P\,X'$, pour obtenir :

$$Y' = \big(P^{-1}\,M\,P\big)\,X',$$

ce qui prouve que la matrice M' de l'homorphisme f dans la base $\mathcal{B}'_E$ est :

$$\boxed{M' = P^{-1}\,M\,P.}$$

Cette relation résout le problème posé.

12. Transpositions de matrices

Soient deux entiers $m \geqslant 1$ et $n \geqslant 1$. À présent, considérons des matrices qui ne sont pas forcément carrées.

Définition 12.1. On appelle *transposée* d'une matrice $M \in \mathscr{M}_{m,n}(\mathbb{K})$ comportant $m \geqslant 1$ lignes et $n \geqslant 1$ colonnes la matrice, notée :

$${}^{\mathtt{t}}M \in \mathscr{M}_{n,m}(\mathbb{K}),$$

qui est obtenue à partir de M en échangeant les lignes et les colonnes, de telle sorte que ${}^{\mathtt{t}}M$ comporte à l'inverse n lignes et m colonnes.

Par exemple :

$$\begin{pmatrix} 1 & 0 & 5 \\ 2 & -4 & 7 \end{pmatrix} \xrightarrow{\text{transposer}} \begin{pmatrix} 1 & 2 \\ 0 & -4 \\ 5 & 7 \end{pmatrix}.$$

En général, si donc on note la matrice considérée :

$$M = \big(a_{i,j}\big)_{1\leqslant i\leqslant m \text{ lignes}}^{1\leqslant j\leqslant n \text{ colonnes}},$$

et si on donne un nom aux éléments de la matrice transposée :

$${}^{\mathtt{t}}M =: \big(a^{\mathtt{t}}_{j,i}\big)_{1\leqslant j\leqslant n \text{ lignes}}^{1\leqslant i\leqslant m \text{ colonnes}},$$

on doit avoir par définition :

$$\forall\,1 \leqslant i \leqslant m \quad \forall\,1 \leqslant j \leqslant n : \qquad a^{\mathtt{t}}_{j,i} := a_{i,j}.$$

On pourra ainsi écrire :

$${}^{\mathtt{t}}\Big[\big(a_{i,j}\big)_{1\leqslant i\leqslant m \text{ lignes}}^{1\leqslant j\leqslant n \text{ colonnes}}\Big] = \big(a_{i,j}\big)_{1\leqslant j\leqslant n \text{ lignes}}^{1\leqslant i\leqslant m \text{ colonnes}}.$$

Cette opération de *transposition* définit donc une application :

$$\begin{aligned} {}^{\mathtt{t}}(\bullet)\colon \quad \mathscr{M}_{m,n}(\mathbb{K}) &\longrightarrow \mathscr{M}_{n,m}(\mathbb{K}) \\ M &\longmapsto {}^{\mathtt{t}}M, \end{aligned}$$

qui est naturellement bijective et inverse d'elle-même :

$${}^{\mathtt{t}}\big({}^{\mathtt{t}}M\big) = M.$$

En particulier, si on se restreint aux matrices carrées d'ordre n, on obtient la propriété suivante.

Observation 12.2. *La transposition $M \longmapsto {}^{\mathtt{t}}M$ est une involution de $\mathscr{M}_n(\mathbb{K})$.* □

Cet opérateur possède quelques propriétés simples vis-à-vis de la structure d'anneau de $\mathscr{M}_n(\mathbb{K})$.

Lemme 12.3. *La transposition est un isomorphisme linéaire pour l'addition et pour la multiplication par un scalaire :*

$$\begin{aligned} {}^{\mathtt{t}}\big[M+N\big] &= {}^{\mathtt{t}}M + {}^{\mathtt{t}}N, \\ {}^{\mathtt{t}}\big[\lambda\, M\big] &= \lambda\, {}^{\mathtt{t}}M. \end{aligned}$$

Preuve. Soient donc deux matrices de $\mathscr{M}_{m,n}(\mathbb{K})$:

$$M = \big(a_{i,j}\big) \qquad \text{et} \qquad N = \big(b_{i,j}\big).$$

Le calcul suivant justifie la première affirmation :

$$\begin{aligned} {}^{\mathtt{t}}\big[M+N\big] &= {}^{\mathtt{t}}\Big[\big(a_{i,j}+b_{i,j}\big)_{1\leqslant i\leqslant m \text{ lignes}}^{1\leqslant j\leqslant n \text{ colonnes}}\Big] = \Big(a_{i,j}+b_{i,j}\Big)_{1\leqslant j\leqslant n \text{ lignes}}^{1\leqslant i\leqslant m \text{ colonnes}} \\ &= \Big(a_{i,j}\Big)_{1\leqslant j\leqslant n \text{ lignes}}^{1\leqslant i\leqslant m \text{ colonnes}} + \Big(b_{i,j}\Big)_{1\leqslant j\leqslant n \text{ lignes}}^{1\leqslant i\leqslant m \text{ colonnes}} \\ &= {}^{\mathtt{t}}M + {}^{\mathtt{t}}N, \end{aligned}$$

tandis que la seconde se justifie de manière analogue :

$$\begin{aligned} {}^{\mathtt{t}}\big[\lambda\, M\big] &= {}^{\mathtt{t}}\Big[\Big(\lambda\, a_{i,j}\Big)_{1\leqslant i\leqslant m \text{ lignes}}^{1\leqslant j\leqslant n \text{ colonnes}}\Big] = \Big(\lambda\, a_{i,j}\Big)_{1\leqslant j\leqslant n \text{ lignes}}^{1\leqslant i\leqslant m \text{ colonnes}} \\ &= \lambda\Big(a_{i,j}\Big)_{1\leqslant j\leqslant n \text{ lignes}}^{1\leqslant i\leqslant m \text{ colonnes}} \\ &= \lambda\, {}^{\mathtt{t}}M. \end{aligned}$$

□

Par conséquent, l'application $M \longmapsto {}^{\mathtt{t}}M$ est un isomorphisme linéaire de l'espace vectoriel $\mathscr{M}_{m,n}(\mathbb{K})$ sur l'espace vectoriel $\mathscr{M}_{n,m}(\mathbb{K})$.

Fait remarquable : vis-à-vis du produit, la transposition *intervertit* l'ordre des facteurs.

Lemme 12.4. *Pour tous entiers $m, n, p \geqslant 1$, toute matrice $M \in \mathscr{M}_{m,n}(\mathbb{K})$, et toute matrice $N \in \mathscr{M}_{n,p}(\mathbb{K})$, on a :*

$$\boxed{{}^{\mathtt{t}}\big[M\cdot N\big] = {}^{\mathtt{t}}N \cdot {}^{\mathtt{t}}M.}$$

Observons, et il est important de le dire, que la matrice ${}^{\mathtt{t}}N$ *a le droit* de multiplier la matrice ${}^{\mathtt{t}}M$, puisque le nombre n de colonnes de :

$${}^{\mathtt{t}}N \in \mathscr{M}_{p,n}(\mathbb{K})$$

est bel et bien égal au nombre de lignes n de :

$${}^{\mathtt{t}}M \in \mathscr{M}_{n,m}(\mathbb{K}).$$

Démonstration. Soient donc deux matrices :

$$M = \Big(a_{i,j}\Big)_{1\leqslant i\leqslant m \text{ lignes}}^{1\leqslant j\leqslant n \text{ colonnes}} \qquad \text{et} \qquad N = \Big(b_{j,k}\Big)_{1\leqslant j\leqslant n \text{ lignes}}^{1\leqslant k\leqslant p \text{ colonnes}},$$

qui ont pour produit :

$$M\cdot N = \bigg(\sum_{j=1}^{n} a_{i,j}\, b_{j,k}\bigg)_{1\leqslant i\leqslant m \text{ lignes}}^{1\leqslant k\leqslant p \text{ colonnes}},$$

d'où, par transposition, un premier résultat :

$${}^{\mathtt{t}}\big[M\cdot N\big] = \bigg(\sum_{j=1}^{n} a_{i,j}\, b_{j,k}\bigg)_{1\leqslant k\leqslant p \text{ lignes}}^{1\leqslant i\leqslant m \text{ colonnes}}.$$

Par ailleurs, les transposées de M et de N sont :

$${}^{\mathtt{t}}M = \Big(a_{j,i}^{\mathtt{t}}\Big)_{1\leqslant j\leqslant n \text{ lignes}}^{1\leqslant i\leqslant m \text{ colonnes}} \qquad \text{et} \qquad {}^{\mathtt{t}}N = \Big(b_{k,j}^{\mathtt{t}}\Big)_{1\leqslant k\leqslant p \text{ lignes}}^{1\leqslant j\leqslant n \text{ colonnes}},$$

avec bien sûr :

$$a^{\mathtt{t}}_{j,i} := a_{i,j} \qquad\qquad \text{et} \qquad\qquad b^{\mathtt{t}}_{k,j} := b_{j,k},$$

et elles ont pour produit autorisé un deuxième résultat :

$$\begin{aligned} {}^{\mathtt{t}}N \cdot {}^{\mathtt{t}}M &= \left(\sum_{j=1}^{n} b^{\mathtt{t}}_{k,j}\, a^{\mathtt{t}}_{j,i}\right)^{1\leqslant i\leqslant m \text{ colonnes}}_{1\leqslant k\leqslant p \text{ lignes}} \\ &= \left(\sum_{j=1}^{n} b_{j,k}\, a_{i,j}\right)^{1\leqslant i\leqslant m \text{ colonnes}}_{1\leqslant k\leqslant p \text{ lignes}}, \end{aligned}$$

et puisque la multiplication dans $\mathbb{K}$ est commutative :

$$a_{i,j}\, b_{j,k} = b_{j,k}\, a_{i,j},$$

nous concluons bien en regardant les deux résultats obtenus que :

$${}^{\mathtt{t}}\big[M \cdot N\big] = {}^{\mathtt{t}}N \cdot {}^{\mathtt{t}}M. \qquad \square$$

Faisons remarquer que si on se restreint à l'ensemble $\mathscr{M}_n(\mathbb{K})$ des matrices carrées d'ordre n, l'involution $M \longmapsto {}^{\mathtt{t}}M$ est un automorphisme pour l'addition, mais ce n'est pas un automorphisme pour la multiplication, car il y a changement de l'ordre des facteurs d'un produit.

Lemme 12.5. *Pour qu'une matrice carrée soit inversible, il faut et il suffit que sa transposée ${}^{\mathtt{t}}M$ le soit.*

Preuve. Soit donc $M \in \mathscr{A}_n(\mathbb{K})$. Alors M^{-1} existe, et :

$$\begin{aligned} M \cdot M^{-1} &= I_n, \\ M^{-1} \cdot M &= I_n. \end{aligned}$$

Passons aux transposées, en appliquant le lemme qui précède et en observant que la transposée de la matrice identité est encore la matrice identité :

$$\begin{aligned} {}^{\mathtt{t}}\big[M \cdot M^{-1}\big] &= {}^{\mathtt{t}}\big[M^{-1}\big] \cdot {}^{\mathtt{t}}M = {}^{\mathtt{t}}I_n = I_n, \\ {}^{\mathtt{t}}\big[M^{-1} \cdot M\big] &= {}^{\mathtt{t}}M \cdot {}^{\mathtt{t}}\big[M^{-1}\big] = {}^{\mathtt{t}}I_n = I_n, \end{aligned}$$

ce qui donne deux équations signifiant par définition que la matrice ${}^{\mathtt{t}}M$ est inversible, et a pour matrice inverse :

$$\big({}^{\mathtt{t}}M\big)^{-1} = {}^{\mathtt{t}}\big[M^{-1}\big].$$

La réciproque est immédiate : si ${}^{\mathtt{t}}M \in \mathscr{A}_n(\mathbb{K})$ est inversible, alors en appliquant à ${}^{\mathtt{t}}M$ la démonstration précédente, on trouve que :

$${}^{\mathtt{t}}\big[{}^{\mathtt{t}}M\big] \in \mathscr{A}_n(\mathbb{K}),$$

c'est-à-dire que $M \in \mathscr{A}_n(\mathbb{K})$ est inversible. $\square$

Ce lemme signifie que la restriction à $\mathscr{A}_n(\mathbb{K})$ de l'involution $M \longmapsto {}^{\mathtt{t}}M$ de $\mathscr{M}_n(\mathbb{K})$ est une involution de $\mathscr{A}_n(\mathbb{K})$.

13. Exercices

Exercice 1. Calculer le produit des deux matrices suivantes dans $\mathscr{M}_{2,3}(\mathbb{Q})$ et $\mathscr{M}_{3,4}(\mathbb{Q})$:

$$\begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 0 & 1 & 0 \\ 1 & 1 & 2 & -2 \\ 1 & 0 & 1 & 0 \end{pmatrix}.$$

Exercice 2. EE

Produit scalaire, produit vectoriel, produit mixte

François DE MARÇAY
Département de Mathématiques d'Orsay
Université Paris-Sud, France

1. Introduction

2. Produit scalaire dans l'espace vectoriel euclidien $V_{\mathbb{R}^3}$

L'espace $\mathbb{R}^3$ est un espace de points. On lui associe l'espace vectoriel $V_{\mathbb{R}^3}$. Les éléments de $V_{\mathbb{R}^3}$ sont tous les vecteurs d'origine le point $(0,0,0) \in \mathbb{R}^3$ et d'extrémité un point quelconque $(x_1, x_2, x_3) \in \mathbb{R}^3$.

On notera donc parfois les vecteurs $\vec{x} \in V_{\mathbb{R}^3}$ sous forme d'une matrice colonne :

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

et plus souvent aussi, sous forme horizontale $\vec{x} = (x_1, x_2, x_3)$.

En fait, avec la base canonique de $V_{\mathbb{R}^3}$, constituée des trois vecteurs :

$$\begin{aligned} \vec{e}_1 &:= (1,0,0), \\ \vec{e}_2 &:= (0,1,0), \\ \vec{e}_3 &:= (0,0,1), \end{aligned}$$

il est clair que :

$$\vec{x} = x_1\,\vec{e}_1 + x_2\,\vec{e}_2 + x_3\,\vec{e}_3.$$

Maintenant, comment parler de la *longueur* d'un vecteur quelconque $\vec{x}$? La réponse à cette question est bien connue, elle date de l'Antiquité, et l'on sait bien que la quantité :

$$\big\|\vec{x}\big\| := \sqrt{x_1^2 + x_2^2 + x_3^2},$$

définit une *norme* sur l'espace vectoriel $V_{\mathbb{R}^3}$, au sens où les deux propriétés évidentes suivantes sont satisfaites :

$$\begin{aligned} \big\|\vec{x}\big\| &\geqslant 0 && (\forall\,\vec{x} \in V_{\mathbb{R}^3}), \\ \big\|\lambda\,\vec{x}\big\| &= |\lambda|\,\big\|\vec{x}\big\| && (\forall\,\lambda \in \mathbb{R},\ \forall\,\vec{x} \in V_{\mathbb{R}^3}), \end{aligned}$$

et au sens où on a l'*inégalité triangulaire :*

$$\big\|\vec{x} + \vec{y}\big\| \leqslant \big\|\vec{x}\big\| + \big\|\vec{y}\big\| \qquad (\forall\,\vec{x}, \vec{y} \in V_{\mathbb{R}^3}).$$

On observera que cette norme dite *euclidienne* attribue la longueur 1 aux trois vecteurs de base :

$$1 = \big\|\vec{e}_1\big\| = \big\|\vec{e}_2\big\| = \big\|\vec{e}_3\big\|.$$

On observera aussi que la *norme au carré :*

$$\big\|\vec{x}\big\|^2 = x_1^2 + x_2^3 + x_3^2$$

satisfait le *Théorème de Pythagore :*

$$\begin{aligned}\big\|x_1\,\vec{e}_1+x_2\,\vec{e}_2+x_3\,\vec{e}_3\big\|^2 &= x_1^2\,\big\|\vec{e}_1\big\|^2+x_2^2\,\big\|\vec{e}_2\big\|^2+x_3^2\,\big\|\vec{e}_3\big\|^2\\ &= x_1^2\,1+x_2^2\,1+x_3^2\,1\\ &= x_1^2+x_2^2+x_3^2,\end{aligned}$$

ce qui s'explique par le fait que les trois vecteurs de base $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$ sont *orthogonaux* entre eux.

Vous avez dit *orthogonaux* ? Oui, je l'ai dit. Car dans l'espace vectoriel physique $V_{\mathbb{R}^3}$ à 3 dimensions, entre deux vecteurs quelconques $\vec{x}\in V_{\mathbb{R}^3}$ et $\vec{y}\in V_{\mathbb{R}^3}$, il est bien connu qu'existe le *produit scalaire euclidien*, défini par :

$$\vec{x}\cdot\vec{y} := x_1\,y_1+x_2\,y_2+x_3\,y_3.$$

Nous laissons au lecteur-étudiant la tâche aisée de vérifier dans les moindres détails la véracité du

Lemme 2.1. *Le produit scalaire euclidien* $(\bullet)\cdot(\bullet)$ *satisfait une propriété de bilinéarité par rapport à ses deux arguments :*

$$\begin{aligned}\big(\lambda\,\vec{x}+\lambda'\,\vec{x}'\big)\cdot\vec{y} &= \lambda\,\vec{x}\cdot\vec{y}+\lambda'\,\vec{x}'\cdot\vec{y},\\ \vec{x}\cdot\big(\mu\,\vec{y}+\mu'\,\vec{y}'\big) &= \mu\,\vec{x}\cdot\vec{y}+\mu'\,\vec{x}\cdot\vec{y}',\end{aligned}$$

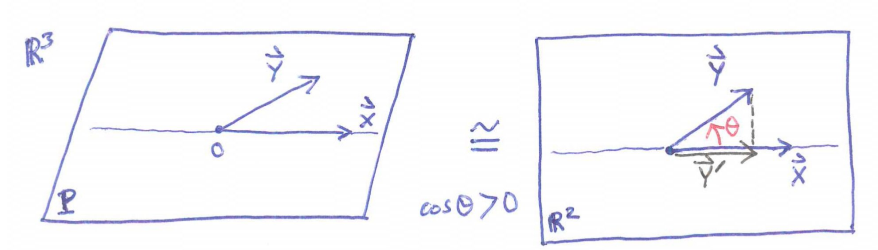
où $\vec{x},\vec{x}',\vec{y},\vec{y}'\in V_{\mathbb{R}^3}$ *sont des vecteurs quelconques, et où* $\lambda,\lambda',\mu,\mu'\in\mathbb{R}$ *sont des scalaires arbitraires.* □

En fait, ces deux propriétés sont équivalentes, car on aura déjà remarqué que le produit scalaire est *symétrique :*

$$\vec{x}\cdot\vec{y} = \vec{y}\cdot\vec{x} \qquad\qquad (\forall\,\vec{x}\in V_{\mathbb{R}^3},\ \forall\,\vec{y}\in V_{\mathbb{R}^3}).$$

De plus, le produit scalaire possède une définition géométrique plus éclairante que la formule $x_1\,y_1+x_2\,y_2+x_3\,y_3$ ci-dessus, et nous pouvons maintenant rappeler cette définition géométrique, connue depuis la classe maternelle, cela va sans dire.

Si $\vec{x}=0$ ou si $\vec{y}=0$, on déclare que $\vec{x}\cdot\vec{y}:=0$ — rien de plus.



Nous pouvons donc supposer que $\vec{x}\neq 0$ et que $\vec{y}\neq 0$, ce qui sera commode pour dessiner des figures. Dans les figures, on va même supposer que $\vec{x}$ et $\vec{y}$ ne sont pas colinéaires, de telle sorte que le couple $\big(\vec{x},\vec{y}\big)$ définit un plan 2-dimensionnel dans l'espace $V_{\mathbb{R}^3}$, plan que nous noterons P.

Ce plan P est alors isomorphe au plan euclidien $V_{\mathbb{R}^2}$ que nous connaissons bien, et pour parler du produit scalaire, on peut alors raisonner entièrement dans ce plan, comme on a appris à le faire en classe de mathématiques maternelles.

Alors si $\vec{y}'$ est la projection orthogonale du vecteur $\vec{y}$ sur la droite que dirige $\vec{x}$, on définit le produit scalaire par la formule :

$$\begin{aligned}\vec{x}\cdot\vec{y} &:= \overline{x}\cdot\overline{y}'\\ &= \vec{x}\cdot\vec{y}',\end{aligned}$$

où $\overline{x}$ et $\overline{y}'$ désignent les longueurs *algébriques*, sur la droite dirigée par $\vec{x}$, des vecteurs $\vec{x}$ et $\vec{y}'$, la valeur de ce produit $\overline{x}\cdot\overline{y}'$ ne dépendant pas d'une orientation choisie sur cette droite, puisque $(-1)\cdot(-1)=1$.

En introduisant l'angle :

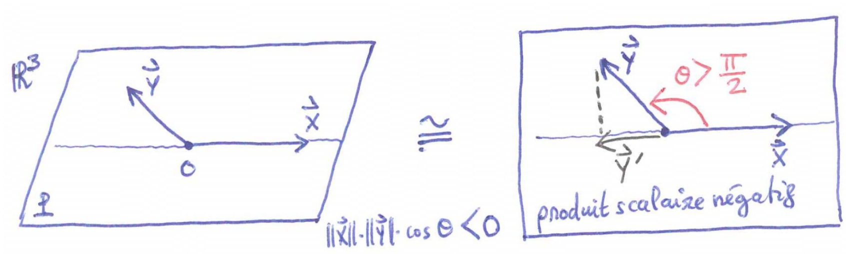
$$\theta \;:=\; \mathsf{Angle}\big(\vec{x},\, \vec{y}\,'\big),$$

nous constatons que la définition alternative :

$$\vec{x}\cdot\vec{y} \;:=\; \big\|\vec{x}\big\|\,\big\|\vec{y}\big\|\,\cos\theta,$$

est équivalente.

La première figure ci-dessus illustrait un cas où $0 < \theta < \frac{\pi}{2}$. En voici une autre qui illustre le cas où $\frac{\pi}{2} < \theta < \pi$ dans lequel le produit scalaire est *négatif*.



Évidemment :

$$\vec{x}\cdot\vec{x} \;=\; \big\|\vec{x}\big\|^2 \qquad\qquad (\forall\,\vec{x}\in V_{\mathbb{R}^3}).$$

3. Présentation des deux (seules) orientations dans l'espace $V_{\mathbb{R}^3}$

L'espace euclidien 3-dimensionnel peut être muni d'une *orientation*, tout comme l'espace 2-dimensionnel, *cf.* ce qu'on appelle le *sens trigonométrique*.

Nous considérerons en effet comme déjà connu le fait que dans $\mathbb{R}^2$, ou dans $V_{\mathbb{R}^2}$, il existe exactement *deux* orientations, opposées l'une de l'autre.

Dans cette section, nous allons présenter une manière de définir les deux orientations possibles — il y en a exactement deux, aussi — de l'espace euclidien $\mathbb{R}^3$, ou de l'espace vectoriel associé $V_{\mathbb{R}^3}$, qui repose sur un procédé dit d'*orthonormalisation de Gram-Schmidt*.

Soient trois vecteurs quelconques :

$$\vec{u} \,\in\, V_{\mathbb{R}^3}, \qquad\qquad \vec{v} \,\in\, V_{\mathbb{R}^3}, \qquad\qquad \vec{w} \,\in\, V_{\mathbb{R}^3},$$

que nous supposerons *linéairement indépendants*. Il faut se les imaginer se baladant dans l'espace, comme trois doigts de la main qui vole. On considère le triplet *ordonné :*

$$\big(\vec{u},\, \vec{v},\, \vec{w}\big).$$

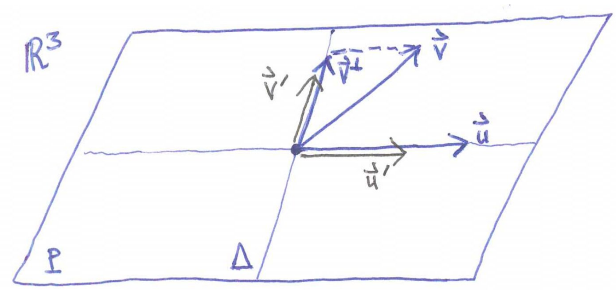
Pour toutes les opérations qui suivent, l'ordre $\big(\vec{u}, \vec{v}, \vec{w}\big)$ devra être respecté.

Commençons par remplacer $\vec{u}$ par le vecteur renormalisé :

$$\vec{u}\,' \;:=\; \frac{\vec{u}}{\|\vec{u}\|},$$

qui devient de norme unité :

$$\big\|\vec{u}\,'\big\| \;=\; \bigg\|\frac{\vec{u}}{\|\vec{u}\|}\bigg\| \;=\; \frac{\|\vec{u}\|}{\|\vec{u}\|} \;=\; 1.$$



Ensuite, considérons le plan vectoriel V_P engendré par les deux premiers vecteurs — on s'occupera de $\vec{w}$ dans quelques instants — :

$$V_P \;:=\; \mathsf{Vect}\,\big(\vec{u},\,\vec{v}\big) \;\cong\; V_{\mathbb{R}^2}.$$

On distingue V_P, constitué de *vecteurs* d'origine 0 contenus dans le plan P, du plan $P \subset \mathbb{R}^3$ lui-même, qui est constitué de *points*.

Comme cela est illustré sur la figure, en travaillant dans ce plan P, projetons orthogonalement le deuxième vecteur $\vec{v}$ sur la droite, notée $\Delta \subset P$, qui est orthogonale à la droite engendrée par le premier vecteur $\vec{u}$.

On obtient ainsi un certain vecteur $\vec{v}^{\perp} \in V_{\Delta}$, qui n'est *pas* nul, puisqu'on a supposé que la famille $\big\{\vec{u}, \vec{v}\big\}$ est linéairement indépendante.

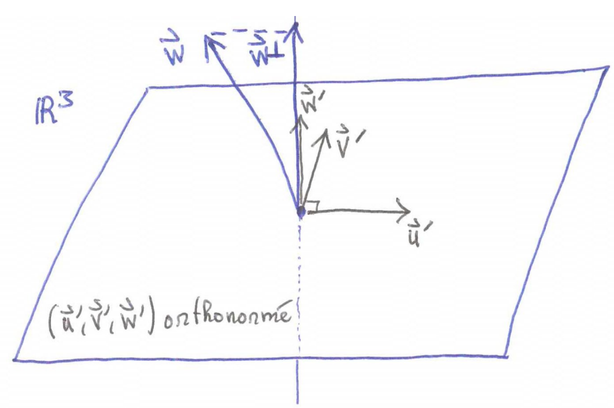
Comme $\vec{v}^{\perp}$ n'est pas forcément de norme unité, renormalisons-le en introduisant :

$$\vec{v}' \;:=\; \frac{\vec{v}^{\perp}}{\|\vec{v}^{\perp}\|},$$

vecteur qui devient de norme $1 = \big\|\vec{v}'\big\|$.

Voilà ! dans le plan vectoriel V_P, on est content ! on a remplacé le couple *ordonné* de vecteurs indépendants $\big(\vec{u}, \vec{v}\big)$ par un nouveau couple ordonné $\big(\vec{u}', \vec{v}'\big)$ qui constitue maintenant une *base orthonormale* du plan vectoriel V_P. Si on travaillait en dimension 2, c'est-à-dire dans $V_{\mathbb{R}^2}$, on s'arrêterait là.

Mais on travaille en dimension 3, et il nous reste encore à malmener $\vec{w}$ — et d'ailleurs, que va-t-on lui faire ?



Soit D la droite orthogonale au plan P. Comme le lecteur l'a deviné, projetons notre troisième et dernier vecteur $\vec{w}$ orthogonalement sur D.

Nous obtenons ainsi un certain vecteur $\vec{w}^{\perp} \in V_D$, qui n'est *pas* nul, parce qu'on a supposé que la famille $\big\{\vec{u}, \vec{v}, \vec{w}\big\}$ est linéairement indépendante.

Comme $\vec{w}^\perp$ n'est pas forcément de norme unité, renormalisons-le en introduisant :

$$\vec{w}' := \frac{\vec{w}^\perp}{\|\vec{w}^\perp\|},$$

vecteur qui devient de norme $1 = \|\vec{w}'\|$.

Voilà ! on a atteint notre objectif ! on a remplacé le triplet *ordonné* de vecteurs indépendants $(\vec{u}, \vec{v}, \vec{w})$ par un nouveau triplet ordonné $(\vec{u}', \vec{v}', \vec{w}')$ qui constitue maintenant une *base orthonormale ordonnée* de l'espace vectoriel euclidien $V_{\mathbb{R}^3}$. En termes moins élégants, on a « *Gram-Schmidté* » nos trois vecteurs $(\vec{u}, \vec{v}, \vec{w})$.

Or $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ constitue aussi une base orthonormale ordonnée de $V_{\mathbb{R}^3}$, et c'est une référence à laquelle on peut donc comparer $(\vec{u}', \vec{v}', \vec{w}')$.

Pour effectuer une telle comparaison, il suffit de déplacer la main dans l'espace (sans oublier d'emporter les trois doigts importants avec lesquels on mange le couscous).

Théorème 3.1. **[Admis]** *Après application du procédé d'orthonormalisation de Gram-Schmidt à une famille libre quelconque* ordonnée *de* 3 *vecteurs dans l'espace* $V_{\mathbb{R}^3}$ *:*

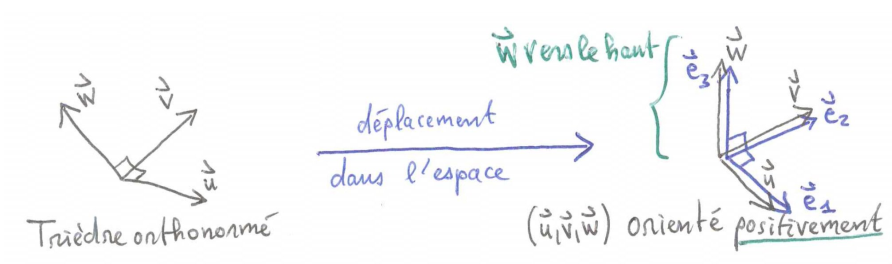
$$\begin{array}{rcl} (\vec{u}, \vec{v}, \vec{w}) & \rightsquigarrow & (\vec{u}', \vec{v}', \vec{w}') \\ V_{\mathbb{R}^3} = \mathsf{Vect}(\vec{u}, \vec{v}, \vec{w}) & & 1 = \|\vec{u}'\| = \|\vec{v}'\| = \|\vec{w}'\| \\ & & 0 = \vec{u}' \cdot \vec{v}' = \vec{u}' \cdot \vec{w}' = \vec{v}' \cdot \vec{w}', \end{array}$$

si on déplace dans l'espace $\vec{u}'$ pour l'amener à coïncider avec $\vec{e}_1$, et si en même temps, on déplace $\vec{v}'$ pour l'amener à coïncider avec $\vec{e}_2$, alors exactement deux situations peuvent se produire concernant $\vec{w}'$:

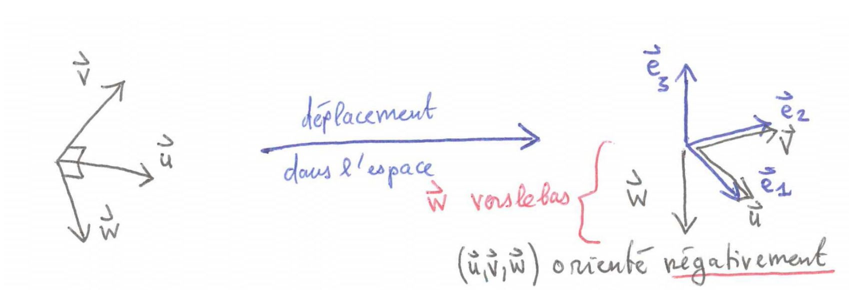
$$\vec{w}' = \vec{e}_3 \qquad\qquad \textit{ou} \qquad\qquad \vec{w}' = -\vec{e}_3. \qquad \square$$

Tête au-dessus, ou tête en-dessous : tout est là !

Définition 3.2. Dans le premier cas, on dira que la famille libre $(\vec{u}, \vec{v}, \vec{w})$ est orientée *positivement*, ou encore, que le trièdre $(\vec{u}, \vec{v}, \vec{w})$ est *direct*.



Dans le deuxième cas, on dira que la famille libre $(\vec{u}, \vec{v}, \vec{w})$ est orientée *négativement*, ou encore, que le trièdre $(\vec{u}, \vec{v}, \vec{w})$ est *indirect*.



4. Produit vectoriel dans $V_{\mathbb{R}^3}$

Maintenant, le Théorème 3.1 s'applique pour faire voir que l'espace vectoriel $V_{\mathbb{R}^3}$ peut être muni d'exactement *deux* orientations opposées.

En effet, nous avons implicitement convenu que le trièdre $\big(\vec{e}_1, \vec{e}_2, \vec{e}_3\big)$ était positif. Une application dudit théorème montre qu'alors le trièdre $\big(\vec{e}_2, \vec{e}_1, \vec{e}_3\big)$ est négatif. Or nous aurions tout à fait pu commencer la théorie en déclarant que c'est *ce* trièdre $\big(\vec{e}_2, \vec{e}_1, \vec{e}_3\big)$ qui est le trièdre de référence positif, et alors, ce qui était positif devient négatif, et vice-versa.

Je laisse en exercice de visualisation géométrique la vérification du fait que les trois trièdres :

$$\big(\vec{e}_1, \vec{e}_2, \vec{e}_3\big), \qquad \big(\vec{e}_2, \vec{e}_3, \vec{e}_1\big), \qquad \big(\vec{e}_3, \vec{e}_1, \vec{e}_2\big), \tag{4.1}$$

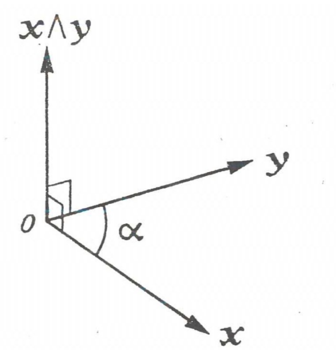
sont toujours en même temps positifs, ou en même temps négatifs, et de même pour les trois trièdres :

$$\big(\vec{e}_2, \vec{e}_1, \vec{e}_3\big), \qquad \big(\vec{e}_1, \vec{e}_3, \vec{e}_2\big), \qquad \big(\vec{e}_3, \vec{e}_2, \vec{e}_1\big).$$

Pour commencer la théorie du produit vectoriel, il faut donc seulement faire un choix parmi les deux orientations possibles de l'espace $V_{\mathbb{R}^3}$. Nous avons fait le choix de déclarer que $\big(\vec{e}_1, \vec{e}_2, \vec{e}_3\big)$ était *positif* : c'est le choix le plus standard et le plus fréquent, donc nous le maintiendrons, tout en répétant qu'il est aussi possible de déclarer au début que $\big(\vec{e}_1, \vec{e}_2, \vec{e}_3\big)$ est *négatif*, ce qui changerait tous les signes partout !

Le point important à retenir, c'est qu'il y a un choix d'orientation de l'espace à effectuer au départ.

Venons-en maintenant à un concept qu'on apprenait, au siècle précédent (*i.e.* dans l'ancien millénaire), en classe de 2^{nde} au Lycée !



Définition 4.2. À tout couple ordonné $\big(\vec{x}, \vec{y}\big)$ de vecteurs de $V_{\mathbb{R}^3}$, on associe un vecteur noté :

$$\vec{x} \wedge \vec{y} \qquad\qquad (\in V_{\mathbb{R}^3}),$$

nommé *produit vectoriel de $\vec{x}$ par $\vec{y}$*, et qui est défini en plusieurs étapes.

(1) Lorsque $\vec{x}$ et $\vec{y}$ sont colinéaires, on déclare que $\vec{x} \wedge \vec{y} := \vec{0}$.

(2) Si $\vec{x}$ et $\vec{y}$ ne sont pas colinéaires, alors $\vec{x} \wedge \vec{y}$ est défini de manière unique par les trois conditions suivantes.

- **(a)** $\vec{x} \wedge \vec{y}$ est perpendiculaire $\vec{x}$ et à $\vec{y}$;
- **(b)** le trièdre $\big(\vec{x}, \vec{y}, \vec{x} \wedge \vec{y}\big)$ est *positif ;*
- **(c)** en termes de la mesure d'angle :

$$\alpha := \mathsf{Angle}\big(\vec{x}, \vec{y}\big),$$

la norme de $\vec{x} \wedge \vec{y}$ doit valoir :

$$\big\|\vec{x} \wedge \vec{y}\big\| := \big\|\vec{x}\big\| \, \big\|\vec{y}\big\| \sin \alpha.$$

Dans peu de temps, le Théorème 4.4 ci-dessous va mieux expliquer pourquoi ces trois conditions **(a)**, **(b)**, **(c)** définissent bien de manière unique le vecteur $\vec{x} \wedge \vec{y}$.

Pour l'instant, constatons que le symbole '$\wedge$' constitue une application :

$$\begin{aligned} V_{\mathbb{R}^3} \times V_{\mathbb{R}^3} &\longrightarrow V_{\mathbb{R}^3} \\ (\vec{x}, \vec{y}) &\longmapsto \vec{x} \wedge \vec{y}. \end{aligned}$$

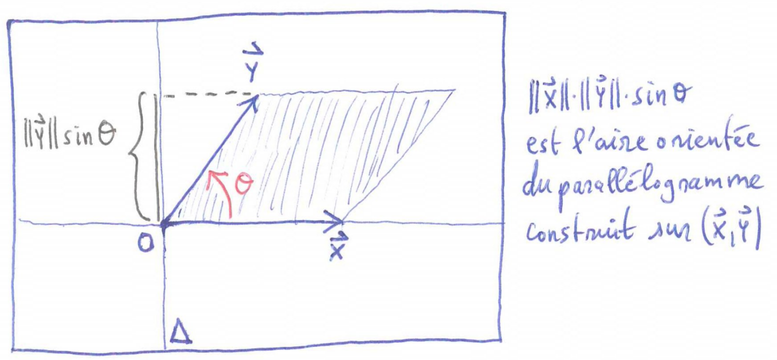
Observons aussi que le produit vectoriel $\vec{x} \wedge \vec{y}$ n'est nul que lorsque $\vec{x}$ et $\vec{y}$ sont colinéaires, ce qui correspond à $\alpha \equiv 0 \operatorname{mod} \pi$.

Dans le cas où $\vec{x}$ et $\vec{y}$ ne sont pas colinéaires, il existe alors une unique droite D orthogonale au plan vectoriel qu'ils engendrent :

$$\operatorname{Vect}(\vec{x}, \vec{y}),$$

et alors la condition **(a)** demande que le vecteur défini soit dirigé par cette droite perpendiculaire :

$$\vec{x} \wedge \vec{y} \in V_D.$$



Il est très important de dire aussi que la longueur (norme) de ce vecteur :

$$\|\vec{x} \wedge \vec{y}\| := \|\vec{x}\| \, \|\vec{y}\| \sin \alpha.$$

est définie comme étant la valeur de l'*aire orientée* du parallélogramme construit sur le couple $(\vec{x}, \vec{y})$, ce qu'explique la figure.

Proposition 4.3. *Quels que soient les vecteurs $\vec{x}$ et $\vec{y}$ dans l'espace, on a :*

$$\vec{y} \wedge \vec{x} = -\vec{x} \wedge \vec{y}.$$

En particulier, pour tout *vecteur $\vec{x}$, on a :*

$$\vec{x} \wedge \vec{x} = \vec{0}.$$

Démonstration. Si $\vec{x}$ et $\vec{y}$ sont colinéaires, c'est évident, car $0 = -0$ est bien vrai.

Quand $\vec{x}$ et $\vec{y}$ ne sont pas colinéaires, la droite orthogonale D et les normes des vecteurs sont les mêmes dans les deux côtés de l'équation à vérifier.

Mais l'orientation du trièdre change d'un côté à l'autre, car comme on s'en convainc en visualisant les choses dans l'espace (ou en tournant le pouce du haut vers le bas), on a :

$$(\vec{x}, \vec{y}, \vec{x} \wedge \vec{y}) \text{ positif} \qquad \Longrightarrow \qquad (\vec{x}, \vec{y}, -\vec{x} \wedge \vec{y}) \text{ négatif.}$$

Enfin, en choisissant $\vec{y} := \vec{x}$, l'identité que nous venons de démontrer conclut :

$$\vec{x} \wedge \vec{x} = -\vec{x} \wedge \vec{x} \qquad \Longrightarrow \qquad 2\,\vec{x} \wedge \vec{x} = \vec{0}. \qquad \square$$

Théorème 4.4. *Dans l'espace vectoriel euclidien 3-dimensionnel $V_{\mathbb{R}^3}$, soit un vecteur non nul $\vec{x} \neq 0$, et soit le plan orthogonal :*

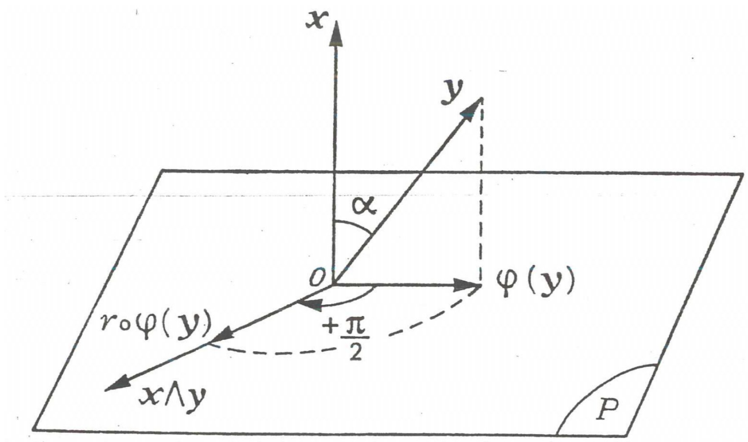
$$P := (\operatorname{Vect} \vec{x})^{\perp}.$$

Alors en termes des 3 *transformations :*

$$\begin{array}{ll} \varphi & \textit{le projecteur orthogonal sur } V_P, \\ \boldsymbol{r} & \textit{la rotation d'axe } \vec{x} \text{ et d'angle } +\frac{\pi}{2}, \\ h & \textit{l'homothétie de rapport } \|\vec{x}\|, \end{array}$$

le produit vectoriel par $\vec{x}$ *de tout vecteur* $\vec{y}$ *s'exprime comme :*

$$h \circ \boldsymbol{r} \circ \varphi(\vec{y}).$$



L'intérêt de cet énoncé, c'est qu'il *redéfinit* le produit vectoriel d'une manière beaucoup plus géométrique que précédemment : c'est une vraie recette !

Démonstration. Lorsque $\vec{y}$ est colinéaire à $\vec{x}$, sa projection $\varphi(\vec{y}) = \vec{0}$ sur P est nulle, et donc on retrouve bien ce qu'on avait déclaré dans la Définition 4.2.

Supposons donc $\vec{y}$ non colinéaire à $\vec{x}$. Alors le trièdre — attention ! la figure ci-dessous a choisi l'orientation opposée — :

$$\big(\vec{x},\, \varphi(\vec{y}),\, \vec{x}\wedge\vec{y}\big) \qquad \text{est trirectangle positif.} \tag{4.5}$$

Constatons que ce trièdre est de même signe que le trièdre $\big(\vec{x}, \vec{y}, \vec{x}\wedge\vec{y}\big)$, grâce à une propriété élémentaire de la projection orthogonale φ sur P parallèlement à $\vec{x}$; cette constatation est d'ailleurs en accord avec la présentation que nous avons faite du procédé de Gram-Schmidt pour déterminer le sens, direct ou indirect, d'un trièdre.

D'autre part, si α désigne la mesure de l'angle $\mathsf{Angle}\big(\vec{x}, \vec{y}\big)$, on a évidemment :

$$\big\|\varphi(\vec{y})\big\| = \|\vec{y}\| \sin\alpha,$$

et par conséquent :

$$\big\|\vec{x}\wedge\vec{y}\big\| = \|\vec{x}\|\, \big\|\varphi(\vec{y})\big\|. \tag{4.6}$$

Comme ci-dessus, soit ensuite $\boldsymbol{r}$ la rotation d'axe $\vec{x}$ et d'angle $+\frac{\pi}{2}$. Elle envoie le vecteur $\varphi(\vec{y})$, orthogonal à $\vec{x}$, sur le vecteur $r\circ\varphi(\vec{y})$ tel que :

$$\Big(\vec{x},\, \varphi(\vec{y}),\, \boldsymbol{r}\circ\varphi(\vec{y})\Big) \qquad \text{est trirectangle positif,}$$

en préservant les normes :

$$\big\|\boldsymbol{r}\circ\varphi(\vec{y})\big\| = \big\|\varphi(\vec{y})\big\|.$$

Enfin, soit h l'homothétie de rapport strictement positif égal à $\|\vec{x}\|$. Elle applique le vecteur $\boldsymbol{r}\circ\varphi(\vec{y})$ sur le vecteur :

$$\vec{z} := h\circ\boldsymbol{r}\circ\varphi(\vec{y}),$$

tel que :

$$\big(\vec{x},\, \varphi(\vec{y}),\, \vec{z}\big) \quad \text{est trirectangle positif,} \tag{4.7}$$

et que :

$$\big\|\vec{z}\big\| \;=\; \big\|\vec{x}\big\|\;\big\|\varphi(\vec{y})\big\| \tag{4.8}$$

En comparant alors (4.5) et (4.7) d'une part, (4.6) et (4.8) d'autre part, nous obtenons bien :

$$\vec{z} \;=\; \vec{x} \wedge \vec{y}. \qquad \square$$

Proposition 4.9. *Quels que soient les vecteurs $\vec{x}$ et $\vec{y}$ ainsi que le nombre réel $a \in \mathbb{R}$, on a :*

$$\big(a\,\vec{x}\big) \wedge \vec{y} \;=\; \vec{x} \wedge \big(a\,\vec{y}\big) \;=\; a\,\big(\vec{x} \wedge \vec{y}\big).$$

Démonstration. La propriété est évidente pour $\vec{x} = 0$. Supposons donc $\vec{x} \neq \vec{0}$.

Puisque les applications $\varphi, \boldsymbol{r}, h$ qui interviennent dans le Théorème 4.4 sont des endomorphismes linéaires de l'espace vectoriel $V_{\mathbb{R}^3}$, alors la composée :

$$h \circ \boldsymbol{r} \circ \varphi$$

est aussi un endomorphisme linéaire de $V_{\mathbb{R}^3}$. Par suite, on a :

$$h \circ \boldsymbol{r} \circ \varphi\big(a\,\vec{y}\big) \;=\; a\,h \circ \boldsymbol{r} \circ \varphi(\vec{y}),$$

c'est-à-dire, en appliquant le Théorème 4.4 :

$$\vec{x} \wedge \big(a\,\vec{y}\big) \;=\; a\,\big(\vec{x} \wedge \vec{y}\big).$$

Enfin, grâce à l'antisymétrie du produit vectoriel, ce premier résultat permet d'obtenir sans effort l'autre résultat comme suit :

$$\big(a\,\vec{x}\big) \wedge \vec{y} \;=\; -\,\vec{y} \wedge \big(a\,\vec{x}\big) \;=\; -\,a\,\big(\vec{y} \wedge \vec{x}\big) \;=\; a\,\big(\vec{x} \wedge \vec{y}\big). \qquad \square$$

Montrons maintenant que le produit vectoriel est distributif par rapport à l'addition vectorielle.

Proposition 4.10. *Quels que soient les vecteurs $\vec{x}, \vec{y}, \vec{z} \in V_{\mathbb{R}^3}$, on a :*

$$\begin{aligned} \big(\vec{x} + \vec{y}\big) \wedge \vec{z} \;&=\; \vec{x} \wedge \vec{z} + \vec{y} \wedge \vec{z}, \\ \vec{x} \wedge \big(\vec{y} + \vec{z}\big) \;&=\; \vec{x} \wedge \vec{y} + \vec{x} \wedge \vec{z}. \end{aligned}$$

Démonstration. Grâce à l'antisymétrie du produit vectoriel qui vient d'être exploitée, la première propriété de distributivité se déduit de la deuxième sans effort aussi.

Cette deuxième propriété est évidente lorsque $\vec{x} = \vec{0}$. Supposons donc $\vec{x} \neq \vec{0}$.

Puisque $h \circ \boldsymbol{r} \circ \varphi$ est un endomorphisme linéaire de $V_{\mathbb{R}^3}$, on a, pour tous vecteurs $\vec{y}, \vec{z}$:

$$h \circ \boldsymbol{r} \circ \varphi\big(\vec{y} + \vec{z}\big) \;=\; h \circ \boldsymbol{r} \circ \varphi(\vec{y}) + h \circ \boldsymbol{r} \circ \varphi(\vec{z}).$$

En appliquant alors le Théorème 4.4, nous concluons bien que :

$$\vec{x} \wedge \big(\vec{y} + \vec{z}\big) \;=\; \vec{x} \wedge \vec{y} + \vec{x} \wedge \vec{z}. \qquad \square$$

Maintenant, que se passe-t-il avec les vecteurs de base $\vec{e}_1$, $\vec{e}_2$, $\vec{e}_3$? Comme la base $\big\{\vec{e}_1, \vec{e}_2, \vec{e}_3\big\}$ est orthonormée, on a unitarité :

$$1 \;=\; \big\|\vec{e}_1\big\| \;=\; \big\|\vec{e}_2\big\| \;=\; \big\|\vec{e}_3\big\|,$$

ainsi qu'orthogonalité :

$$\vec{e}_2 \perp \vec{e}_3, \qquad\qquad \vec{e}_3 \perp \vec{e}_1, \qquad\qquad \vec{e}_1 \perp \vec{e}_2,$$

c'est-à-dire :

$$0 \;=\; \vec{e}_2 \cdot \vec{e}_3 \;=\; \vec{e}_3 \cdot \vec{e}_1 \;=\; \vec{e}_1 \cdot \vec{e}_2.$$

Comme la base $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ et celles (4.1) qui s'en déduisent pas permutation circulaire sont toutes positives, alors les relations précédentes entraînent :

$$\boxed{\begin{aligned} \vec{e}_2 \wedge \vec{e}_3 &= \vec{e}_1, \\ \vec{e}_3 \wedge \vec{e}_1 &= \vec{e}_2, \\ \vec{e}_1 \wedge \vec{e}_2 &= \vec{e}_3. \end{aligned}} \tag{4.11}$$

Considérons maintenant deux vecteurs quelconques $\vec{x}$ et $\vec{y}$, donnés par leurs coordonnées dans la base de référence :

$$\begin{aligned} \vec{x} &= x_1\,\vec{e}_1 + x_2\,\vec{e}_2 + x_3\,\vec{e}_3, \\ \vec{y} &= y_1\,\vec{e}_1 + y_2\,\vec{e}_2 + y_3\,\vec{e}_3. \end{aligned}$$

Grâce aux Propositions 4.9 et 4.10, le produit vectoriel $\vec{x} \wedge \vec{y}$ se calcule aisément :

$$\begin{aligned} \vec{x} \wedge \vec{y} &= \big(x_1\,\vec{e}_1 + x_2\,\vec{e}_2 + x_3\,\vec{e}_3\big) \wedge \big(y_1\,\vec{e}_1 + y_2\,\vec{e}_2 + y_3\,\vec{e}_3\big) \\ &= x_1\,y_1\,\underline{\vec{e}_1 \wedge \vec{e}_1}_{\circ} + x_1\,y_2\,\vec{e}_1 \wedge \vec{e}_2 + x_1\,y_3\,\vec{e}_1 \wedge \vec{e}_3 \\ &\quad + x_2\,y_1\,\vec{e}_2 \wedge \vec{e}_1 + x_2\,y_2\,\underline{\vec{e}_2 \wedge \vec{e}_2}_{\circ} + x_2\,y_3\,\vec{e}_2 \wedge \vec{e}_3 \\ &\quad + x_3\,y_1\,\vec{e}_3 \wedge \vec{e}_1 + x_3\,y_2\,\vec{e}_3 \wedge \vec{e}_2 + x_3\,y_3\,\underline{\vec{e}_3 \wedge \vec{e}_3}_{\circ}. \end{aligned}$$

Comme d'habitude, nous avons souligné les termes qui s'annulent. En utilisant :

$$\vec{e}_{i_1} \wedge \vec{e}_{i_2} = -\,\vec{e}_{i_2} \wedge \vec{e}_{i_1} \qquad\qquad (1 \leqslant i_1, i_2 \leqslant 3),$$

nous pouvons alors écrire :

$$\begin{aligned} \vec{x} \wedge \vec{y} = 0 &\qquad\qquad + x_1\,y_2\,\vec{e}_1 \wedge \vec{e}_2 - x_1\,y_3\,\vec{e}_3 \wedge \vec{e}_1 \\ - x_2\,y_1\,\vec{e}_1 \wedge \vec{e}_2 &+ 0 \qquad\qquad + x_2\,y_3\,\vec{e}_2 \wedge \vec{e}_3 \\ + x_3\,y_1\,\vec{e}_3 \wedge \vec{e}_1 &- x_3\,y_2\,\vec{e}_2 \wedge \vec{e}_3 + 0. \end{aligned}$$

et enfin regrouper par paires les 6 termes restants pour obtenir un résultat fondamental.

Théorème 4.12. *Dans la base orthonormale canonique $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ de l'espace, le produit vectoriel entre deux vecteurs quelconques :*

$$\begin{aligned} \vec{x} &= x_1\,\vec{e}_1 + x_2\,\vec{e}_2 + x_3\,\vec{e}_3, \\ \vec{y} &= y_1\,\vec{e}_1 + y_2\,\vec{e}_2 + y_3\,\vec{e}_3, \end{aligned}$$

vaut :

$$\vec{x} \wedge \vec{y} = \big(x_2y_3 - y_2x_3\big)\,\vec{e}_1 + \big(x_3y_1 - y_3x_1\big)\,\vec{e}_2 + \big(x_1y_2 - y_1x_2\big)\,\vec{e}_3. \qquad \square$$

5. Applications bilinéaires

Nous savons que le produit scalaire euclidien est une application bilinéaire :

$$\begin{aligned} V_{\mathbb{R}^3} \times V_{\mathbb{R}^3} &\longrightarrow \mathbb{R} \\ \big(\vec{x},\, \vec{y}\big) &\longmapsto \vec{x} \cdot \vec{y}, \end{aligned}$$

à savoir qu'il satisfait, comme l'a déjà énoncé le Lemme 2.1 :

$$\begin{aligned} \big(\lambda\,\vec{x} + \lambda'\,\vec{x}'\big) \cdot \vec{y} &= \lambda\,\vec{x} \cdot \vec{y} + \lambda'\,\vec{x}' \cdot \vec{y}, \\ \vec{x} \cdot \big(\mu\,\vec{y} + \mu'\,\vec{y}'\big) &= \mu\,\vec{x} \cdot \vec{y} + \mu'\,\vec{x} \cdot \vec{y}'. \end{aligned}$$

On parle de *bilinéarité*, car quand $\vec{y}$ est fixé, l'application $\vec{x} \longmapsto \vec{x} \cdot \vec{y}$ est linéaire, et de manière symétrique, quand $\vec{x}$ est fixé, l'application $\vec{y} \longmapsto \vec{x} \cdot \vec{y}$ est linéaire.

À travers les Proposition 4.9 et 4.10, nous venons de voir que le produit vectoriel :

$$\begin{aligned} V_{\mathbb{R}^3} \times V_{\mathbb{R}^3} &\longrightarrow \mathbb{R} \\ \big(\vec{x},\, \vec{y}\big) &\longmapsto \vec{x} \wedge \vec{y}, \end{aligned}$$

est aussi une application bilinéaire, c'est-à-dire qu'il satisfait :

$$\begin{aligned}\big(\lambda\,\vec{x}+\lambda'\,\vec{x}'\big)\wedge\vec{y} &= \lambda\,\vec{x}\wedge\vec{y}+\lambda'\,\vec{x}'\wedge\vec{y},\\ \vec{x}\wedge\big(\mu\,\vec{y}+\mu'\,\vec{y}'\big) &= \mu\,\vec{x}\wedge\vec{y}+\mu'\,\vec{x}\wedge\vec{y}'.\end{aligned}$$

Ainsi, quand $\vec{y}$ est fixé, l'application $\vec{x}\longmapsto\vec{x}\wedge\vec{y}$ est linéaire, et de manière symétrique, quand $\vec{x}$ est fixé, l'application $\vec{y}\longmapsto\vec{x}\wedge\vec{y}$ est linéaire.

Toutes ces observations motivent une conceptualisation générale « *en abstraction* » de ces propriétés. Revenons à un corps arbitraire $\mathbb{K}$ — penser $\mathbb{K}=\mathbb{Q}$, $\mathbb{R}$, ou $\mathbb{C}$.

Définition 5.1. Soient E, F, G trois espaces vectoriels sur le même corps $\mathbb{K}$. Une application f de $E\times F$ dans G sera dite *bilinéaire* si, quels que soient les vecteurs :

$$\vec{x},\,\vec{x}'\in E \qquad\qquad \text{et} \qquad\qquad \vec{y},\,\vec{y}'\in F,$$

et quel que soit le scalaire $a\in\mathbb{K}$, elle satisfait :

$$\begin{aligned}f\big(a\,\vec{x},\,\vec{y}\big) &= f\big(\vec{x},\,a\,\vec{y}\big) = a\,f\big(\vec{x},\,\vec{y}\big),\\ f\big(\vec{x}+\vec{x}',\,\vec{y}\big) &= f\big(\vec{x},\,\vec{y}\big)+f\big(\vec{x}',\,\vec{y}\big),\\ f\big(\vec{x},\,\vec{y}+\vec{y}'\big) &= f\big(\vec{x},\,\vec{y}\big)+f\big(\vec{x},\,\vec{y}'\big).\end{aligned}$$

En d'autres termes, l'application $(\vec{x},\vec{y})\longmapsto f(\vec{x},\vec{y})$ est, pour $\vec{y}$ fixé, linéaire selon la variable $\vec{x}$, et pour $\vec{x}$ fixé, linéaire selon la variable $\vec{y}$. C'est bien pourquoi f est nommée *bi*linéaire.

Terminologie 5.2. Maintenant, lorsque l'espace d'arrivée G où f prend ses valeurs coïncide avec le corps $\mathbb{K}$ lui-même, f est nommée *forme bilinéaire*.

Par exemple, le produit scalaire est une forme bilinéaire de $V_{\mathbb{R}^3}\times V_{\mathbb{R}^3}$ dans $\mathbb{R}$. À ce sujet, remarquons que le produit scalaire $f:=\cdot$ est de plus commutatif :

$$f(\vec{x},\vec{y}) = f(\vec{y},\vec{x}),$$

ce qui n'est *pas* le cas pour toute application bilinéaire.

Terminologie 5.3. Quand $f(\vec{x},\vec{y})=f(\vec{y},\vec{x})$ pour tous vecteurs $\vec{x}$, $\vec{y}$, on dit que la forme bilinéaire f est *symétrique*. Et quand $E=F$ avec :

$$f(\vec{y},\vec{x}) = -\,f(\vec{x},\vec{y}) \qquad (\forall\,\vec{x}\in E,\ \forall\,\vec{y}\in E),$$

on dit que la forme f est *antisymétrique*.

Supposons donc maintenant que $E=F$.

Définition 5.4. Une application bilinéaire de $E\times E$ dans G est dite *alternée* si, quel que soit le vecteur $\vec{x}\in E$, elle satisfait :

$$f(\vec{x},\vec{x}) = \vec{0}.$$

Par exemple, le produit vectoriel est une application bilinéaire alternée de $V_{\mathbb{R}^3}\times V_{\mathbb{R}^3}$ dans $V_{\mathbb{R}^3}$, car nous savons que :

$$\vec{x}\wedge\vec{x} = \vec{0} \qquad (\forall\,\vec{x}\in V_{\mathbb{R}^3}).$$

Cependant, le produit scalaire euclidien n'est *pas* alterné, car on sait que :

$$\vec{x}\cdot\vec{x} = \big\|\vec{x}\big\|^2 \qquad (\forall\,\vec{x}\in V_{\mathbb{R}^3}),$$

et on sait que :

$$0 = \big\|\vec{x}\big\| \qquad \Longrightarrow \qquad \vec{x} = \vec{0}.$$

Il est évident qu'une application antisymétrique est alternée, car :

$$f(\vec{x},\vec{x}) = -\,f(\vec{x},\vec{x}) \qquad \Longrightarrow \qquad 2\,f(\vec{x},\vec{x}) = \vec{0}.$$

Mais le fait d'être alterné est équivalent au fait d'être antisymétrique, comme le fait voir la

Proposition 5.5. *Si f est une application bilinéaire alternée de $E \times E$ dans G, alors quels que soient les vecteurs $\vec{x}$ et $\vec{y}$ de E, on a :*

$$f(\vec{x}, \vec{y}) \,=\, -\, f(\vec{y}, \vec{x}).$$

Démonstration. Partons de :

$$\vec{0} \,=\, f\big(\vec{x} + \vec{y},\, \vec{x} + \vec{y}\big),$$

utilisons la bilinéarité pour développer :

$$\vec{0} \,=\, \underline{f\big(\vec{x}, \vec{x}\big)}_{\circ} + f\big(\vec{x}, \vec{y}\big) + f\big(\vec{y}, \vec{x}\big) + \underline{f\big(\vec{y}, \vec{y}\big)}_{\circ},$$

et appliquons l'hypothèse que f est alternée pour obtenir effectivement :

$$\vec{0} \,=\, f\big(\vec{x}, \vec{y}\big) + f\big(\vec{y}, \vec{x}\big).$$

□

Nous allons maintenant étudier les applications bilinéaires alternées dans le cas où E et de dimension 2 sur le corps $\mathbb{K}$, l'espace vectoriel G étant quelconque. Ceci va nous permettre d'introduire pour la toute première fois des objets nouveaux dont nous reparlerons énormément au chapitre suivant, les *déterminants*, ici dans le cas prototypique des déterminants de taille 2×2, les plus simples d'entre tous.

Théorème 5.6. *Soient E et G deux espaces vectoriels sur un même corps $\mathbb{K}$, avec :*

$$\dim_{\mathbb{K}} E \,=\, 2.$$

Pour toute base $\big\{\vec{e}_1, \vec{e}_2\big\}$ de E, et pour tout vecteur $\vec{k} \in G$, il existe une application bilinéaire alternée, et une seule, $f \colon E \times E \longrightarrow G$ telle que :

$$f\big(\vec{e}_1, \vec{e}_2\big) \,=\, \vec{k}.$$

Démonstration. Toute paire de vecteurs $\vec{x}$, $\vec{y}$ sera rapportée à la base choisie :

$$\begin{aligned} \vec{x} &\,=\, x_1\, \vec{e}_1 + x_2\, \vec{e}_2, \\ \vec{y} &\,=\, y_1\, \vec{e}_1 + y_2\, \vec{e}_2. \end{aligned}$$

Commençons par établir l'*unicité* de $f \colon E \times E \longrightarrow G$. Comme f est supposée alternée, on a :

$$\begin{aligned} \vec{0} \,=\, f\big(\vec{e}_1, \vec{e}_1\big) &\,=\, f\big(\vec{e}_2, \vec{e}_2\big), \\ f\big(\vec{e}_1, \vec{e}_2\big) &\,=\, -\, f\big(\vec{e}_2, \vec{e}_1\big). \end{aligned}$$

Comme f est bilinéaire, on peut développer :

$$\begin{aligned} f(\vec{x}, \vec{y}) &\,=\, f\Big(x_1\, \vec{e}_1 + x_2\, \vec{e}_2,\ y_1\, \vec{e}_1 + y_2\, \vec{e}_2\Big) \\ &\,=\, x_1\, y_1\, \underline{f\big(\vec{e}_1, \vec{e}_1\big)}_{\circ} + x_1\, y_2\, f\big(\vec{e}_1, \vec{e}_2\big) + x_2\, y_1\, f\big(\vec{e}_2, \vec{e}_1\big) + x_2\, y_2\, \underline{f\big(\vec{e}_2, \vec{e}_2\big)}_{\circ} \\ &\,=\, \big(x_1\, y_2 - y_1\, x_2\big)\, f\big(\vec{e}_1, \vec{e}_2\big). \end{aligned}$$

Ceci prouve que, si f existe, elle coïncide nécessairement avec l'application uniquement déterminée :

$$f(\vec{x}, \vec{y}) \,:=\, \big(x_1\, y_2 - y_1\, x_2\big)\, \vec{k}.$$

Ensuite, pour faire voir l'*existence réelle* de f, prouvons que l'application ainsi définie en termes de $\vec{k}$ est bien bilinéaire alternée.

Pour $\vec{y} \in E$ fixé, on a en effet, quels que soient $\vec{x}, \vec{x}\,' \in E$ et quels que soient $a, a' \in \mathbb{K}$:

$$\begin{aligned} f\,\big(a\, \vec{x} + a'\, \vec{x}\,',\ \vec{y}\big) &\,=\, \Big[\big(a\, x_1 + a'\, x_1'\big)\, y_2 - y_1\, \big(a\, x_2 + a'\, x_2'\big)\Big]\, \vec{k} \\ &\,=\, a\, \big(x_1\, y_2 - y_1 x_2\big)\, \vec{k} + a'\, \big(x_1'\, y_2 - y_1 x_2'\big)\, \vec{k} \\ &\,=\, a\, f(\vec{x}, \vec{y}) + a'\, f(\vec{x}\,', \vec{y}). \end{aligned}$$

Ainsi, $f(\vec{x}, \vec{y})$ est bien linéaire en $\vec{x}$. On prouverait de même qu'elle est aussi linéaire en $\vec{y}$.

Enfin, f est trivialement alternée, car :

$$f(\vec{x},\vec{x}) \;=\; \big(x_1\,x_2 - x_1\,x_2\big)\,\vec{k} \;=\; \vec{0}. \qquad \square$$

Nous pouvons maintenant introduire les déterminants d'ordre 2. Plaçons-nous dans les hypothèses du Théorème 5.6, avec de plus :

$$G \;:=\; \mathbb{K}.$$

Choisissons comme vecteur $\vec{k}$ le « vecteur » 1 de l'espace vectoriel $\mathbb{K}$. Ledit théorème affirme alors qu'il existe une unique *forme* bilinéaire alternée f telle que :

$$f\big(\vec{e}_1,\vec{e}_2\big) \;=\; 1.$$

Terminologie 5.7. Une telle forme se nomme alors *déterminant des vecteurs $\vec{x}$ et $\vec{y}$ par rapport à la base* $\big\{\vec{e}_1,\vec{e}_2\big\}$.

On note ce déterminant :

$$x_1\,y_2 - y_1\,x_2 \;=:\; \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}.$$

Ce tableau encadré de deux barres verticales — belle escorte ! — est constitué de par les deux colonnes des coordonnées des vecteurs $\vec{x}$ et $\vec{y}$ dans la base considérée.

Attention ! Il ne faut *absolument pas* confondre cette notation avec celle d'une matrice carrée d'ordre 2, pour laquelle les escortes sont deux grandes parenthèses.

D'ailleurs, si on part d'une matrice de $\mathscr{M}_{2\times 2}(\mathbb{K})$:

$$M \;:=\; \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

on appelle *déterminant de M* la quantité notée :

$$\mathsf{det}\,M \;:=\; \begin{vmatrix} a & b \\ c & d \end{vmatrix} \;=\; ad - bc.$$

Observons au passage que d'après cette définition concrète, on voit immédiatement — exercice visuel, crayon non interdit en cas de panne de cervelle — que :

$$\mathsf{det}\,\big({}^{\mathsf{t}}M\big) \;=\; \mathsf{det}\,M.$$

De plus, on constate que si on échange deux lignes, le déterminant change de signe, et pareillement si on change deux colonnes.

Le prochain chapitre montrera cela à l'envi, et beaucoup, beaucoup plus !

6. Produit mixte

Revenons à l'espace vectoriel $V_{\mathbb{R}^3}$ de la géométrie euclidienne.

Définition 6.1. À tout triplet ordonné $\vec{x}$, $\vec{y}$, $\vec{z}$ de trois vecteurs de $V_{\mathbb{R}^3}$, on associe un nombre réel, nommé *produit mixte* des trois vecteurs, noté et défini par :

$$\big(\vec{x}\,|\,\vec{y}\,|\,\vec{z}\big) \;:=\; \big(\vec{x}\wedge\vec{y}\big)\cdot\vec{z}.$$

C'est le produit scalaire de $\vec{z}$ par le produit vectoriel $\vec{x}\wedge\vec{y}$.

Le produit mixte établit donc une application :

$$\begin{aligned} V_{\mathbb{R}^3}\times V_{\mathbb{R}^3}\times V_{\mathbb{R}^3} &\longrightarrow \mathbb{R} \\ \big(\vec{x},\vec{y},\vec{z}\big) &\longmapsto \big(\vec{x}\,|\,\vec{y}\,|\,\vec{z}\big). \end{aligned}$$

Par construction, le produit mixte est nul si et seulement si l'une des trois éventualités suivantes est satisfaite.

(1) L'un des vecteurs $\vec{x}$, $\vec{y}$, $\vec{z}$ est nul.

(2) Les vecteurs $\vec{x}$ et $\vec{y}$ sont colinéaire, propriété équivalente à $\vec{0} = \vec{x}\wedge\vec{y}$.

(3) Les vecteurs $\vec{x} \wedge \vec{y}$ et $\vec{z}$ sont non nuls et orthogonaux. Dans ce cas, le vecteur $\vec{z}$ appartient alors au sous-espace à deux dimensions V_P défini par $\vec{x}$ et $\vec{y}$, où $P := \mathsf{Vect}(\vec{x}, \vec{y})$.

Définition 6.2. Trois vecteurs $\vec{x}, \vec{y}, \vec{z}$ sont dits *coplanaires* lorsqu'il existe un sous-espace V_P à deux dimensions qui les contient tous les trois.

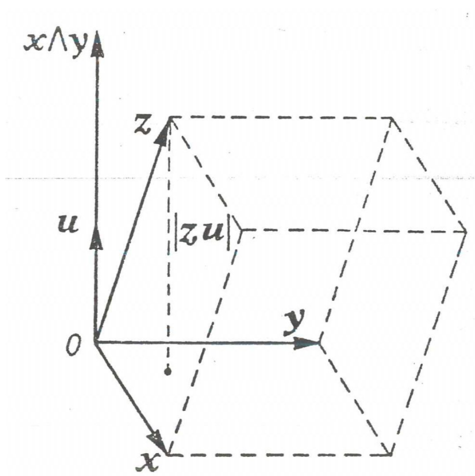
En d'autres termes, $\vec{x}$, $\vec{y}$, $\vec{z}$ sont coplanaires s'ils sont tous parallèles à un même plan vectoriel. D'après ce qui précède, on peut donc énoncer en résumé une

Proposition 6.3. *Pour que trois vecteurs soient coplanaires, il faut et il suffit que leur produit mixte soit nul.* □

Maintenant, à quoi sert au juste le produit mixte ? Possède-t-il une signification géométrique éclairante ? Ah, que oui, petzi !

En effet, on peut interpréter le produit mixte $(\vec{x} \,|\, \vec{y} \,|\, \vec{z})$ comme le *volume signé* du parallélépipède construit sur $\vec{x}$, $\vec{y}$, $\vec{z}$.

On sait déjà que $\|\vec{x} \wedge \vec{y}\|$ est la mesure de l'aire (positive) du parallélogramme construit sur les deux vecteurs $\vec{x}$ et $\vec{y}$.



Produit mixte: $(x \,|\, y \,|\, z)$

Désignons alors par $\vec{u}$ le vecteur unitaire unique tel que :

$$\vec{x} \wedge \vec{y} = \|\vec{x} \wedge \vec{y}\| \, \vec{u} \qquad \Longleftrightarrow \qquad \vec{u} := \frac{\vec{x} \wedge \vec{y}}{\|\vec{x} \wedge \vec{y}\|}.$$

On a alors, pour le produit mixte :

$$(\vec{x} \,|\, \vec{y} \,|\, \vec{z}) = \|\vec{x} \wedge \vec{y}\| \, (\vec{u} \cdot \vec{z}).$$

La valeur absolue $|\vec{u} \cdot \vec{z}|$ du produit scalaire de $\vec{u}$ par $\vec{z}$ est la mesure de la hauteur du parallélépipède, relative à la base constituée par $\mathsf{Vect}(\vec{x}, \vec{y})$. Par conséquent on a un

Théorème 6.4. *Le volume du parallélépipède construit sur trois vecteurs $\vec{x}$, $\vec{y}$, $\vec{z}$ dans l'espace euclidien $V_{\mathbb{R}^3}$ vaut :*

$$\mathsf{volume} = \big|(\vec{x} \,|\, \vec{y} \,|\, \vec{z})\big|.$$

On peut, si l'on veut, interpréter le signe du produit mixte en introduisant la notion de *volume signé*. En effet, on dira qu'un parallélépipède construit sur trois vecteurs linéairement indépendants $\vec{x}$, $\vec{y}$, $\vec{z}$ est *positif* ou *négatif* suivant que le trièdre est positif ou négatif, et on posera, sans valeur absolue :

$$\mathsf{volume\ signé} := (\vec{x} \,|\, \vec{y} \,|\, \vec{z}).$$

Maintenant, comment calculer le produit mixte en coordonnées ? À cette fin, choisissons la base orthonormée canonique $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ que nous avons plusieurs fois utilisée. Dans cette base, soient (x_1, x_2, x_3) les coordonnées de $\vec{x}$, soient (y_1, y_2, y_3) les coordonnées de $\vec{y}$, soient (z_1, z_2, z_3) les coordonnées de $\vec{z}$.

Nous savons que les trois coordonnées de $\vec{x} \wedge \vec{y}$ sont :

$$\big(x_2y_3 - y_2x_3\big), \qquad \big(x_3y_1 - y_3x_1\big), \qquad \big(x_1y_2 - y_1x_2\big).$$

Par conséquent, la valeur du produit mixte en question est :

$$\begin{aligned}\big(\vec{x}\,|\,\vec{y}\,|\,\vec{z}\big) &= \big(\vec{x} \wedge \vec{y}\big) \cdot \vec{z} \\ &= \big(x_2y_3 - y_2x_3\big)\, z_1 + \big(x_3y_1 - y_3x_1\big)\, z_2 + \big(x_1y_2 - y_1x_2\big)\, z_3,\end{aligned}$$

et si nous développons et réorganisons les termes, nous obtenons une formule finalisée et harmonieuse :

$$\big(\vec{x}\,|\,\vec{y}\,|\,\vec{z}\big) = x_1y_2z_3 + x_2y_3z_1 + x_3y_1z_2 - x_3y_2z_1 - x_2y_1z_3 - x_1y_3z_2.$$

En effet, remarquons que les produits $x_iy_jz_k$ sont précédés du signe '$+$' ou du signe '$-$' selon que $\{i, j, k\}$ est l'image de $\{1, 2, 3\}$ par une permutation *paire* ou *impaire* :

$$\begin{array}{rccc} \text{permutations paires :} & \big\{1,2,3\big\}, & \big\{2,3,1\big\}, & \big\{3,1,2\big\}, \\ \text{permutations impaires :} & \big\{2,1,3\big\}, & \big\{1,3,2\big\}, & \big\{3,2,1\big\}. \end{array}$$

Il importe de remarquer que, d'après sa définition même, le produit mixte est indépendant de la base orthonormée choisie.

Proposition 6.5. *Le produit mixte $\big(\vec{x}\,|\,\vec{y}\,|\,\vec{z}\big)$ est linéaire en chacun des trois vecteurs $\vec{x}$, $\vec{y}$, $\vec{z}$.*

Démonstration. Puisque nous savons que le produit vectoriel est bilinéaire, nous avons en particulier pour tous vecteurs $\vec{x}, \vec{x}', \vec{y} \in V_{\mathbb{R}^3}$ et tous scalaires $a, a' \in \mathbb{R}$:

$$\big(a\,\vec{x} + a'\,\vec{x}'\big) \wedge \vec{y} = a\,\big(\vec{x} \wedge \vec{y}\big) + a'\,\big(\vec{x}' \wedge \vec{y}\big).$$

En multipliant scalairement cela par $\vec{z}$, nous obtenons, puisque le produit scalaire $\big(\vec{v}, \vec{z}\big) \longmapsto \vec{v}\cdot\vec{z}$ est linéaire en $\vec{v}$:

$$\big(a\,\vec{x} + a'\,\vec{x}'\,\big|\,\vec{y}\,\big|\,\vec{z}\big) = a\,\big(\vec{x}\,|\,\vec{y}\,|\,\vec{z}\big) + a'\,\big(\vec{x}'\,|\,\vec{y}\,|\,\vec{z}\big).$$

Donc le produit mixte est linéaire en son premier argument $\vec{x}$.

On prouverait de même qu'il est linéaire en $\vec{y}$ et en $\vec{z}$: ceci résulte du fait que le produit scalaire et le produit vectoriel sont tous deux bilinéaires. □

Proposition 6.6. *Le produit mixte change de signe quand on échange deux vecteurs quelconques.*

Démonstration. La démonstration est immédiate si l'on considère le parallélépipède orienté construit sur le triplet ordonné $\big(\vec{x}, \vec{y}, \vec{z}\big)$.

Si l'on échange deux vecteurs du triplet, le parallélépipède change simplement d'orientation, et son volume signé change alors de signe. □

7. Applications trilinéaires alternées

Prenons deux espaces vectoriels E et G sur le même corps $\mathbb{K}$.

Définition 7.1. Une application :

$$\begin{aligned} f\colon \quad E \times E \times E &\longrightarrow G \\ \big(\vec{x}, \vec{y}, \vec{z}\big) &\longmapsto f\big(\vec{x}, \vec{y}, \vec{z}\big) \end{aligned}$$

est dite *trilinéaire alternée* si :

(1) elle est linéaire en chacun des vecteurs $\vec{x}$, $\vec{y}$, $\vec{z}$;

(2) $f(\vec{x}, \vec{y}, \vec{z}) = \vec{0}$ dès que deux parmi trois vecteurs coïncident.

En d'autres termes, si l'on fixe l'un quelconque des trois vecteurs, l'application f est *bilinéaire alternée* en les deux autres vecteurs.

Par conséquent :

Observation 7.2. *$f(\vec{x}, \vec{y}, \vec{z})$ change de signe dès que l'on échange deux vecteurs.* □

D'autre part :

Observation 7.3. *Si l'un des trois vecteurs est combinaison linéaire des deux autres, alors $f(\vec{x}, \vec{y}, \vec{z}) = \vec{0}$.* □

Par exemple, avec :

$$\vec{z} = a\,\vec{x} + b\,\vec{y},$$

on trouve l'annulation complète :

$$\begin{aligned} f(\vec{x}, \vec{y}, \vec{z}) &= a\,\underline{f(\vec{x}, \vec{y}, \vec{x})} + b\,\underline{f(\vec{x}, \vec{y}, \vec{y})} \\ &= 0 \qquad\qquad\quad + 0. \end{aligned}$$

Il en résulte un énoncé très important dans la pratique.

Proposition 7.4. *Tout application trilinéaire alternée $f(\vec{x}, \vec{y}, \vec{z})$ ne change pas quand on ajoute à l'un des vecteurs une combinaison linéaire quelconque des deux autres :*

$$\boxed{f\big(\vec{x},\ \vec{y},\ \vec{z} + a\,\vec{x} + b\,\vec{y}\big) = f(\vec{x}, \vec{y},\ \vec{z}).}$$ □

Terminologie 7.5. Lorsque l'espace vectoriel G coïncide avec le corps $\mathbb{K}$, une application trilinéaire prend le nom de *forme trilinéaire*.

Par exemple, le produit mixte est une forme trilinéaire alternée de $V_{\mathbb{R}^3}$ à valeurs dans $\mathbb{R}$.

Nous allons maintenant étudier les applications trilinéaires alternées dans le cas où E est de dimension 3 sur le corps $\mathbb{K}$, toujours avec un $\mathbb{K}$-espace vectoriel G quelconque. Ceci nous permettra d'introduire de manière naturelle les déterminants de taille 3×3.

Théorème 7.6. *Soient E et G deux $\mathbb{K}$-espaces vectoriels, avec $\dim_{\mathbb{K}} E = 3$. Pour toute base $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ de E et tout vecteur $\vec{k} \in G$, il existe une unique application trilinéaire $f \colon E \times E \times E \longrightarrow G$ telle que :*

$$f\big(\vec{e}_1, \vec{e}_2, \vec{e}_3\big) = \vec{k}.$$

Démonstration. Commençons par établir l'*unicité* de $f \colon E \times E \times E \longrightarrow G$. Décomposons dans la base :

$$\vec{x} = \sum_{i=1}^{3} x_i\,\vec{e}_i, \qquad\qquad \vec{y} = \sum_{j=1}^{3} y_j\,\vec{e}_j, \qquad\qquad \vec{z} = \sum_{k=1}^{3} z_k\,\vec{e}_k.$$

Puisque f est trilinéaire :

$$\begin{aligned} f\big(\vec{x}, \vec{y}, \vec{z}\big) &= f\Big(\sum_i x_i\,\vec{e}_i,\ \sum_j y_j\,\vec{e}_j,\ \sum_k z_k\,\vec{e}_k\Big) \\ &= \sum_i \sum_j \sum_k f\big(\vec{e}_i, \vec{e}_j, \vec{e}_k\big). \end{aligned}$$

Ici, la somme est étendue à tous les indices i, j, k dans $\{1, 2, 3\}$, donc il y a au total $3^3 = 27$ termes.

Or puisqu'on sait que $f\big(\vec{e}_i, \vec{e}_j, \vec{e}_k\big) = \vec{0}$ dès que deux indices sont égaux, il ne reste à considérer que les termes de la sommation où les trois indices i, j, k sont distincts, c'est-à-dire les cas où l'ensemble $\{i, j, k\}$ est l'image de $\{1, 2, 3\}$ par une permutation.

Il y a au total $3! = 6$ termes de cette espèce. De plus, ce que nous venons de dire garantit que les $27 - 6 = 21$ autres termes sont *tous* nuls.

Assertion 7.7. *Les six $f\big(\vec{e}_i, \vec{e}_j, \vec{e}_k\big)$ restants s'expriment tous en fonction de :*

$$\vec{k} \,=\, f\big(\vec{e}_1, \vec{e}_2, \vec{e}_3\big).$$

Preuve. En effet, puisque f est alternée, on a :

$$\begin{aligned}
f\big(\vec{e}_2, \vec{e}_1, \vec{e}_3\big) \,&=\, -\,\vec{k}, \\
f\big(\vec{e}_1, \vec{e}_3, \vec{e}_2\big) \,&=\, -\,\vec{k} \,=\, -\,f\big(\vec{e}_3, \vec{e}_1, \vec{e}_2\big), \\
f\big(\vec{e}_3, \vec{e}_2, \vec{e}_1\big) \,&=\, -\,\vec{k} \,=\, -\,f\big(\vec{e}_2, \vec{e}_3, \vec{e}_1\big).
\end{aligned}$$

Ici, nous voyons les 6 permutations de $\{1, 2, 3\}$, ce qui finit. □

On remarquera que $f\big(\vec{e}_1, \vec{e}_2, \vec{e}_3\big)$ vaut $\vec{k}$ ou $-\,\vec{k}$ selon que la permutation :

$$\big\{1, 2, 3\big\} \,\longmapsto\, \big\{\vec{e}_1,\, \vec{e}_2,\, \vec{e}_3\big\}$$

est paire ou impaire.

Nous obtenons par conséquent :

$$f\big(\vec{x}, \vec{y}, \vec{z}\big) \,=\, \Big(x_1y_2z_3 + x_2y_3z_1 + x_3y_1z_2 - x_3y_2z_1 - x_2y_1z_3 - x_1y_3z_2\Big)\,\vec{k}, \tag{7.8}$$

ce qui établit l'*unicité* de l'application trilinéaire f recherchée.

Il reste à démontrer son *existence*, car ce n'est pas parce que le *Monstre du Loch Ness* est unique qu'il existe — soit dit en passant !

À cette fin, prouvons que l'application f de $E \times E \times E$ à valeurs dans G qui est définie par (7.8) est bien trilinéaire alternée.

Établissons, par exemple, que f est trilinéaire en $\vec{x}$, les deux autres vecteurs $\vec{y}$ et $\vec{z}$ restant fixes. Comme les coordonnées de $\vec{x}$ interviennent une fois et une fois seulement dans chacun des six termes de la somme (7.8), et que les coordonnées de $a\,\vec{x}$ sont $a\,x_1$, $a\,x_2$, $a\,x_3$, alors, en remplaçant dans (7.8), nous obtenons :

$$f\big(a\,\vec{x}, \vec{y}, \vec{z}\big) \,=\, a\,f\big(\vec{x}, \vec{y}, \vec{z}\big).$$

Comme les coordonnesses — oh la faute, oh la monstresse ! — de $\vec{x} + \vec{x}\,'$ sont $x_1 + x_1'$, $x_2 + x_2'$, $x_3 + x_3'$, alors, en remplaçant dans (7.8), nous obtenons :

$$f\big(\vec{x} + \vec{x}\,',\, \vec{y}, \vec{z}\big) \,=\, f\big(\vec{x}, \vec{y}, \vec{z}\big) + f\big(\vec{x}\,', \vec{y}, \vec{z}\big).$$

On prouverait de même la linéarité de f en $\vec{y}$ ou en $\vec{z}$.

Enfin, si deux des trois vecteurs $\vec{x}$, $\vec{y}$, $\vec{z}$ sont égaux, alors la relation (7.8) montre immédiatement que $f(\vec{x}, \vec{y}, \vec{z}) = \vec{0}$.

Ceci conclut la démonstration du Théorème 7.6. □

8. Déterminants d'ordre 3

Plaçons-nous dans les hypothèses du Théorème 7.6, avec :

$$G \,:=\, \mathbb{K}.$$

Choisissons pour vecteur $\vec{k}$ le vecteur 1 du corps $\mathbb{K}$. Il existe alors une unique forme trilinéaire alternée f telle que $f\big(\vec{e}_1, \vec{e}_2, \vec{e}_3\big) = 1$.

Terminologie 8.1. Une telle forme f se nomme *déterminant des vecteurs $\vec{x}$, $\vec{y}$, $\vec{z}$ par rapport à la base $\big\{\vec{e}_1, \vec{e}_2, \vec{e}_3\big\}$*, et se note :

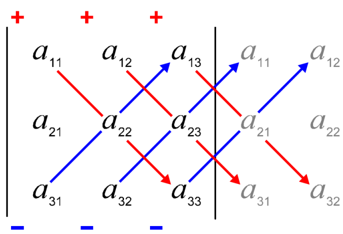
$$\begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}.$$

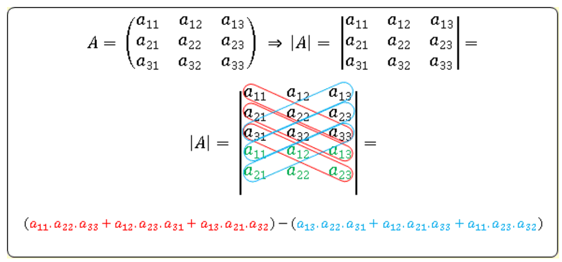
D'après (7.8), ce déterminant vaut :

$$\begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix} = x_1y_2z_3 + x_2y_3z_1 + x_3y_1z_2 - x_3y_2z_1 - x_2y_1z_3 - x_1y_3z_2.$$

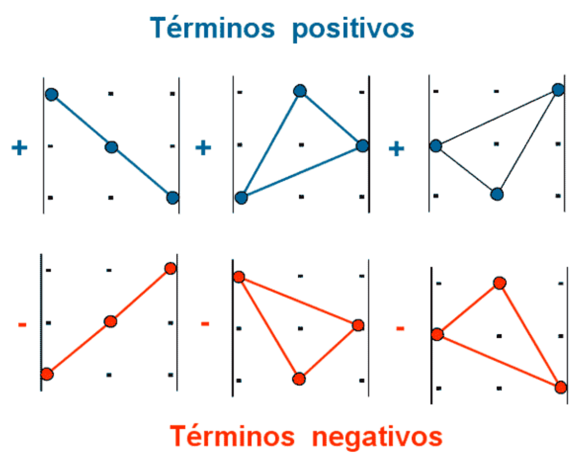
On dit que l'on a *développé* le déterminant selon la *règle de Sarrus :* les trois produits précédés du signe '$+$' contiennent soit les termes de la diagonale principale, soit deux termes d'une parallèle à cette diagonale. Pour les trois produits précédés du signe '$-$', il y a une règle analogue en remplaçant la diagonale principale par l'autre diagonale.

Donnons deux premiers diagrammes colorés :

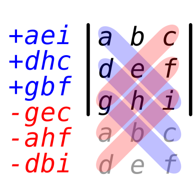


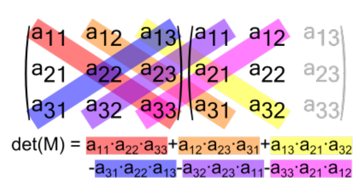


En fait, l'idée principale concernant les signes est illustrée comme suit :



Allez ! Encore deux autres illustrations bien flashys !





Soit une matrice carrée d'ordre 3 :

$$M := \begin{pmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{pmatrix}.$$

Définition 8.2. On appelle *déterminant* de M le scalaire :

$$\begin{aligned} \det M &:= \begin{vmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{vmatrix} \\ &:= ab'c'' + a'b''c + a''bc' - a''b'c - ab''c' - a'bc''. \end{aligned}$$

De cette définition, il résulte immédiatement que le déterminant de la matrice transposée ${}^{\mathrm{t}}M$ est le même que le déterminant de M :

$$\det\left({}^{\mathrm{t}}M\right) = \det M.$$

Par conséquent, toute propriété de $\det M$ valable pour les lignes est aussi valable pour les colonnes.

Formulons précisément ces propriétés qui découlent de celles d'une application trilinéaire alternée.

Proposition 8.3. **(1)** *Avec des déterminants* 3×3*, on peut factoriser un facteur commun aux éléments d'une rangée :*

$$\begin{vmatrix} \lambda a & a' & a'' \\ \lambda b & b' & b'' \\ \lambda c & c' & c'' \end{vmatrix} = \lambda \begin{vmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{vmatrix}.$$

(2) *Si les éléments d'une rangée sont des sommes, on peut mettre le déterminant sous la forme d'une somme de déterminants :*

$$\begin{vmatrix} a_1 + a_2 & a' & a'' \\ b_1 + b_2 & b' & b'' \\ c_1 + c_2 & c' & c'' \end{vmatrix} = \begin{vmatrix} a_1 & a' & a'' \\ b_1 & b' & b'' \\ c_1 & c' & c'' \end{vmatrix} + \begin{vmatrix} a_2 & a' & a'' \\ b_2 & b' & b'' \\ c_2 & c' & c'' \end{vmatrix}.$$

(3) *Quand on échange deux lignes (ou deux colonnes), le déterminant change de signe :*

$$\begin{vmatrix} b & b' & b'' \\ a & a' & a'' \\ c & c' & c'' \end{vmatrix} = - \begin{vmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{vmatrix}.$$

(4) *Si une ligne (ou une colonne) est combinaison linéaire des autres, le déterminant est nul :*

$$\begin{vmatrix} a & a' & \lambda a + \mu a' \\ b & b' & \lambda b + \mu b' \\ c & c' & \lambda c + \mu c' \end{vmatrix} = 0. \qquad \square$$

Maintenant, dans la formule de la Définition 8.2, portons notre attention sur les éléments a', b', c' de la seconde colonne, par exemple. Groupons les termes contenant a', puis ceux contenant b', et enfin ceux contenant c', pour obtenir :

$$\begin{vmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{vmatrix} = a'\left(b''c - bc''\right) + b'\left(ac'' - a''c\right) + c'\left(a''b - ab''\right).$$

On peut alors écrire cela :

$$\begin{vmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{vmatrix} = -a' \begin{vmatrix} b & b'' \\ c & c'' \end{vmatrix} + b' \begin{vmatrix} a & a'' \\ c & c'' \end{vmatrix} - c' \begin{vmatrix} a & a'' \\ b & b'' \end{vmatrix}.$$

On dit que l'on a développé le déterminant selon les termes de sa seconde colonne.

Les coefficients de a', de b', de c' sont nommés les *cofacteurs* de ces éléments. On a en effet reconnu trois déterminants 2×2. Le *cofacteur* d'un élément est un déterminant de taille 2×2 affecté d'un certain signe :

(a) le déterminant est obtenu en supprimant la ligne et la colonne de l'élément considéré ;

(b) le signe du cofacteur est '$+$' ou '$-$' selon que la somme des rangs de la ligne et de la colonne de l'élément considéré est paire ou impaire.

On peut obtenir des développements du déterminant, analogues au précédent, selon les termes d'une colonne quelconque, ou d'une ligne quelconque.

9. Exercices

Exercice 1. EE

Exercice 2. EE

Déterminants

François DE MARÇAY
Département de Mathématiques d'Orsay
Université Paris-Sud, France

1. Introduction

2. Applications multilinéaires

Soit E un espace vectoriel sur un corps $\mathbb{K}$. Pour tout entier $p \geqslant 1$, on notera en abrégé le produit cartésien de p copies de E :

$$E^p := \underbrace{E \times \cdots \times E}_{p \text{ facteurs}}.$$

Soit aussi F un deuxième $\mathbb{K}$-espace vectoriel. La définition suivante généralise les notions d'application bilinéaire ($p = 2$) et trilinéaire ($p = 3$) vues au chapitre précédent.

Définition 2.1. Une application de E^p à valeurs dans F :

$$\big(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p\big) \longmapsto f\big(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p\big)$$

sera dite p-*linéaire* si elle est linéaire en chacun des vecteurs $\vec{x}_1$, $\vec{x}_2$, ..., $\vec{x}_p$.

En d'autres termes, pour tout indice $i \in \{1, 2, \dots, p\}$ et pour tout choix de vecteurs fixés :

$$\vec{x}_1, \; \dots, \; \vec{x}_{i-1}, \; \vec{x}_{i+1}, \; \dots, \; \vec{x}_p,$$

l'application partielle :

$$x_i \longmapsto f\big(\vec{x}_1, \dots, \vec{x}_i, \dots, \vec{x}_p\big)$$

est *linéaire* en $\vec{x}_i$.

Proposition 2.2. *Quels que soient les scalaires* $a_1, a_2, \dots, a_p \in \mathbb{K}$, *on a :*

$$f\big(a_1\vec{x}_1, \, a_2\vec{x}_2, \, \dots, \, a_p\vec{x}_p\big) = a_1 a_2 \cdots a_p \, f\big(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p\big).$$

Preuve. En effet, puisque f est linéaire en $\vec{x}_1$, on a :

$$f\big(a_1\vec{x}_1, \, a_2\vec{x}_2, \, \dots, \, a_p\vec{x}_p\big) = a_1 \, f\big(\vec{x}_1, \, a_2\vec{x}_2, \, \dots, \, a_p\vec{x}_p\big),$$

et ensuite, puisque f est linéaire en $\vec{x}_2$:

$$f\big(a_1\vec{x}_1, \, a_2\vec{x}_2, \, a_3\vec{x}_3, \, \dots, \, a_p\vec{x}_p\big) = a_1 \, a_2 \, f\big(\vec{x}_1, \, \vec{x}_2, \, a_3\vec{x}_3, \, \dots, \, a_p\vec{x}_p\big),$$

et ainsi de suite jusqu'au p-ième vecteur. □

Supposons maintenant que les p vecteurs $\vec{x}_1$, $\vec{x}_2$, ..., $\vec{x}_p$ soient des combinaisons linéaires de n autres vecteurs donnés :

$$\vec{u}_1,\ \vec{u}_2,\ \dots,\ \vec{u}_n,$$

c'est-à-dire que pour tout indice $1 \leqslant i \leqslant p$, supposons qu'il existe des scalaires $a_{i,1}, a_{i,2}, \dots, a_{i,n}$ tels que :

$$\vec{x}_i \,=\, a_{i,1}\,\vec{u}_1 + a_{i,2}\,\vec{u}_2 + \cdots + a_{i,n}\,\vec{u}_n,$$

ce que l'on note aussi :

$$\vec{x}_i \,=\, \sum_{j=1}^{n} a_{i,j}\,\vec{u}_j \qquad\qquad (1 \leqslant i \leqslant p).$$

Tout d'abord, puisque f est linéaire selon le premier vecteur, on peut développer :

$$f\bigg(\sum_{j=1}^{n} a_{1,j}\,\vec{u}_j,\ \vec{x}_2,\dots,\vec{x}_p\bigg) \,=\, \sum_{j=1}^{n} a_{1,j}\,f\big(\vec{u}_j,\,\vec{x}_2,\dots,\vec{x}_p\big).$$

Ensuite, puisque f est linéaire selon le second vecteur, on peut à nouveau dévélopper chacun des termes obtenus :

$$f\bigg(\vec{u}_j,\,\sum_{k=1}^{n} a_{2,k}\,\vec{u}_k,\ \vec{x}_3,\dots,\vec{x}_p\bigg) \,=\, \sum_{k=1}^{n} a_{2,k}\,f\big(\vec{u}_j,\,\vec{u}_k,\,\vec{x}_3,\,\dots,\,\vec{x}_p\big).$$

Ainsi, en remplaçant dans la relation précédente, nous voyons apparaître une double somme :

$$f\bigg(\sum_{j=1}^{n} a_{1,j}\,\vec{u}_j,\ \sum_{k=1}^{n} a_{2,k}\,\vec{u}_k,\ \vec{x}_3,\dots,\vec{x}_p\bigg) \,=\, \sum_{j=1}^{n}\sum_{k=1}^{n} a_{1,j}\,a_{2,k}\,f\big(\vec{u}_j,\,\vec{u}_k,\,\vec{x}_3,\,\dots,\,\vec{x}_p\big).$$

Dans cette double sommation, il y a donc $n \times n = n^2$ termes, puisque le système de deux indices $\{j, k\}$ parcourt l'ensemble $\{1, 2, \dots, n\}$.

Nous voulons maintenant raisonner par récurrence sur l'entier p. Nous noterons alors les systèmes d'indices de manière plus précise, en les *indiciant* eux-mêmes :

$$j_1,\ j_2,\ \dots,\ j_{p-1},\ j_p.$$

Au lieu des indices j et k, cela signifie que nous écrivons dorénavant j_1 et j_2.

En raisonnant par récurrence sur p, supposons donc que nous ayons déjà obtenu :

$$\begin{aligned} f\bigg(\sum_{j_1=1}^{n} a_{1,j_1}\,\vec{u}_{j_1},\ \sum_{j_2=1}^{n} a_{2,j_2}\,\vec{u}_{j_2},\ \dots\dots,\ \sum_{j_{p-1}=1}^{n} a_{p-1,j_{p-1}}\,\vec{u}_{j_{p-1}},\ \vec{x}_p\bigg) \,&=\\ =\, \sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\cdots\sum_{j_{p-1}=1}^{n} a_{1,j_1}a_{2,j_2}\cdots a_{p-1,j_{p-1}}\,f\big(\vec{u}_{j_1},\,\vec{u}_{j_2},\,\dots,\,\vec{u}_{j_{p-1}},\,\vec{x}_p\big).& \end{aligned}$$

Alors comme f est linéaire selon $\vec{x}_p$, on peut, après avoir remplacé $\vec{x}_p$ en fonction des $\vec{u}_{j_p}$, développer chacun des termes qui viennent d'apparaître :

$$f\bigg(\vec{u}_{j_1},\,\vec{u}_{j_2},\,\dots,\,\vec{u}_{j_{p-1}},\,\sum_{j_p=1}^{n} a_{p,j_p}\,\vec{u}_{j_p}\bigg) \,=\, \sum_{j_p=1}^{n} a_{p,j_p}\,f\big(\vec{u}_{j_1},\,\vec{u}_{j_2},\,\dots,\,\vec{u}_{j_{p-1}},\,\vec{u}_{j_p}\big).$$

En reportant cette somme dans la ligne qui précède, nous obtenons une collection de p sommes :

$$\sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\cdots\sum_{j_{p-1}=1}^{n}\sum_{j_p=1}^{n},$$

avec au total n^p termes, et le résultat que nous atteignons peut s'énoncer comme suit.

Proposition 2.3. *Si f est une application p-linéaire, alors :*

$$f\bigg(\sum_{j_1=1}^{n} a_{1,j_1}\,\vec{u}_{j_1},\,\sum_{j_2=1}^{n} a_{2,j_2}\,\vec{u}_{j_2},\,\dots,\,\sum_{j_{p-1}=1}^{n} a_{p-1,j_{p-1}}\,\vec{u}_{j_{p-1}},\,\sum_{j_p=1}^{n} a_{p,j_p}\,\vec{u}_{j_p}\bigg) \;=$$

$$= \sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\cdots\sum_{j_{p-1}=1}^{n}\sum_{j_p=1}^{n} a_{1,j_1}\,a_{2,j_2}\cdots a_{p-1,j_{p-1}}\,a_{p,j_p}\,f\big(\vec{u}_{j_1},\,\vec{u}_{j_2},\,\dots,\,\vec{u}_{j_{p-1}},\,\vec{u}_{j_p}\big). \qquad \square$$

3. Applications multilinéaires alternées

Soient E et F deux espaces vectoriels sur le même corps $\mathbb{K}$.

Définition 3.1. Une application p-linéaire :

$$f\colon\quad E^p \longrightarrow F$$

est dite *aternée* si elle s'annule quand deux de ses vecteurs en argument sont égaux :

$$\Big(\vec{x}_i = \vec{x}_j \quad \text{avec} \quad i \neq j\Big) \qquad\qquad \Longrightarrow \qquad\qquad f\big(\vec{x}_1,\dots,\vec{x}_i,\dots,\underbrace{\vec{x}_j}_{=\,\vec{x}_i},\dots\vec{x}_p\big) \;=\; \vec{0}.$$

En d'autres termes, une application p-linéaire f est alternée si elle est *bilinéaire alternée* par rapport à *tout* couple $(\vec{x}_i,\vec{x}_j)$ de deux variables distinctes, les $p-2$ autres variables étant fixées.

Proposition 3.2. *Pour toute application p-linéaire alternée f, quand on échange deux des p vecteurs $\vec{x}_1,\dots,\vec{x}_p$, alors $f\big(\vec{x}_1,\dots,\vec{x}_p\big)$ change de signe.*

Démonstration. Désignons par $\vec{x}_i$ et $\vec{x}_j$ les deux vecteurs que l'on échange, où on peut supposer que $i<j$. On doit prouver :

$$f\big(\vec{x}_1,\dots,\vec{x}_i,\dots,\vec{x}_j,\dots,\vec{x}_p\big) \;=\; -\,f\big(\vec{x}_1,\dots,\vec{x}_j,\dots,\vec{x}_i,\dots,\vec{x}_p\big).$$

À cette fin, il suffit de considérer l'application bilinéaire partielle :

$$\varphi\big(\vec{x}_i,\vec{x}_j\big) \;=\; f\big(\vec{x}_1,\dots,\vec{x}_i,\dots,\vec{x}_j,\dots,\vec{x}_p\big),$$

à deux variables $\vec{x}_i$ et $\vec{x}_j$, les $p-2$ autres variables/vecteurs restants fixes. La propriété résulte alors de ce qui a été vu au chapitre précédent, au moment où nous avons étudié les applications bilinéaires alternées. $\square$

Proposition 3.3. *Pour toute application p-linéaire alternée f, si φ est une permutation de l'ensemble $\big\{1,2,3,\dots,p\big\}$ dont la signature est $\mathsf{sign}(\varphi)$, alors :*

$$f\big(\vec{x}_{\varphi(1)},\,\vec{x}_{\varphi(2)},\,\dots,\,\vec{x}_{\varphi(p)}\big) \;=\; \mathsf{sign}(\varphi)\,f\big(\vec{x}_1,\vec{x}_2,\dots,\vec{x}_p\big).$$

Démonstration. Nous savons (ou nous admettons) que toute permutation $\varphi \in \mathscr{P}_p$ est une composée de transpositions :

$$\varphi \;=\; \tau_q \circ \tau_{q-1} \circ \cdots \circ \tau_2 \circ \tau_1,$$

q étant le nombre de ces transpositions, et la signature $\mathsf{sign}(\varphi)$ ayant pour valeur :

$$\mathsf{sign}(\varphi) \;=\; (-1)^q.$$

Or d'après la Proposition 3.2 précédente, pour toute transposition $\tau \in \mathscr{P}_p$ appliquée aux indices des vecteurs $\vec{x}_i$, la valeur de l'application f change de signe.

Par conséquent, en appliquant successivement les transpositions $\tau_1,\dots,\tau_q$ qui constituent φ, on obtient exactement q changements de signes, et donc on a bien :

$$f\big(\vec{x}_{\varphi(1)},\,\vec{x}_{\varphi(2)},\,\dots,\,\vec{x}_{\varphi(p)}\big) \;=\; (-1)^q\,f\big(\vec{x}_1,\vec{x}_2,\dots,\vec{x}_p\big). \qquad \square$$

Maintenant, énonçons une propriété cruciale de dégénérescence des applications multilinéaires alternées : elles ne *voient pas*, ou *écrasent*, tout ce qui est lié.

Proposition 3.4. *Pour toute application p-linéaire alternée $f\colon E^p \longrightarrow F$, si la famille $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p\}$ est liée dans E, alors :*

$$f\big(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p\big) = \vec{0}.$$

Démonstration. En effet, supposons que la famille $\{\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p\}$ est liée. Alors il existe au moins un indice i avec $1 \leqslant i \leqslant p$ tel que $\vec{x}_i$ soit combinaison linéaire des $p-1$ autres vecteurs. En changeant au besoin l'ordre de la numérotation, on peut supposer $i = 1$, car grâce à la proposition qui précède, changer l'ordre des arguments revient seulement à multiplier l'expression par ± 1.

Il existe alors des scalaires $a_2, \dots, a_p \in \mathbb{K}$ tels que :

$$\vec{x}_1 = \sum_{j=2}^{p} a_j\, \vec{x}_j,$$

où ici, la somme commence à $j = 2$. Or nous avons vu que :

$$f\bigg(\sum_{j=2}^{p} a_j\, \vec{x}_j,\, \vec{x}_2, \dots, \vec{x}_p\bigg) = \sum_{j=2}^{p} a_j\, f\big(\vec{x}_j, \vec{x}_2, \dots, \vec{x}_p\big).$$

Puisque f est alternée, alors, pour tous indices $2 \leqslant j \leqslant p$, on a annulation :

$$f\big(\vec{x}_j, \vec{x}_2, \dots, \vec{x}_p\big) = \vec{0},$$

en vertu du fait que deux vecteurs sont ici toujours égaux dans les arguments de f. La somme ci-dessus est donc nulle, conclusion ! □

Corollaire 3.5. *Si E est de dimension finie sur $\mathbb{K}$, et si :*

$$\mathsf{dim}_{\mathbb{K}}\, E \leqslant p - 1,$$

alors quel que soit l'espace vectoriel F sur $\mathbb{K}$, toute application p-linéaire alternée $f\colon E^p \longrightarrow F$ est identiquement nulle.

Démonstration. En effet, $\mathsf{dim}\, E < p$ entraîne que toute famille $\{\vec{x}_1, \dots, \vec{x}_p\}$ de p vecteurs de E est liée. La propriété précédente donne alors :

$$\forall\,\big(\vec{x}_1, \dots, \vec{x}_p\big) \in E^p \qquad\qquad f\big(\vec{x}_1, \dots, \vec{x}_p\big) = \vec{0},$$

donc $f \equiv 0$ est l'application identiquement nulle. □

Proposition 3.6. *Pour toute application p-linéaire alternée f, si p vecteurs $\vec{x}_1$, $\vec{x}_2$, …, $\vec{x}_p$ sont des combinaisons linéaires de p autres vecteurs donnés $\vec{u}_1$, $\vec{u}_2$, …, $\vec{u}_p$:*

$$\vec{x}_j = \sum_{i=1}^{p} a_{i,j}\, \vec{u}_i \qquad\qquad (1 \leqslant j \leqslant p),$$

alors :

$$f\big(\vec{x}_j, \vec{x}_2, \dots, \vec{x}_p\big) = \bigg[\sum_{\varphi \in \mathscr{P}_p} \mathsf{sign}(\varphi)\, a_{\varphi(1),1}\, a_{\varphi(2),2} \cdots a_{\varphi(p),p}\bigg]\, f\big(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p\big),$$

la sommation étant étendue à l'ensemble $\mathscr{P}_p$ des $p!$ permutations φ de $\{1, 2, 3, \dots, p\}$.

Démonstration. Supposons en effet que nos p vecteurs soient combinaisons linéaires de p vecteurs $\vec{u}_j$:

$$\vec{x}_j = \sum_{i=1}^{p} a_{i,j}\, \vec{u}_i \qquad\qquad (1 \leqslant j \leqslant p).$$

Appliquons la Proposition 2.3 :

$$f\big(\vec{x}_1, \dots, \vec{x}_p\big) \;=\; \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_p=1}^{n} a_{i_1,1}\, a_{i_2,2} \,\cdots\, a_{i_p,p}\, f\big(\vec{u}_{i_1}, \vec{u}_{i_2}, \dots, \vec{u}_{i_p}\big), \tag{3.7}$$

la sommation étant étendue aux p^p sytèmes $\big\{i_1, i_2, \dots, i_p\big\}$ d'indices i_k parcourant tous l'ensemble $\big\{1, 2, 3, \dots, p\big\}$.

Mais puisque f est alternée, dès que deux indices i_k et i_ℓ sont égaux, on a annulation :

$$f\big(\vec{u}_{i_1}, \vec{u}_{i_2}, \dots, \vec{u}_{i_p}\big) \;=\; \vec{0}.$$

Par conséquent, dans la sommation (3.7), il ne reste plus qu'à considérer les systèmes d'indices $\big\{i_1, i_2, \dots, i_p\big\}$ *deux à deux distincts*, c'est-à-dire les images de $\big\{1, 2, 3, \dots, p\big\}$ par les permutations $\varphi \in \mathscr{P}_p$.

On sait qu'il y a $p!$ systèmes de cette sorte, et qu'il y a aussi $p!$ permutations $\varphi \in \mathscr{P}_p$.

Les $p^p - p!$ autres systèmes d'indices $\big\{i_1, i_2, \dots, i_p\big\}$ *non* tous deux à deux distincts ont certainement une image nulle par f, puisque f est alternée, donc s'annule quand deux vecteurs sont égaux.

Les $p!$ systèmes d'indices intéressants s'écrivent donc $\big\{\varphi(1), \varphi(2), \dots, \varphi(p)\big\}$ pour une certaine permutation $\varphi \in \mathscr{P}_p$, et tous les systèmes d'indices deux à deux distincts s'obtiennent en faisant parcourir à φ toutes les $p!$ permutations de $\big\{1, 2, \dots, p\big\}$ appartenant à $\mathscr{P}_p$.

Enfin, grâce à la Proposition 3.3 :

$$f\big(u_{\varphi(1)},\, \vec{u}_{\varphi(2)},\, \dots,\, \vec{u}_{\varphi(p)}\big) \;=\; \mathsf{sign}(\varphi)\, f\big(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p\big),$$

où $\mathsf{sign}(\varphi)$ est la signature de la permutation φ. Le vecteur $f\big(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_p\big)$ se met donc en facteur dans la sommation. Ceci achève la démonstration de la proposition. □

Maintenant, étudions le cas spécial où la dimension de E est égale à p, le premier cas intéressant, car nous avons vu que pour $\mathsf{dim}\, E \leqslant p - 1$, toutes les applications multilinéaires alternées sont identiquement nulles. Nous supposons toujours le $\mathbb{K}$-espace vectoriel F quelconque.

Rappelons que dans le chapitre précédent, nous avons déjà expliqué les cas $p = 2$ et $p = 3$ du théorème suivant, qui ne doit donc paraître ni mystérieux, ni difficile d'accès.

Théorème 3.8. [Fondamental !] *Soient E et F deux espaces vectoriels sur le même corps $\mathbb{K}$, avec :*

$$\mathsf{dim}_{\mathbb{K}}\, E \;=\; p \;\geqslant\; 1.$$

Alors pour toute base $\big\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p\big\}$ de E, et tout vecteur fixé $\vec{k} \in F$, il existe une unique application p-linéaire alternée $f\colon E^p \longrightarrow F$ telle que :

$$f\big(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p\big) \;=\; \vec{k}.$$

Démonstration. Commençons par établir l'*unicité* de f. Dans la base $\big\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p\big\}$ de E, pour tout indices $1 \leqslant i \leqslant p$, désignons par $a_{i,1}$, $a_{i,2}$, …, $a_{i,p}$ les coordonnées des vecteurs $\vec{x}_i$, c'est-à-dire écrivons :

$$\vec{x}_j \;=\; \sum_{i=1}^{p} a_{i,j}\, \vec{e}_i \qquad\qquad (1 \leqslant j \leqslant p).$$

Soit $f\colon E^p \longrightarrow F$ une application p-linéaire alternée. Alors grâce à la Proposition 3.6 qui précède, nous pouvons développer :

$$f\big(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p\big) \;=\; f\big(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_p\big) \sum_{\varphi \in \mathscr{P}_p} \mathsf{sign}(\varphi)\, a_{\varphi(1),1}\, a_{\varphi(2),2} \,\cdots\, a_{\varphi(p),p},$$

la sommation étant étendue à toutes les permutations φ de $\big\{1, 2, \dots, p\big\}$, et $\mathsf{sign}(\varphi)$ désignant la signature de la permutation φ.

Par conséquent, si l'application p-linéaire alternée f de l'énoncé existe, c'est nécessairement l'application de E^p dans F définie, relativement à la base donnée, par :

$$f\big(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p\big) \ = \ \vec{k} \sum_{\varphi \in \mathscr{P}_p} \mathsf{sign}(\varphi)\, a_{\varphi(1),1}\, a_{\varphi(2),2} \cdots a_{\varphi(p),p}. \tag{3.9}$$

L'unicité est donc établie.

Traitons à présent l'*existence*. Soit alors $f\colon E^p \longrightarrow F$ l'application p-linéaire alternée définie par (3.9).

Pour tout tout indice $1 \leqslant j \leqslant p$, nous devons démontrer que f est linéaire en $\vec{x}_j$, les autres vecteurs restants étant fixes. Nous procéderons en trois étapes.

Étape 1. Si :

$$\vec{x}_j \ = \ \vec{x}_j^{\,\prime} + \vec{x}_j^{\,\prime\prime},$$

alors pour tout indice $1 \leqslant i \leqslant p$, les coordonnées correspondantes vérifient la relation :

$$a_{i,j} \ = \ a_{i,j}' + a_{i,j}''.$$

Considérons alors la valeur :

$$f\big(\vec{x}_1, \dots, \vec{x}_j^{\,\prime} + \vec{x}_j^{\,\prime\prime}, \dots, \vec{x}_p\big), \tag{3.10}$$

où l'on a remplacé $\vec{x}_j$ par $\vec{x}_j^{\,\prime} + \vec{x}_j^{\,\prime\prime}$, sans toucher aux autres vecteurs.

Le terme général de la sommation (3.9) correspondante est :

$$\mathsf{sign}(\varphi)\, a_{\varphi(1),1} \cdots \big(a_{\varphi(j),j}' + a_{\varphi(j),j}''\big) \cdots a_{\varphi(p),p},$$

ce qui s'écrit, par distributivité dans $\mathbb{K}$:

$$\mathsf{sign}(\varphi)\, a_{\varphi(1),1} \cdots a_{\varphi(j),j}' \cdots a_{\varphi(p),p} + \mathsf{sign}(\varphi)\, a_{\varphi(1),1} \cdots a_{\varphi(j),j}'' \cdots a_{\varphi(p),p},$$

et donc, par associativité et commutativité de la sommation $\sum$, le résultat de cette sommation est :

$$f\big(\vec{x}_1, \dots, \vec{x}_j^{\,\prime}, \dots, \vec{x}_p\big) + f\big(\vec{x}_1, \dots, \vec{x}_j^{\,\prime\prime}, \dots, \vec{x}_p\big). \tag{3.11}$$

Par conséquent, les valeurs (3.10) et (3.11) sont égales.

Étape 2. Si on prend maintenant :

$$\vec{x}_j \ = \ b\, \vec{x}_j^{\,\prime}$$

avec un scalaire fixé $b \in \mathbb{K}$, alors les coordonnées vérifient, pour tout $1 \leqslant i \leqslant p$:

$$a_{i,j} \ = \ b\, a_{i,j}'.$$

Considérons la valeur :

$$f\big(\vec{x}_1, \dots, b\, \vec{x}_j^{\,\prime}, \dots, \vec{x}_p\big), \tag{3.12}$$

où l'on a remplacé $\vec{x}_j$ par $b\, \vec{x}_j^{\,\prime}$, sans toucher aux autres vecteurs.

Le terme général de la sommation (3.9) correspondante devient, par commutativité de la multiplication dans $\mathbb{K}$:

$$\mathsf{sign}(\varphi)\, a_{\varphi(1),1} \cdots b\, a_{\varphi(j),j} \cdots a_{\varphi(p),p} \ = \ b\, \mathsf{sign}(\varphi)\, a_{\varphi(1),1} \cdots a_{\varphi(j),j} \cdots a_{\varphi(p),p},$$

et par distributivité pour la sommation $\sum$, le résultat de cette sommation est :

$$b\, f\big(\vec{x}_1, \dots, \vec{x}_j^{\,\prime}, \dots, \vec{x}_p\big). \tag{3.13}$$

Les valeurs (3.12) et (3.13) sont donc égales.

Le résultat conjoint des Étapes 1 et 2, c'est que f est p-linéaire.

Étape 3. Prouvons enfin que f est alternée.

Pour que l'application f soit alternée, nous savons qu'il suffit qu'elle change de signe quand on échange deux vecteurs quelconques.

Soit donc $j_1 < j_2$, et échangeons $\vec{x}_{j_1}$ et $\vec{x}_{j_2}$.

Pour les coordonnées de ces deux vecteurs, posons, pour tout indice $1 \leqslant i \leqslant p$:

$$a'_{i,j_1} := a_{i,j_2} \qquad \text{et} \qquad a'_{i,j_2} := a_{i,j_1}.$$

Considérons la valeur :

$$f\big(\vec{x}_1, \dots, \vec{x}_{j_2}, \dots, \vec{x}_{j_1}, \dots, \vec{x}_p\big). \tag{3.14}$$

Le terme général de la sommation (3.9) correspondante devient :

$$\mathsf{sign}(\varphi)\, a_{\varphi(1),1} \cdots a'_{\varphi(j_1),j_1} \cdots a'_{\varphi(j_2),j_2} \cdots a_{\varphi(p),p}.$$

Maintenant, soit $\tau \in \mathscr{P}_p$ la transposition qui échange j_1 et j_2. Posons :

$$\psi := \varphi \circ \tau.$$

On a alors :

$$\psi(j_1) = \varphi(j_2) \qquad \text{et} \qquad \psi(j_2) = \varphi(j_1),$$

tandis que, pour tous autres indices j' différents de j_1 et de j_2, on a :

$$\psi(j') = \varphi(j').$$

D'autre part, grâce à la propriété (connue ou admise) de multiplicativité des signatures :

$$\mathsf{sign}(\psi) = \mathsf{sign}(\varphi)\,\mathsf{sign}(\tau) = -\,\mathsf{sign}(\varphi),$$

le terme général de la valeur (3.14) s'écrit donc :

$$-\,\mathsf{sign}(\psi)\, a_{\psi(1),1} \cdots a'_{\psi(j_2),j_1} \cdots a'_{\psi(j_1),j_2} \cdots a_{\psi(p),p}.$$

Or comme :

$$a'_{\psi(j_2),j_1} = a_{\psi(j_2),j_2} \qquad \text{et} \qquad a'_{\psi(j_1),j_2} = a_{\psi(j_1),j_1},$$

ce terme général s'écrit alors, par commutativité de la multiplication dans $\mathbb{K}$:

$$\begin{aligned} &-\,\mathsf{sign}(\psi)\, a_{\psi(1),1} \cdots a_{\psi(j_2),j_2} \cdots a_{\psi(j_1),j_1} \cdots a_{\psi(p),p} = \\ &= -\,\mathsf{sign}(\psi)\, a_{\psi(1),1} \cdots a_{\psi(j_1),j_1} \cdots a_{\psi(j_2),j_2} \cdots a_{\psi(p),p}. \end{aligned} \tag{3.15}$$

Maintenant, remarquons que :

$$\psi = \varphi \circ \tau \qquad \Longleftrightarrow \qquad \varphi = \psi \circ \tau,$$

de sorte que la correspondance $\varphi \longmapsto \psi$ est une permutation bijective de $\mathscr{P}_p$. Par conséquent, sommer (3.15) quand φ parcourt $\mathscr{P}_p$ revient à sommer (3.15) quand ψ parcourt $\mathscr{P}_p$.

Or le résultat de cette sommation est :

$$-\,f\big(\vec{x}_1, \dots, \vec{x}_{j_1}, \dots, \vec{x}_{j_2}, \dots, \vec{x}_p\big).$$

Par conséquent, f est bien alternée.

Les trois étapes sont achevées, et le théorème est complètement démontré. □

4. Déterminants

L'objectif est d'introduire les *formes* p-linéaires alternées, qui nous conduiront aux *déterminants*. À partir de maintenant, nous changerons de notation :

$$p \rightsquigarrow n.$$

Plaçons-nous dans les hypothèses du Théorème 3.8 fondamental, avec de plus :

$$F := \mathbb{K}.$$

Soit donc E un $\mathbb{K}$-espace vectoriel, soit $n \geqslant 1$ un entier, et soit $\big\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\big\}$ une base de E.

Si on choisit pour $\vec{k} \in G$ le « vecteur » $1 \in \mathbb{K}$, alors ce théorème garantit qu'il existe une unique application n-multilinéaire alternée f satisfaisant :

$$f\big(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\big) \,=\, 1.$$

Définition 4.1. On appelle *déterminant* de n vecteurs quelconques :

$$\vec{x}_1,\, \vec{x}_2,\, \dots,\, \vec{x}_n \;\in\; E$$

dans la base $\big\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\big\}$ la valeur sur eux de cette unique forme n-linéaire alternée :

$$\boldsymbol{D}\,\big(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\big) \,:=\, f\big(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\big).$$

Ainsi :

$$\boldsymbol{D}\,\big(\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\big) \,=\, 1.$$

Si :

$$\vec{x}_j \,=\, \sum_{i=1}^{n} a_{i,j}\, \vec{e}_i \qquad\qquad (1 \leqslant j \leqslant n),$$

alors le déterminant de $\vec{x}_1, \dots, \vec{x}_n$ est :

$$\boldsymbol{D}\big(\vec{x}_1, \dots, \vec{x}_n\big) \,=\, \sum_{\varphi \in \mathscr{P}_n} \mathsf{sign}(\varphi)\, a_{\varphi(1),1} \cdots a_{\varphi(n),n}.$$

Comme conséquence du Théorème 3.8, nous obtenons le

Corollaire 4.2. *Toute autre forme n-linéaire alternée de E^n à valeurs dans $\mathbb{K}$ est du type :*

$$\lambda\, \boldsymbol{D}\big(\bullet, \dots, \bullet\big),$$

avec une constante $\lambda \in \mathbb{K}$. □

Maintenant, nous pouvons enfin introduire la notion de *déterminant* d'une matrice carrée de taille $n \times n$ quelconque, généralisant en cela ce que nous avons vu à la fin du chapitre précédent pour $n = 2$ et pour $n = 3$.

Soit donc M une matrice carrée d'ordre n, à coefficients dans $\mathbb{K}$:

$$M \,=\, \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}.$$

Définition 4.3. On appelle *déterminant de la matrice* M, et on note :

$$\mathsf{det}\, M \,:=\, \begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{vmatrix},$$

le *scalaire* défini par :

$$\mathsf{det}\, M \,=\, \sum_{\varphi \in \mathscr{P}_n} \mathsf{sign}(\varphi)\, a_{\varphi(1),1} \cdots a_{\varphi(n),n}.$$

Le déterminant ne change pas quand on transpose la matrice.

Proposition 4.4. *On a :*

$$\mathsf{det}\,\big({}^{\mathsf{t}}M\big) \,=\, \mathsf{det}\, M.$$

Démonstration. Soient :

$$M = (a_{i,j}) \qquad \text{et} \qquad {}^{\mathtt{t}}M = (a^{\mathtt{t}}_{i,j}),$$

avec :

$$a^{\mathtt{t}}_{i,j} = a_{j,i}.$$

Par définition :

$$\det\left({}^{\mathtt{t}}M\right) = \sum_{\varphi\in\mathscr{P}_n} \mathsf{sign}(\varphi)\, a^{\mathtt{t}}_{\varphi(1),1} \cdots a^{\mathtt{t}}_{\varphi(n),n}.$$

Par commutativité dans $\mathbb{K}$, on peut écrire le terme général de façon à ce que $\{\varphi(1), \dots, \varphi(n)\}$ revienne à l'ordre naturel $\{1, 2, \dots, n\}$. On passe du premier au second ensemble par la permutation $\psi := \varphi^{-1}$.

Si l'on pose, pour tout $1 \leqslant j \leqslant n$:

$$\varphi(j) =: k,$$

d'où :

$$\psi(k) = j,$$

il vient par suite :

$$a^{\mathtt{t}}_{\varphi(j),j} = a^{\mathtt{t}}_{k,\psi(k)} = a_{\psi(k),k},$$

de sorte que le terme général s'écrit, en remarquant que $\mathsf{sign}(\varphi) = \mathsf{sign}(\varphi^{-1})$:

$$\mathsf{sign}(\psi)\, a_{\psi(1),1}\, a_{\psi(2),2} \cdots a_{\psi(n),n}.$$

D'autre part, sommer lorsque φ parcourt $\mathscr{P}_n$ revient à sommer quand $\varphi^{-1} = \psi$ parcourt $\mathscr{P}_n$. Par conséquent, nous atteignons la conclusion :

$$\det\left({}^{\mathtt{t}}M\right) = \sum_{\psi\in\mathscr{P}_n} \mathsf{sign}(\psi)\, a_{\psi(1),1} \cdots a_{\psi(n),n} = \det M. \qquad \square$$

Il en résulte que toute propriété concernant les lignes d'un déterminant est valable aussi pour les colonnes.

Terminologie 4.5. Lignes et colonnes seront nommées *rangées*.

On peut alors énoncer les propriétés qui découlent du fait que le déterminant est une fonction n-linéaire alternée.

Proposition 4.6. [Fondamentale !] (1) *Si on échange deux rangées parallèles d'une matrice, son déterminant change de signe.*

(2) *Si deux rangées parallèles d'une matrice carrée sont identiques, son déterminant est nul.*

(3) *Si on multiplie par $\lambda \in \mathbb{K}$ tous les éléments d'une rangée d'une matrice carrée, son déterminant est multiplié par λ.*

(4) *Si une rangée d'une matrice carrée est une combinaison linéaire de rangées parallèles, son déterminant est nul.* □

5. Déterminant du produit de deux matrices de $\mathscr{M}_n(\mathbb{K})$

Soient E un espace vectoriel de dimension n sur un corps $\mathbb{K}$, rapporté à une base $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$, et soit $\boldsymbol{D}(\bullet, \dots, \bullet)$ le déterminant de E^n sur $\mathbb{K}$ associé à cette base.

Théorème 5.1. *Pour tout endomorphisme $u \in \mathscr{L}(E)$, la composée $\boldsymbol{D} \circ u$ est n-linéaire alternée, et on a :*

$$\boldsymbol{D}\big[u(\vec{x}_1), \dots, u(\vec{x}_n)\big] = \det A \cdot \boldsymbol{D}(\vec{x}_1, \dots, \vec{x}_n),$$

où $A := \mathsf{Mat}(u)$ est la matrice associée à l'endomorphisme u.

Remarquons d'abord que, pour ne pas multiplier les notations, nous avons désigné par la même lettre u l'endomorphisme donné de E et l'endomorphisme suivant de E^n à valeurs dans E^n :

$$\big(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n\big) \longmapsto \big(u(\vec{x}_1), u(\vec{x}_2), \dots, u(\vec{x}_n)\big).$$

Démonstration. Commençons par établir que $\boldsymbol{D} \circ u$ est n-linéaire alternée.

À cet effet, prouvons que, pour tout indice $1 \leqslant j \leqslant n$, l'application $\boldsymbol{D} \circ u$ est linéaire en $\vec{x}_j$, quand les autres vecteurs $u(\vec{x}_k)$ sont fixés. Mais cela est évident : comme $u(\vec{x}_j)$ est linéaire en $\vec{x}_j$, et comme $\boldsymbol{D}$ est aussi linéaire, la composée $\boldsymbol{D} \circ u$ est bien linéaire en $\vec{x}_j$.

Prouvons ensuite que cette composée $\boldsymbol{D} \circ u$ est alternée.

En effet, si, pour deux indices quelconques distincts j_1 et j_2, on suppose que $\vec{x}_{j_1} = \vec{x}_{j_2}$, alors $u(\vec{x}_{j_1}) = u(\vec{x}_{j_2})$, et comme $\boldsymbol{D}$ est alternée, il vient :

$$\boldsymbol{D}\Big(u\big(\vec{x}_1\big), \dots, u\big(\vec{x}_{j_1}\big), \dots, u\big(\vec{x}_{j_2}\big), \dots, u\big(\vec{x}_n\big)\Big) = \vec{0}.$$

Ainsi, la composée $\boldsymbol{D} \circ u$ est bien n-linéaire alternée.

D'après le Théorème 3.8, puisque $\boldsymbol{D}$ et $\boldsymbol{D} \circ u$ sont deux formes n-linéaires alternées de E^n dans $\mathbb{K}$, il existe une constante $\lambda \in \mathbb{K}$ telle que :

$$\boldsymbol{D} \circ u = \lambda \boldsymbol{D},$$

c'est-à-dire que, pour tout $(\vec{x}_1, \dots, \vec{x}_n) \in E^n$, on a :

$$\boldsymbol{D}\big(u(\vec{x}_1), \dots, u(\vec{x}_n)\big) = \lambda \boldsymbol{D}\big(\vec{x}_1, \dots, \vec{x}_n\big).$$

Prenons alors pour ces vecteurs les vecteurs de la base de l'espace vectoriel E :

$$\vec{x}_1 := \vec{e}_1, \; \dots\dots, \; \vec{x}_n := \vec{e}_n,$$

ce qui donne :

$$\begin{aligned} \boldsymbol{D}\big(u(\vec{x}_1), \dots, u(\vec{x}_n)\big) &= \lambda \boldsymbol{D}\big(\vec{e}_1, \dots, \vec{e}_n\big) \\ &= \lambda. \end{aligned}$$

Or comme $u(\vec{e}_1), \dots, u(\vec{e}_n)$ sont les vecteurs-colonnes de la matrice A de l'endomorphisme u dans la base considérée, nous avons :

$$\lambda = \det A,$$

ce qui achève la démonstration. □

Théorème 5.2. *Quelles que soient les matrices carrées $A \in \mathscr{M}_n(\mathbb{K})$ et $B \in \mathscr{M}_n(\mathbb{K})$ d'ordre $n \geqslant 1$ quelconque, on a :*

$$\det\big(A\,B\big) = \det(A)\,\det(B).$$

Démonstration. Soit donc E un espace vectoriel de dimension n sur le corps $\mathbb{K}$, et soit $\boldsymbol{B}_E = \big\{\vec{e}_1, \dots, \vec{e}_n\big\}$ une base de E.

Soient u et v les endomorphismes de E de matrices A et B dans cette base :

$$E \xrightarrow[B]{\;v\;} E \xrightarrow[A]{\;u\;} E.$$

On sait que la matrice produit $A\,B$ représente, dans cette base, la composée $u \circ v$.

Maintenant, appliquons le Théorème 5.1 à l'endomorphisme $u \circ v \in \mathscr{L}(E)$, ce qui nous donne :

$$\boldsymbol{D}\Big(u \circ v(\vec{x}_1), \dots, u \circ v(\vec{x}_n),\Big) = \det\big(A\,B\big)\,\boldsymbol{D}\big(\vec{x}_1, \dots, \vec{x}_n\big). \tag{5.3}$$

Posons :

$$v(\vec{x}_1) =: \vec{y}_1, \; \dots\dots, \; v(\vec{x}_n) =: \vec{y}_n.$$

Appliquons à nouveau le Théorème 5.1 à l'endomorphisme u, ce qui nous donne :

$$\boldsymbol{D}\big(u(\vec{y}_1), \dots, u(\vec{y}_n)\big) = \det A\;\boldsymbol{D}\big(\vec{y}_1, \dots, \vec{y}_n\big). \tag{5.4}$$

Remarquons que les premiers membres de ces deux équations (5.3) et (5.4) sont identiques.

Appliquons enfin le Théorème 5.1 à l'endomorphisme v :

$$\begin{aligned} \boldsymbol{D}\big(\vec{y}_1, \ldots, \vec{y}_n\big) &= \boldsymbol{D}\big(v(\vec{x}_1), \ldots, v(\vec{x}_n)\big) \\ &= \operatorname{det} B \; \boldsymbol{D}\big(\vec{x}_1, \ldots, \vec{x}_n\big). \end{aligned}$$

En remplaçant dans (5.4) et en comparant avec (5.3), on obtient :

$$\operatorname{det}\big(A\,B\big)\; \boldsymbol{D}\big(\vec{x}_1, \ldots, \vec{x}_n\big) \;=\; \operatorname{det} A \; \operatorname{det} B \; \boldsymbol{D}\big(\vec{x}_1, \ldots, \vec{x}_n\big).$$

Pour conclure, il suffit de prendre $\vec{x}_1 := \vec{e}_1$, ..., $\vec{x}_n := \vec{e}_n$ pour que :

$$\boldsymbol{D}\big(\vec{x}_1, \ldots, \vec{x}_n\big) \;=\; \boldsymbol{D}\big(\vec{e}_1, \ldots, \vec{e}_n\big) \;=\; 1,$$

et le théorème est démontré. □

6. Développement d'un déterminant selon les éléments d'une rangée

Nous allons maintenant exposer la méthode de développement d'un déterminant selon les éléments d'une colonne. D'après ce qu'on a vu, cette méthode sera aussi applicable au développement selon les éléments d'une ligne.

Considérons donc une matrice carrée $M \in \mathscr{M}_n(\mathbb{K})$ d'ordre $n \geqslant 1$ quelconque :

$$M \;=\; \Big(a_{i,j}\Big)_{1 \leqslant i \leqslant n \text{ lignes}}^{1 \leqslant j \leqslant n \text{ colonnes}}.$$

Alors son déterminant vaut :

$$\operatorname{det} M \;=\; \sum_{\varphi \in \mathscr{P}_n} \operatorname{sign}(\varphi)\; a_{\varphi(1),1}\, a_{\varphi(2),2} \cdots a_{\varphi(n),n}.$$

Soit E un espace vectoriel sur $\mathbb{K}$, rapporté à une base $\big\{\vec{e}_1, \ldots, \vec{e}_n\big\}$. Si, pour tous indices $1 \leqslant j \leqslant n$:

$$\vec{x}_j \;=\; \sum_{i=1}^{n} a_{i,j}\, \vec{e}_i, \tag{6.1}$$

alors, relativement à la base choisie, on a :

$$\boldsymbol{D}\big(\vec{x}_1, \ldots, \vec{x}_j, \ldots, \vec{x}_n\big) \;=\; \operatorname{det} M.$$

Pour tout indice $1 \leqslant j \leqslant n$, l'application $\boldsymbol{D}$ est linéaire en $\vec{x}_j$, donc si on remplace $\vec{x}_j$ par la combinaison linéaire (6.1), on obtient :

$$\begin{aligned} &\boldsymbol{D}\bigg(\vec{x}_1, \ldots, \vec{x}_{j-1},\; \sum_{i=1}^{n} a_{i,j}\, \vec{e}_i,\; \vec{x}_{j+1}, \ldots, \vec{x}_n\bigg) \;= \\ &= \sum_{i=1}^{n} a_{i,j}\, \boldsymbol{D}\Big(\vec{x}_1, \ldots, \vec{x}_{j-1},\, \vec{e}_i,\, \vec{x}_{j+1}, \ldots, \vec{x}_n\Big). \end{aligned} \tag{6.2}$$

Ceci motive d'introduire :

$$\begin{aligned} \Delta_{i,j} &:= \boldsymbol{D}\Big(\vec{x}_1, \ldots, \vec{x}_{j-1},\, \vec{e}_i,\, \vec{x}_{j+1}, \ldots, \vec{x}_n\Big) \\ &= \operatorname{det} \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i,1} & \cdots & a_{i,j-1} & 1 & a_{i,j+1} & \cdots & a_{i,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}. \end{aligned}$$

Ainsi, $\Delta_{i,j}$ et le déterminant déduit de M en remplaçant le vecteur colonne $\vec{x}_j$ de rang j par le vecteur de base $\vec{e}_i$. En d'autres termes, on remplace la colonne de rang j par des zéros, sauf l'élément de la ligne i — l'élément $a_{i,j}$ — que l'on remplace par 1.

La relation (6.2) s'écrit alors :

$$\begin{aligned}\det M &= \sum_{i=1}^{n} a_{i,j}\,\Delta_{i,j}\\ &= a_{1,j}\,\Delta_{1,j} + a_{2,j}\,\Delta_{2,j} + \cdots + a_{n,j}\,\Delta_{n,j}.\end{aligned}$$

Terminologie 6.3. On dit que l'on a *développé* le déterminant de M selon les éléments de la colonne de rang j.

Pour tous indices $1 \leqslant i, j \leqslant n$, ce déterminant $\Delta_{i,j}$ se nomme *cofacteur* de $a_{i,j}$.

Maintenant, montrons comment calculer ces cofacteurs $\Delta_{i,j}$.

Étape 1 : *Calcul de :*

$$\Delta_{1,1} = \boldsymbol{D}\big(\vec{e}_1, \vec{x}_2, \dots, \vec{x}_n\big) = \det \begin{pmatrix} 1 & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}.$$

Ceci est le déterminant de la matrice déduite de M en remplaçant les éléments de la première colonne par des zéros, sauf l'élément de la première ligne — l'élément $a_{1,1}$ — que l'on remplace par 1.

On obtient donc la valeur de $\Delta_{1,1}$ en remplaçant, dans la sommation donnant le déterminant $\det M$, les $a_{\varphi(1),1}$ par :

$$\begin{aligned} 0 \quad &\text{si } \varphi(1) \neq 1, \\ 1 \quad &\text{si } \varphi(1) = 1, \end{aligned}$$

Par suite, on obtient :

$$\Delta_{1,1} = \sum_{\substack{\varphi \in \mathscr{P}_n \\ \varphi(1)=1}} \operatorname{sign}(\varphi)\, a_{\varphi(2),2}\, a_{\varphi(3),3} \cdots a_{\varphi(n),n},$$

la sommation étant étendue aux permutations φ de l'ensemble $\{1, 2, 3, \dots, n\}$ qui laissent l'élément 1 invariant. La restriction de φ à l'ensemble $\{2, 3, \dots, n\}$ est alors une permutation $\varphi' \in \mathscr{P}_{n-1}$.

Réciproquement, toute permutation $\varphi' \in \mathscr{P}_{n-1}$ de l'ensemble précédent se prolonge d'une manière unique en une permutation φ de $\{1, 2, \dots, n\}$ laissant 1 invariant. De plus, on a égalité des signatures :

$$\operatorname{sign}(\varphi) = \operatorname{sign}(\varphi').$$

En définitive :

$$\Delta_{1,1} = \sum_{\varphi' \in \mathscr{P}_{n-1}} \operatorname{sign}(\varphi')\, a_{\varphi'(2),2} \cdots a_{\varphi'(n),n}.$$

Ensuite, désignons par :

$$M_{1,1} := \begin{pmatrix} a_{2,2} & \cdots & a_{2,n} \\ \vdots & \ddots & \vdots \\ a_{n,2} & \cdots & a_{n,n} \end{pmatrix}$$

la matrice de $\mathscr{M}_{n-1}(\mathbb{K})$ déduite de la matrice considérée M par suppression de la première ligne et suppression de la première colonne. Par définition du déterminant d'une matrice, la formule qui précède exprime alors visiblement que :

$$\Delta_{1,1} = \det M_{1,1}.$$

Étape 2 : *Calcul de :*

$$
\begin{aligned}
\Delta_{i,j} &:= \boldsymbol{D}\Big(\vec{x}_1, \dots, \vec{x}_{j-1},\, \vec{e}_i,\, \vec{x}_{j+1}, \dots, \vec{x}_n\Big) \\
&= \det \begin{pmatrix}
a_{1,1} & \cdots & a_{1,j-1} & 0 & a_{1,j+1} & \cdots & a_{1,n} \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
a_{i-1,1} & \cdots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \cdots & a_{i-1,n} \\
a_{i,1} & \cdots & a_{i,j-1} & 1 & a_{i,j+1} & \cdots & a_{i,n} \\
a_{i+1,1} & \cdots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \cdots & a_{i+1,n} \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
a_{n,1} & \cdots & a_{n,j-1} & 0 & a_{n,j+1} & \cdots & a_{n,n}
\end{pmatrix}.
\end{aligned}
$$

Amenons le vecteur $\vec{e}_i$ de la j-ème colonne à la première colonne, en l'échangeant successivement avec tous les vecteurs qui précèdent ; il y a au total j échanges, et comme la forme n-linéaire $\boldsymbol{D}$ est alternée, on obtient :

$$\Delta_{i,j} \,=\, (-1)^j \,\mathsf{det} \begin{pmatrix} 0 & a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 1 & a_{i,1} & \cdots & a_{i,j-1} & a_{i,j+1} & \cdots & a_{i,n} \\ 0 & a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}.$$

Dans le déterminant obtenu, amenons la ligne de rang i à la première place par échanges successifs avec les lignes qui la précèdent ; il y a au total i échanges de lignes, et comme $\boldsymbol{D}$ est alternée, on obtient :

$$\Delta_{i,j} \,=\, (-1)^{i+j}\,(-1)^j \,\mathsf{det} \begin{pmatrix} 1 & a_{i,1} & \cdots & a_{i,j-1} & a_{i,j+1} & \cdots & a_{i,n} \\ 0 & a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}.$$

Au déterminant qui est obtenu ici, on peut appliquer le résultat trouvé à l'Étape 1. Si on supprime la première ligne et la première colonne de ce déterminant, on obtient donc :

$$\begin{vmatrix} 1 & a_{i,1} & \cdots & a_{i,j-1} & a_{i,j+1} & \cdots & a_{i,n} \\ 0 & a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ 0 & a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix} = \begin{vmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{vmatrix}.$$

Par conséquent :

$$\Delta_{i,j} \,=\, (-1)^{i+j} \,\mathsf{det} \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}.$$

Nous pouvons donc énoncer le théorème récapitulatif suivant.

Théorème 6.4. *Soit une matrice carrée d'ordre $n \geqslant 1$ quelconque :*

$$M \,=\, \Big(a_{i,j}\Big)_{1\leqslant i\leqslant n \text{ lignes}}^{1\leqslant j\leqslant n \text{ colonnes}} \qquad\qquad (a_{i,j} \in \mathbb{K}),$$

c'est-à-dire :

$$M = \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}.$$

Si on désigne par :

$$M_{i,j} := \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}$$

la matrice de $\mathscr{M}_{n-1}(\mathbb{K})$ déduite de M par suppression de la ligne de rang i et suppression de la colonne de rang j, alors on a :

$$\det M = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det\big(M_{i,j}\big). \qquad \square$$

Terminologie 6.5. On parle du *développement du déterminant* de M selon les éléments d'une colonne, celle de rang j.

Théorème 6.6. *Sous les mêmes hypothèses, on a aussi :*

$$\det M = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \det\big(M_{i,j}\big). \qquad \square$$

Terminologie 6.7. On parle du *développement du déterminant* de M selon les éléments d'une ligne, celle de rang i.

Tout cela étant très général et très abstrait, il convient maintenant d'illustrer ces deux procédés de développement sur des exemples simples et concrets.

Exemple 6.8. Développons le déterminant d'une matrice de $\mathscr{M}_4(\mathbb{K})$ à 4 lignes et 4 colonnes selon les éléments de sa *dernière* ligne :

$$\begin{vmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \\ a''' & b''' & c''' & d''' \end{vmatrix} = -a''' \begin{vmatrix} b & c & d \\ b' & c' & d' \\ b'' & c'' & d'' \end{vmatrix} + b''' \begin{vmatrix} a & c & d \\ a' & c' & d' \\ a'' & c'' & d'' \end{vmatrix} - \\ - c''' \begin{vmatrix} a & b & d \\ a' & b' & d' \\ a'' & b'' & d'' \end{vmatrix} + d''' \begin{vmatrix} a & b & c \\ a' & b' & c' \\ a'' & b'' & c'' \end{vmatrix}.$$

Exemple 6.9. Étudions le déterminant d'une matrice triangulaire. Soit donc une matrice triangulaire supérieure d'ordre $n \geqslant 1$:

$$T := \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ 0 & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ 0 & 0 & a_{3,3} & \cdots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n,n} \end{pmatrix}.$$

et proposons-nous de calculer son déterminant :

$$\det T = \begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ 0 & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ 0 & 0 & a_{3,3} & \cdots & a_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n,n} \end{vmatrix}.$$

À cette fin, développons $\det T$ selon les éléments de la première colonne, ce qui donne :

$$\det T = a_{1,1} \begin{vmatrix} a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ 0 & a_{3,3} & \cdots & a_{3,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{n,n} \end{vmatrix}.$$

Recommençons avec la première colonne du déterminant qui apparaît :

$$\det T = a_{1,1}\, a_{2,2} \begin{vmatrix} a_{3,3} & \cdots & a_{3,n} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{n,n} \end{vmatrix},$$

et ainsi de suite, jusqu'à :

$$\begin{aligned} \det T &= a_{1,1}\, a_{2,2} \cdots a_{n-2,n-2} \begin{vmatrix} a_{n-1,n-1} & a_{n,n-1} \\ 0 & a_{n,n} \end{vmatrix} \\ &= a_{1,1}\, a_{2,2} \cdots a_{n-2,n-2}\, a_{n-1,n-1} \left| a_{n,n} \right| \\ &= a_{1,1}\, a_{2,2} \cdots a_{n-2,n-2}\, a_{n-1,n-1}\, a_{n,n}. \end{aligned}$$

Le même phénomène de simplicité des calculs se produit aussi avec des matrices triangulaires inférieures. En résumé :

Proposition 6.10. *Le déterminant d'une matrice triangulaire est égal au produit des éléments de sa diagonale principale.* □

7. Matrice adjointe

Grâce aux déterminants extraits $\det M_{i,j}$, nous allons construire un objet qui permet de calculer l'inverse M^{-1} d'une matrice — quand elle est inversible !

Définition 7.1. Soit $M = \big(a_{i,j}\big)$ une matrice de $\mathscr{M}_n(\mathbb{K})$. Définissons d'abord la matrice $N = \big(b_{i,j}\big)$ de $\mathscr{M}_n(\mathbb{K})$ dont les éléments sont :

$$b_{i,j} := (-1)^{i+j} \det M_{i,j},$$

où $M_{i,j}$ est la matrice déduite de M par suppression de la ligne de rang i et suppression de la colonne de rang j :

$$M_{i,j} := \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}.$$

Terminologie 7.2. On appelle *matrice adjointe*, ou brièvement, *adjointe*, de M, la matrice *transposée* de N :

$$M^* := {}^{\mathtt{t}}N.$$

Exemple 7.3. Dans $\mathscr{M}_3(\mathbb{R})$, soit :

$$M := \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ -1 & 2 & 0 \end{pmatrix}.$$

Ses huit premiers cofacteurs valent :

$$b_{1,1} = +\begin{vmatrix} -1 & 1 \\ 2 & 0 \end{vmatrix} = -2, \qquad b_{1,2} = -\begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix} = -1,$$

$$b_{1,3} = +\begin{vmatrix} 0 & -1 \\ -1 & 2 \end{vmatrix} = -1, \qquad b_{2,1} = -\begin{vmatrix} 0 & -1 \\ 2 & 0 \end{vmatrix} = -2,$$

$$b_{2,2} = +\begin{vmatrix} 1 & -1 \\ -1 & 0 \end{vmatrix} = -1, \qquad b_{2,3} = -\begin{vmatrix} 1 & 0 \\ -1 & 2 \end{vmatrix} = -2,$$

$$b_{3,1} = +\begin{vmatrix} 0 & -1 \\ -1 & 1 \end{vmatrix} = -1, \qquad b_{3,2} = -\begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix} = -1,$$

tandis que le neuvième et petit dernier arrive en clopinant :

$$b_{3,3} = +\begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix} = -1.$$

Par conséquent :

$$\begin{aligned} N &= \begin{pmatrix} b_{1,1} & b_{1,2} & b_{1,3} \\ b_{2,1} & b_{2,2} & b_{2,3} \\ b_{3,1} & b_{3,2} & b_{3,3} \end{pmatrix} \\ &= \begin{pmatrix} -2 & -1 & -1 \\ -2 & -1 & -2 \\ -1 & -1 & -1 \end{pmatrix} \end{aligned}$$

La matrice adjointe de M est par définition la *transposée* de cette matrice, c'est-à-dire :

$$M^* := {}^{\mathtt{t}}N = \begin{pmatrix} -2 & -2 & -1 \\ -1 & -1 & -1 \\ -1 & -2 & -1 \end{pmatrix}.$$

Comme on le voit, rappelons que N est formée en remplaçant chaque élément de la matrice donnée M par son cofacteur dans le développement de $\det M$.

Théorème 7.4. *Quelle que soit la matrice $M \in \mathscr{M}_n(\mathbb{K})$, on a :*

$$M^* M = M M^* = \big(\det M\big) I_n,$$

où I_n est la matrice identité de $\mathscr{M}_n(\mathbb{K})$.

Démonstration. Avec $M = \big(a_{i,j}\big)$, posons :

$$b_{i,j} := (-1)^{i+j} \det\big(M_{i,j}\big),$$

et introduisons :

$$b_{i,j}^{\mathtt{t}} := b_{j,i},$$

de telle sorte que :

$$M^* = \big(b_{i,j}^{\mathtt{t}}\big).$$

Par définition du produit de deux matrices, on a :

$$M^* M = \big(c_{i,k}\big),$$

où :

$$c_{i,k} := \sum_{j=1}^{n} b_{i,j}^{\mathtt{t}}\, a_{j,k}.$$

On a par conséquent :

$$\begin{aligned} c_{i,k} &= \sum_{j=1}^{n} a_{j,k}\, b_{j,i} \\ &= \sum_{j=1}^{n} (-1)^{j+i}\, a_{j,k} \operatorname{det}\big(M_{j,i}\big). \end{aligned}$$

Deux cas se présentent.

Cas 1 : $k = i$. Dans ce cas, le coefficient vaut :

$$\begin{aligned} c_{i,i} &= \sum_{j=1}^{n} (-1)^{j+i}\, a_{j,i} \operatorname{det}\big(M_{j,i}\big) \\ &= \operatorname{det} M, \end{aligned}$$

puisqu'on reconnaît la formule du Théorème 6.4, en échangeant les noms des indices $i \longleftrightarrow j$.

Cas 2 : $k \neq i$. Appelons M' la matrice déduite de M en remplaçant les éléments $a_{1,i}$, $a_{2,i}$, …, $a_{n,i}$ de la colonne i par les éléments $a_{1,k}$, $a_{2,k}$, …, $a_{n,k}$ de la colonne k, ce qui donne :

$$M' := \begin{pmatrix} a_{1,1} & \cdots & a_{1,k} & \cdots & a_{1,k} & \cdots & a_{1,n} \\ a_{2,1} & \cdots & a_{2,k} & \cdots & a_{2,k} & \cdots & a_{2,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,k} & \cdots & a_{n,k} & \cdots & a_{n,n} \end{pmatrix},$$

en supposant $i < k$ pour fixer les idées dans cette représentation.

Or comme cette matrice M' a deux colonnes égales, son déterminant s'annule :

$$0 = \operatorname{det} M'.$$

Mais en développant ce déterminant selon la colonne i, on obtient une formule :

$$0 = \operatorname{det} M' = \sum_{j=1}^{n} (-1)^{j+i}\, a_{j,k} \operatorname{det}\big(M_{j,i}\big),$$

qui est exactement celle obtenue plus haut, ce qui nous donne :

$$c_{i,k} = 0 \qquad\qquad (\forall\, i \neq k).$$

Autrement dit, en utilisant le symbole de Kronecker :

$$c_{i,k} = \delta_{i,k} \operatorname{det} M,$$

et par conséquent :

$$M^* M = \big(\operatorname{det} M\big)\, I_n.$$

On prouverait de même que :

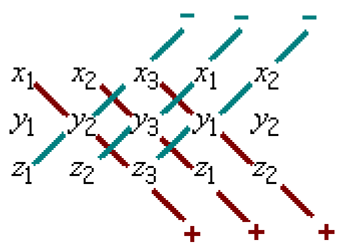
$$M\, M^* = \big(\operatorname{det} M\big)\, I_n.$$ □

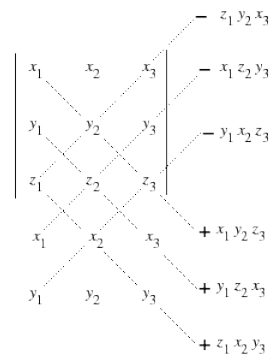
Exemple 7.5. Nous avons vu plus haut que la matrice adjointe de la matrice :

$$M := \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ -1 & 2 & 0 \end{pmatrix}.$$

est :

$$M^* := {}^{\mathrm{t}}N = \begin{pmatrix} -2 & -2 & -1 \\ -1 & -1 & -1 \\ -1 & -2 & -1 \end{pmatrix}.$$





La règle de Sarrus donne la non-annulation des déterminants :

$$\begin{aligned}
\det M &= 0+0+0-(-1)(-1)(-1)-0-1\cdot 1\cdot 2 \\
&= -1, \\
\det M^* &= (-2)(-1)(-1)+(-2)(-1)(-1)+(-1)(-1)(-2)-(-1)(-1)(-1)-(-2)(-1)(-1)-(-2)(-1)(-2) \\
&= -2-2-2+1+2+4 \\
&= -1,
\end{aligned}$$

et le lecteur vérifiera que l'on a bien :

$$\begin{aligned}
M^*\,M = M\,M^* &= \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \\
&= \det M \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.
\end{aligned}$$

8. Critère pour l'invertibilité des matrices

Maintenant, nous pouvons appliquer ce qui précède aux matrices inversibles de l'anneau $\mathscr{M}_n(\mathbb{K})$.

Théorème 8.1. *Pour qu'une matrice $M \in \mathscr{M}_n(\mathbb{K})$ soit inversible, il faut et il suffit que :*

$$\det M \neq 0.$$

Dans ce cas, on a alors :

$$\boxed{M^{-1} = \frac{1}{\det M}\,M^*,}\,.$$

Démonstration. Premièrement, établissons la nécessité. Supposons donc M inversible. Ainsi, il existe une matrice $M^{-1} \in \mathscr{M}_n(\mathbb{K})$ telle que :

$$M^{-1}\,M = M\,M^{-1} = I_n.$$

Appliquons le Théorème 5.2 :

$$\det\big(M^{-1}\,M\big) = \det M^{-1}\,\det M.$$

Comme nous savons que $\det I_n = 1$, nous obtenons :

$$1 = \det M^{-1} \det M, \tag{8.2}$$

ce qui fait voir que, dans le corps $\mathbb{K}$:

$$\det M \neq 0.$$

Deuxièmement, établissons la suffisance. D'après le Théorème 7.4, nous avons :

$$M^* M = \det M \; I_n.$$

Supposons donc $\det M \neq 0$. Alors :

$$\frac{M^*}{\det M} M = I_n.$$

Par conséquent, M est inversible, avec, comme annoncé :

$$M^{-1} = \frac{M^*}{\det M}.$$

□

Corollaire 8.3. *Si $M \in \mathscr{M}_n(\mathbb{K})$ est inversible, alors :*

$$\det M^{-1} = \frac{1}{\det M}.$$

□

Démonstration. En effet, nous avons vu cela en chemin dans l'équation (8.2). □

Exemple 8.4. La matrice :

$$M = \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ -1 & 2 & 0 \end{pmatrix}$$

vue plus haut est inversible, car $\det M = -1$. Son inverse est :

$$M^{-1} = \frac{M^*}{\det M} = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix}.$$

Terminons ce chapitre en introduisant le concept de *déterminant* d'un endomorphisme linéaire $f \colon E \longrightarrow E$.

Soit donc E un espace vectoriel de dimension $n \geqslant 1$ sur le corps $\mathbb{K}$. Soit aussi une base de E :

$$\boldsymbol{B}_E = \{\vec{e}_1, \dots, \vec{e}_n\},$$

et soit la matrice de f dans cette base :

$$M := \mathsf{Mat}_{\boldsymbol{B}_E \boldsymbol{B}_E}(f).$$

Soit par ailleurs une *autre* base :

$$\boldsymbol{B}'_E = \{\vec{e}\,'_1, \dots, \vec{e}\,'_n\},$$

d'où la deuxième matrice :

$$M' := \mathsf{Mat}_{\boldsymbol{B}'_E \boldsymbol{B}'_E}(f).$$

Si P désigne la matrice de passage qui envoie la base $\boldsymbol{B}_E$ sur la base $\boldsymbol{B}'_E$, alors nous avons vu dans un chapitre qui précède que :

$$M' = P^{-1} M P.$$

En appliquant le Théorème 5.2, on en déduit :

$$\begin{aligned} \det M' &= \det\big(P^{-1} M P\big) \\ &= \det P^{-1} \det M \det P \\ &= \det M, \end{aligned}$$

grâce au Corollaire 8.3.

Proposition 8.5. *Le déterminant de la matrice associée à un endomorphisme linéaire $f \in \mathscr{L}(E, E)$ d'un $\mathbb{K}$-espace vectoriel E de dimension $n \geqslant 1$ est indépendant de la base choisie.* $\square$

Ceci justifie la

Définition 8.6. Pour tout endomorphisme $f \in \mathscr{L}(E)$, on appelle *déterminant* de f, et on note :

$$\det f,$$

le déterminant de la matrice de f dans une base quelconque de E.

9. Exercices

Exercice 1. EE

Exercice 2. EE

Théorie du rang
Systèmes linéaires

François DE MARÇAY
Département de Mathématiques d'Orsay
Université Paris-Sud, France

1. Introduction

2. Rang d'une application linéaire

Soit $\mathbb{K}$ un corps, par exemple $\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$. Soient E et F deux espaces vectoriels de dimensions finies :

$$\dim_{\mathbb{K}} E =: n \geqslant 1 \qquad \text{et} \qquad \dim_{\mathbb{K}} F =: m \geqslant 1.$$

Soit aussi une application linéaire :

$$E \xrightarrow{\ f\ } F.$$

Définition 2.1. On appelle *rang* de f la dimension du sous-espace vectoriel $f(E) \subset F$, image de E par f :

$$\mathsf{rang}(f) := \dim \mathsf{Im}\, f.$$

Si on munit E d'une base $\{\vec{e}_1, \dots, \vec{e}_n\}$, alors $f(E)$ est engendré par la famille des n vecteurs images :

$$\{f(\vec{e}_1), \dots, f(\vec{e}_n)\}.$$

De cette famille, on peut extraire une *base* de $f(E)$, laquelle aura exactement un nombre de vecteur égal à $\mathsf{rang}(f)$, d'après la définition du rang, et d'après un théorème vu au début du développement de la théorie.

On a alors évidemment :

$$\begin{aligned} \mathsf{rang}(f) &\leqslant n = \dim E, \\ \mathsf{rang}(f) &\leqslant \dim F = m, \end{aligned}$$

d'où toujours :

$$\mathsf{rang}(f) \leqslant \min\big(\dim E, \dim F\big).$$

Ensuite, introduisons la notion de *rang* d'une famille de vecteurs.

Définition 2.2. Dans un espace vectoriel F sur un corps $\mathbb{K}$, on appelle *rang* d'une famille de vecteurs :

$$\{\vec{y}_1, \vec{y}_2, \dots, \vec{y}_n\} \subset F,$$

la dimension du sous-espace vectoriel qu'elle engendre :

$$\mathsf{rang}\,\{\vec{y}_1, \dots, \vec{y}_n\} := \dim \mathsf{Vect}\,\big(\vec{y}_1, \dots, \vec{y}_n\big).$$

Le rang d'une telle famille $\{\vec{y}_1, \dots, \vec{y}_n\}$ est donc égal au cardinal — nombre d'éléments — de toute base que l'on peut extraire du sous-espace vectoriel :

$$F' := \mathsf{Vect}(\vec{y}_1, \dots, \vec{y}_n) \subset F.$$

Or puisque l'espace engendré ne dépend pas de l'ordre des vecteurs dans la famille, on peut supposer, en changeant au besoin l'ordre de la numérotation, qu'une base extraite

$$F' = \mathsf{Vect}\,(\vec{y}_1, \dots, \vec{y}_r, \dots, \vec{y}_n) = \mathsf{Vect}\,(\vec{y}_1, \dots, \vec{y}_r)$$

est constituée des r premiers vecteurs, r étant le rang de la famille considérée.

Autrement dit, les vecteurs $\vec{y}_1, \dots, \vec{y}_r$ sont linéairement indépendants, et aussi, tout autre vecteur :

$$\vec{y}_j \qquad \text{d'indice} \qquad r+1 \leqslant j \leqslant n,$$

est combinaison linéaire des vecteurs $\vec{y}_1, \dots, \vec{y}_r$ de cette base.

Observons au passage que la définition du rang d'une famille de vecteurs s'applique à toute famille finie de formes linéaires de l'espace dual :

$$E^* := \mathscr{L}(E, \mathbb{K}),$$

et alors, on appelle *rang* d'une famille de formes linéaires :

$$\{f_1, f_2, \dots, f_n\} \subset E^*,$$

la dimension du sous-espace vectoriel de E^* qu'elles engendrent.

Enfin, introduisons la notion de *rang* d'une matrice. Ici, il faut s'imaginer que $E := \mathbb{K}^n$ et que $F := \mathbb{K}^m$, comme nous allons l'argumenter dans un instant.

Définition 2.3. Étant donné une matrice $M \in \mathscr{M}_{m,n}(\mathbb{K})$ à $m \geqslant 1$ lignes et à $n \geqslant 1$ colonnes :

$$M = \begin{array}{c} \begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} \\ \begin{array}{ccc} \scriptstyle\vec{y}_1 & \scriptstyle\cdots & \scriptstyle\vec{y}_n \end{array} \end{array},$$

on appelle *rang* de M, et on note :

$$\mathsf{rang}(M),$$

le rang de la famille des n vecteurs-colonnes de cette matrice, envisagés comme vecteurs appartenant à l'espace vectoriel $\mathbb{K}^m$.

Si on introduit la base canonique de $\mathbb{K}^m$:

$$\{\vec{\varepsilon}_1, \dots, \vec{\varepsilon}_m\},$$

constituée des m vecteurs ayant la coordonnée 1 à la i-ème place, tandis qu'il n'y a que des 0 ailleurs ;

$$\varepsilon_i := (0, \dots, 0, 1, 0, \dots, 0) \qquad\qquad (1 \leqslant i \leqslant m),$$

alors les n vecteurs-colonnes dans $\mathbb{K}^m$ associés à cette matrice s'écrivent :

$$\vec{y}_j := \sum_{i=1}^{m} a_{i,j}\, \vec{\varepsilon}_i \qquad\qquad (1 \leqslant j \leqslant n),$$

et cette matrice $M = (a_{i,j})$ représente alors une application linéaire bien définie $g \colon \mathbb{K}^n \longrightarrow \mathbb{K}^m$. Dans ces conditions, observons que d'après les Définitions 2.1, 2.2, 2.3, on peut écrire :

$$\mathsf{rang}(g) = \mathsf{rang}(M).$$

L'énoncé suivant clarifie encore mieux le lien entre les diverses notions de rang.

Théorème 2.4. *Soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions respectives $n \geqslant 1$ et $m \geqslant 1$, rapportés à deux bases respectives :*

$$\boldsymbol{B}_E = \{\vec{e}_1, \dots, \vec{e}_n\} \qquad \text{et} \qquad \boldsymbol{B}_F = \{\vec{f}_1, \dots, \vec{f}_m\}.$$

Si, relativement à ces deux bases, une application linéaire $f \in \mathscr{L}(E, F)$ est représentée par sa matrice :

$$M := \mathsf{Mat}_{\boldsymbol{B}_E \boldsymbol{B}_F}(f) \in \mathscr{M}_{m,n}(\mathbb{K}),$$

alors :

$$\mathsf{rang}(M) = \mathsf{rang}(f).$$

Démonstration. Dans la base $\boldsymbol{B}_E$, soit l'application qui constitue et définit les coordonnées $x_1, x_2, \dots, x_n$ d'un vecteur quelconque $\vec{x} \in E$:

$$\varphi(x_1, x_2, \dots, x_n) := x_1\, \vec{e}_1 + x_2\, \vec{e}_2 + \cdots + x_n\, \vec{e}_n = \vec{x}.$$

En particulier, pour tout indice $1 \leqslant j \leqslant n$, le vecteur $\vec{\varepsilon}_j$ de la base canonique de $\mathbb{K}^n$ défini ci-dessus est envoyé par φ sur le vecteur $\vec{e}_j$ de la base $\boldsymbol{B}_E$

$$\varphi(\vec{\varepsilon}_j) = \vec{e}_j \qquad (1 \leqslant j \leqslant n).$$

Élaborons un diagramme commençant :

$$\begin{tikzcd} E \arrow[r, "f"] & F \\ \mathbb{K}^n \arrow[u, "\varphi"] & \mathbb{K}^m. \arrow[u, "\psi"'] \end{tikzcd}$$

De même, il existe, relativement à la base $\boldsymbol{B}_F = \{\vec{f}_1, \dots, \vec{f}_m\}$ de F un isomorphisme et un seul de $\mathbb{K}^m$ sur F défini par :

$$\psi(y_1, y_2, \dots, y_m) := y_1\, \vec{f}_1 + y_2\, \vec{f}_2 + \cdots + y_m\, \vec{f}_m.$$

Complétons alors notre schéma :

$$\begin{tikzcd} E \arrow[r, "f"] & F \\ \mathbb{K}^n \arrow[u, "\varphi"] \arrow[r, "g", "\psi^{-1}\circ f\circ\varphi"'] & \mathbb{K}^m, \arrow[u, "\psi"'] \end{tikzcd}$$

en introduisant l'application composée $g\colon \mathbb{K}^n \longrightarrow \mathbb{K}^m$ qui est définie indirectement en parcourant l'arche par le haut :

$$g := \psi^{-1} \circ f \circ \varphi.$$

Proposons-nous maintenant de rechercher la matrice de g dans les bases canoniques de $\mathbb{K}^n$ et de $\mathbb{K}^m$.

Pour tout vecteur $\vec{\varepsilon}_j$ de la base de $\mathbb{K}^n$ avec $1 \leqslant j \leqslant n$, on sait que $\varphi(\vec{\varepsilon}_j) = \vec{e}_j$.

Soit maintenant $M = (a_{i,j})$ la matrice de $f \in \mathscr{L}(E, F)$ relativement aux bases $\boldsymbol{B}_E$ et $\boldsymbol{B}_F$, c'est-à-dire :

$$f(\vec{e}_j) = a_{1,j}\, \vec{f}_1 + a_{2,j}\, \vec{f}_2 + \cdots + a_{m,j}\, \vec{f}_m \qquad (1 \leqslant j \leqslant n).$$

Par conséquent, nous déduisons, toujours avec $1 \leqslant j \leqslant n$ quelconque :

$$\begin{aligned} g(\vec{\varepsilon}_j) = \psi^{-1}\big(f(\varphi(\vec{\varepsilon}_j))\big) &= \psi^{-1}\big(f(\vec{e}_j)\big) \\ &= \big(a_{1,j},\, a_{2,j},\, \dots,\, a_{m,j}\big) \in \mathbb{K}^m. \end{aligned}$$

Le vecteur-colonne de rang j de la matrice de g coïncide donc précisément le vecteur de coordonnées $\big(a_{1,j}, \dots, a_{m,j}\big)$, et ainsi, la matrice de g est :

$$\begin{aligned}\mathsf{Mat}_{B_{\mathbb{K}^n} B_{\mathbb{K}^m}}(g) &= \Big(a_{i,j}\Big)_{1\leqslant i\leqslant m \text{ lignes}}^{1\leqslant j\leqslant n \text{ colonnes}} \\ &= M.\end{aligned}$$

Or nous avons vu que :

$$\mathsf{rang}(M) = \mathsf{rang}(g),$$

et il ne reste plus qu'à prouver que :

$$\mathsf{rang}(g) = \mathsf{rang}(f),$$

c'est-à-dire que :

$$\mathsf{dim}\, g(\mathbb{K}^n) = \mathsf{dim}\, f(E).$$

À cette fin, prouvons que ψ établit un isomorphisme entre $g(\mathbb{K}^n)$ et $f(E)$, ce qui conclura.

En effet, comme ψ est un isomorphisme entre :

$$\mathbb{K}^m \supset g(\mathbb{K}^n) \qquad\qquad \text{et} \qquad\qquad F \supset f(E),$$

il suffit de prouver que la restriction de ψ à $g(\mathbb{K}^n)$ est une surjection de $g(\mathbb{K}^n)$ sur $f(E)$, c'est-à-dire que :

$$\psi\big(g(\mathbb{K}^n)\big) = f(E).$$

Mais comme :

$$g = \psi^{-1} \circ f \circ \varphi,$$

il vient :

$$\psi \circ g = f \circ \varphi.$$

Il ne reste donc plus qu'à se convaincre que :

$$f \circ \varphi(\mathbb{K}^n) = f(E),$$

mais cela est évident, puisque φ est surjective :

$$\varphi(\mathbb{K}^n) = E.$$

Le théorème est donc démontré. □

Exprimons une conséquence importante de ces raisonnement.

Comme précédemment, soit F un espace vectoriel de dimension $m \geqslant 1$ sur le corps $\mathbb{K}$, rapporté à une base $\boldsymbol{B}_F = \big\{\vec{g}_1, \dots, \vec{g}_m\big\}$. Soit une famille de $n \geqslant 1$ vecteurs :

$$\big\{\vec{y}_1, \dots, \vec{y}_n\big\},$$

exprimés dans la base $\boldsymbol{B}_F$ comme :

$$\vec{y}_j = \sum_{i=1}^{m} a_{i,j}\, \vec{f}_i \qquad (1 \leqslant j \leqslant n).$$

Clairement, la matrice :

$$M := \begin{array}{c}\begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \cdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \\ \begin{array}{cccc} \vec{y}_1 & \vec{y}_2 & \cdots & \vec{y}_n \end{array}\end{array}$$

de taille $m \times n$ est la matrice des coordonnées des n vecteurs $\vec{y}_j$ dans la base $\boldsymbol{B}_F$ de F.

Considérons alors l'endomorphisme linéaire $f \in \mathscr{L}(\mathbb{K}^n, F)$ qui envoie les vecteurs $\vec{\varepsilon}_1, \dots, \vec{\varepsilon}_n$ de la base canonique de $\mathbb{K}^n$ sur les vecteurs $\vec{y}_1, \dots, \vec{y}_n$:

$$f\big(\vec{\varepsilon}_j\big) = \vec{y}_j \qquad\qquad (1 \leqslant j \leqslant n).$$

La matrice de f, relativement aux bases $\big\{\vec{\varepsilon}_1, \dots, \vec{\varepsilon}_n\big\}$ et $\big\{\vec{f}_1, \dots, \vec{f}_m\big\}$ respectives de $\mathbb{K}^n$ et de F est précisément :

$$\mathsf{Mat}(f) = M = \big(a_{i,j}\big).$$

D'après le Théorème 2.4, on a :

$$\mathsf{rang}(f) = \mathsf{rang}(M).$$

Or d'après la Définition 2.2, le rang de la famille $\big\{\vec{y}_1, \dots, \vec{y}_n\big\}$ est la dimension du sous-espace vectoriel qu'elle engendre, espace qui, d'après la construction de f, coïncide ici avec $\mathsf{Im}\, f$. Donc le rang de cette famille est égal à $\mathsf{rang}(f)$, c'est-à-dire à $\mathsf{rang}(M)$.

Nous pouvons donc récapituler tous les raisonnements qui précèdent sous la forme d'un théorème synthétique utile dans la pratique.

Théorème 2.5. *Le rang d'une famille de vecteurs d'un espace vectoriel F de dimension finie est égal au rang de la matrice des coordonnées de ces vecteurs dans une base quelconque de F.* □

Terminons cette section en formulant un critère d'indépendance linéaire, valable quand le nombre de vecteurs est égal à la dimension de l'espace vectoriel ambiant.

Théorème 2.6. [Fondamental !] *Dans un espace vectoriel de dimension $n \geqslant 1$ sur un corps $\mathbb{K}$, les trois propriétés suivantes sont équivalentes concernant n vecteurs :*

$$\vec{x}_1, \dots, \vec{x}_n \in E.$$

(i) *La famille $\{\vec{x}_1, \dots, \vec{x}_n\}$ est libre.*

(ii) *La matrice M des coordonnées de $\vec{x}_1, \dots, \vec{x}_n$ dans une base quelconque de E est inversible.*

(iii) *Le rang de M est égal à $n = \mathsf{rang}(M)$.*

Démonstration. Les arguments consistent à faire la synthèse de tout ce qui a été dit dans les pages et dans les chapitres qui précèdent, et donc, il n'y a presque rien à écrire.

Le lecteur-étudiant est fortement invité à revenir en arrière afin de se convaincre qu'il a bien compris pourquoi ces équivalences sont conséquences directes de toute la téorie qui a été élaborée jusqu'à présent. □

3. Matrices extraites d'une matrice

Soient deux entiers $m \geqslant 1$ et $n \geqslant 1$, et soit une matrice de taille $m \times n$:

$$M = \Big(a_{i,j}\Big)_{1 \leqslant i \leqslant m \text{ lignes}}^{1 \leqslant j \leqslant n \text{ colonnes}}.$$

Soit aussi un entier $p \geqslant 1$ avec :

$$p \leqslant m,$$
$$p \leqslant n,$$

d'où :

$$1 \leqslant p \leqslant \mathsf{min}\,(m, n).$$

Définition 3.1. Pour tous choix de deux collections de p indices intermédiaires :

$$1 \leqslant i_1 < \cdots < i_p \leqslant m,$$
$$1 \leqslant j_1 < \cdots < j_p \leqslant n,$$

on appelle *matrice extraite* associée la matrice carrée de taille $p \times p$:

$$M_{i_1,\dots,i_p}^{j_1,\dots,j_p} := \begin{pmatrix} a_{i_1,j_1} & a_{i_1,j_2} & \cdots & a_{i_1,j_p} \\ a_{i_2,j_1} & a_{i_2,j_2} & \cdots & a_{i_2,j_p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_p,j_1} & a_{i_p,j_2} & \cdots & a_{i_p,j_p} \end{pmatrix},$$

et on appelle : *déterminant extrait* :

$$\Delta_{i_1,\dots,i_p}^{j_1,\dots,j_p} := \det M_{i_1,\dots,i_p}^{j_1,\dots,j_p} = \begin{vmatrix} a_{i_1,j_1} & a_{i_1,j_2} & \cdots & a_{i_1,j_p} \\ a_{i_2,j_1} & a_{i_2,j_2} & \cdots & a_{i_2,j_p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i_p,j_1} & a_{i_p,j_2} & \cdots & a_{i_p,j_p} \end{vmatrix}.$$

Nous admettrons sans démonstration le

Théorème 3.2. *Si une matrice $M \in \mathscr{M}_{m,n}(\mathbb{K})$ n'est pas identiquement nulle, alors il existe un entier unique :*

$$1 \leqslant r \leqslant \min(m,n)$$

tel que, pour tout entier $p \geqslant r+1$ strictement supérieur, tous les déterminants extraits s'annulent

$$0 = \Delta_{i_1,\dots,i_p}^{j_1,\dots,j_p},$$

tandis qu'il existe au moins un choix d'indices :

$$1 \leqslant i_1 < \cdots < i_r \leqslant m,$$
$$1 \leqslant j_1 < \cdots < j_r \leqslant n,$$

pour lesquels le déterminant extrait correspondant ne s'annule pas :

$$0 \neq \Delta_{i_1,\dots,i_r}^{j_1,\dots,j_r}$$

De plus, cet entier vaut :

$$\mathrm{rang}(M) = r. \qquad \square$$

Évidemment, quand la matrice est identiquement nulle, son rang vaut 0.

Grâce à ce critère général, on peut (en principe) déterminer de manière algorithmique le rang d'une matrice explicite donnée.

Exemple 3.3. Comme illustration élémentaire, dans l'espace $\mathbb{R}^4$ de la physique à quatre dimensions — celle dans laquelle nous faisons des mathématiques avec notre petite cervelle en ébullition —, soient les trois vecteurs :

$$\vec{x}_1 := \big(1,0,1,0\big), \qquad \vec{x}_2 := \big(0,1,0,1\big), \qquad \vec{x}_3 := \big(1,-1,0,2\big),$$

La matrice de ces vecteurs dans la base canonique de $\mathbb{R}^4$ s'écrit :

$$M := \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 2 \end{pmatrix}.$$

Désignons en abrégé par M_3 la matrice formée des trois premières lignes de M. On trouve aisément (exercice) que :

$$\det M_3 = -1 \neq 0,$$

et par conséquent, $\mathrm{rang}(M) = 3$, donc la famille $\big\{\vec{x}_1, \vec{x}_2, \vec{x}_3\big\}$ est libre.

Formulons maintenant un critère général d'indépendance linéaire pour des collections quelconques de vecteurs.

Dans un $\mathbb{K}$-espace vectoriel F à $m \geqslant 1$ dimensions rapporté à une base :

$$\boldsymbol{B}_F \,=\, \big\{\vec{f}_1, \dots, \vec{f}_m\big\},$$

considérons une famille $\big\{\vec{x}_1, \dots, \vec{x}_n\big\}$ de $n \geqslant 1$ vecteurs arbitraires. Introduisons aussi la matrice :

$$A \,=\, \Big(a_{i,j}\Big)_{1 \leqslant i \leqslant m \text{ lignes}}^{1 \leqslant j \leqslant n \text{ colonnes}},$$

des coordonnées de ces vecteurs dans la base $\boldsymbol{B}_F$.

Pour que cette famille $\big\{\vec{x}_1, \dots, \vec{x}_n\big\}$ de vecteurs soit libre, il faut et il suffit que son rang soit égal au nombre n des vecteurs, c'est-à-dire d'après le Théorème 3.4, que $\mathsf{rang}\, A = n$.

En appliquant alors le Théorème 3.2, on peut énoncer le critère général suivant.

Théorème 3.4. *Dans un espace vectoriel F de dimension $m \geqslant 1$, pour qu'une famille de $n \geqslant 1$ vecteurs soit libre, il faut et il suffit que l'on puisse extraire, de la matrice de ces vecteurs dans une base quelconque, une matrice carrée d'ordre n qui soit inversible,* i.e. *de déterminant non nul.*

4. Systèmes d'équations linéaires

On considère un *système linéaire* de $m \geqslant 1$ à $n \geqslant 1$ inconnues :

$$\left\{\begin{aligned} a_{1,1}\, x_1 + a_{1,2}\, x_2 + \cdots + a_{1,n}\, x_n \,&=\, b_1, \\ a_{2,1}\, x_1 + a_{2,2}\, x_2 + \cdots + a_{2,n}\, x_n \,&=\, b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots&\cdots\cdots \\ a_{m,1}\, x_1 + a_{m,2}\, x_2 + \cdots + a_{m,n}\, x_n \,&=\, b_m, \end{aligned}\right. \tag{4.1}$$

dont les éléments $a_{i,j} \in \mathbb{K}$ sont nommés *coefficients*, ainsi que $b_i \in \mathbb{K}$, appelés *seconds membres*.

Ici évidemment, ce sont les inconnues $x_1, \dots, x_n$ qu'il s'agit de déterminer. On appelle alors *solution* du système tout n-uplet :

$$\big(x_1, x_2, \dots, x_n\big) \,\in\, \mathbb{K}^n$$

qui vérifie les m équations du système.

Résoudre le système, c'est déterminer l'*ensemble* de *toutes* les solutions $(x_1, \dots, x_n)$ — ensemble qui peut éventuellement être vide.

On appelle *matrice* du système la matrice des coefficients du membre de gauche :

$$A \,:=\, \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}.$$

Si donc nous introduisons les deux matrices unicolonnes :

$$X \,:=\, \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad\qquad \text{et} \qquad\qquad B \,:=\, \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

à n lignes et à m lignes, la définition du produit matriciel nous permet d'abréger l'écriture d'un tel système linéaire sous la forme :

$$A\, X \,=\, B. \tag{4.2}$$

Résoudre le système (4.1), c'est donc résoudre cette équation matricielle. Les matrices $A \in \mathscr{M}_{m,n}(\mathbb{K})$ et $B \in \mathscr{M}_{m,1}(\mathbb{K})$ étant données, il s'agit de déterminer l'ensemble des matrices X de $\mathscr{M}_{n,1}(\mathbb{K})$ vérifiant $A\,X = B$.

De plus, un tel système admet une *interprétation vectorielle* qui est absolument essentielle.

En effet, soient E et F deux $\mathbb{K}$-espaces vectoriels de dimensions respectives $n \geqslant 1$ et $m \geqslant 1$ sur le corps $\mathbb{K}$, munis de deux bases :

$$\boldsymbol{B}_E = \{\vec{e}_1, \dots, \vec{e}_n\} \qquad \text{et} \qquad \boldsymbol{B}_F = \{\vec{f}_1, \dots, \vec{f}_m\}.$$

Soit aussi $f\colon E \longrightarrow F$ l'application linéaire qui admet pour matrice la matrice $A = \big(a_{i,j}\big)$ du système. Si on introduit les vecteurs :

$$\vec{x} := x_1\,\vec{e}_1 + \cdots + x_n\,\vec{e}_n \qquad \text{et} \qquad \vec{b} := b_1\,\vec{f}_1 + \cdots + b_m\,\vec{f}_m,$$

le système linéaire (4.1) équivaut alors à :

$$f(\vec{x}) = \vec{b}. \tag{4.3}$$

Résoudre ce système, c'est donc déterminer l'ensemble des vecteurs $\vec{x} \in E$ qui ont pour image le vecteur fixé $\vec{b} \in E$, que nous noterons :

$$\mathsf{Sol} := \big\{\vec{x} \in E\colon\ f(\vec{x}) = \vec{b}\big\}.$$

Dans cette interprétation, il est aisé de délimiter les grandes lignes de la discussion.

Premier cas. Supposons que $\vec{b}$ n'appartienne *pas* à l'image de E par f :

$$\vec{b} \notin \mathsf{Im}\,f.$$

Alors l'équation (4.3) n'a évidemment pas de solution. Dans ce cas, on a donc vacuité $\mathsf{Sol} = \emptyset$ de l'ensemble des solutions.

Deuxième cas. Supposons au contraire maintenant que $\vec{b}$ *appartienne* à l'image de E par f :

$$\vec{b} \in \mathsf{Im}\,f.$$

Il existe alors au moins un vecteur $\vec{x}_0 \in E$ tel que :

$$f(\vec{x}_0) = \vec{b}. \tag{4.4}$$

Pour toute solution $\vec{x} \in \mathsf{Sol}$ de l'équation vectorielle (4.3), on aura alors, en retranchant membre à membre (4.3) et (4.4) :

$$f(\vec{x}) - f(\vec{x}_0) = \vec{0},$$

et comme f est linéaire, il vient :

$$f\big(\vec{x} - \vec{x}_0\big) = \vec{0}.$$

Par suite, si $\mathsf{Ker}\,f$ désigne le noyau de f, ceci veut dire que :

$$\vec{x} - \vec{x}_0 \in \mathsf{Ker}\,f.$$

Réciproquement, tout vecteur $\vec{x}$ de la forme :

$$\vec{x} = \vec{x}_0 + \vec{y} \qquad \text{avec} \qquad \vec{y} \in \mathsf{Ker}\,f \quad \text{quelconque},$$

est solution du problème, car il est clair que :

$$\begin{aligned} f(\vec{x}) &= f(\vec{x}_0) + \underline{f(\vec{y})} \\ &= f(\vec{x}_0) + 0 \\ &= \vec{b}. \end{aligned}$$

Par conséquent, si $\vec{x}_0$ est une certaine solution (particulière) de l'équation vectorielle (4.3), alors l'ensemble de *toutes* les solutions de cette équation est exactement égal à :

$$\mathsf{Sol} = \Big\{\vec{x} \in E\colon\ \exists\,\vec{y} \in \mathsf{Ker}\,f, \quad \vec{x} = \vec{x}_0 + \vec{y}\Big\}.$$

Il est *paramétré* par le noyau de f. Énonçons ce résultat très important sous une forme synthétique, en prenant pour fixer les idées :

$$E := \mathbb{K}^n \qquad \text{et} \qquad F := \mathbb{K}^m.$$

Théorème 4.5. *Étant donné un système de $m \geqslant 1$ équations linéaires à $n \geqslant 1$ inconnues :*

$$\sum_{j=1}^{n} a_{i,j}\, x_j \,=\, b_i \qquad\qquad (1 \leqslant i \leqslant m),$$

interprété comme équation vectorielle :

$$f(\vec{x}) \,=\, \vec{b},$$

au moyen de l'application linéaire $f \in \mathscr{L}(\mathbb{K}^n, \mathbb{K}^m)$ qui a pour matrice $\big(a_{i,j}\big)$ dans les bases canoniques de $\mathbb{K}^m$ et de $\mathbb{K}^n$, et avec $\vec{b} \in \mathbb{K}^m$ fixé, l'ensemble complet de ses solutions :

$$\mathsf{Sol} \,:=\, \big\{\vec{x} \in \mathbb{K}^n \colon\ f(\vec{x}) = \vec{b}\big\},$$

est le suivant.

(1) *Quand $\vec{b} \notin \mathsf{Im}\, f$, il est vide $\mathsf{Sol} = \emptyset$.*

(2) *Quand $\vec{b} \in \mathsf{Im}\, f$, il admet au moins une solution $\vec{x}_0 \in \mathsf{Sol}$, et alors l'ensemble complet de ses solutions est :*

$$\mathsf{Sol} \,=\, \Big\{\vec{x} \in E \colon\ \exists\, \vec{y} \in \mathsf{Ker}\, f, \quad \vec{x} = \vec{x}_0 + \vec{y}\Big\}. \qquad \square$$

Trois cas spéciaux méritent d'être explicités.

(a) Quand f est *surjective*, ce qui exige $m \leqslant n$, c'est-à-dire quand :

$$\mathsf{Im}\, f \,=\, F,$$

tout vecteur $\vec{b} \in F$ appartient à $\mathsf{Im}\, f$, et alors le système $A\,X \,=\, B$ admet toujours des solutions, *quel que soit le choix d'un second membre fixé B*.

(b) Quand f est est *injective*, un théorème vu dans un chapitre qui précède a montré que :

$$\dim \mathsf{Im}\, f \,=\, \dim E \,=\, n,$$

et l'inclusion $\mathsf{Im}\, f \subset F$ force $n \leqslant m = \dim F$. En prenant alors un second membre B tel que le vecteur associé $\vec{b} \in \mathsf{Im}\, f$ appartienne à l'image de f, on voit que dans ce cas, le système linéaire admet une solution *unique*, parce que $\mathsf{Ker}\, f = \{0\}$, donc les vecteurs $\vec{y}$ qui paramétrisent l'espace Sol des solutions sont tous nuls.

(c) Enfin, quand f est *bijective*, c'est-à-dire simultanément injective et surjective, ce qui exige $n = m$, le système admet une solution par **(a)**, et une seule par **(b)**. Nous allons maintenant étudier plus en détail ce cas très intéressant.

5. Systèmes de Cramér

Dans le système (4.1) linéaire $A\,X \,=\, B$ réinterprété comme équation (4.3) vectorielle, nous supposons donc dans cette section que deux conditions sont réalisées :

(1) $n = m$;

(2) la matrice $A = \big(a_{i,j}\big)$ du système est inversible.

Ainsi, en partant de :

$$A\,X \,=\, B, \tag{4.2}$$

nous pouvons multiplier par A^{-1} ce qui fournit instantanément :

$$X \,=\, A^{-1} B, \tag{5.1}$$

la solution unique X du système.

Or nous avons vu dans le chapitre précédent une formule donnant cette matrice inverse :

$$A^{-1} = \frac{1}{\det A} A^*,$$

où A^* est la matrice adjointe de A. Alors l'équation (5.1) qui donne la solution s'écrit de manière équivalente :

$$\det A \cdot X = A^* B, \tag{5.2}$$

Or la matrice adjointe a pour entrées, après transposition :

$$a^*_{i,j} := (-1)^{i+j} \det \big(A_{j,i}\big) \qquad {\scriptstyle (1 \leqslant i, j \leqslant n)},$$

où $A_{i,j}$ est la matrice déduite de A par suppression de la ligne i et de la ligne j. Par conséquent, le produit matriciel $A^* B$, qui est une matrice colonne, a pour éléments :

$$\sum_{j=1}^n a^*_{i,j}\, b_j \qquad {\scriptstyle (1 \leqslant i \leqslant n)}.$$

En identifiant terme à terme les deux matrices de $\mathscr{M}_{1,n}(\mathbb{K})$ qui interviennent dans la relation (5.2), on obtient donc :

$$\big(\det A\big)\, x_i = \sum_{j=1}^n (-1)^{i+j} \det \big(A_{j,i}\big)\, b_j \qquad {\scriptstyle (1 \leqslant i \leqslant n)}.$$

Et là, si on est très intelligent, on voit quelque chose de crucial à deviner.

Observation 5.3. *Le second membre ci-dessus est le développement du déterminant de la matrice A' déduite de A en remplaçant la colonne de rang j de A par les seconds membres $b_1, \dots, b_n$.* □

Ceci conduit à énoncer un magnifique

Théorème 5.4. [Formules de Cramér] *La solution unique $(x_1, \dots, x_n)$ d'un système de $n \geqslant 1$ équations linéaire à $n \geqslant 1$ inconnues :*

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

dont la matrice $A = \big(a_{i,j}\big)$ est inversible*, c'est-à-dire a un déterminant est non nul :*

$$0 \neq \det A \qquad (\textit{Hypothèse}),$$

possède n coordonnées x_i donnée par les formules de Cramér *:*

$$x_i = \frac{\begin{vmatrix} a_{1,1} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1,n} \\ a_{2,1} & \cdots & a_{2,i-1} & b_2 & a_{2,i+1} & \cdots & a_{2,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{n,n} \end{vmatrix}}{\begin{vmatrix} a_{1,1} & \cdots & a_{1,i-1} & a_{1,i} & a_{1,i+1} & \cdots & a_{1,n} \\ a_{2,1} & \cdots & a_{2,i-1} & a_{2,i} & a_{2,i+1} & \cdots & a_{2,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,i-1} & a_{n,i} & a_{n,i+1} & \cdots & a_{n,n} \end{vmatrix}} \qquad {\scriptstyle (1 \leqslant i \leqslant n)},$$

dans lesquelles on divise toujours par $\det A \neq 0$. □

En dimension $n = 2$, la solution unique du système :

$$\begin{aligned} a\,x + b\,y &= F, \\ c\,x + d\,y &= G, \end{aligned}$$

est donc, sous l'hypothèse qu'on puisse effectivement diviser par le déterminant, qu'on suppose non nul :

$$x = \frac{\begin{vmatrix} F & b \\ G & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}, \qquad\qquad y = \frac{\begin{vmatrix} a & F \\ c & G \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}.$$

En dimension $n = 3$, un système de Cramér général :

$$\begin{aligned} a\,x + a'\,y + a''\,z &= F, \\ b\,x + b'\,y + b''\,z &= G, \\ c\,x + c'\,y + c''\,z &= H, \end{aligned}$$

dont le déterminant est non nul a pour solution unique :

$$x = \frac{\begin{vmatrix} F & a' & a'' \\ G & b' & b'' \\ H & c' & c'' \end{vmatrix}}{\begin{vmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{vmatrix}}, \qquad y = \frac{\begin{vmatrix} a & F & a'' \\ b & G & b'' \\ c & H & c'' \end{vmatrix}}{\begin{vmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{vmatrix}}, \qquad z = \frac{\begin{vmatrix} a & a' & F \\ b & b' & G \\ c & c' & H \end{vmatrix}}{\begin{vmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{vmatrix}}.$$

Exemple 5.5. Proposons-nous de résoudre le système :

$$\begin{aligned} 3\,x + 2\,y - z &= 1, \\ x\, - \,y\, + \,z &= 0, \\ x + y - 2\,z &= -1, \end{aligned}$$

de $m = 3$ équations à $n = 3$ inconnues. La matrice de ses coefficients est :

$$A = \begin{pmatrix} 3 & 2 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$

La règle de Sarrus nous permet aisément (exercise) de calculer :

$$\det A = 7.$$

Comme ce déterminant est non nul, nous sommes en présence d'un système de Cramér. Appliquons alors les formules du théorème qui précède pour obtenir les trois cooordonnées de la solution unique :

$$\begin{aligned} 7\,x &= \begin{vmatrix} 1 & 2 & -1 \\ 0 & -1 & 1 \\ -1 & 1 & -2 \end{vmatrix} = 0, \\ 7\,y &= \begin{vmatrix} 3 & 1 & -1 \\ 1 & 0 & 1 \\ 1 & -1 & -2 \end{vmatrix} = 7, \\ 7\,z &= \begin{vmatrix} 3 & 2 & 1 \\ 1 & -1 & 0 \\ 1 & 1 & -1 \end{vmatrix} = 7. \end{aligned}$$

La solution du système est donc :

$$x = 0, \qquad y = 1, \qquad z = 1.$$

En toutes circonstances, l'étudiant qui passe un examen doit vérifier si la solution qu'il trouve est bien une solution, en la remplaçant dans le système donné.

6. Résolution d'un système linéaire général

Sans supposer $m = n$, considérons un système (4.1) de $m \geqslant 1$ équations à $n \geqslant 1$ inconnues, où m et n sont des entiers quelconques. Désignons pas r le rang de la matrice A de ses coefficients :

$$r := \operatorname{rang} A = \operatorname{rang} \big(a_{i,j}\big)_{1 \leqslant i \leqslant m \text{ lignes}}^{1 \leqslant j \leqslant n \text{ colonnes}}.$$

D'après le Théorème 3.2, nous savons qu'il existe une sous-matrice carrée A_r de taille $r \times r$ extraite de A qui est inversible, *i.e.* satisfait :

$$0 \neq \det A_r.$$

En renumérotant au besoin les lignes et les colonnes de A, ce qui revient à modifier l'ordre des équations et a changer le nom des inconnues, nous pouvons supposer que cette sous-matrice $A_r \subset A$ est constituée des r premières lignes et des r premières colonnes de A.

Ainsi, le système linéaire que nous devons résoudre s'écrit sous une forme :

$$\tag{6.1} \begin{cases} a_{1,1}\, x_1 + \cdots + a_{1,r}\, x_r + \cdots + a_{1,n}\, x_n = b_1, \\ a_{2,1}\, x_1 + \cdots + a_{2,r}\, x_r + \cdots + a_{2,n}\, x_n = b_2, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{r,1}\, x_1 + \cdots + a_{r,r}\, x_r + \cdots + a_{r,n}\, x_n = b_r, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m,1}\, x_1 + \cdots + a_{m,r}\, x_r + \cdots + a_{m,n}\, x_n = b_m, \end{cases}$$

dans laquelle nous faisons apparaître les éléments de la sous-matrice inversible $A_r \subset A$, dont le déterminant est non nul :

$$\det A_r = \begin{vmatrix} a_{1,1} & \cdots & a_{1,r} \\ \vdots & \ddots & \vdots \\ a_{r,1} & \cdots & a_{r,r} \end{vmatrix} \neq 0.$$

Terminologie 6.2. Les inconnues $x_1, \dots, x_r$ se nomment *inconnues principales*.

Les r premières équations se nomment *équations principales*.

Deux cas se présentent nécessairement, que nous envisageons à la suite l'un de l'autre.

Premier cas : *Le rang de la matrice est égal au nombre des équations du système :*

$$r = m.$$

Revenons au point de vue vectoriel, d'après lequel le système (6.1) équivaut à :

$$\tag{4.3} f(\vec{x}) = \vec{b}.$$

On a ici :

$$r = \dim \operatorname{Im} f = \dim F = m,$$

ce qui force :

$$\operatorname{Im} f = F,$$

c'est-à-dire que f est *surjective*. Ainsi, quel que soit le vecteur $\vec{b}$, l'équation vectorielle (4.3) a des solutions.

Si $\vec{x}_0$ est l'une de ces solutions, toutes les autres solutions sont de la forme $\vec{x}_0 + \vec{y}$, quand $\vec{y}$ parcourt $\operatorname{Ker} f$.

Or d'après la formule du rang vue dans un chapitre qui précède, on a toujours :

$$\dim \operatorname{Ker} f + \dim \operatorname{Im} f \;=\; \dim E \;=\; n,$$

donc nous déduisons que :

$$\dim \operatorname{Ker} f \;=\; n - r.$$

Par conséquent, les vecteurs $\vec{x}_0 + \vec{y}$ qui sont solutions de (4.3) dépendent de $n - r$ scalaires arbitraires, par exemple les coordonnées de $\vec{y}$ dans une certaine base de $\operatorname{Ker} f$.

Observons que si $r = n$, la surjection f est aussi une injection, c'est-à-dire que $\operatorname{Ker} f = \{\vec{0}\}$, et alors on est en présence d'un système de Cramér, objet que nous avons déjà étudié.

Expliquons maintenant comment réaliser le calcul effectif des solutions du système, lequel s'écrit puisque $r = m$:

$$\left\{\begin{aligned} a_{1,1}\,x_1 + \cdots + a_{1,m}\,x_m + a_{1,m+1}\,x_{m+1} + \cdots + a_{1,n}\,x_n \;&=\; b_1, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m,1}\,x_1 + \cdots + a_{m,m}\,x_m + a_{m,m+1}\,x_{m+1} + \cdots + a_{m,n}\,x_n \;&=\; b_m, \end{aligned}\right.$$

Attribuons des valeurs arbitraires dans $\mathbb{K}$ aux $n - m$ inconnues non principales :

$$x_{m+1},\;\dots,\;x_n.$$

Faisons-les toutes passer à droite des équations du système :

$$\left\{\begin{aligned} a_{1,1}\,x_1 + \cdots + a_{1,m}\,x_m \;&=\; b_1 - a_{1,m+1}\,x_{m+1} - \cdots - a_{1,n}\,x_n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m,1}\,x_1 + \cdots + a_{m,m}\,x_m \;&=\; b_m - a_{m,m+1}\,x_{m+1} - \cdots - a_{m,n}\,x_n, \end{aligned}\right.$$

et appelons en abrégé b_i' ces nouveaux membres de droite :

$$\left\{\begin{aligned} a_{1,1}\,x_1 + \cdots + a_{1,m}\,x_m \;&=\; b_1', \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m,1}\,x_1 + \cdots + a_{m,m}\,x_m \;&=\; b_m'. \end{aligned}\right.$$

Maintenant, nous sommes en présence d'un système de Cramér, puisque le nombre d'équations du système est égal au nombre d'inconnues, et puisque par hypothèse, le déterminant de la matrice extraite A_r que nous voyons bien ici est non nul. On résout alors ce système, et on constate qu'il dépend des :

$$n - m \quad \text{paramètres } \; x_{m+1}, \dots, x_n.$$

Deuxième cas : *Le rang de la matrice est strictemen inférieur au nombre des équations du système :*

$$r \;<\; m.$$

Nous savons que l'équation vectorielle associée $f(\vec{x}) = \vec{b}$ possède ou n'a pas de solutions selon que $\vec{b} \in \operatorname{Im} f$ ou que $\vec{b} \notin \operatorname{Im} f$.

Or l'image $\operatorname{Im} f$ est engendrée par les r premières colonnes de la matrice A, c'est-à-dire par les images par f des r premiers vecteurs de la base $f(\vec{e}_1), \dots, f(\vec{e}_r)$.

Ensuite, pour que $\vec{b} \in \operatorname{Im} f$, il faut et il suffit que la famille :

$$\big\{f(\vec{e}_1), \dots, f(\vec{e}_r),\, \vec{b}\big\},$$

soit *liée* dans F. Or d'après le critère général d'indépendance linéaire que le Théorème 3.4 a formulé, il faut et il suffit que la matrice des coordonnées de ces vecteurs dans une base de F soit de rang strictement inférieur à $r + 1$.

Mais comme cette matrice est :

$$A' = \begin{pmatrix} a_{1,1} & \cdots & a_{1,r} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{r,1} & \cdots & a_{r,r} & b_r \\ \vdots & \ddots & \vdots & \vdots \\ a_{m,1} & \cdots & a_{m,r} & b_m \end{pmatrix},$$

et comme par hypothèse le sous-déterminant de cette matrice :

$$\det A_r = \begin{vmatrix} a_{1,1} & \cdots & a_{1,r} \\ \vdots & \ddots & \vdots \\ a_{r,1} & \cdots & a_{r,r} \end{vmatrix} \neq 0$$

est non nul, pour que $\vec{b} \in \operatorname{Im} f$, il faut et il suffit que l'on ait aussi :

$$\operatorname{rang}(A') = r,$$

et ceci équivaut, d'après le Théorème 3.2, à ce que *toutes* les matrices carrées extraites de taille $(r+1) \times (r+1)$ soient de déterminat nul.

Et ici, puisque A' n'a que $r+1$ colonnes, cela équivaut à ce que, pour tout entier $k = r + 1, \dots, m$, on ait annulation de :

$$\begin{vmatrix} a_{1,1} & \cdots & a_{1,r} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{r,1} & \cdots & a_{r,r} & b_r \\ a_{k,1} & \cdots & a_{k,r} & b_k \end{vmatrix} = 0 \qquad (r+1 \leqslant k \leqslant m).$$

Terminologie 6.3. Ces $m-r$ déterminants seront appelés *déterminants caractéristiques* du système.

En résumé, pour que le système (6.1) ait des solutions, il faut et il suffit que *tous les déterminants caractéristiques soient nuls*.

Dans ce cas, si $(x_1, \dots, x_n)$ est une solution des r (premières) équations principales du système, alors elle est solution de toute équation non principale.

En conclusion de cette étude, nous pouvons énoncer un résultat synthétique.

Théorème 6.4. [de Fontené-Rouché] *Soit un système $A\,X = B$ de m équations à n inconnues, de rang r.*

(1) *Si $r = m$, alors le système a des solutions, obtenues en attribuant des valeurs arbitraires aux $n-r$ inconnues non principales, et en résolvant le système de Cramér associé de $r = m$ équations à $r = m$ inconnues.*

(2) *Si $r < m$, et si l'un des déterminants caractéristiques est différent de zéro, alors le système n'a aucune solution.*

(3) *Si $r < m$ et si tous les déterminants caractéristiques sont nuls, alors le système se réduit aux r équations principales, en supprimant les $m-r$ dernières équations, et le système réduit se résout en appliquant la méthode du* **(1)**. □

Exemple 6.5. Proposons-nous de résoudre dans $\mathbb{R}$ le système suivant :

$$\begin{cases} a\,x + y + z + t = 1, \\ x + ay + z + t = 1, \\ x + y + az + t = 1, \\ x + y + z + at = 1, \end{cases}$$

en discutant les cas possibles selon la valeur du paramètre $a \in \mathbb{R}$.

Commençons par calculer le déterminant de la matrice de ce système :

$$\begin{aligned}\Delta := \begin{vmatrix} a & 1 & 1 & 1 \\ 1 & a & 1 & 1 \\ 1 & 1 & a & 1 \\ 1 & 1 & 1 & a \end{vmatrix} &= \begin{vmatrix} a+3 & 1 & 1 & 1 \\ a+3 & a & 1 & 1 \\ a+3 & 1 & a & 1 \\ a+3 & 1 & 1 & a \end{vmatrix} \\ &= \big(a+3\big) \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & a & 1 & 1 \\ 1 & 1 & a & 1 \\ 1 & 1 & 1 & a \end{vmatrix},\end{aligned}$$

où nous avons déduit le deuxième déterminant du premier en ajoutant les 3 colonnes suivantes à la première.

Dans le dernier déterminant obtenu, retranchons la première colonne successivement aux trois autres colonnes suivantes :

$$\begin{aligned}\Delta &= \big(a+3\big) \begin{vmatrix} 1 & 0 & 0 & 0 \\ 1 & a-1 & 0 & 0 \\ 1 & 0 & a-1 & 0 \\ 1 & 0 & 0 & a-1 \end{vmatrix} \\ &= \big(a+3\big)\big(a-1\big)^3,\end{aligned}$$

et développons le déterminant obtenu par rapport à sa première colonne pour terminer le calcul. Envisageons alors plusieurs cas.

Premier cas :

$$\big(a+3\big)\big(a-1\big) \neq 0.$$

Le système proposé est alors de type Cramér. Pour en obtenir la solution (unique), il serait ici maladroit d'appliquer directement les formules de Cramér, qui exigeraient de calculer 4 autres déterminants de taille 4×4.

Puisqu'on est astucieux, il vaut mieux remarquer que, en ajoutant membre à membre les quatre équations proposées, on obtient :

$$(a+3)\big(x+y+z+t\big) = 4,$$

d'où puisque nous supposons $a \neq -3$:

$$x+y+z+t = \frac{4}{a+3}.$$

En retranchant cette équation successivement de chaque équation proposée, on obtient :

$$\begin{aligned}\big(a-1\big)\,x = \big(a-1\big)\,y = \big(a-1\big)\,z = \big(a-1\big)\,t &= 1 - \frac{4}{a+3} \\ &= \frac{1}{a-1}\,\frac{a+3-4}{a+3},\end{aligned}$$

et enfin, puisque nous supposons $a-1 \neq 0$:

$$x = y = z = t = \frac{1}{a+3}.$$

C'est bien la solution unique du système, comme on le vérifie en remplaçant ces valeurs dans le système.

Deuxième cas :

$$a = -3.$$

Le système est alors de rang 3, car :

$$\begin{vmatrix} -3 & 1 & 1 \\ 1 & -3 & 1 \\ 1 & 1 & -3 \end{vmatrix} = -16 \neq 0.$$

Il y a un seul déterminant caractéristique :

$$\begin{vmatrix} -3 & 1 & 1 & 1 \\ 1 & -3 & 1 & 1 \\ 1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -3 \end{vmatrix} = \begin{vmatrix} -3 & 1 & 1 & 0 \\ 1 & -3 & 1 & 0 \\ 1 & 1 & -3 & 0 \\ 1 & 1 & 1 & 4 \end{vmatrix}$$
$$= -16 \cdot 4,$$

où le second déterminant est déduit du premier en ajoutant les trois premières colonnes à la quatrième, et où on l'a développé par rapport à sa quatrième colonne.

Comme cet unique déterminant caractéristique est non nul, le système n'a pas de solution.

Troisième cas :

$$a = 1.$$

Le système est alors constitué de quatre équations identiques à la première, qui se réduit à une seule équation principale :

$$x + y + z + t = 1.$$

On peut fixer arbitrairement les valeurs de trois inconnues (non principales) et en déduire la valeur de l'inconnue principale.

Remarquons pour terminer que la matrice du système a tous ses éléments égaux. Son rang est évidemment égal à 1.

7. Exercices

Exercice 1. EE

Exercice 2. EE