Calcul vectoriel et matriciel de première année

Janvier 2011

J. Féjoz

J. Féjoz
Université Paris – Dauphine.
Url : `http://ceremade.dauphine.fr/~fejoz/`
E-mail : `fejoz@ceremade.dauphine.fr`

AVERTISSEMENT AU LECTEUR ÉTUDIANT

Dans un cours de mathématiques supérieures, une difficulté peut venir de la place accordée aux démonstrations. Celles-ci sont l'extraordinaire expérience de pensée inventée dans l'Antiquité grecque, notamment par EUCLIDE (fondateur des mathématiques modernes, trois siècles avant notre ère) dans ses *Éléments* pour établir la validité d'une formule ou d'une propriété, par déduction à partir d'autres propriétés déjà connues.

Les démonstrations élargissent considérablement l'horizon mathématique (penser au nombre dérisoire de formules immédiatement évidentes, par rapport à toutes celles que l'on peut démontrer par le calcul), et sont devenues l'essence des mathématiques. Encore plus que les concepts liés au Calcul vectoriel et matriciel, c'est plus généralement le raisonnement mathématique lui-même qui est au coeur de ce cours.

Par ailleurs, à supposer qu'il faille le rappeler, l'appropriation d'un nouveau cours exige un travail personnel assidu. Cela passe par la lecture du cours et sa relecture, ainsi que par la résolution des exercices proposés.

Ses efforts sont-ils suffisants ? Il faut pouvoir redonner sans aide les définitions et énoncés du cours (lemmes, propositions, propriétés, théorèmes, etc., compléments en petits caractères mis à part), rédiger leur démonstration, et trouver des exemples et contre-exemples. Il faut pouvoir résoudre non seulement des problèmes semblables à ceux proposés, mais aussi *d'autres problèmes*, qui se résolvent avec le même appareil conceptuel, mais dont l'habillage peut être très différent.

Le jeu est à la fois plus difficile et plus intéressant que s'il s'agissait de régurgiter des données apprises par coeur et de résoudre des exercices ciblés. Si l'on n'y prend pas goût et si l'on ne perçoit que d'abstraites élucubrations, mieux vaut changer de cours ! J'espère au contraire que ce cours suscitera de l'intérêt, voire même parfois du plaisir ?

... Bon travail !

Merci à toute l'équipe enseignante, dont R. Anton, B. Collas, C. David, A. Deloro, A. Ducros, L. Sabbagh, S. Morier-Genoud, J.-M. Trépreau, M. Wolff et L. Zapponi ; ainsi qu'aux étudiants dont A. Michail et T. L. Mamou, pour leurs nombreuses observations.

TABLE DES MATIÈRES

CHAPITRE 1

LE PLAN COMPLEXE

Mots-clef du chapitre. — Nombre complexe, forme cartésienne, opération, conjugaison, inverse, démonstration par récurrence, argument, exponentielle complexe, formules d'addition, forme polaire, formules d'Euler, polynôme, racine, multiplicité, racine n-ième, théorème fondamental de l'algèbre, discriminant, translation, rotation, homothétie, similitude

ALORS que les Babyloniens avaient compris quinze siècles avant notre ère comment résoudre des équations du second degré telles que $x^2 - 4x + 3 = 0$, longtemps il n'y eu pas d'avancée significative sur les équations du troisième degré. Le déblocage vint de la découverte, au XVIe siècle italien, d'une mystérieuse quantité $\sqrt{-1}$ avec laquelle on pouvait calculer comme avec les nombres réels sans pour autant qu'elle en soit un (voir § 1.4).

C'est l'un des objectifs de ce chapitre que de définir le nombre $\sqrt{-1}$, noté i depuis L. EULER (mathématicien suisse, 1707–1783), ainsi, plus généralement, que les *nombres complexes* (ou *imaginaires*), de la forme $a + ib$. De même qu'on ne peut pas poser -2 carottes sur une table, on ne peut pas construire de segment de longueur $1 + i$. Mais contrairement à ce que la terminologie semble indiquer, les nombres complexes ne sont pas plus complexes, ni plus imaginaires, ni moins réels (au sens commun), que les nombres réels ou entiers.

1.1. Nombres complexes

Définitions. — Un *nombre complexe* est un couple $z = (x, y)$ de nombres réels ou, ce qui est équivalent, un point du plan cartésien Oxy (d'après le nom de R. DESCARTES, mathématicien, physicien et philosophe français, 1596–1650). On appelle *plan complexe* et on note $\mathbb{C}$ l'ensemble de ces nombres. Deux complexes $z = (x, y)$ et $z' = (x', y')$ sont égaux et l'on note $z = z'$ si et seulement si $x = x'$ et $y = y'$.

Définissons les opérations d'*addition* et de *multiplication par un réel* :

$$(x, y) + (x', y') := (x + x', y + y'), \qquad (x, y)a := a(x, y) := (xa, ya).$$

Le quadrilatère de sommets $O = (0, 0)$, (x, y), $(x + x', y + y')$ et (x', y') est un *parallélogramme.*

Nous noterons

$$\mathbf{1} := (1, 0) \quad \text{et} \quad i := (0, 1)$$

(la notation « := » indique qu'on définit le membre de gauche, et qu'il ne faut donc pas chercher à déduire ces égalités en fonction de ce qui précède), et omettrons en réalité de noter **1** explicitement :

$$(x, y) = \mathbf{1}x + iy = x + iy\ ;$$

l'écriture de $x+iy$ est la *forme cartésienne* de ce nombre. Voilà donc i défini ! Les nombres réels x et y sont respectivement les *parties réelle et imaginaire* de z et nous les noterons $x =: \Re(z)$ et $y =: \Im(z)$. Ce sont les coordonnées du point du plan identifié à z.

Le complexe *conjugué* de z est $\bar{z} := x - iy$; c'est le point symétrique de z par rapport à l'axe des abscisses.

On peut de plus définir une opération de *multiplication* entre complexes, telle que

$$i^2 = -1,$$

puisqu'en utilisant cette égalité ainsi que les mêmes règles de calcul que dans $\mathbb{R}$, on trouve :

$$\begin{aligned} zz' &= (x+iy)(x'+iy') \\ &= x(x'+iy') + iy(x'+iy') \quad \text{(distributivité)} \\ &= xx' + xiy' + iyx' + iyiy' \quad \text{(distributivité)} \\ &= xx' + ixy' + iyx' - yy' \quad \text{(commutativité de la multiplication et } i^2 = -1\text{)} \\ &= xx' - yy' + ixy' + iyx' \quad \text{(commutativité de l'addition)} \\ &= xx' - yy' + i(xy' + x'y) \quad \text{(distributivité)} \end{aligned}$$

(il n'est pas essentiel de mémoriser ce résultat, mais il faut savoir refaire le calcul rapidement). On constate alors que le produit

$$z\bar{z} = \bar{z}z = x^2 + y^2$$

est un nombre réel positif, en l'occurence le carré de la distance du point z à l'origine du plan, et on appelle *module* de z le réel ≥ 0

$$|z| = \sqrt{z\bar{z}}.$$

Proposition. — Tout complexe $z \neq 0$ admet un inverse, c'est-à-dire un complexe z^{-1} tel que $z^{-1}z = 1$, et cet inverse est unique.

Démonstration – D'après le calcul précédent, le nombre

$$z^{-1} = \frac{\bar{z}}{|z|^2}$$

est un inverse de z.

De plus, si w est un inverse quelconque de z : $wz = 1$, il s'avère en multipliant les deux membres par z^{-1} que $w = z^{-1}$, ce qui démontre l'unicité de l'inverse. □

Exercice. — Quelle sont les parties réelle et imaginaire, le module, le conjugué et l'inverse de $z = (1+i) + i(1-i)$, de $z' = (1+i)(2-i)$ et de $z'' = \frac{1+i}{1-i}$?

Solution. — On a $z = 1 + i + i(1-i) = 2 + 2i$, donc $\Re z = \Im(z) = 2$. Attention à ne pas commettre l'erreur de penser que $\Re z = 1 + i$ et $\Im z = 1 - i$; $\Re z$ et $\Im z$, par définition, sont réels. Il vient alors $|z| = \sqrt{2^2 + 2^2} = 2\sqrt{2}$, $\bar{z} = 2 - 2i$ et $z^{-1} = \frac{1}{4}(1-i)$.

De plus,

$$\begin{aligned} z' &= (1+i)(2-i) \\ &= 1(2-i) + i(2-i) \\ &= (2-i) + (2i - i^2), \quad i^2 = -1 \\ &= 3 + i, \end{aligned}$$

donc $\Re z' = 3$, $\Im z' = 1$, $|z'| = \sqrt{10}$, $\bar{z}' = 3 - i$ et $z^{-1} = \frac{1}{10}(3-i)$.
Enfin,

$$z'' = \frac{(1+i)^2}{2} = i$$

donc $\Re z'' = 0$, $\Im z'' = 1$, $\bar{z}'' = -i$ et $z''^{-1} = -i$.

Exercice. — Montrer que la distance entre deux points $a, z \in \mathbb{C}$ vaut $|z - a|$. Quelle est l'équation dans $\mathbb{C}$ du cercle de centre $a \in \mathbb{C}$ et de rayon $r \geq 0$?
Solution. — Soient $z = x + iy$ et $a = u + iv$ les formes cartésiennes de z et a. On a

$$|z - a|^2 = (x - u)^2 + (y - v)^2,$$

ce qui, d'après le théorème de Pythagore est bien le carré de la distance entre a et z. L'équation du cercle est donc $|z - a| = r$.

1.2. Trigonométrie

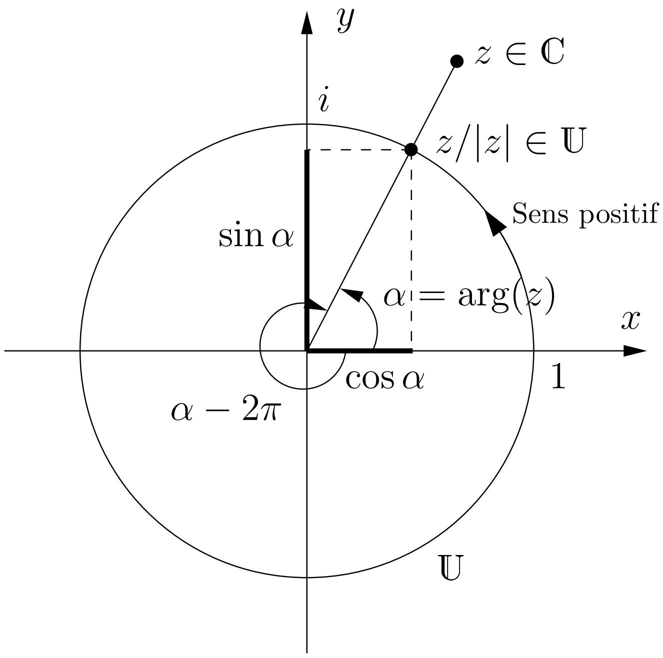
Nous allons écrire les nombres complexes sous une forme correspondant aux coordonnées polaires dans le plan.
Notons

$$\mathbb{U} := \{z \in \mathbb{C},\ |z| = 1\}$$

le *cercle trigonométrique* (en grec, *trigonon* = triangle, *metria* = science de la mesure).

Définitions. — L'*argument* d'un complexe z non nul, est l'angle orienté $\widehat{Ox, Oz}$; on le note $\arg(z)$. Rappelons qu'il s'agit de la longueur orientée d'un arc quelconque de $\mathbb{U}$ joignant 1 à $z/|z|$.
Rappelons que le cercle $\mathbb{U}$ a pour longueur 2π (c'est une définition de $\pi \simeq 3,14$ comme déjà le savant grec Archimède de Syracuse le savait trois siècles avant notre ère). Donc $\arg(z)$ n'est bien défini qu'à un multiple entier de 2π près — on dit *modulo* 2π. (On trouvera la nomenclature des lettres grecques, très utilisées en Mathématiques, à l'adresse `http://fr.wikipedia.org/wiki/Alphabet_grec`.) Par exemple, les arguments suivants sont tous égaux modulo 2π parce qu'ils correspondent au même point, -1, sur $\mathbb{U}$:

$$... = -3\pi = -\pi = \pi = 3\pi = ... \pmod{2\pi},$$



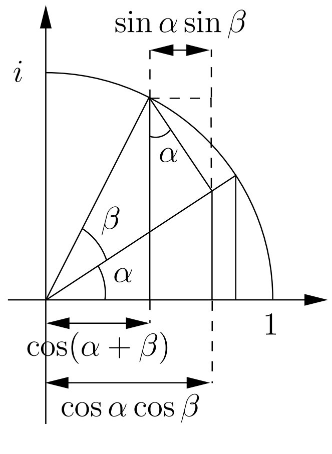
Rappel : formules d'addition de cos *et* sin. — Si $\alpha \in \mathbb{R}$, le *cosinus* et le *sinus* de α sont les coordonnées du point de $\mathbb{U}$ d'argument α (voir la figure ci-dessus) ; elle sont notées $\cos\alpha$ et $\sin\alpha$ et, en particulier, elles sont 2π-périodiques :

$$\cos(\alpha + 2\pi) = \cos\alpha, \quad \sin(\alpha + 2\pi) = \sin\alpha.$$

De plus, les formules d'addition suivantes :

$$\begin{cases} \cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta \\ \sin(\alpha + \beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta \end{cases}$$

se lisent sur le dessin suivant et son analogue pour le sinus :



On pourra comparer cette démonstration avec la devise mise en exergue par J. L. Lagrange (mathématicien et astronome européen, 1736–1813) dans sa *Mécanique analytique* : « On ne trouvera pas de figures dans cet ouvrage. Les méthodes que j'y expose ne demandent ni constructions, ni raisonnements géométrique ou mécanique, mais seulement des opérations algébriques assujetties à une marche régulière et uniforme. »

Exercice. — Retrouver les sinus et cosinus de π/n, avec $n = 1, 2, 3, 4, 6$.

Solution. — Par symétrie le point de $\mathbb{U}$ d'argument π (= la moitié de la longueur de $\mathbb{U}$) est -1, donc $\cos\pi = -1$ et $\sin\pi = 0$. De même, le point de $\mathbb{U}$ d'argument $\pi/2$ (= le quart de la longueur de $\mathbb{U}$) est i, donc $\cos\frac{\pi}{2} = 0$ et $\sin\frac{\pi}{2} = 1$.

Quant au point de $\mathbb{U}$ d'argument $\pi/4$, par symétrie il est sur la première diagonale, donc $x := \cos\frac{\pi}{4} = \sin\frac{\pi}{4}$ et, d'après le théorème de Pythagore, $2x^2 = 1$ donc $x = \frac{1}{\sqrt{2}}$.

Le triangle dont les sommets sont 0, 1 et le point de $\mathbb{U}$ dont l'argument est $\pi/3$ est isocèle. Comme la somme des angles d'un triangle vaut π, il est même équilatéral. Donc $\cos\frac{\pi}{3} = \frac{1}{2}$ et, d'après le théorème de Pythagore, $\sin\frac{\pi}{3} = \sqrt{3}/2$. Pour $\pi/6$, par symétrie (par rapport à la première diagonale) le sinus et le cosinus sont inversés : $\sin\frac{\pi}{3} = \frac{1}{2}$ et $\cos\frac{\pi}{3} = \sqrt{3}/2$.

Définition. — Rappelons que e est l'unique nombre réel dont le logarithme vaut 1 : $\ln e = 1$, $e \simeq 2.71828$. On suppose connue l'exponentielle d'un nombre réel.

Définissons l'exponentielle d'un nombre complexe $z = x + iy$ (avec $x, y \in \mathbb{R}$) par

$$e^z := e^x\,(\cos y + i\sin y)\ \ ;$$

si $z = x$, cette définition coïncide bien avec l'exponentielle réelle (et il n'y a donc pas conflit de définitions).

En particulier,

$$e^{iy} = \cos y + i\sin y$$

et

$$e^0 = 1, \quad e^{i2\pi} = 1, \quad e^{i\pi} + 1 = 0, \quad e^{i\pi/2} = i.$$

Propriété fondamentale. — $e^z\,e^{z'} = e^{z+z'}$

Démonstration –

$$\begin{aligned}
e^z\,e^{z'} &= e^x\,e^{x'}\,(\cos y + i\sin y)\,(\cos y' + i\sin y') \\
&\qquad \text{(définition de l'exponentielle complexe)} \\
&= e^x\,e^{x'}\,\left[(\cos y\cos y' - \sin y\sin y') + i(\sin y\cos y' + \cos y\sin y')\right] \\
&\qquad \text{(distributivité)} \\
&= e^{x+x'}\left[\cos(y+y') + i\sin(y+y')\right] \\
&\qquad \text{(formules d'addition de l'exponentielle réelle, cos et sin)} \\
&= e^{(x+x')+i(y+y')} = e^{z+z'}.
\end{aligned}$$

□

Corollaire immédiat. — Pour tout $z \in \mathbb{C}$, l'inverse de e^z est e^{-z}.

De plus, comme avec les réels on peut définir la puissance entière d'un complexe quelconque de la façon suivante : on pose $w^0 = 1$ (conventionnellement ; avec l'exception que certains mathématiciens préfèrent ne pas définir 0^0), $w^1 = w$, et supposant w^n connu pour un certain entier $n \geq 1$, on définit $w^{n+1} = w\,w^n$. Ceci permet de calculer w^n pour un entier naturel n arbitraire, en n étapes. On pose de plus $w^{-n} = (w^{-1})^n$, où w^{-1} est l'inverse précédemment introduit de w. La propriété fondamentale implique que, pour tout $n \in \mathbb{Z}$,

$$(e^z)^n = e^{nz}.$$

Démonstration de cette formule par récurrence – Elle est vraie pour $n = 0$ et pour $n = 1$. Supposons l'avoir démontrée à un rang $n \geq 1$ et déduisons-en la formule au rang suivant. Avec la propriété fondamentale on voit alors que

$$(e^z)^{n+1} = (e^z)^n \, e^z = e^{nz} e^z = e^{(n+1)z},$$

ce qui est la formule au rang suivant. Donc la formule est vraie pour tout $n \geq 0$.

Remarque : Ceci est un archétype de démonstration par récurrence. La dernière affirmation de la démonstration vient du fait que si au contraire la propriété n'était pas vraie pour tout entier $n \geq 2$, en considérant le plus petit entier n pour laquelle elle est violée, l'induction ci-dessus appliquée au rang $n-1$ aboutirait à une absurdité.

Il reste à démontrer la formule pour $n < 0$. Mais ceci est immédiat, puisqu'alors en posant $m = -n \geq 0$ on obtient

$$(e^z)^n = (e^{-z})^m = e^{-mz} = e^{nz}.$$

□

Exercice. — Que penser de cette suite d'égalités :

$$-1 = (-1)^{\frac{2}{2}} = ((-1)^2)^{\frac{1}{2}} = 1^{\frac{1}{2}} = 1 \ ?$$

Solution. — La deuxième égalité ne rentre pas dans le cadre de ce qui précède parce que l'exposant $1/2$ n'est pas entier. On verra dans le paragraphe § 1.3 quel sens donner à la « fonction racine n-ième » d'un nombre complexe ; dans ce sens la dernière égalité ci-dessus est fausse parce qu'elle choisit précisément, parmi les *deux* racines carrées de 1 (1 et -1), celle qu'il ne faut pas.

Exercice. — Que penser de cette suite d'égalités :

$$e^{i\alpha} = \left(e^{i\alpha}\right)^{\frac{2\pi}{2\pi}} = \left(e^{i2\pi}\right)^{\frac{\alpha}{2\pi}} = 1^{\frac{\alpha}{2\pi}} = 1 \ ?$$

Solution. — La deuxième égalité est fausse (et c'est la seule !). La propriété $(x^a)^b = x^{ab}$, vraie pour des réels > 0, se démontre avec la fonction logarithme. Mais la fonction logarithme ne jouit pas dans le complexe de propriétés aussi simples que dans le réel (c'est une fonction *multivaluée*, comme arg) et la propriété correspondante est fausse, comme les égalités précédentes le montrent par l'absurde. C'est le concept de *surface de Riemann*, du nom de B. Riemann (mathématicien allemand, 1826–1866) en analyse complexe, qui permet d'élucider totalement ces questions.

Lemme. — Soient ρ et θ deux réels tels que $\rho > 0$, et $z = \rho e^{i\theta}$. Alors

$$|z| = \rho \quad \text{et} \quad \arg z = \theta \pmod{2\pi}.$$

Démonstration – On a

$$|z|^2 = z\bar{z} = \rho^2 e^{i(\theta-\theta)} = \rho^2$$

donc, comme $|z|$ et ρ sont tous deux ≥ 0, $|z| = \rho$. Ensuite, par définition l'argument de z est l'argument de $z/|z| = e^{i\theta}$, soit $\theta \pmod{2\pi}$. □

N.B. : Si on avait $\rho < 0$, alors on aurait

$$z = (-\rho)e^{i(\pi+\theta)}$$

donc, d'après le lemme, $|z| = -\rho$ et $\arg(z) = \pi + \theta \pmod{2\pi}$.

Un corollaire important du lemme est que pour tout complexe $z \neq 0$ on a

$$z = |z|\, e^{i \arg(z)} \ ;$$

c'est la *forme polaire* de z.

Corollaire. — Le produit de deux complexes non nuls z et z' a pour module $|z|\,|z'|$ et pour argument $\arg z + \arg z'$.

Démonstration – Avec des notations polaires évidentes,

$$zz' = |z|e^{i\theta}\,|z'|e^{i\theta'} = |z|\,|z'|\,e^{i(\theta+\theta')}.$$

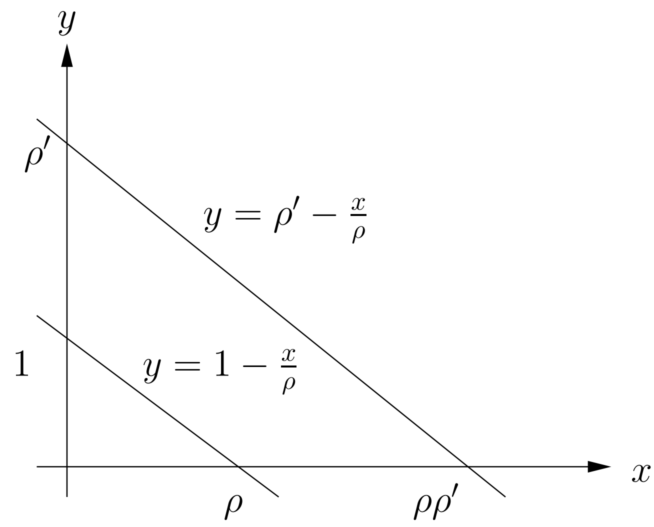
□

Exercice. — Étant donné deux complexes z et z', construire leur produit zz' à la règle et au compas (c'est-à-dire avec des instruments permettant de tracer des droites et des cercles de centre et de rayon donnés).

Solution. — On peut supposer z et z' non nuls, sans quoi la construction est triviale. Notons $z = \rho e^{i\alpha}$ et $z' = \rho' e^{i\alpha'}$, avec $\rho, \rho' > 0$. D'après le corollaire § 1.2, zz' est le complexe de module $\rho\rho'$ et d'argument $\alpha + \alpha'$.

– La construction de l'angle $\alpha + \alpha'$ est simple et est laissée à la sagacité du lecteur.

– La construction de la longueur $\rho\rho'$ est résumée par la figure suivante :



Sans calcul, on peut aussi invoquer le théorème de Thalès [**Gra97**, p. 76], attribué à Thalès de Milet, mathématicien grec vers -600 .

Exercice. — Démontrer la *formule de Moivre* (A. de Moivre, mathématicien français, 1667–1754), où $\theta \in \mathbb{R}$:

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

Solution. — Si l'on pose $z = e^{i\theta}$,

$$(\cos\theta + i\sin\theta)^n = z^n = e^{in\theta} = \cos n\theta + i\sin n\theta.$$

Formules d'Euler. — $\cos\theta = \dfrac{e^{i\theta}+e^{-i\theta}}{2}$ et $\sin\theta = \dfrac{e^{i\theta}-e^{-i\theta}}{2i}$.

Démonstration – On a

$$\begin{cases} e^{i\theta} &= \cos\theta + i\sin\theta \\ e^{-i\theta} &= \cos\theta - i\sin\theta. \end{cases}$$

Ces deux équations peuvent être vues comme un système de deux équations linéaires à deux inconnues, $\cos\theta$ et $\sin\theta$. En faisant la somme des deux équations on élimine $\sin\theta$ et on obtient l'expression voulue de $\cos\theta$; en faisant la différence des deux équations on élimine $\cos\theta$ et on obtient l'expression voulue de $\sin\theta$. □

Ces formules permettent de retrouver facilement de nombreuses identités trigonométriques.

Exercice. — Trouver les coefficients $a, b, c \in \mathbb{R}$ tels que

$$\cos^3\theta = a\cos 3\theta + b\cos\theta + c$$

pour tout $\theta \in \mathbb{R}$; le membre de droite, combinaison linéaire de 1, $\cos\theta$ et $\cos 3\theta$ (i.e. ces quantités interviennent à la puissance un), s'appelle la *linéarisation* de $\cos^3\theta$.

Solution. — Posons $z = e^{i\theta}$. Alors

$$\begin{aligned} \cos^3\theta &= \left(\frac{z+z^{-1}}{2}\right)^3 \\ &= \frac{1}{8}\left(z^3+3z+3z^{-1}+z^{-3}\right) \\ &= \frac{1}{4}\cos 3\theta + \frac{3}{4}\cos\theta \quad \text{(à justifier).} \end{aligned}$$

Pour élever $e^{i\theta} \pm e^{-i\theta}$ à une puissance élevée, on aura intérêt, plutôt qu'à développer la puissance pas à pas, à utiliser la très utile formule du binôme de NEWTON (§ 1.4).

Exercice. — Calculer la dérivée de la fonction $f : \mathbb{R} \to \mathbb{C}$, $t \mapsto e^{i\omega t}$, où ω est un paramètre réel. En déduire une solution de l'équation différentielle $y'' + \omega^2 y = 0$.

Solution. — On a

$$f'(t) = (\cos\omega t + i\sin\omega t)' = \omega\left(-\sin\omega t + i\cos\omega t\right) = i\omega f(t).$$

En dérivant une fois de plus, on voit que f est solution de l'équation différentielle donnée.

1.3. Factorisation des polynômes

Définitions. — Un *polynôme* est une expression de la forme

$$P(z) = a_n z^n + a_{n-1}z^{n-1} + ... + a_1 z + a_0,$$

où les coefficients a_j sont des nombres complexes. Si $a_n \neq 0$, l'entier n s'appelle le *degré* de P. Une *racine* de P est un complexe a tel que $P(a) = 0$.

Exercice. — Montrer qu'il existe un polynôme T_3 de degré trois tel que $\cos(3\theta) = T_3(\cos\theta)$. (Plus généralement, on appelle *n-ième polynôme de Tchebytchev* le polynôme T_n tel que $\cos n\theta = T_n(\cos\theta)$.)

Solution. — On a calculé dans un exercice précédent que

$$\cos^3\theta = \frac{1}{4}\cos 3\theta + \frac{3}{4}\cos\theta.$$

Donc

$$\cos 3\theta = 4\cos^3\theta - 3\cos\theta,$$

et le polynôme

$$P(z) = 4z^3 - 3$$

convient (on peut d'ailleurs montrer que c'est le seul).

. — S'il existe un polynôme $Q(z)$ et un complexe a tels que $P(z) = (z-a)Q(z)$, alors a est évidemment racine de P. La réciproque est vraie :

Lemme admis. — Si a est une racine du polynôme $P(z)$, il existe un polynôme $Q(z)$ tel que $P(z) = (z-a)Q(z)$.

Ce lemme sera démontré dans un cours d'algèbre ultérieur avec l'algorithme de division euclidienne des polynômes.

La *multiplicité* de la racine a est le plus grand entier m tel qu'il existe un polynôme $R(z)$ tel que $P(z) = (z-a)^m R(z)$.

Exercice. — Notons $j = e^{i2\pi/3}$. Montrer que $1+j+j^2 = 0$; pour cela on pourra calculer $(1-j)(1+j+j^2)$.

Solution. — On a $(1-j)(1+j+j^2) = 1-j^3 = 0$. Comme $j \neq 1$, forcément $1+j+j^2 = 0$.

Définition. — S'il est difficile de déterminer les racines d'un polynôme en général, certains polynômes particulièrement simples sont bien compris et jouent un rôle important. C'est le cas des polynômes de la forme $z^n - 1$.

Une *racine n-ième de l'unité* est un complexe z tel que $z^n = 1$.

Injectons la forme polaire de z : $z = r\,e^{i\theta}$ dans l'équation :

$$r^n\,e^{in\theta} = 1.$$

En égalant module et argument, on voit que $r^n = 1$ (équation analogue à l'équation de départ, mais maintenant dans $\mathbb{R}^+$!) et que $n\theta = 0 \pmod{2\pi}$. L'équation portant sur r admet une unique solution réelle positive, $r = 1$. Celle portant sur θ équivaut à ce qu'il existe $k \in \mathbb{Z}$ tel que

$$n\theta = k2\pi,$$

soit

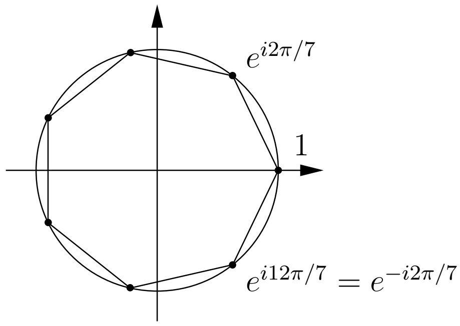
$$\theta = k\frac{2\pi}{n}.$$

On obtient ainsi une infinité de solutions $(1, k2\pi/n)$, $k \in \mathbb{Z}$. Mais beaucoup de ces solutions correspondent en réalité au même complexe z, puisque $re^{i\theta}$ est 2π-périodique par rapport à θ. Précisément, deux solutions sont égales modulo 2π : $k2\pi/n = l2\pi/n + m2\pi$ (pour un certain $m \in \mathbb{Z}$) si et seulement si $k-l = mn$ (pour un certain m), i.e. $k-l$ est multiple de n. Donc on obtient toutes les solutions une fois et une seule en laissant k prendre les valeurs $0, ..., n-1$ (au-delà, k est forcément égal à une valeur déjà prise modulo n).

Conclusion. — Les racines n-ièmes de l'unité sont les n nombres de la forme

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

Géométriquement, ce sont les point du cercle trigonométrique obtenus à partir de 1 par rotations successives de $2\pi/n$, c'est-à-dire les sommets du polygone régulier à n côtés, comme le représente la figure suivante dans le cas particulier $n = 7$.



En particulier, pour tout entier $n \geq 0$ le polynôme $z^n - 1$ admet n racines complexes, alors qu'il admet au plus deux racines réelles (parmi 1 et -1, puisque les racines sont de module unité).

Un théorème remarquable affirme qu'il en est de même de *tous* les polynômes de même degré.

Théorème fondamental de l'algèbre (admis). — Tout polynôme $P(z) = a_n z^n + a_{n-1} z^{n-1} + ... + a_1 z + a_0$ de degré $n > 0$ admet n racines complexes (comptées avec leur multiplicité). Autrement dit, P s'écrit

$$P(z) = a_n (z - \alpha_1)^{m_1} \cdots (z - \alpha_p)^{m_p},$$

où les α_j sont les racines de P, où m_j est la multiplicité de α_j, et où $m_1 + \cdots + m_p = n$.

Il existe de nombreuses démonstrations de ce théorème (cf. `http://fr.wikipedia.org/wiki/Theoreme_` mais il est frappant que toutes utilisent au moins un outil analytique et donc qu'aucune n'est purement algébrique.

Exercice. — Trouver les racines n-ièmes de -2.

Solution. — On cherche les racines n-ièmes de -2 sous leur forme polaire $re^{i\theta}$: $r^n e^{in\theta} = 2e^{i\pi}$, i.e. $r = 2^{1/n}$ et $\theta = \pi/n + k2\pi/n$, avec $k \in \mathbb{Z}$. En prenant $k \in \{0, ..., n-1\}$ on obtient déjà toutes les solutions :

$$z = 2^{1/n} e^{i\pi(1+2k)/n}, \quad 0 \leq k \leq n-1.$$

Les tels z s'obtiennent à partir des racines n-ièmes de l'unité par rotation de π puis par homothétie d'un facteur $2^{1/n}$.

Exercice. — Montrer que le polynôme de degré 2

$$P(z) = az^2 + bz + c$$

a pour racines

$$z_{\pm} = \frac{1}{2a} \left(-b \pm \delta\right),$$

où δ est l'une quelconque des deux racines opposées du *discriminant* $\Delta := b^2 - 4ac$.

Solution. — La méthode pour trouver les racines complexes est la même que pour les racines réelles. On commence par mettre le polynôme sous forme canonique :

$$\begin{aligned} P(z) &= a\left[z^2 + \frac{b}{a}z + \frac{c}{a}\right] \\ &= a\left[\left(z + \frac{b}{2a}\right)^2 - \frac{\Delta}{4a^2}\right], \quad \Delta = (\pm\delta)^2. \end{aligned}$$

À ce stade, la discussion est simplifiée par rapport au cas où l'on cherche les solutions réelles d'une équation à coefficients réels, puisque Δ possède toujours deux racines carrées complexes $(\pm\delta)$. L'identité remarquable $a^2 - b^2 = (a-b)(a+b)$ permet alors de factoriser $P(z)$:

$$P(z) = a(z - z_-)(z - z_+),$$

ce qui montre que l'ensemble des racines de P est $\{z_-, z_+\}$.

Exercice. — Calculer

$$S(\theta) := \sum_{0 \le k \le n} \cos k\theta := 1 + \cos\theta + \cos 2\theta + ... + \cos n\theta.$$

Comment faut-il choisir θ pour que $S(\theta)$ soit nul ? Expliquer géométriquement pourquoi.

Solution. — En remarquant que

$$\sum \cos k\theta = \Re \sum e^{ik\theta} = \Re \sum (e^{i\theta})^k,$$

on est ramené à utiliser l'identité

$$\sum_{0 \le k \le n} r^n = \frac{1 - r^{n+1}}{1 - r}$$

vérifiée pour tout complexe $r \neq 1$ (il suffit de multiplier à gauche et à droite par $1 - r$ pour le voir).

On peut en effet écarter le cas exceptionnel où $e^{i\theta} = 1$, i.e. $\theta = 0 \pmod{2\pi}$, qui est explicite : $S(0) = n$. (*i.e.* est l'abréviation de la locution latine *id est*, qui signifie *c'est-à-dire* tout en étant plus rapide à écrire. On l'utilise entre deux propriétés pour signifier qu'elles sont équivalentes l'une à l'autre.)

Si donc $e^{i\theta} \neq 1$,

$$S(\theta) = \frac{1 - e^{i(n+1)\theta}}{1 - e^{i\theta}}.$$

Cette somme est nulle si et seulement si $e^{i(n+1)\theta} = 1$, i.e. $(n+1)\theta = 2\pi \pmod{2\pi}$.

Exercice. — Soit P un polynôme réel. Montrer que si λ est une racine de P, il en est de même de $\bar{\lambda}$. En déduire que si P est de degré trois, il possède nécessairement une racine réelle.

Solution. — Si

$$P(\lambda) = a_n\lambda^n + \cdots + a_0 = 0,$$

alors

$$P(\bar{\lambda}) = \overline{a_n\lambda^n + \cdots + a_0} = a_n\bar{\lambda}^n + \cdots + a_1\bar{\lambda} + a_0 = 0.$$

Si P est de degré 3, il possède trois racines :

$$P(z) = a_n(z - \lambda_1)(z - \lambda_2)(z - \lambda_3), \quad \lambda_p \in \mathbb{C}.$$

Supposons pas l'absurde que P ne possède pas de racine réelle : pour $j = 1, 2, 3$, $\lambda_j = \alpha_j + i\beta_j$, avec $\beta_j \neq 0$. D'après ce qui précède, il existe j tel que $\bar{\lambda}_1 = \lambda_j$. Comme $\beta_1 \neq 0$, $j \neq 1$. Donc $j = 2$ ou 3. Quitte à renuméroter les racines, supposons que $j = 2$. Mais alors le terme constant (par exemple) du polynôme réel

$$\frac{P(z)}{a_3} = (z - \lambda_1)(z - \lambda_2)(z - \lambda_3)$$

vaut $\lambda_1 \bar{\lambda}_1 \lambda_3 = |\lambda_1|^2 \lambda_3$ et n'est donc pas réel, alors que P est supposé réel, ce qui est absurde.

1.4. Similitudes

Définition de transformations élémentaires. — La *translation de vecteur* $v \in \mathbb{C}$ est l'application

$$\tau_v : \mathbb{C} \to \mathbb{C}, \quad z \mapsto z + v.$$

(Ici, *vecteur* est utilisé comme synonyme de nombre complexe, ce qui est justifié par la généralisation à venir aux translations dans $\mathbb{R}^3$.)

La *rotation de centre l'origine et d'angle* $\alpha \in \mathbb{R}$ est l'application

$$\rho_{0,\alpha} : \mathbb{C} \to \mathbb{C}, \quad z \mapsto z\, e^{i\alpha} \;;$$

autrement dit, $\rho_{a,\alpha}(z)$ est le complexe z' tel que $|z'| = |z|$ et, si $z \neq 0$, $\arg z'/z = \alpha \pmod{2\pi}$.

L'*homothétie de centre l'origine et de rapport* $\kappa \in \mathbb{R} \setminus \{0\}$ est l'application

$$\eta_{0,\kappa} : \mathbb{C} \to \mathbb{C}, \quad z \mapsto \kappa z \;;$$

autrement dit, $\eta_{0,\kappa}(z)$ est le complexe z' tel que $\arg z = \arg z' \pmod{\pi}$ et l'abscisse de z' le long de l'axe Oz muni de l'unité z vaut κ.

Exercice. — Identifier $\tau_v \circ \tau_{v'}$, $\rho_{0,\alpha} \circ \rho_{0,\alpha'}$ et $\eta_{0,\kappa} \circ \eta_{0,\kappa'}$.

Solution. — On a $\tau_v \circ \tau_{v'}(z) = z + v + v'$ donc $\tau_v \circ \tau_{v'} = \tau_{v+v'}$, et de même $\rho_{0,\alpha} \circ \rho_{0,\alpha'} = \rho_{0,\alpha+\alpha'}$ et $\eta_{0,\kappa} \circ \eta_{0,\kappa'} = \eta_{0,\kappa\kappa'}$.

Définitions complémentaires. — Pour définir la *rotation de centre* $a \in \mathbb{C}$ *et d'angle* $\alpha \in \mathbb{R}$, notons z' l'image du nombre complexe z par une telle rotation. Pour se ramener à l'origine il suffit de poser $Z = z - a$ et $Z' = z' - a$. On dit que la transformation $Z \mapsto Z'$ est *conjuguée* à $z \mapsto z'$ par la translation $z \mapsto Z$:

$$\begin{array}{ccc} \mathbb{C}, z & \xrightarrow{\rho_{a,\alpha}} & \mathbb{C}, z' \\ {\scriptstyle +a}\big\uparrow & & \big\uparrow{\scriptstyle +a} \\ \mathbb{C}, Z & \xrightarrow{\rho_{0,\alpha}} & \mathbb{C}, Z' \end{array}$$

(la *conjugaison* dans ce sens ne doit pas être confondue avec la conjugaison des nombres complexes). D'après la définition précédente, on a $Z' = Z\, e^{i\alpha}$, soit $z' = a + (z - a)\, e^{i\alpha}$:

$$\rho_{a,\alpha} : \mathbb{C} \to \mathbb{C}, \quad z \mapsto a + (z - a)\, e^{i\alpha}.$$

On vérifie que par exemple $\rho_{a,\alpha}(a) = a$: a est un point fixe de $\rho_{a,\alpha}$.

De même, l'*homothétie de centre* $a \in C$ *et de rapport* $\kappa \in \mathbb{R} \setminus \{0\}$ est

$$\eta_{a,\kappa} : \mathbb{C} \to \mathbb{C}, \quad z \mapsto a + \kappa(z - a).$$

Il est évident que la composée de deux rotations (par exemple) de même centre est encore une rotation. Mais si les deux rotations n'ont pas même centre, le résultat est beaucoup moins simple à imaginer ; heureusement le calcul permet de dépasser l'intuition.

Exercice. — Montrer qu'une transformation de $\mathbb{C}$ de la forme $f(z) = a + ze^{i\alpha}$ est une rotation ou une translation ; l'identifier en fonction de a et de α. En déduire que la composée de deux rotations d'angles α et β est :
– une translation si $\alpha + \beta = 0 \pmod{2\pi}$
– une rotation sinon.
Quel est l'affirmation analogue pour des homothéties ?

Solution. — Si $\alpha = 0 \pmod{2\pi}$, $f = \tau_a$. Sinon, l'équation $f(z) = z$ possède un unique point fixe $A = \frac{a}{1-e^{i\alpha}}$ et l'on voit que

$$f(z) = A\left(1 - e^{i\alpha}\right) + ze^{i\alpha} = A + (z - A)e^{i\alpha}$$

est la rotation $\rho_{A,\alpha}$.
Donc la composition

$$\begin{aligned}\rho_{a,\alpha} \circ \rho_{b,\beta}(z) &= a + [\rho_{b,\beta}(z) - a]\, e^{i\alpha} \\ &= a + \Big[b + (z - b)e^{i\beta} - a\Big]\, e^{i\alpha} \\ &= a + (b - a)e^{i\alpha} - be^{i(\alpha+\beta)} + ze^{i(\alpha+\beta)}\ ;\end{aligned}$$

est :
– la translation de vecteur $a + (b - a)e^{i\alpha} - be^{i\beta}$ si $\alpha + \beta = 0 \pmod{2\pi}$
– la rotation d'angle $\alpha+\beta$ et de centre $A = (a+(b-a)e^{i\alpha} - be^{i\beta})(1-e^{i(\alpha+\beta)})^{-1}$ sinon.
Remarquons que quand $\alpha + \beta$ tend vers $0 \pmod{2\pi}$, alors $|A|$ tend vers l'infini, ce qui permet de voir les translations comme des « rotations autour de l'infini ».

De même (remplacer $e^{i\alpha}$ et $e^{i\beta}$, dans le calcul précédent, par des réels non nuls κ et λ) la composition de deux homothéties de rapports κ et λ est :
– la translation de vecteur $a + (b - a)\kappa - b\lambda$ si $\kappa\lambda = 1$
– l'homothétie de rapport $\kappa\lambda$ et de centre $A = (a + (b - a)\kappa - b\lambda)(1 - \kappa\lambda)^{-1}$ sinon.

Des transformations plus générales. — Une transformation de $\mathbb{C}$ de la forme $z \mapsto az + b$ ou $a\bar{z} + b$, avec $a \neq 0$, est une *similitude*. Si elle est de la forme $az + b$, elle est *directe*, et sinon elle est *indirecte*.

Exercice. — Montrer que la composée de deux similitudes est une similitude.

Solution. — La composée de deux similitudes directes $z \mapsto az + b$ et de $z \mapsto a'z + b'$ (avec $a, a' \neq 0$) est

$$z \mapsto a(a'z + b') + b = aa'z + (ab' + b),$$

qui est une similitude directe. De même, la composée de deux similitudes indirectes est une similitude directe et la composée de deux similitudes de types opposés est une similitude indirecte (autrement dit les similitudes suivent la règle des signes).

Exercice. — Identifier les transformations suivantes :

$$z \mapsto z' = iz + 2, \quad z \mapsto z'' = 2\bar{z} + 1.$$

Solution. — Les points fixes $a \in \mathbb{C}$ de la première transformation vérifient $a = ia + 2$, soit $a = 1 + i$. Donc la première transformation s'écrit

$$z' = a + i(z - a) \ ;$$

c'est une rotation de centre a et d'angle $\pi/2$.

La seconde transformation est une similitude indirecte : s'obtient par composition de la symétrie par rapport à l'axe des x, $z \mapsto \bar{z}$, et de l'homothétie de centre $a = -1$ et de rapport 2, i.e. $z \mapsto 2(z + 1) - 1$.

Exercice. — Des complexes $a, b, c, d \in \mathbb{C}$ étant donnés, soit la fonction

$$\varphi : z \mapsto \frac{az + b}{cz + d}$$

définie pour $cz + d \neq 0$. Supposons $ad - bc \neq 0$ et $c \neq 0$. Montrer que la fonction φ est une bijection de $\mathbb{C} \setminus \{-d/c\}$ sur $\mathbb{C} \setminus \{a/c\}$ (i.e. tout élément de l'ensemble image possède un unique antécédent), et que sa bijection réciproque est

$$\varphi^{-1} : w \mapsto \frac{dw - b}{-cw + a}.$$

La fonction φ s'appelle une *homographie* en raison de la similarité entre elle-même et son inverse.

Solution. — Il suffit de résoudre l'équation

$$w = \frac{az + b}{cz + d}.$$

N.B. : On note $\bar{\mathbb{C}}$ la réunion $\mathbb{C} \cup \{\infty\}$ de $\mathbb{C}$ et d'un élément quelconque $\notin \mathbb{C}$ noté ∞ et appelé *infini*. On prolonge φ en une fonction $\bar{\mathbb{C}} \to \bar{\mathbb{C}}$ en posant $\varphi(-d/c) = \infty$ et $\varphi(\infty) = a/c$. On montre alors que la fonction obtenue est continue, dans un sens qui sera justifié dans le cours d'analyse.

Exercice. — Soit $f : \mathbb{C} \to \mathbb{C}$, $z \mapsto z^2$. Combien un point $w \in \mathbb{C}$ a-t-il d'antécédents par f ? Quelle est l'image par f d'une ellipse centrée en l'origine ?

Solution. — Tout complexe non nul possède deux antécédents par f : si $re^{i\theta}$ est sa forme polaire, ses deux antécédents sont $\sqrt{r}e^{i\theta}$ et $\sqrt{r}e^{i(\theta+\pi)}$. En revanche, 0 possède un unique antécédent, à savoir lui-même.

Par ailleurs, une ellipse centrée à l'origine admet un paramétrage de la forme

$$z = e^{i\alpha} \left(a \cos\theta + b \sin\theta\right), \quad 0 \leq \theta < 2\pi$$

(voir [**Gra97**]). Son image par f admet le paramétrage suivant :

$$\begin{aligned} z^2 &= e^{i\alpha} \left(a \cos\theta + b \sin\theta\right)^2 \\ &= e^{i\alpha} \left(\frac{a^2 - b^2}{2} + \frac{a^2 - b^2}{2} \cos 2\theta + iab \sin 2\theta\right), \end{aligned}$$

dont l'image est bien une ellipse, centrée au point $e^{i\alpha}\dfrac{a^2 - b^2}{2}$, de demi axes $\frac{|a^2-b^2|}{2}$ et ab.

Compléments

Un calcul audacieux. — Soit à résoudre

(1) $$x^3 - 15x - 4 = 0.$$

Cherchons une solution sous la forme $x = u + v$:

$$(u+v)^3 - 15(u+v) - 4 = 0,$$

ce qu'on peut écrire

$$u^3 + v^3 - 4 + 3(u+v)(uv-5) = 0.$$

Cette équation est satisfaite en particulier si

$$u^3 + v^3 = 4 \quad \text{et} \quad u^3v^3 = 5^3 = 125,$$

c'est-à-dire si u^3 et v^3 sont solutions de l'équation auxiliaire

$$(w - u^3)(w - v^3) = w^2 - (u^3 + v^3)w + u^3v^3 = w^2 - 4w + 125 = 0,$$

soit, si on essaye de ne faire apparaître w qu'une seule fois,

$$(w-2)^2 + 11^2 = 0.$$

Pour ensuite utiliser l'identité $a^2 - b^2 = (a-b)(a+b)$, il serait commode de disposer d'un nombre dont le carré vaut -1 puisqu'alors, en notant $\sqrt{-1}$ ce nombre, l'équation auxiliaire deviendrait

$$(w-2)^2 - (11\sqrt{-1})^2 = (w - 2 + 11\sqrt{-1})(w - 2 - 11\sqrt{-1}) = 0,$$

et ses solutions seraient

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

Mais alors $u = 2 - \sqrt{-1}$ et $v = 2 + \sqrt{-1}$ conviendraient puisque

$$(2 - \sqrt{-1})^3 = 2 - 11\sqrt{-1} \quad \text{et} \quad (2 + \sqrt{-1})^3 = 2 + 11\sqrt{-1},$$

et une solution de l'équation initiale serait donc

$$x = u + v = 4.$$

Il est maintenant aisé de vérifier que le nombre $x = 4$ est bel et bien une solution de l'équation voulue ! Il a cependant fallu une audace intellectuelle immense aux algébristes italiens pour imaginer un tel calcul reposant de façon cruciale sur la manipulation d'un nombre imaginaire $\sqrt{-1}$, sorti d'on ne sait où.

Exercice. — Dans le calcul du § 1.4, quelles sont les étapes où l'on ajoute des solutions ? Où l'on en perd ? Pourquoi ce calcul garde-t-il tout son intérêt pour trouver les solutions de l'équation initiale ?

Solution. — Le fait de chercher une solution a priori sous la forme $x = u + v$ n'est pas une véritable restriction parce que tout nombre réel peut toujours s'écrire comme la somme de deux nombres réels ($x = x + 0$ par exemple). En revanche, le fait de remplacer l'équation

$$u^3 + v^3 - 4 + 3(u+v)(uv-5) = 0$$

par la paire d'équations

$$u^3 + v^3 = 4, \qquad u^3v^3 = 5^3 = 125,$$

fait perdre a priori des solutions (si $A = B = 0$, certainement $A + B = 0$, mais l'implication inverse n'est pas vraie, puisque par exemple $(1) + (-1) = 0$ alors que $1 \neq 0$ et $-1 \neq 0$). Il en est de même quand on passe de $u^3 = 2 - 11\sqrt{11}$ à $u = 2 - \sqrt{-1}$; un calcul direct permet en effet de vérifier que $(2 - \sqrt{-1})^3 = 2 - 11\sqrt{11}$ (en admettant que les règles de calcul des nombres réels s'appliquent à $\sqrt{-1}$), tandis que l'implication inverse n'est évidente (et le cours montrera plus loin qu'elle est fausse).

C'est donc calcul *par conditions suffisantes* : on passe d'une équation à la suivante en perdant des solutions sans jamais en ajouter. La solution $x = 4$ qu'on trouve à la fin est donc bien une solution de l'équation initiale, ce qui permet, par factorisation de $x - 4$, de se ramener à une équation de degré deux, qu'on sait résoudre plus facilement.

Mais qu'est-ce qu'un couple ?— Un *couple* $z = (x, y)$ peut être vu comme une application $z : \{1, 2\} \to \mathbb{R}$, exceptionnellement notée $z = (x, y)$ au lieu de $z : j \mapsto z(j)$. L'inconvénient de cette caractérisation est que, en théorie des ensembles, la définition même d'une application utilise le concept de couple...

Un *couple* peut être défini directement à partir de la notion d'ensemble. Rappelons d'abord qu'un ensemble à un ou deux éléments s'appelle respectivement un *singleton* ou une *paire*. Par exemple, $\{0, 0\} = \{0\}$ est un singleton et $\{0, 1\} = \{1, 0\}$ est une paire. On peut alors définir un *couple*, aussi appelé une *paire ordonnée*, par la formule $(a, b) := \{a, \{a, b\}\}$.

De façon analogue on définit les *triplets*, les *quadruplets*, ou, plus généralement, les n-uplets.

Exercice. — En utilisant la propriété de la relation d'appartenance $\in$ selon laquelle on n'a jamais $a \in x \in a$, vérifier la propriété que deux paires (a, b) et (a', b') sont égales si et seulement si $a = a'$ et $b = b'$ (voir [**Kri07**]).

Solution. — Le sens indirect de l'équivalence est immédiat (si $a = a'$ et $b = b'$, $(a, b) = (a', b')$). Réciproquement, supposons que $(a, b) = (a', b')$. Avec la définition ci-dessus, ceci nous dit

$$\{a, \{a, b\}\} = \{a', \{a', b'\}\},$$

donc a priori que :
- soit $a = a'$ et $\{a, b\} = \{a', b'\}$, donc aussi $b = b'$ (ce que l'on veut montrer)
- soit $a = \{a', b'\}$ et $a' = \{a, b\}$ (ce que l'on veut infirmer).

Supposons par l'absurde que $a = \{a', b'\}$ et $b = a'$. Alors

$$a' \in \{a', b'\} = a \in \{a, b\} = a,$$

ce qui est impossible effectivement.

L'exponentielle comme somme infinie. — La définition donnée de l'argument suppose connue la notion d'angle, donc de longueur d'arc. Ces notions sont familières depuis longtemps (l'angle se mesure en enroulant une ficelle le long du cercle trigonométrique puis en mesurant la longueur de la ficelle). Mais elles n'en sont pas moins *transcendantes*, au sens que leur définition rigoureuse exige un passage à la limite [**Gab01**, Partie C].

Dans les cours plus avancés de mathématiques (voir par exemple [**Car61**]), on trouve plus élégant de définir l'exponentielle directement comme une limite (dans un sens à justifier par l'analyse) :

$$e^z := \sum_{0 \leq n < \infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots,$$

puis d'en déduire les notions transcendantes liées aux angles, dont les fonctions sinus et cosinus, *définies* par l'égalité

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

Structure de corps. — Les opérations que nous avons définies sur $\mathbb{C}$ satisfont un certain nombres de propriétés qui, prises abstraitement, définissent une structure de *corps*. Ces propriétés sont :

— Concernant l'addition :
- commutativité : $a + b = b + a$ (sous-entendu : *pour tous* a, b)
- associativité : $a + (b + c) = (a + b) + c$ (ce qui permet d'ôter les parenthèses sans ambiguité sur le résultat)
- existence d'un zéro : $a + 0 = a$
- existence d'un opposé : $a + (-a) = 0$

— Concernant la multiplication :
- commutativité : $ab = ba$
- associativité : $a(bc) = (ab)c$

– existence d'un élément neutre : $a1 = a$
– existence d'un inverse pour tout élément non nul : $a\frac{\bar{a}}{|a|^2} = 1$

— Concernant l'addition et la multiplication :
– distributivité : $a(b + c) = ab + ac$

Exercice. — Pourquoi les multiplications successives suivantes ne font pas de $\mathbb{C}$ un corps (avec l'addition habituelle) ?
– $z \cdot z' = xx' + yy'$ (*produit scalaire*)
– $z \wedge z' = xy' - yx'$ (*produit vectoriel*)
– $z * z' = xx' + i\, yy'$ (*produit composante par composante*),

Solution. — Dans chacun des trois cas, il suffit de trouver un axiome dans la définition d'un corps, qui soit violé par la multiplication envisagée.

Le produit scalaire de deux nombres complexe n'est pas associatif, puisque

$$(1 \cdot i) \cdot i = 0 \neq 1 \cdot (i \cdot i) = 1.$$

Le produit vectoriel est anti-commutatif : $z \wedge z' = -z' \wedge z$.

Pour le produit composante par composante, l'élément neutre est $1 + i$ mais les réels ($x \in \mathbb{R}$) et les imaginaires purs (iy, avec $y \in \mathbb{R}$) n'ont pas d'inverse.

Exercice. — Les ensembles suivants sont-ils des corps : $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$?

Solution. — $\mathbb{Q}$ et $\mathbb{R}$ sont des corps. $\mathbb{N}$ n'en est pas un parce que par exemple pour son addition seul 0 possède un opposé. $\mathbb{Z}$ n'en est pas un non plus parce que par exemple pour sa multiplication seuls ± 1 possèdent un inverse.

Formule du binôme de Newton (du nom du physicien, mathématicien et astronome anglais Sir I. Newton*, 1643–1727).* — Soient $w, z \in \mathbb{C}$ (on pourrait remplacer $\mathbb{C}$ par un anneau commutatif quelconque) et $n \geq 0$. Rappelons que $k! = k \times (k-1) \times \cdots \times 2 \times 1$. Alors

$$(w+z)^n = \sum_{0 \leq k \leq n} C_n^k w^k z^{n-k}, \quad C_n^k = \frac{n!}{k!(n-k)!}.$$

Cette formule est immédiate, puisque C_n^k est le nombre de façon de choisir k objets parmi n, ici en l'occurrence le nombre de façon de choisir, parmi les n facteurs du produit

$$(w+z)(w+z)\cdots(w+z),$$

les k facteurs où l'on utilise w plutôt que z.

Les nombres C_n^k, appelés *nombres de combinaisons*, satisfont la récurrence suivante :

$$C_{n+1}^{k+1} = C_n^k + C_n^{k+1} \ ;$$

ceci traduit le fait qu'il y a autant de façon de choisir $k+1$ objets parmi $n+1$, qu'il y en a de choisir k objets parmi n *ou* $k+1$ objets parmi n (les $k+1$ objets choisis pouvant a priori être dans les n premiers ou pas).

Ajoutée au fait que $C_n^n = C_n^1 = 1$ pour tout n, cette identité permet de recalculer facilement les premiers nombres de combinaisons, qu'on écrit habituellement sous la forme du *triangle de Pascal* (d'après le mathématicien, physicien et philosophe français B. Pascal, 1623–1662) :

$$\begin{array}{ccccc}
1 & & & & \\
1 & 2 & 1 & & \\
1 & 3 & 3 & 1 & \\
1 & 4 & 6 & 4 & 1 \\
\vdots & & \cdots & & \ddots
\end{array}$$

On voit ainsi facilement, par exemple, que

$$(w+z)^4 = w^4 + 4w^3z + 6w^2z^2 + 4wz^3 + z^4.$$

Exercice. — Linéariser $\sin^5\theta$.

Solution. — Notons $z = e^{i\theta}$. En calculant une ligne de plus dans le triangle de Pascal que ci-dessus, on voit que

$$(a+b)^5 = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5,$$

donc que

$$\left(z - z^{-1}\right)^5 = z^5 - 5z^3 + 10z - 10z^{-1} + 5z^{-3} - z^{-5}$$

(remarquer que les puissances de z décroissent de deux unités d'un terme au suivant, et que les signes alternent). Donc

$$\begin{aligned}
\sin^5\theta &= \frac{1}{(2i)^5}\left(z - z^{-1}\right)^5 \\
&= \frac{1}{2^5 i}\left(z^5 - 5z^3 + 10z - 10z^{-1} + 5z^{-3} - z^{-5}\right) \\
&= \frac{1}{2^4}\left(\sin 5\theta - 5\sin 3\theta + 10\sin\theta\right).
\end{aligned}$$

Exercice. — Les ensembles

$$\mathscr{T} := \{\tau_v,\ v \in \mathbb{C}\}, \quad \mathscr{R}_0 := \{\rho_{0,\alpha},\ \alpha \in \mathbb{R}\} \quad \text{et} \quad \mathscr{H}_0 := \{\eta_{0,\kappa},\ \kappa \in \mathbb{R}\setminus\{0\}\}$$

munis de la loi de composition des applications, sont des *groupes commutatifs*, au sens que, si G est l'un quelconque de ces ensembles, la composition définit une multiplication $G \times G \to G$, $(g,h) \mapsto g \circ h$ vérifiant les axiomes suivants :

- Commutativité : $g \circ h = h \circ g$;
- Associativité : $g \circ (h \circ k) = (g \circ h) \circ k$;
- Existence d'un élément neutre e ($g \circ e = e \circ g = g$) ;
- Existence d'un inverse g^{-1} pour tout g ($g \circ g^{-1} = g^{-1} \circ g = e$).

Solution. — La commutativité et l'associativité découlent de l'exercice du § 1.4. Dans les trois cas, l'élément neutre est l'application *identité* $\mathrm{id} : \mathbb{C} \to \mathbb{C}$, $z \mapsto z$. Toujours d'après le même exercice, l'inverse d'un élément quelconque est, respectivement :

$$(\tau_v)^{-1} = \tau_{-v}, \quad (\rho_{0,\alpha})^{-1} = \rho_{0,-\alpha} \quad \text{et} \quad (\eta_{0,\kappa})^{-1} = \eta_{0,\kappa^{-1}}.$$

Exercice. — L'exercice § 1.4 montre que le monde des rotations n'est pas clos, puisque par composition on obtient parfois une translation. Montrer que l'ensemble

$$\mathscr{D} := \{f : \mathbb{C} \to \mathbb{C}, z \mapsto a + ze^{i\alpha}\}_{a\in\mathbb{C},\alpha\in\mathbb{R}}$$

est un groupe (non commutatif) ; c'est le *groupe des déplacements.*

Solution. — L'exercice précédent montre que la composition des applications définit une multiplication $\mathscr{D} \times \mathscr{D} \to \mathscr{D}$ (le fait important est que l'application composée de deux éléments de $\mathscr{D}$ est bien dans $\mathscr{D}$). On a déjà vu que la composition est associative et que l'élément neutre est $\mathrm{id} = \tau_0 : z \mapsto z$, et que les translations et les rotations sont inversibles dans $\mathscr{D}$ (on dit que ces applications sont des *bijections*).

CHAPITRE 2

L'ESPACE RÉEL À TROIS DIMENSIONS

Mots-clef du chapitre. — Vecteur de $\mathbb{R}^3$, base canonique, combinaison linéaire, vecteurs libres ou liés, quantificateur, implication logique, droites et plans vectoriels et affines, produit scalaire, inégalité de Cauchy-Schwarz, orthogonalité, volume orienté, déterminant, multi-linéarité, algorithme fang-cheng, formule de développement, produit vectoriel, raisonnement par analyse-synthèse

DEPUIS l'Antiquité égyptienne, les branches de l'Algèbre et de la Géométrie étaient conçues comme distinctes. Or, au XVIIe siècle des savants tels que G. GALILEI (1564–1642) et R. DESCARTES (1596–1650) découvrirent les fondements de la Géométrie analytique, en représentant les points du plan ou de l'espace par des couples ou des triplets de nombres réels. Immédiatement, les opérations connues sur les nombres conduisirent à des propriétés géométriques nouvelles :

> Tous les Problèmes de Géométrie se peuvent facilement réduire à tels termes, qu'il n'est besoin par après que de connaître la longueur de quelques lignes droites, pour les construire. R. DESCARTES, Premier Livre de *La Géométrie*, en appendice au *Discours de la Méthode* (1637)
> (`http://www.bibnum.education.fr/mathematiques/geometrie/le-livre-premier-de-la-geometrie-de-descartes`)

La Géométrie analytique réduisit à des nombres non seulement les points, mais aussi les courbes et les surfaces. On comprit notamment que les droites, les cercles et les coniques n'étaient que des exemples importants de courbes beaucoup plus générales que les nouvelles méthodes algébriques permettaient dorénavant d'appréhender. Les succès furent spectaculaires et ouvrirent la voie au Calcul différentiel.

Ici, nous introduisons les bases de la Géométrie analytique dans l'espace, en complétant les concepts du chapitre précédent pour le plan.

2.1. Vecteurs de $\mathbb{R}^3$

Définitions. — Un *vecteur (réel) à* 3 *composantes* est un triplet $v = (x, y, z)$ de nombres réels, ou, ce qui est équivalent, un point de l'espace cartésien $\mathbb{R}^3 = Oxzy$. Les réels x, y, z sont les *composantes* de v.

Deux vecteurs $v = (x, y, z)$ et $v' = (x', y', z')$ sont égaux si et seulement si toutes leurs composantes sont égales : $x = x'$, $y = y'$ et $z = z'$.

Définissons les opérations d'addition et de multiplication par un réel comme on l'avait fait pour $\mathbb{C} = \mathbb{R}^2$:

$$(x, y, z) + (x', y', z') := (x + x', y + y', z + z'), \qquad a(x, y, z) := (ax, ay, az).$$

Deux petites différences avec le plan complexe : l'habitude est de noter le réel *à gauche* du vecteur ; et l'on ne donnera pas de nom particulier à $\mathbb{R}^3$ comme on l'avait fait en notant $\mathbb{C} = \mathbb{R}^2$. Ceci est lié à une différence majeure : il n'est pas possible de définir sur $\mathbb{R}^3$ une multiplication analogue de la multiplication des complexes, pour une raison algébrique profonde exprimée par le théorème de Frobenius.

Notations. —

$$0 := 0_{\mathbb{R}^3} := (0, 0, 0), \quad -v := (-1)v, \quad v - w := v + (-w).$$

Les trois *vecteurs de la base canonique* (la notion de *base* sera expliquée ultérieurement) :

$$i := (1, 0, 0), \quad j := (0, 1, 0), \quad k := (0, 0, 1).$$

Attention : la lettre i désigne ici tout autre chose que le nombre complexe i (il faut se faire à ce que les notations changent avec le contexte ; dans $\mathbb{R}^4$, on note $1, i, j, k$ les vecteurs de la base canonique !).

Exemple. — $(x, y, z) = xi + yj + zk$ (cette somme de *trois* vecteurs étant justifiée par l'associativité de l'addition : $(xi + yj) + zk = xi + (yj + zk)$, qui permet sans ambiguité de supprimer les parenthèses dans les sommes de plusieurs vecteurs).

Exercice. — Calculer à chaque instant $t \in \mathbb{R}$, le vecteur-vitesse $v(t)$ de la courbe paramétrée $\gamma : t \mapsto (\cos t, \sin t, \cos(2t))$. Quand ce vecteur est-il horizontal (troisième composante nulle) ?

Solution. — On a

$$v(t) = \gamma'(t) = (-\sin t, \cos t, -2\sin(2t)).$$

Ce vecteur est horizontal quand $\sin 2t = 0$, soit $t = 0 \pmod{\pi/2}$.

2.2. Vecteurs libres

Définitions. — Une *combinaison linéaire* de deux vecteurs (ou plus) v et w est une expression de la forme

$$av + bw,$$

où a et b sont deux réels, appelés les *coefficients* de la combinaison linéaire.

Les vecteurs v et w sont *liés par une relation linéaire* s'il existe une combinaison linéaire de v et w qui, quand on l'évalue, est nulle. Il existe toujours la *relation linéaire triviale* entre v et w :

$$0v + 0w = 0_{\mathbb{R}^3},$$

mais cette relation, étant universellement satisfaite (quelle que soient v et w), ne dit justement rien sur v et w.

Un exemple de relation linéaire non triviale est :

$$2(1, 2, 3) - (2, 4, 6) = 0.$$

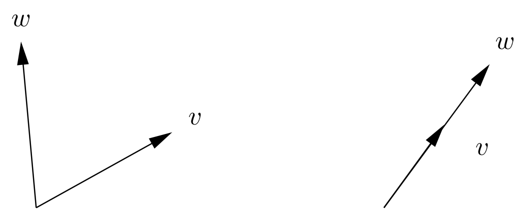
Deux vecteurs v et w sont *liés* (sous-entendu : par une relation linéaire non triviale) s'il existe deux scalaires a et b non tous deux nuls tels que $av + bw = 0$.

Par exemple, $0_{\mathbb{R}^3}$ est lié à n'importe quel vecteur v de $\mathbb{R}^3$, puisque $1 \cdot 0_{\mathbb{R}^3} + 0 \cdot v = 0$ est une relation linéaire non triviale.

Remarque. — Supposons v et w liés : $av + bw = 0$ avec a et b non tous deux nuls.
- Si $a \neq 0$, $v = -\frac{b}{a}w$.
- Sinon, $a = 0$ donc $b \neq 0$, et $w = -\frac{a}{b}v$.

On voit que v et w sont liés si et seulement si ils sont proportionnels l'un à l'autre. Donc *liés* est synonyme de *proportionnels* ou de *colinéaires* (attention, ceci n'est vrai que pour *deux* vecteurs !). Mais *lié* est une propriété symétrique par rapport aux deux vecteurs (peu importe si l'un est nul) et se généralisera plus facilement à plus de deux vecteurs.

Définition. — Les vecteurs v et w sont *libres* si ils ne sont pas liés.



Vecteurs libres Vecteurs liés

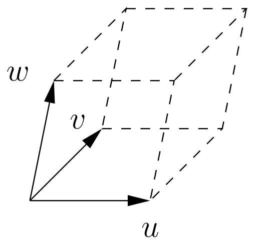
Ici il est commode d'introduire certains symboles de la Logique mathématique :
- le *connecteur d'implication* $\Rightarrow$, utilisé entre deux assertions logiques pour signifier que la première implique la seconde
- le *connecteur d'équivalence* $\Leftrightarrow$, utilisé entre deux assertions logiques pour signifier qu'elles sont équivalentes l'une à l'autre (i.e. chacune des deux implique la seconde)
- le *quantificateur universel* $\forall$ = « quel(s) que soi(en)t »
- le *quantificateur existentiel* $\exists$ = « il existe ».

C'est à la fin du XIXe siècle que G. Frege (mathématicien, logicien et philosophe allemand, 1848–1925), comprit la nécessité des connecteurs logiques et des quantificateurs pour découvrir la structure des propositions mathématiques, et fonder le *calcul des prédicats*.

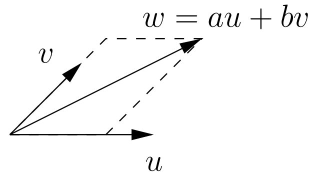
Voici quelques exemples d'implications mathématiques vraies :
- $x = 1 \Rightarrow 2x = 2$
- $0 = 1 \Rightarrow 0 = 2$ (la première assertion est fausse et donc implique n'importe quoi)
- $0 = 0 \Rightarrow 1 = 1$
- $0 = 1 \Rightarrow 1 = 1$ (la seconde assertion est vraie et donc n'importe quoi l'implique)
- Neige en novembre $\Rightarrow$ Noël en décembre (même explication).

Ces exemples devraient convaincre qu'il faut limiter l'usage de ces symboles, et s'interdire leur insertion dans des phrases en français pour éviter des contre-sens logiques. En particulier, il faut proscrire l'utilisation du signe $\Rightarrow$ à la place de la conjonction « donc » .

. — Formalisons les propriétés d'être libre ou lié. Pour l'exercice, nous le faisons avec trois vecteurs au lieu de deux.



Trois vecteurs libres



Trois vecteurs liés

u, v et w sont liés $\Leftrightarrow$ ($\exists a, b, c \in \mathbb{R}^3$ non tous nuls tels que $au + bv + bw = 0$)

ou, de façon équivalente,

u, v et w sont libres $\Leftrightarrow$ non ($\exists a, b, c \in \mathbb{R}^3$ non tous nuls tels que $au + bv + bw = 0$)

Comment simplifier cette expression logique ? Il suffit de réfléchir au sens des mots ! Exercice : Se convaincre que nier une proposition logique revient à permuter les quantificateurs ($\forall$ et $\exists$) et nier la conclusion.

Donc : u, v et w sont libres $\Leftrightarrow \forall a, b, c \in \mathbb{R}$ non tous nuls, $au + bv + bw \neq 0$,

soit

u, v et w sont libres $\Leftrightarrow (\forall a, b, c \in \mathbb{R} \quad (au + bv + bw = 0 \Rightarrow a = b = c = 0))$.

Dans la pratique, c'est cette dernière proposition logique qu'on utilise le plus souvent pour voir si deux vecteurs sont libres ou liés. Comprendre que la proposition logique à droite du signe d'équivalence n'est pas une simple équation ; elle est elle-même une assertion !

Remarque. — Le *contraire* de "Paul aime Virginie" est que "Paul n'aime pas Virginie", tandis que sont *opposé* est "Paul hait Virginie".

De façon similaire, le contraire de "$\forall x\ P(x)$" est "$\exists x$ non $P(x)$", tandis que son opposé est "$\forall x$ non $P(x)$". Nier la proposition "$\forall x\ P(x)$" revient donc à en prendre le contraire, et non l'opposé ! Il convient de ne pas confondre ces deux notions.

Exemple. — Les vecteurs $v = (1, 2, 3)$ et $w = (3, 2, 1)$ sont-ils libres ou liés ? Soit $(a, b) \in \mathbb{R}^2$ tel que $av + bw = 0$. Nous allons simplifier cette équation de façon à voir explicitement quelles valeurs de a et b sont possibles :

- soit la seule possibilité est $a = b = 0$, et alors la caractérisation encadrée de deux vecteurs libres est vérifiée : v et w sont libres ;
- soit il existe des solutions (a, b) différentes de $(0, 0)$, correspondant à autant de relations linéaires non triviales entre v et w, et alors v et w sont liés.

L'équation vectorielle $av + bw = 0$ équivaut au système de trois équations linéaires scalaires suivant :

$$\begin{cases} a + 3b = 0 \\ 2a + 2b = 0 \\ 3a + b = 0. \end{cases}$$

Si l'on note ℓ_1, ℓ_2 et ℓ_3 les lignes de ce système, la combinaison $2\ell_1 - 3\ell_2$ permet d'éliminer b et montre que $-4a = 0$, soit $a = 0$. Ensuite, n'importe laquelle des trois équations de départ implique que $b = 0$. La proposition logique encadrée étant vérifiée, v et w sont libres.

N.B. : Dans le chapitre suivant, on verra une méthode pour résoudre les systèmes d'équations linéaires de façon systématique (algorithme fang-cheng).

Exercice. — Soient u et v les deux vecteurs de $\mathbb{R}^2$ définis par $u = (1, 2)$ et $v = (2, 3)$. La famille (u, v) est-elle libre ?

Solution. — Résolvons l'équation

$$au + bv = 0,$$

équivalente au système de deux équations scalaires

$$\begin{cases} a + 2b = 0 \\ 2a + b = 0. \end{cases}$$

Si l'on soustrait deux fois la seconde équation à la première, b s'élimine et on trouve $a = 0$. Il en découle que $b = 0$. Donc les deux vecteurs u et v sont libres.

Exercice. — Montrer que des vecteurs v_j sont liés si et seulement si l'un d'eux est une combinaison linéaire des autres.

Solution. — Supposons que les v_j sont liés. Il existe des réels α_j non tous nuls tels que $\sum_j \alpha_j v_j = 0$. Comme les α_j ne sont pas tous nuls, il existe k tel que $\alpha_k \neq 0$. Alors

$$v_k = -\frac{1}{\alpha_k} \sum_{j \neq k} \alpha_j v_j.$$

Réciproquement, supposons qu'il existe un indice k et des réels β_j pour $j \neq k$ tels que

$$v_k = \sum_{j \neq k} \beta_j v_j.$$

Posons $\alpha_j = \beta_j$ pour $j \neq k$ et $\alpha_k = -1$. Alors les α_j ne sont pas tous nuls (puisque $\alpha_k \neq 0$) et

$$\sum_j \alpha_j v_j = 0.$$

Donc les v_j sont liés.

Remarques :

- Ceci vaut pour un nombre quelconque de vecteurs (et en dimension quelconque !).
- Dans cette rédaction de la solution, k est un indice fixé, à la différence de j, qui prend pour cette raison le nom d'*indice muet*.

2.3. Droites et plans

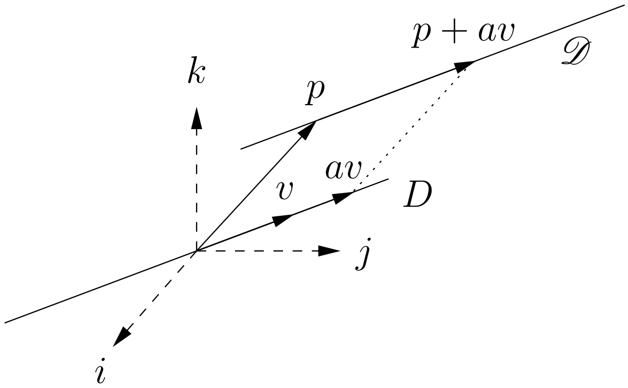
Définitions. — La *droite vectorielle* D engendrée par un vecteur non nul v est

$$D = \mathbb{R}v = \{av \; ; \; a \in \mathbb{R}\} \; ;$$

c'est l'ensemble des vecteurs w liés à v ($w - av = 0$).

La *droite affine* $\mathscr{D}$ passant par un point $p \in \mathbb{R}^3$ et dirigée par v (ou par D) est

$$\mathscr{D} = p + D = \{p + av \; ; \; a \in R\}.$$



Exemple. — La droite affine passant par $p = (1, 1, 1)$ et dirigée par $v = (1, 2, 3)$ est l'ensemble des $q = (x, y, z)$ tels qu'il existe un réel a tel que

$$q = p + av, \quad \text{i.e.} \quad \begin{cases} x = 1 + a \\ y = 1 + 2a \\ z = 1 + 3a. \end{cases}$$

Ce système, qui donne les composantes de w comme fonction d'un paramètre réel a, est un *paramétrage* de $\mathscr{D}$.

La première équation du paramétrage peut être « résolue en a », i.e. peut servir à exprimer a en fonction du reste : $a = x - 1$, et l'on peut injecter cette égalité dans les deux autres équations :

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

Le fait qu'il existe a tel que ces trois équations soient satisfaites équivaut à ce que les deux dernières équations soient satisfaites :

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

(x, y, z et a étant donnés, le premier système implique certainement le second, obtenu en négligeant la première équation ; réciproquement, le second système, implique le premier à condition de poser justement $a = x - 1$). Le système obtenu de deux équations scalaires entre les 3 composantes de q s'appelle une *équation (cartésienne)* de $\mathscr{D}$. On est passé d'un paramétrage à une équation en *éliminant* le paramètre a.

Inversement, on peut passer d'une équation à un paramétrage en introduisant un paramètre, opération pour laquelle il y a beaucoup de choix possibles (correspondant à tous les points $p \in \mathscr{D}$ et tous les vecteurs v dirigeant $\mathscr{D}$).

Exercice. — Montrer qu'une droite affine qui passe par 0 est une droite vectorielle.

Solution. — Si $0 \in \mathscr{D}$, il existe a_0 tel que $p + a_0 v = 0$. Alors, pour tout a on a

$$p + av = (p + av) - (p + a_0 v) = (a - a_0)v,$$

i.e. $\mathscr{D}$ est la droite vectorielle engendrée par v.

Exercice. — Soit $f : I = [0, 1] \to \mathbb{R}^3$ la courbe paramétrée définie par

$$f(t) = (t, t^2, t^3).$$

Calculer à chaque « instant » t l'équation de la droite $\mathscr{D}_t$ tangente au mouvement, i.e. la droite passant par le point $f(t)$ et dirigée par la vitesse $f'(t)$.

Solution. — Un paramétrage de $\mathscr{D}_t$ est

$$p = f(t) + sf'(t) \quad (s \in \mathbb{R}),$$

soit, puisque $f'(t) = (1, 2t, 3t^2)$,

$$\begin{cases} x = t + s \\ y = t^2 + 2st \\ z = t^3 + 3st^2. \end{cases}$$

La première équation permet d'éliminer s : $s = x - t$, et donc une équation de $\mathscr{D}_t$ est

$$y = t^2 + 2t(x - t), \quad z = t^3 + 3t(x - t),$$

soit, après simplification,

$$y = 2tx - t^2, \quad z = 3tx - 2t^3.$$

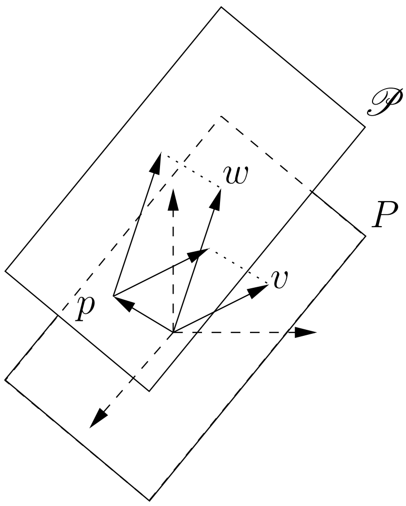
Définitions. — Le *plan vectoriel* P engendrée par deux vecteurs libres v et w est

$$P = \mathbb{R}v + \mathbb{R}w = \{av + bw \ ; \ a, b \in \mathbb{R}\} \ ;$$

c'est l'ensemble des vecteurs obtenus par combinaisons linéaires de v et de w.
Le *plan affine* $\mathscr{P}$ passant par un point $p \in \mathbb{R}^3$ et dirigé par v et w (ou par P) est

$$\mathscr{P} = p + P = \{p + av + bw \ ; \ a, b \in \mathbb{R}\}.$$

De même que pour les droites, un plan affine passant par l'origine est un plan vectoriel.



Exercice. — Pourquoi suppose-t-on, dans cette définition, que v et w sont libres ?

Solution. — Si v et w étaient liés, l'un d'eux serait proportionnel à l'autre, disons par exemple $w = cv$, et toute combinaison linéaire des deux serait en réalité proportionnelle à v : $av + bw = (a + bc)v$. Alors P serait une droite vectorielle.

Exemple. — Le plan affine passant par $p = (1, 1, 1)$ et dirigé par les vecteurs libres $v = (1, 2, 3)$ et $w = (3, 2, 1)$ est l'ensemble des $q = (x, y, z)$ tels qu'il existe a et b tels que

$$a = p + av + bw \quad \text{i.e.} \quad \begin{cases} x = 1 + a + 3b \\ y = 1 + 2a + 2b \\ z = 1 + 3a + b. \end{cases}$$

Ce système d'équations donnant q en fonction d'un paramètre est un *paramétrage* de $\mathscr{P}$. De même que pour les droites, en éliminant les paramètres on obtient une *équation* de $\mathscr{P}$. Pour un plan, comme il y a deux paramètres à éliminer, il ne restera plus qu'une seule équation scalaire.

La combinaison linéaire $\ell_2 - 2\ell_3$ permet d'éliminer b :

$$y - 2z = -1 - 4a,$$

soit, en résolvant l'équation par rapport à a,

$$a = \frac{1}{4}(-1 - y + 2z).$$

Grâce à la symétrie particulière de cet exemple ($\mathscr{P}$ est invariant par la symétrie échangeant x et z) on voit que

$$b = \frac{1}{4}(-1 - y + 2x).$$

On peut maintenant injecter ces expressions de a et de b dans l'une quelconque des trois équations de départ (et vérifier qu'on obtient des équations équivalentes dans chacun des trois cas). Par exemple, la deuxième équation donne

$$x - 2y + z = 0.$$

Il s'avère que $\mathscr{P}$ passe par $0_{\mathbb{R}^3}$ et qu'il est donc vectoriel.

Exercice. — Soit $f : I^2 = [0, 1]^2 \to \mathbb{R}^3$ la surface paramétrée $\mathscr{S}$ définie par

$$f(s, t) = (1 + t, s + t^2, s^2 + t^3).$$

Calculer, pour chaque valeur du paramètre (s, t), l'équation du plan $\mathscr{P}_{s,t}$ tangent à $\mathscr{S}$, i.e. du plan passant par $f(s, t)$ et dirigé par $\frac{\partial f}{\partial s}(s, t)$ et $\frac{\partial f}{\partial t}(s, t)$. Pour quelles valeurs du paramètre $\mathscr{P}_{s,t}$ est-il horizontal ? Parallèle à Ox ?

Solution. — Un paramétrage de $\mathscr{P}_{s,t}$ est

$$\begin{cases} x = 1 + t + v \\ y = s + t^2 + u + 2tv \\ z = s^2 + t^3 + 2su + 3tv. \end{cases}$$

Les deux premières équations permettent d'éliminer u et v :

$$\begin{cases} v = x - 1 - t \\ u = y - s - t^2 - 2 - v = -x + y - s + t - t^2 - 1, \end{cases}$$

et une équation de $\mathscr{P}_{s,t}$ est donc

$$z = s^2 + t^3 + 2s(-x + y - s + t - t^2 - 1) + 3 + (x - 1 - t),$$

soit, après simplification,

$$(1 + 2s)x + 2sy - z - s^2 + t^3 + 2s(t - t^2 - 1) - 3t - 3 = 0.$$

$\mathscr{P}_{s,t}$ est horizontal si le coefficient de x et celui de y sont nuls : ceci n'arrive jamais simultanément. $\mathscr{P}_{s,t}$ est parallèle à Ox si le coefficient de x est nul, c'est-à-dire si $s = -1/2$.

2.4. Produit scalaire

Définition. — Le *produit scalaire* de deux vecteurs v w est le scalaire (réel)

$$v \cdot w = (v_1, v_2, v_3) \cdot (w_1, w_2, w_3) = v_1 w_1 + v_2 w_2 + v_3 w_3.$$

La *norme (euclidienne)* de v est

$$\|v\| = \sqrt{v \cdot v} = \sqrt{v_1^2 + v_2^2 + v_3^2}.$$

Un vecteur est *unitaire* si sa norme vaut 1. Deux vecteurs sont *orthogonaux* si leur produit scalaire est nul.

Propriétés. — Produit scalaire :

- Symétrie : $v \cdot w = w \cdot v$
- Linéarité en le premier argument : $(av + a'v') \cdot w = a(v \cdot w) + a'(v' \cdot w)$
- Linéarité en le second argument : $v \cdot (bw + b'w') = b(v \cdot w) + b'(v \cdot w')$

Norme :

- Positivité : $\|v\| \geq 0$, et $\|v\| = 0 \Leftrightarrow v = 0$
- Homogénéité : $\|av\| = |a| \, \|v\|$
- Inégalité triangulaire : $\|v + w\| \leq \|v\| + \|w\|$.

Ces propriétés découlent directement des définitions, sauf la dernière qui est une conséquence du résultat suivant.

Inégalité de Cauchy-Schwarz. — $|v \cdot w| \leq \|v\| \, \|w\|$, avec égalité si et seulement v et w sont liés (A. L. Cauchy, mathématicien français, 1789–1857 ; H. A. Schwarz, mathématicien allemand, 1843–1921).

Démonstration – On peut supposer, par exemple, que $w \neq 0$, parce que sinon $v \neq 0$ et la démonstration dans ce cas est similaire. L'idée astucieuse est de remarquer que la fonction « norme au carré » est positive en particulier sur la droite affine $v + \mathbb{R}w$, où $\mathbb{R}w = \{aw \; ; \; a \in \mathbb{R}\}$ est l'ensemble des vecteurs liés à w :

$$(\forall a \in \mathbb{R}) \quad \|v + aw\|^2 = \|v\|^2 + 2a(v \cdot w) + a^2 \|w\|^2 \geq 0 \; ;$$

le discriminant de ce polynôme en a est donc ≤ 0, d'où l'inégalité.

Si $|v \cdot w| = \|v\| \, \|w\|$, le discriminant est nul, et donc il existe a tel que $v + aw = 0$: v et w sont liés. Réciproquement, si v et w sont liés, quitte à échanger éventuellement v et w on peut supposer qu'il existe a tel que $v = aw$, ce qui implique bien $|v \cdot w| = |a| w \cdot w = \|v\| \, \|w\|$. □

On peut maintenant démontrer l'inégalité triangulaire :

$$\begin{aligned} \|v + w\|^2 &= \|v\|^2 + 2v \cdot w + \|w\|^2 \\ &\leq \|v\|^2 + 2 \|v\| \, \|w\| + \|w\|^2 = (\|v\| + \|w\|)^2. \end{aligned}$$

Proposition. — $v \cdot w = \|v\| \, \|w\| \cos(\widehat{v, w})$

Idée de la démonstration – D'abord, grâce à la bilinéarité du produit scalaire, il suffit de démontrer l'égalité quand $\|v\| = \|w\| = 1$.

Ensuite on montre que si R est une rotation de $\mathbb{R}^3$ (que l'on définira précisément plus tard dans le cours) elle préserve le produit scalaire : $v \cdot w = (Rv) \cdot (Rw)$ et les angles : $\cos(\widehat{v, w}) = \cos(\widehat{Rv, Rw})$.

Enfin on choisit une rotation qui envoie v sur le premier vecteur i de la base canonique, et qui envoie w dans le plan $Oxy \subset \mathbb{R}^3$. Alors Rv et Rw sont respectivement de la forme $(1, 0, 0)$ et $(w_1', w_2', 0)$ avec ${w_1'}^2 + {w_2'}^2 = 1$, de sorte que

$$v \cdot w = w_1' = \cos(\widehat{v, w}).$$

□

Proposition. — L'ensemble des vecteurs orthogonaux à un vecteur $v \neq 0$ est un plan vectoriel appelé le *plan orthogonal* à v.

Démonstration – Soit P l'ensemble des vecteurs orthogonaux à $v \neq 0$. Notons $v = (v_1, v_2, v_3)$ et supposons par exemple que $v_3 \neq 0$, les autres cas étant similaires. L'ensemble P a pour équation

$$v_1 x + v_2 y + v_3 z = 0,$$

i.e.

$$z = -\frac{1}{v_3}(v_1 x + v_2 y) \ ;$$

l'équation est satisfaite si et seulement si il existe deux scalaires a et b tels que

$$\begin{cases} x = a \\ y = b \\ z = -\frac{1}{v_3}(v_1 a + v_2 b). \end{cases}$$

Donc P est le plan vectoriel engendré par les vecteurs $(1, 0, -v_1/v_3)$ et $(0, 1, -v_2/v_3)$. □

Exercice. — Trouver une équation du plan affine $\mathscr{P}$ passant par $p = (1, 1, 1)$ et orthogonal à $v = (1, 2, 3)$ (i.e. dirigé par le plan vectoriel orthogonal à v).

Solution. — $q = (x, y, z)$ appartient à $\mathscr{P}$ si et seulement si $(q - p) \cdot v = 0$, i.e.

$$(x - 1) + 2(y - 1) + 3(z - 1) = 0,$$

soit

$$x + 2y + 3z = 6.$$

Exercice. — Montrer que trois vecteurs de $\mathbb{R}^3$ contenus dans le plan vectoriel $z = 0$ sont liés. (On verra plus tard qu'il en est de même de trois vecteurs contenus dans un plan vectoriel quelconque.)

Solution. — Les trois vecteurs sont de la forme

$$u = (u_1, u_2, 0), \quad v = (v_1, v_2, 0) \quad \text{et} \quad w = (w_1, w_2, 0).$$

On veut montrer qu'il existe une relation linéaire non triviale entre eux. On peut supposer que $u \neq 0$, parce que sinon c'est évident $(u + 0v + 0w = 0)$, et même que $u_1 \neq 0$ (les autres cas étant similaires).

– *Analyse.* Supposons qu'ils satisfont une relation linéaire

$$au + bv + cw = 0, \quad i.e. \quad \begin{cases} au_1 + bv_1 + cw_1 = 0 & (\ell_1) \\ au_2 + bv_2 + cw_2 = 0, & (\ell_2) \end{cases}$$

avec $a, b, c \in \mathbb{R}$. Comme $u_1 \neq 0$, on peut éliminer a de la deuxième équation en remplaçant la ligne ℓ_2 par la combinaison $u_1\ell_2 - u_2\ell_1$:

$$(u_1v_2 - u_2v_1)b + (u_1, w_2 - u_2w_1)c = 0.$$

Premier cas : $u_1v_2 - u_2v_1 = 0$. Alors $v = \frac{v_1}{u_1}u$.

Second cas : $u_1v_2 - u_2v_1 \neq 0$. Alors

$$b = -\frac{u_1, w_2 - u_2w_1}{u_1v_2 - u_2v_1}c.$$

– *Synthèse.* Dans chacun des deux cas distingués, on voit que u, v et w sont liés :

Premier cas : $\frac{v_1}{u_1}u - v = 0$.

Second cas : on peut choisir c arbitrairement (non nul), par exemple $c = 1$. Alors il suffit de poser

$$b = -\frac{u_1, w_2 - u_2w_1}{u_1v_2 - u_2v_1}$$

puis

$$a = -\frac{1}{u_1}(bv_1 + cw_1),$$

pour que $au + bv + cw = 0$.

Définition. — Soit $\gamma : I = [0, 1] \to \mathbb{R}^3$ une courbe paramétrée de classe C^1 (ceci signifie que γ est dérivable et que sa dérivée est continue). Soit aussi f une fonction continue définie sur la courbe $\Gamma = \gamma(I)$; en physique, f est par exemple une masse par unité de longueur (imaginer une cordelette non uniforme).

L'*intégrale curviligne* de f le long de la courbe Γ est

$$\int_\Gamma f := \int_0^1 f(\gamma(t)) \, \|\gamma'(t)\| \, dt.$$

Exemple. — La longueur de Γ est l'intégrale curviligne

$$L(\Gamma) = \int_\Gamma 1 = \int_0^1 \|\gamma'(t)\| \, dt.$$

Cette formule est intuitive. En effet, d'après la formule de Taylor,

$$\gamma(t + dt) = \gamma(t) + \gamma'(t) \, dt + o(dt).$$

Donc, au premier ordre quand dt tend vers 0^+, pendant l'intervalle de temps $[t, t + dt]$ le déplacement du point mobile γ est le vecteur $\gamma'(t) \, dt$, de norme $\|\gamma'(t)\| \, dt$. La somme de ces contributions, à la limite quand $dt \to 0$, est donnée par l'intégrale ci-dessus (ce passage à la limite est justifié dans le cours d'intégration).

Exercice. — Calculer la longueur de l'hélice

$$\gamma : t \in [0, 1] \to (\cos t, \sin t, t).$$

Solution. — À l'instant t, la vitesse est

$$\gamma'(t) = (-\sin t, \cos t, 1),$$

et sa norme

$$\|\gamma'(t)\| = \sqrt{\sin^2 t + \cos^2 t + 1} = \sqrt{2},$$

donc la longueur de l'hélice vaut

$$\int_0^1 \sqrt{2}\,dt = \sqrt{2}.$$

2.5. Déterminant en petite dimension

Comment calculer la longueur d'un vecteur, l'aire d'un parallélogramme ou le volume d'un parallélépipède ? Ces quantités s'avèrent plus faciles à calculer quand on les munit d'un signe $\pm$ (de façon cohérente, de façon notamment que le signe change quand on permute deux côtés) ; on les qualifie alors d'*orientées*, et alors ce sont toutes des incarnations, dans des dimensions particulères, de la fonction *déterminant*.

Par exemple, la longueur orientée du « vecteur » x de $\mathbb{R}$ est x lui-même (une expression polynomiale), tandis que la longueur (positive) est $|x|$. Mais c'est en dimensions 2 et 3 que nous allons comprendre comment faire ces calculs en général.

Axiomes intuitifs de l'aire orientée $\det(v,w)$. — Considérons l'aire orientée du parallélogramme $P = \{av+bw\ ;\ 0 \leq a \leq 1,\ 0 \leq b \leq 1\}$ engendré par deux vecteurs $v = (v_1, v_2)$ et $w = (w_1, w_2)$.

– *Homogénéité*

$$\det(\alpha v, w) = \alpha \det(v,w) = \det(v, \alpha w)$$

– *Additivité*

$$\det(v+v', w) = \det(v,w) + \det(v', w)$$

(pour s'en convaincre, décomposer v' en un vecteur parallèle à v et un vecteur parallèle à w, et faire un dessin plus précis que ci-dessous)

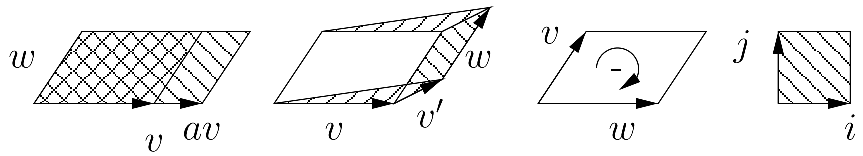
– *Antisymétrie*

$$\det(v,w) = -\det(w,v)$$

(l'aire orientée change de signe quand on change l'ordre des vecteurs)

– *Normalisation*

$$\det(i,j) = 1, \quad \text{où} \quad i = (1,0),\ j = (0,1).$$



L'homogénéité, l'additivité et l'antisymétrie montrent que $\det(v,w)$ est *bilinéaire*, i.e. *linéaire* en chacun de ses deux arguments :

– $\det(av + a'v', w) = a\det(v,w) + a'\det(v', w)$
– $\det(v, bw + b'w') = b\det(v,w) + b'\det(v, w')$.

Généralisation au volume orienté. — De même, le *volume orienté* serait une fonction $\det : (\mathbb{R}^3)^3 \to \mathbb{R}$, $(u,v,w) \mapsto \det(u,v,w)$,

– *trilinéaire* :
$$\begin{aligned}\det(au + a'u', v, w) &= a\det(u,v,w) + a'\det(u',v,w)\\ \det(u, bv + b'v', w) &= b\det(u,v,w) + b'\det(u,v',w)\\ \det(u, v, cw + c'w') &= c\det(u,v,w) + c'\det(u,v,w')\end{aligned}$$

– antisymétrique : $\det(u', v', w') = -\det(u, v, w)$ si le triplet de vecteurs (u', v', w') s'obtient à partir de (u, v, w) par permutation de deux des trois vecteurs
– normalisée : $\det(i, j, k) = 1$.

Remarques sur les parallélépipèdes aplatis. —
– Si $u = 0$, par linéarité par rapport à u on voit que $\det(u, v, w) = \det(2u, v, w) = 2\det(u, v, w)$ donc $\det(u, v, w) = 0$, et de même pour v et w.
– Si deux des vecteurs parmi u, v et w sont égaux, on voit que $\det(u, v, w) = 0$; en effet, si par exemple $u = v$, par antisymétrie $\det(u, v, w) = \det(v, u, w) = -\det(u, v, w)$.
– Plus généralement, si u, v et w sont liés (par une relation linéaire non triviale : $au + bv + cw = 0$ avec $(a, b, c) \neq (0, 0, 0)$), $\det(u, v, w) = 0$. En effet, supposons par exemple que $a \neq 0$ (les autres cas étant analogues). On peut résoudre la relation linéaire en u : $u = -\frac{b}{a}v - \frac{c}{a}w$, puis injecter cette égalité dans $\det(u, v, w)$ et utiliser la linéarité de det par rapport à son premier argument :

$$\det(u, v, w) = -\frac{b}{a}\det(v, v, w) - \frac{c}{a}\det(w, v, w) \ ;$$

mais chacun des deux termes du membre de droite valent 0 d'après la remarque précédente.

Nous allons voir que ces règles de calcul suffisent pour calculer $\det(u, v, w)$, sans ambiguïté.

Mais nous supposerons dans ce cours qu'une telle fonction det existe. (Son existence n'est en effet pas évidente : pourquoi obtient-on le même résultat en manipulant les vecteurs de différentes façons ?)

Définitions. — Considérons le tableau

$$M = (u, v, w) = \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}$$

des neufs composantes de u, v et w (on aurait aussi pu écrire les vecteurs en lignes les uns au-dessus des autres) ; un tel tableau s'appelle aussi une *matrice*.

Une *opération élémentaire* sur les colonnes de M est l'une des opérations suivantes :

– *Permutation-transposition*. On permute deux colonnes, par exemple en remplaçant u par $u' = v$ et v par $v' = u$:

$$(u, v, w) = \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix} \mapsto (v, u, w) = \begin{pmatrix} v_1 & u_1 & w_1 \\ v_2 & u_2 & w_2 \\ v_3 & u_3 & w_3 \end{pmatrix}.$$

– *Dilatation (élémentaire)*. On multiplie une colonne par un réel $a \neq 0$, par exemple en remplaçant u par $u' = au$:

$$(u, v, w) = \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix} \mapsto (au, v, w) = \begin{pmatrix} au_1 & v_1 & w_1 \\ au_2 & v_2 & w_2 \\ au_3 & v_3 & w_3 \end{pmatrix}.$$

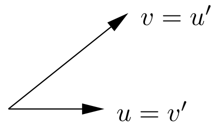
– *Transvection (élémentaire)*. On ajoute à une colonne donnée une autre colonne (distincte) multipliée par un réel a quelconque, par exemple en remplaçant u par

$u' = u + av$:

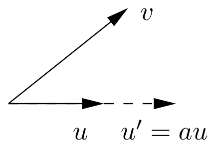
$$(u, v, w) = \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix} \mapsto (u + av, v, w) = \begin{pmatrix} u_1 + av_1 & v_1 & w_1 \\ u_2 + av_2 & v_2 & w_2 \\ u_3 + av_3 & v_3 & w_3 \end{pmatrix}.$$

(Attention : au § 3.3, on définira la matrice transposée d'une matrice ; cette opération ne doit pas être confondue avec l'opération de transposition *de deux lignes ou colonnes*, qui porte pourtant le même nom.)

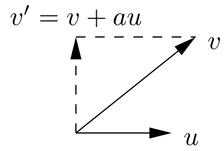
Les opérations analogues en dimension 2 se visualisent facilement.



Transposition



Dilatation



Transvection

Les opérations élémentaires sont inversibles : on peut revenir à la matrice initiale par une autre opération élémentaire, dans les exemples donnés en posant respectivement

– $u = v'$ et $v = u'$
– $u = u'/a$ (d'où l'importance ici d'avoir $a \neq 0$)
– $u = u' - av$.

Lemme. — Le volume orienté est modifié de la façon suivante par chacune des opérations élémentaires :

– $\det(v, u, w) = -\det(u, v, w)$ (antisymétrie)
– $\det(au, v, w) = a \det(u, v, w)$ (homogénéité)
– $\det(u + av, v, w) = \det(u, v, w)$ (trilinéarité et antisymétrie).

Nous allons voir une méthode systématique pour, à partir d'un triplet de vecteurs (u, v, w) quelconque, nous ramener par une suite d'opérations élémentaires à un triplet de vecteurs dont on connaît trivialement le volume (soit que ce volume s'annule, soit qu'il vaille 1 d'après la condition de normalisation).

Exercice. — Calculer $\det(v, w, u)$, $\det(w, u, v)$ en fonction de $\det(u, v, w)$.

Solution. — On a

$$\det(v, w, u) = -\det(v, u, w) = \det(u, v, w)$$

et

$$\det(w, u, v) = -\det(u, w, v) = \det(u, v, w).$$

Algorithme fang-cheng. — L'*algorithme fang-cheng* (=algorithme sur les modèles rectangulaires, en chinois) est un procédé de calcul couramment appelé *algorithme de Gauss* en Occident, d'après le nom du « prince des mathématiciens » C. F. GAUSS (1777–1855), mais découvert en Chine au IIᵉ siècle avant notre ère [**Gab01**].

(1) *Un premier cas d'arrêt de l'algorithme.* Si tous les coefficients de la première ligne de M sont nuls, les vecteurs u, v et w sont dans le plan Oyz, donc liés (exercice § 2.4), et $\det(u, v, w) = 0$.

(2) *Choix du premier pivot.* Sinon, quitte à permuter deux colonnes on peut supposer que $u_1 \neq 0$, et même, quitte à dilater u par le facteur $1/u_1$, que $u_1 = 1$:

$$M = (u, v, w) = \begin{pmatrix} \boxed{1} & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}.$$

Le coefficient à l'intersection de la première ligne et première colonne ainsi obtenu, égal à 1, s'appelle le *pivot* (pour l'opération suivante).

(3) *Élimination des coefficients à droite du pivot.* En remplaçant v et w par

$$v' = v - v_1 u \quad \text{et} \quad w' = w - w_1 u,$$

on est ramené à une matrice de la forme suivante :

$$M = \begin{pmatrix} \boxed{1} & 0 & 0 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}.$$

(4) *Deuxième cas d'arrêt de l'algorithme.* À ce stade, si $v_2 = w_2 = 0$ (en plus du fait que $v_1 = w_1 = 0$), v et w appartiennent à la droite $x = y = 0$ et sont liés, donc $\det(u, v, w) = 0$.

(5) *Itération.* Sinon, de même que précédemment, par une permutation-dilatation on se ramène à

$$M = \begin{pmatrix} \boxed{1} & 0 & 0 \\ u_2 & \boxed{1} & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}$$

et, par une transvection, à

$$M = \begin{pmatrix} \boxed{1} & 0 & 0 \\ u_2 & \boxed{1} & 0 \\ u_3 & v_3 & w_3 \end{pmatrix}.$$

(6) *Troisième cas d'arrêt de l'algorithme.* Si $w_3 = 0$, la dernière colonne est nulle donc $\det(u, v, w) = 0$.

(7) *Élimination des coefficients à gauche des trois pivots.* Sinon, $w_3 \neq 0$, soit, après dilatation de w_3, $w_3 = 1$:

$$M = \begin{pmatrix} \boxed{1} & 0 & 0 \\ u_2 & \boxed{1} & 0 \\ u_3 & v_3 & \boxed{1} \end{pmatrix}.$$

Maintenant, en remplaçant v par $v' = v - v_3 w$, puis u par $u' = u - u_2 v' - u_3 w$, on aboutit à la matrice des trois vecteurs de la base canonique :

$$M = \begin{pmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{1} \end{pmatrix},$$

dont le déterminant vaut 1.

Remarques :

– Dans tous les cas, on aboutit à un déterminant connu (0 ou 1), par une suite d'opérations élémentaires dont on connaît l'effet sur le déterminant ; donc le déterminant initial s'en déduit.

– De cet algorithme il découle que le volume orienté est une fonction polynomiale des composantes des vecteurs.
– Cet algorithme et ses variantes sont utiles dans beaucoup d'autres problèmes que le calcul des déterminants, et le lecteur aura remarqué qu'il se généralise directement à n vecteurs de $\mathbb{R}^n$.

Proposition. — Le déterminant d'une matrice triangulaire est égal au produit des termes diagonaux :

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix} = \det \begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33}.$$

En particulier, on voit que, quand on calcule un déterminant par l'algorithme fang-cheng, il est inutile de calculer les dernières transvections qui éliminent les coefficients à gauche des pivots.

Démonstration – Montrons la formule par exemple pour la matrice

$$A := \begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Si l'un des coefficients diagnonaux est nul, le déterminant est nul. En effet, c'est évident $a_{33} = 0$, puisqu'alors le troisième vecteur colonne de A est nul. Si $a_{22} = 0$ et $a_{33} \neq 0$, la transvection remplaçant A_2 par $A_2' = A_2 - \frac{a_{32}}{a_{33}} A_3$, montre que le déterminant de A vaut

$$\begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & 0 & 0 \\ a_{31} & 0 & a_{33} \end{pmatrix},$$

soit 0 puisque le deuxième vecteur-colonne de cette matrice est nul. Si enfin $a_11 = 0$ tandis que a_22 et a_{33} sont non nuls, de même par deux transvections on peut éliminer les deux coefficients restants de la première colonne, de sorte que

$$\det A = \det \begin{pmatrix} 0 & & \\ 0 & a_{22} & \\ 0 & a_{32} & a_{33} \end{pmatrix} = 0.$$

Supposons sinon que a_{11}, a_{22} et a_{33} sont tous non nuls :

$$\det A = a_{11}a_{22}a_{33} \begin{pmatrix} 1 & & \\ * & 1 & \\ * & * & 1 \end{pmatrix}.$$

La dernière étape de l'algorithme fang-cheng permet d'éliminer les trois coefficients à gauche des pivots, par trois transvections élémentaires successives. Or une transvection ne modifie pas le déterminant. Donc on a bien

$$\det A = a_{11}a_{22}a_{33}.$$

□

Formule générale de l'aire orientée d'un parallélogramme. — Soit à calculer l'aire A du parallélogramme engendré par deux vecteurs

$$v = \begin{pmatrix} a \\ b \end{pmatrix} \quad \text{et} \quad w = \begin{pmatrix} c \\ d \end{pmatrix},$$

c'est-à-dire le déterminant $\det M$ de la matrice

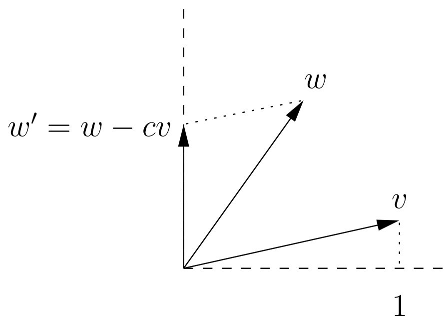
$$M = \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

L'algorithme de fang-cheng en deux dimensions est parfaitement similaire à sa version en trois dimensions.

Supposons pour commencer que $a \neq 0$, i.e. que v n'est pas vertical. Dilatons le premier vecteur d'un facteur $1/a$:

$$A = a \det \begin{pmatrix} \boxed{1} & c \\ b/a & d \end{pmatrix}.$$

Remplaçons w par $w - cv$:



en utilisant la linéarité par rapport à w et le fait que $\det(v, v) = 0$:

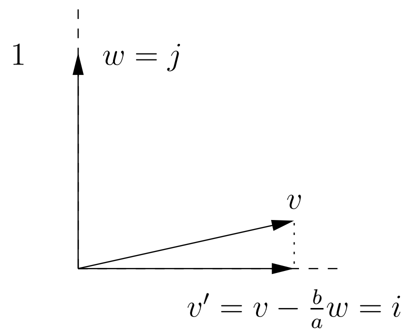
$$A = a \det \begin{pmatrix} 1 & 0 \\ b/a & d - bc/a \end{pmatrix}.$$

Si $d - bc/a = 0$, la seconde colonne est nulle et $A = 0$.

Sinon,

$$A = a\left(d - \frac{bc}{a}\right) \det \begin{pmatrix} 1 & 0 \\ b/a & 1 \end{pmatrix},$$

soit, après la transvection qui substitue $v - \frac{b}{a}w$ à v :



on obtient

$$A = (ad - bc) \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = ad - bc.$$

On voit que cette formule donne encore le bon résultat, $A = 0$, quand $d - b/a = 0$.

Par ailleurs, on avait exclu le cas où $a = 0$. Nous sommes en effet passé par des intermédiaires de calcul qui n'avaient pas de sens dans ce cas. Mais, dans le cours d'analyse, on montre que les fonctions polynomiales sont continues. Comme le déterminant est polynomial en a, b, c et d, en particulier il dépend continûment de a. Donc l'expression $ad - bc$, qui est continue par rapport à a, reste valable quand $a = 0$.

Finalement, dans tous les cas on a

$$\det \begin{pmatrix} a & c \\ b & d \end{pmatrix} = ad - bc.$$

Exercice. — Soit $z : [-1, 1] \to \mathbb{C}$ une courbe paramétrée assez dérivable. Soit $A(t)$ l'aire balayée par le vecteur $z(t)$ à partir du temps $t = 0$. Montrer que sa dérivée vérifie

$$2A' = \det(z, z') = \Im\left(\bar{z}z'\right);$$

le membre de droite est appelé *moment cinétique* en Mécanique et *wronskien* en Théorie des équations différentielles du second ordre. On pourra utiliser la formule de Taylor du premier ordre pour $z(t)$.

Supposons maintenant que $z(t)$ vérifie une équation de Newton à force centrale, c'est-à-dire une équation de la forme

$$z''(t) = f(z(t))\, z(t),$$

où $f :]0, +\infty[\to \mathbb{R}$ est une fonction donnée. Montrer que dans ce cas l'aire balayée par $z(t)$ croît linéairement avec le temps. Ce comportement remarquable, qui dicte la loi horaire des planètes autour du Soleil, s'appelle la *deuxième loi de Kepler*. Il fut découvert au début du XVIIe siècle par le mathématicien, astronome et astrologue allemand J. KEPLER (1571–1630), dans son étude du mouvement de Mars autour du Soleil, avant d'être déduit des équations de la Mécanique par I. NEWTON.

Solution. — La formule de Taylor du premier ordre appliquée à la courbe $z(t)$ entre les instants t et $t + dt$ montre que

$$z(t + dt) = z(t) + z'(t)\, dt + o(dt).$$

L'aire du triangle (=demi- parallélogramme) de côtés $z(t)$ et $z(t + dt)$ vaut donc

$$A(t + dt) - A(t) = \frac{1}{2} \det\left(z(t), z(t) + z'(t)\, dt + o(dt)\right) = \frac{1}{2} \det(z(t), z'(t))\, dt + o(dt).$$

Donc

$$A'(t) = \lim_{dt \to 0} \frac{A(t + dt) - A(t)}{dt} = \frac{1}{2} \det(z(t), z'(t)) = \frac{1}{2} \Im(\bar{z}(t) z'(t)).$$

Faisons maintenant l'hypothèse sur l'équation de Newton. Il s'agit de montrer que la dérivée seconde $A''(t)$ est identiquement nulle. On a

$$2A''(t) = \det(z'(t), z'(t)) + \det(z(t), z''(t)).$$

Le premier terme est toujours nul, et le second l'est parce que l'accélération $z''(t)$ est parallèle à la position $z(t)$.

Formule du développement par rapport à une colonne. —

$$\det \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix} = w_1 \det \begin{pmatrix} u_2 & v_2 \\ u_3 & v_3 \end{pmatrix} - w_2 \det \begin{pmatrix} u_1 & v_1 \\ u_3 & v_3 \end{pmatrix} + w_3 \det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix}.$$

Le facteur qui apparaît devant w_r, $r = 1, 2, 3$, est le déterminant de la matrice obtenue à partir de la matrice de départ en supprimant la ligne et la colonne de w_r, affecté d'un signe qui change à chaque terme.

Par antisymétrie, on en déduit des formules analogues pour développer le déterminant par rapport aux deux autres colonnes.

Ces formules sont commodes à utiliser notamment quand la matrice possède un ou plusieurs coefficients nuls. Mais on tâchera de ne pas se tromper dans les signes ; le savant allemand G. W. Leibniz (1646–1716) lui-même se trompa, dans un mémoire de 1678, certes pour les 24 termes du déterminant d'une matrice 4×4 [**Gab01**, p. 56].

Démonstration – Notons Δ le déterminant à calculer. On a

$$w = w_1 i + w_2 j + w_3 k.$$

Par linéarité de Δ par rapport à w,

$$\Delta = w_1 \det(u, v, i) + w_2 \det(u, v, j) + w_3 \det(u, v, k),$$

où

$$\det(u, v, i) = \det \begin{pmatrix} u_1 & v_1 & 1 \\ u_2 & v_2 & 0 \\ u_3 & v_3 & 0 \end{pmatrix} = \det \begin{pmatrix} 0 & 0 & 1 \\ u_2 & v_2 & 0 \\ u_3 & v_3 & 0 \end{pmatrix}, \quad \text{etc.}$$

Mais ce dernier déterminant est le volume orienté d'un parallélépipède rectangle dont la base a pour aire orientée

$$\det \begin{pmatrix} u_2 & v_2 \\ u_3 & v_3 \end{pmatrix}$$

et une hauteur unité ; donc

$$\det \begin{pmatrix} 0 & 0 & 1 \\ u_2 & v_2 & 0 \\ u_3 & v_3 & 0 \end{pmatrix} = \det \begin{pmatrix} u_2 & v_2 \\ u_3 & v_3 \end{pmatrix}$$

(cette égalité peut aussi être obtenue en calculant chacun des deux déterminants par l'algorithme fang-cheng et en constant que le résultat obtenu, ainsi d'ailleurs que les opérations élémentaires effectuées, sont identifques). Les deux autres termes se calculent de la même façon. □

Exercice. — Calculer le volume V du parallélépipède de sommets

$$A = (1, 1, 1),\ B = (2, 3, 4),\ C = (3, 4, 2) \quad \text{et} \quad D = (4, 2, 3).$$

Ces quatre points sont-ils dans un même plan ?

Solution. — Le volume orienté du parallélépipède engendré par les vecteurs $\overrightarrow{AB}$, $\overrightarrow{AC}$ et $\overrightarrow{D}$ vaut

$$\pm V = \det \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix} = -18,$$

donc $V = 18$. Comme ce volume est non nul, les quatre points ne sont pas coplanaires.

Remarque. — Soient $\phi : U \to V$ une transformation inversible entre deux régions de $\mathbb{R}^3$, et $f : U \to \mathbb{R}$ une fonction continue. Recouvrons U d'un nombre fini de petits parallélépipèdes rectangles de sommets $x^1, ..., x^n$, de même volume $dx_1\, dx_2\, dx_3$. Dans le cours d'intégration, on définit l'intégrale de f sur V comme la limite des sommes

$$\sum f(x^i)\, dx_1\, dx_2\, dx_3,$$

quand la taille des parrallélépipèdes tend vers 0.

Sous certaines conditions de régularité sur ϕ, on montre alors la formule du changement de variable suivante :

$$\int_V f\left(\phi^{-1}(y)\right)\, dy_1\, dy_2\, dy_3 = \int_U f(x)\, |\det d\phi(x)|\, dx_1\, dx_2\, dx_3,$$

où $d\phi(x)$ désigne la matrice des dérivées partielles de la transformation ϕ, par rapport aux trois coordonnées (les deux autres étant fixées) :

$$d\phi = \left(\frac{\partial\phi}{\partial x_1}, \frac{\partial\phi}{\partial x_3}, \frac{\partial\phi}{\partial x_3}\right) = \begin{pmatrix} \partial_1\phi_1 & \partial_2\phi_1 & \partial_3\phi_1 \\ \partial_1\phi_2 & \partial_2\phi_1 & \partial_3\phi_2 \\ \partial_1\phi_3 & \partial_2\phi_1 & \partial_3\phi_3 \end{pmatrix}.$$

Cette formule est naturelle. En effet, d'après la formule de Taylor,

$$\phi(x + dx) = \phi(x) + d\phi(x) \cdot dx + \cdots,$$

donc un petit parallélépipède de volume orienté $dx_1\, dx_2\, dx_3$ est envoyé par ϕ, au premier ordre, sur un parallélépipède de volume orienté

$$\det d\phi(x)\, dx_1\, dx_2\, dx_3.$$

2.6. Produit vectoriel

Définition. — Le *produit vectoriel* de deux vecteurs $u, v \in \mathbb{R}^3$ est l'unique vecteur, noté $u \wedge v$, tel que pour tout vecteur $w \in \mathbb{R}^3$ on ait

$$\det(u, v, w) = (u \wedge v) \cdot w.$$

Démonstration de l'existence et de l'unicité de $u \wedge v$ – Unicité. Appliquons successivement la formule voulue à $w = i$, j et k, en notant $u \wedge v = (x_1, x_2, x_3)$. D'après la formule de développement du déterminant par rapport à la dernière colonne,

$$x_1 = \det\begin{pmatrix} u_2 & v_2 \\ u_3 & v_3 \end{pmatrix}, \quad x_2 = -\det\begin{pmatrix} u_1 & v_1 \\ u_3 & v_3 \end{pmatrix} \text{ et } \quad x_3 = \det\begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix},$$

soit

$$x_1 = u_2v_3 - u_3v_2, \quad x_2 = u_3v_1 - u_1v_3 \quad \text{et} \quad x_3 = u_1v_2 - u_2v_1.$$

Donc $u \wedge v$ est unique.

Existence. Définissons $u \wedge v = (x_1, x_2, x_3)$ par les formules précédentes. D'après la formule du développement du déterminant par rapport à la troisième colonne, pour tout vecteur w on a bien $\det(u, v, w) = (u \wedge v) \cdot w$, ce qui montre l'existence d'un vecteur satisfaisant l'égalité voulue. □

Remarque. Il peut sembler bizarre dans la démonstration d'existence et d'unicité d'un objet mathématique quelconque de commencer par montrer son unicité quand on ne connaît pas encore son existence. De ce point de vue, il serait plus naturel d'inverser l'ordre de présentation. Mais la démonstration de l'unicité permet d'analyser ce que doit être le vecteur $u \wedge v$ (par conditions nécessaires), et la démonstration de l'existence fait la synthèse de cette analyse. Du point de vue de la logique, tout y est, malgré l'ordre inhabituel, et cela raccourcit la rédaction. Ce type de preuve par *analyse-synthèse* est très courant.

Corollaire. — Le produit vectoriel est bi-linéaire :

$$\begin{cases} (au + a'u') \wedge v = a(u \wedge v) + a'(u' \wedge v) \\ u \wedge (bv + b'v') = bu \wedge v + b'u \wedge v' \end{cases}$$

et antisymétrique :

$$u \wedge v + v \wedge u = 0.$$

Ceci découle directement de la trilinéarité et de l'antisymétrie du déterminant.

Exercice. — Montrer que, si u et v sont deux vecteurs du plan horizontal $z = 0$ de $\mathbb{R}^3$,

$$u \wedge v = \det(u, v)\ k.$$

Solution. — Notons $z(u, v) = (u \wedge v) \cdot k$. On veut montrer que $z(u, v) = \det(u, v)$. D'après le § 2.6, z est une fonction $\mathbb{R}^2 \to \mathbb{R}$ bilinéaire antisymétrique. De plus, $z(i, j) = 1$. Donc z satisfait aux axiomes de définition de l'aire orientée (§ 2.5). Par unicité (admise) de l'aire orientée, on a bien $z(u, v) = \det(u, v)$. Voir aussi l'exercice § 2.6.

Lemme : formule du double produit vectoriel. —

$$(u \wedge v) \wedge w = (u \cdot w)v - (v \cdot w)u$$

Non démonstration – En écrivant $u = u_1 i + u_2 j + u_3 k$, et de même pour v et w, on se ramène à vérifier la formule pour les vecteurs de la base canonique, soit $3^3 = 27$ vérifications, que nous laissons à la charge du lecteur scrupuleux. $\square$

Plus tard dans le cours, quand nous aurons étudié les applications orthogonales, nous pourrons ramener la démonstration à deux vérifications seulement.

Proposition. — Le produit vectoriel $u \wedge v$ est nul si et seulement si u et v sont liés. Sinon, c'est un vecteur orthogonal à u et v, et

$$\|u \wedge v\| = \|u\|\, \|v\|\, |\sin(\widehat{u, v})|.$$

Démonstration – Si u et v sont liés, $\det(u, v, w) = 0$ pour tout w donc $u \wedge v = 0$. Réciproquement, si $u \wedge v = 0$, les formules en composantes ci-dessus montrent que u et v sont liés.

Supposons maintenant $u \wedge v \neq 0$. Par définition,

$$(u \wedge v) \cdot u = \det(u, v, u) = 0,$$

donc $u \wedge v$ est orthogonal à u et, pour la même raison, à v aussi.

Quant à la dernière formule :

$$\begin{aligned}
\|u \wedge v\|^2 &= \det(u, v, u \wedge v) \quad \text{(définition du produit vectoriel)} \\
&= \det(u \wedge v, u, v) \quad \text{(antisymétrie)} \\
&= ((u \wedge v) \wedge u) \cdot v \quad \text{(définition du produit vectoriel)} \\
&= (u^2 v - (u \cdot v) u) \cdot v \quad \text{(formule du double produit)} \\
&= u^2 v^2 - (u \cdot v)^2 \\
&= (1 - \cos^2(\widehat{u, v})) u^2 v^2 \quad \text{(formule § 2.4)} \\
&= \sin^2(\widehat{u, v})) u^2 v^2.
\end{aligned}$$

□

Exemple. — Trouver l'équation de la droite D normale à $x + y + z = 0$ passant par $p = (1, 2, 3)$.

Le plan $P : x + y + z = 0$ est le plan orthogonal de $v = (1, 1, 1)$. On cherche donc l'équation de la droite passant par p et dirigée par v. Une solution serait, comme on l'a déjà fait, d'écrire un paramétrage de la droite, puis d'éliminer le paramètre. Une solution plus directe, qui utilise le produit vectoriel, consiste à remarquer que D est l'ensemble des $q = (x, y, z)$ tels que $q - p$ est parallèle à v, i.e.

$$(q - p) \wedge v = 0, \quad q - p = \begin{pmatrix} x - 1 \\ y - 2 \\ z - 3 \end{pmatrix},$$

i.e.

$$\begin{pmatrix} (y-2) - (z-3) \\ (z-3) - (x-1) \\ (x-1) - (y-2) \end{pmatrix} = \begin{pmatrix} y - z + 1 \\ z - x - 2 \\ x - y + 1 \end{pmatrix} = 0.$$

La petite subtilité ici est que les trois équations scalaires obtenues ne sont pas indépendantes (parce que le produit vectoriel de $q - p$ et de v est a priori orthogonal à v) : la somme de ces trois équations est nulle, en effet.

Donc le système de ces équations est équivalent au système obtenu en ne gardant par exemple que les deux premières équations :

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

Exercice. — Démontrer la formule de l'exercice § 2.6 et interpréter $\|u \wedge v\|$ comme un volume.

Solution. — $u \wedge v$ est perpendiculaire à u et v, donc parallèle à k. Donc

$$\begin{aligned}
u \wedge v &= ((u \wedge v) \cdot k)\, k \\
&= \det(u, v, k)\, k.
\end{aligned}$$

Donc $\|u \wedge v\| = |\det(u, v, k)|$ est le volume du parallélépipède engendré par u, v et le vecteur unitaire orthogonal k.

Définition. — Soit $\gamma : I^2 = [0,1]^2 \to \mathbb{R}^3$ une surface paramétrée de classe C^1. Ceci signifie que (chaque composante de) γ est dérivable par rapport à chacun de ses deux arguments (l'autre argument étant gelé), et que ses *dérivées partielles*, soit

$$\frac{\partial\gamma}{\partial s}(s,t) := \lim_{\sigma\to 0}\frac{\gamma(s+\sigma,t)-\gamma(s,t)}{\sigma} \quad \text{et} \quad \frac{\partial\gamma}{\partial t}(s,t) := \lim_{\tau\to 0}\frac{\gamma(s,t+\tau)-\gamma(s,t)}{\tau},$$

sont continues. Soit aussi f une fonction continue définie sur la surface $\Gamma := \gamma(I^2)$; en physique, f est par exemple une masse par unité de surface.

L'*intégrale de surface* de f le long de la surface Γ est

$$\int_\Gamma f := \int_0^1\int_0^1 f(\gamma(s,t))\left\|\frac{\partial\gamma}{\partial s}(s,t)\wedge\frac{\partial\gamma}{\partial t}(s,t)\right\| ds\,dt.$$

Exemple. — L'*aire* de Γ est l'intégrale curviligne

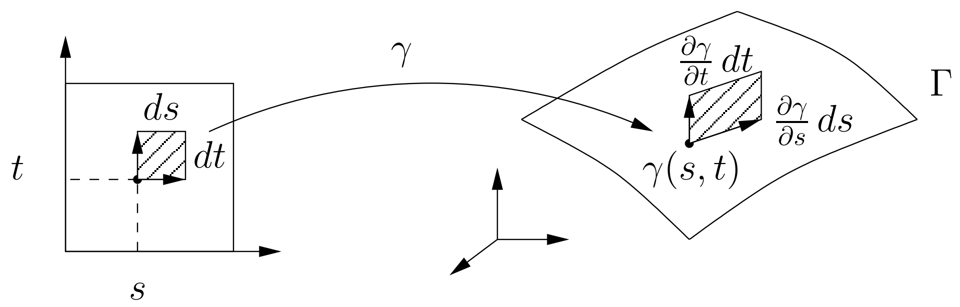
$$A(\Gamma) := \int_\Gamma 1 := \int_0^1\int_0^1 \left\|\frac{\partial\gamma}{\partial s}(s,t)\wedge\frac{\partial\gamma}{\partial t}(s,t)\right\| ds\,dt.$$

Cette formule est intuitive. En effet, d'après la formule de Taylor (pour les fonctions de deux variables),

$$\gamma(s+ds,t+dt) = \gamma(s,t) + \frac{\partial\gamma}{\partial s}(s,t)\,ds + \frac{\partial\gamma}{\partial t}(s,t)\,dt + o(ds,dt).$$

Au premier ordre, un petit rectangle d'aire $ds\,dt$ dans le plan des paramètres (s,t) est envoyé par γ sur un petit parallélogramme engendré par les vecteurs

$$\frac{\partial\gamma}{\partial s}(s,t)\,ds \quad \text{et} \quad \frac{\partial\gamma}{\partial t}(s,t)\,dt.$$



L'aire de ce petit parallélogramme vaut

$$\left|\det\left(\frac{\partial\gamma}{\partial s}(s,t)\,ds, \frac{\partial\gamma}{\partial t}(s,t)\,dt\right)\right| = \left\|\frac{\partial\gamma}{\partial s}(s,t)\wedge\frac{\partial\gamma}{\partial t}(s,t)\right\| ds\,dt$$

et l'intégrale ci-dessus n'est que la somme de ces contributions, à la limite quand ds et dt tendent vers 0 (cette limite étant justifiée dans le cours d'intégration).

Exercice. — Donner une formule pour l'aire du graphe d'une fonction $f : D \to \mathbb{R}$ de classe C^1, où D est une région de $\mathbb{R}^2$.

Solution. — Le graphe de f est la surface de $\mathbb{R}^3$ d'équation $z = f(x,y)$. Il est donc l'image du paramétrage naturel

$$\gamma : D \to \mathbb{R}^3, \quad (x,y) \mapsto (x,y,f(x,y)).$$

On a

$$\frac{\partial\gamma}{\partial x} = \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x}(x,y)\end{pmatrix} \quad \text{et} \quad \frac{\partial\gamma}{\partial y} = \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y}(x,y)\end{pmatrix},$$

donc

$$\left\|\frac{\partial\gamma}{\partial s}(s,t)\wedge\frac{\partial\gamma}{\partial t}(s,t)\right\| ds\,dt = \sqrt{1+\frac{\partial f}{\partial x}(x,y)^2+\frac{\partial f}{\partial y}(x,y)^2},$$

et l'aire du graphe de f vaut donc

$$A(\Gamma) = \int_D \sqrt{1+\frac{\partial f}{\partial x}(x,y)^2+\frac{\partial f}{\partial y}(x,y)^2}\, dx\, dy.$$

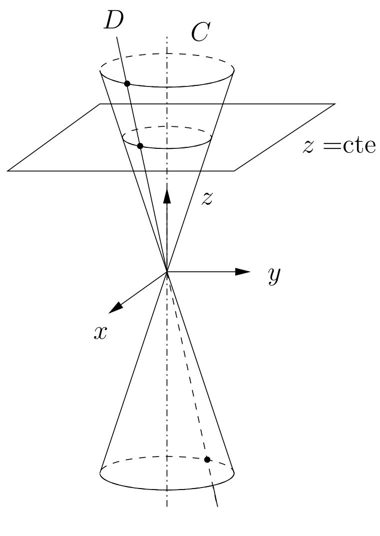
Complément : Sections coniques

Définition. — Soient D une droite vectorielle de $\mathbb{R}^3$ contenue dans le plan $y=0$. Soit θ l'angle de D avec l'axe Oz ($k\cdot v = \pm\cos\theta$, $0\le\theta\le\pi/2$), de sorte que D ait pour équation

$$x = z\tan\theta, \quad y = 0.$$

On suppose que D n'est ni verticale ni horizontale : $0<\theta<\pi/2$.

Le *cône de révolution* C de sommet 0 engendré par D est la surface dont la trace sur chaque plan $z = cte$ (i.e. l'intersection avec ce plan) est le cercle horizontal centré en $(0,0,z)$ et passant par le point où D coupe le plan.



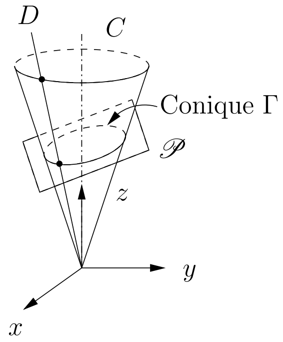
Exercice. — Quelle est l'équation du cône C de sommet O et engendré par la droite $D : x = z\tan\theta$, $y=0$?

Solution. — Le cercle d'intersection de C et du plan d'altitude z a pour rayon $|x| = |z|\tan\theta$, et donc pour équation $x^2+y^2 = z^2\tan^2\theta$. Donc le cône lui-même a pour équation, dans $\mathbb{R}^3$,

$$x^2+y^2 = z^2\tan^2\theta.$$

Définition. — Une *(section) conique* Γ est l'intersection d'un cône de révolution et d'un plan affine quelconque $\mathscr{P}$.

Les *foyers* de Γ sont les points de contact de $\mathscr{P}$ avec chacune des deux sphères inscrites dans C et tangentes à $\mathscr{P}$.



Exercice. — Trouver l'équation d'une conique dans les coordonnées polaires de son plan relatives à l'un de ses deux foyers.

Autres compléments

Structure d'espace vectoriel. — Les opérations d'addition et de multiplication par un réel forment une structure d'espace vectoriel (réel) sur $\mathbb{R}^3$. Abstraitement, une telle structure est définie par les axiomes suivants :
- L'addition est une loi de groupe commutatif
- Le scalaire 1 est neutre : $1v = v$
- Distributivité : $a(v+w) = av + aw$
- Exo-distritributivité : $(a+b)v = av + bv$
- Exo-associativité : $a(bv) = (ab)v$.

Exercice. — Dans $\mathbb{R}^4$ on définit l'addition et la multiplication par un réel de façon analogue. Montrer qu'il existe une unique multiplication (i.e. une application $\mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}^4$, $(x, y) \mapsto xy$), qui soit

– bilinéaire : $(ax + a'x')y = axy + a'x'y$ et $x(by + b'y') = bxy + b'xy'$,

et telle que, si l'on note $\mathbb{1} = (1, 0, 0, 0)$, $i = (0, 1, 0, 0)$, $j = (0, 0, 1, 0)$ et $k = (0, 0, 0, 1)$,

– $\mathbb{1}$ soit l'élément neutre

– et $ij = k,\ jk = i,\ ki = j,\ i^2 = j^2 = k^2 = -1$, et $xy + yx = 0$ pour tous $x, y \in \{i, j, k\}$.

Montrer que l'ensemble $\mathbb{R}^4$ muni de ces opérations est un corps ; comme pour $\mathbb{C}$, on pourra utiliser la *conjugaison*

$$\overline{a + ib + jc + kd} = a - ib - jc - kd$$

et remarquer que $x\bar{x} \neq 0$ si $x \neq 0$.

L'ensemble $\mathbb{R}^4$ muni de ces opération se note $\mathbb{H}$ et s'appelle le *corps des quaternions*. Il fut découvert par W. R. Hamilton (mathématicien, physicien et astronome irlandais, 1805–1865). Pour avoir une idée de comment les quaternions peuvent être utilisés pour étudier les rotations dans l'espace réel à trois dimensions, voir [**Gra97**].

Solution. — En développant les expressions de la forme

$$(a + ib + jc + kd)(a' + ib' + jc' + kd')$$

on voit qu'on peut se ramener à une somme de termes obtenus par l'un des produits spécifiés ci-dessus. Ceci définit bien une unique multiplication. Le fait que cette dernière satisfasse les axiomes de définition d'un corps est une question de simple vérification, mis à part l'inversibilité des éléments non nuls. Pour un tel $x = a + ib + jc + kd \neq 0$, on vérifie que

$$x\bar{x} = a^2 + b^2 + c^2 + d^2 \neq 0,$$

donc que x est bien inversible d'inverse

$$x^{-1} = \frac{\bar{x}}{x\bar{x}}.$$

Structure d'espace affine. — Les droites et plans vectoriels définis ci-dessus sont des espaces vectoriels (cf. § 2.6).

Les droites et plans affines sont des exemples de la structure algébrique suivante. Une structure d'*espace affine* dirigé par un espace vectoriel E sur un ensemble $\mathscr{E}$ est une application $\mathscr{E} \times \mathscr{E} \to E$ notée $(p, q) \mapsto \overrightarrow{pq}$, jouissant des propriétés suivantes :

– pour tout $p \in \mathscr{E}$, l'application $q \mapsto \overrightarrow{pq}$ est bijective
– relation de Chasles : pour tous $p, q, r \in \mathscr{E}$, $\overrightarrow{pq} + \overrightarrow{qr} + \overrightarrow{rp} = 0_E$ (du nom de M. Chasles, mathématicien français, 1793–1880).

Exercice : Montrer qu'un espace vectoriel possède une structure d'espace affine naturelle.

Habituellement, les éléments de $\mathbb{R}^3$ se nomment *points* ou *vecteurs* selon que l'on se réfère aux structures affine ou vectorielle [**Gra97**].

Démonstration par Newton de la loi des aires. — Considérons un *mouvement à force centrale*, i.e. une solution d'une équation différentielle de la forme

$$z'' = f(z)\, z, \quad z \in \mathbb{C},$$

où f est une fonction de z à valeurs réelles. Nous avons vu que l'aire balayée par le vecteur $z \in \mathbb{C} = \mathbb{R}^2$ croît linéairement avec le temps (§ 2.5). Nous allons en voir la célèbre démonstration originale de Newton, telle qu'elle est expliquée dans le Livre I, section II des *Principes mathématiques de philosophie naturelle.*

Considérons deux intervalles de temps $[t, t + dt]$ et $[t + dt, t + 2dt]$. Quand dt tend vers 0, d'après la formule de Taylor justifiée dans le cours d'Analyse, la vitesse $z'(t + dt)$ à l'instant $t + dt$ vaut

$$z'(t + dt) = z'(t) + z''(t)dt + o(dt).$$

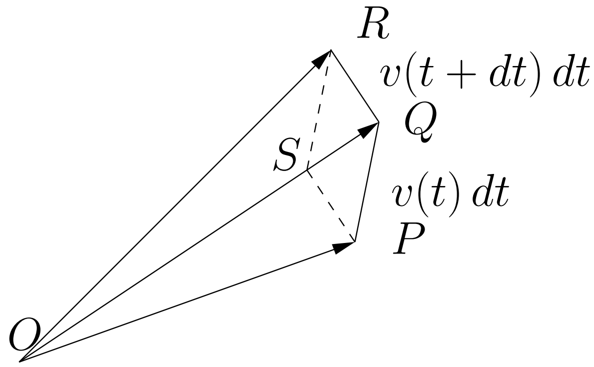
Newton eu l'idée décrire une variante de cette formule, toute aussi correcte, moins naturelle du point de vue de la formule de Taylor, mais beaucoup plus utile ici :

$$z'(t + dt) = z'(t) + z''(t + dt)dt + o(dt).$$

En utilisant l'équation différentielle, on voit alors que si l'on note $P = z(t)$ et $Q = z(t) + z'(t)\, dt$, la différence entre les vitesses aux instants t et $t + dt$ est parallèle à Q, au second ordre près :

$$z'(t + dt) = z'(t) + f(Q)\, Q\, dt + o(dt).$$

Notons de plus $R = Q + z'(t)$ et $S = Q + f(Q)\, Q$. Alors le quadrilatère $PQRS$ est un parallélogramme :



Maintenant on vérifie que les aires des triangles OPQ et OQR coïncident.

CHAPITRE 3

INCARTADE FORMELLE : MATRICES

Mots-clef du chapitre. — Matrice, addition, multiplication, système linéaire, inversion, opération élémentaire, permutation, dilatation, transvection, réduction, pivot, algorithme fang-cheng

Le terme de *matrice*, ainsi que le produit et l'inversion matricielles, furent introduit par le mathématicien britannique A. CAYLEY (1821–1895). Mais la notation matricielle avait déjà été utilisée pour la résolution de systèmes d'équations linéaires par LEIBNIZ et même, plus de deux siècles avant notre ère, par les mathématiciens chinois.

3.1. L'anneau des matrices carrées

Définitions. — Nous avons déjà rencontré des matrices à trois lignes et trois colonnes (§ 2.5). Plus généralement, soient m et n des entiers ≥ 1. Une *matrice* (de format) $m \times n$ est une collection de mn nombres arrangés en tableau :

$$\begin{array}{cc} & n \text{ colonnes} \\ m \text{ lignes} & \begin{pmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \cdots & A_{mn} \end{pmatrix} \end{array}.$$

Par exemple,

$$\begin{pmatrix} 0 & 1 & e \\ \sqrt{2} & -1 & \pi \\ 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

est une matrice 4×3.

Les nombres dans une matrice A sont les *coefficients* ; ils sont souvent notés A_{ij}, avec deux indices entiers tels que $1 \leq i \leq m$ et $1 \leq j \leq n$. Par convention, le premier indice, i, numérote les lignes, et le second, j, les colonnes. Les coefficients *diagonaux* sont ceux pour lesquels $i = j$.

On écrit parfois $A = (A_{ij})$. La matrice A peut être vue comme une application qui, à toute paire d'indice $(i, j) \in \{1, ..., m\} \times \{1, ..., n\}$ associe le nombre $A_{ij} = A(i, j)$.

On note $M_{mn}(\mathbb{R})$ l'ensemble de toutes les matrices $m \times n$ à coefficients réels.

Une matrice $1 \times n$ s'appelle une *matrice-ligne* ou un *vecteur-ligne*. On note alors plus simplement

$$A = (A_1, ..., A_n) \quad \text{au lieu de} \quad (A_{11}, ..., A_{1n}).$$

De même, une matrice $m \times 1$ s'appelle une *matrice-colonne* ou un *vecteur-colonne* et se note

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_m \end{pmatrix} \quad \text{au lieu de} \quad \begin{pmatrix} A_{11} \\ \vdots \\ A_{m1} \end{pmatrix}.$$

Cette simplification se fait au prix de gommer la distinction entre les vecteurs ligne et colonne, mais ne prête généralement pas à confusion.

Opérations matricielles. — Notons $A = (A_{ij})$ et $B = (B_{ij})$. Les deux opérations suivantes généralisent simplement les opérations introduites pour les vecteurs.

– *Addition des matrices de même dimension :*

$$A + B = (A_{ij} + B_{ij}).$$

Par exemple,

$$\begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & 5 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 3 \\ 4 & -3 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 3 \\ 5 & 0 & 6 \end{pmatrix}.$$

– *Multiplication d'une matrice par un scalaire (=nombre réel, ici) :*

$$a(A_{ij}) = (aA_{ij}).$$

Par exemple,

$$2 \begin{pmatrix} 0 & 1 \\ 2 & 3 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 4 & 6 \\ 4 & 2 \end{pmatrix}.$$

L'opération compliquée est celle de multiplication.

– *Multiplication de deux matrices $m \times n$ et $n \times p$* (le nombre de colonnes de la première matrice et le nombre de lignes de la seconde étant donc un même entier n) :

Le produit est la matrice $m \times p$ (on perd trace de la dimension intermédiaire n)

$$AB = (C_{ik})_{1 \leq i \leq m,\ 1 \leq k \leq p}$$

avec

$$C_{ik} = A_{i1}B_{1k} + \cdots + A_{in}B_{nk} = \sum_{1 \leq j \leq n} A_{ij}B_{jk},$$

le signe de sommation $\sum$ étant un raccourci de notation commode.

Autrement dit, le coefficient (i, k) du produit est le produit scalaire du i-ième vecteur-ligne de A et du k-ième vecteur-colonne de B :

$$\begin{pmatrix} \boxed{A_{i1} \quad \cdots \quad A_{in}} \end{pmatrix} \begin{pmatrix} & \boxed{\begin{matrix} B_{1k} \\ \\ \vdots \\ \\ B_{nk} \end{matrix}} & \end{pmatrix} = \begin{pmatrix} & \vdots & \\ \cdots & C_{ik} & \cdots \\ & \vdots & \end{pmatrix}.$$

Par exemple,

$$\begin{pmatrix} 0 & -1 & 2 \\ 3 & 4 & -6 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}. \tag{$*$}$$

Le produit d'une matrice et d'un vecteur-colonne peut s'écrire d'une façon encore légèrement différente, qui aura son intérêt plus tard. Notons $A_1, ..., A_n$ les vecteurs-colonnes de A. Alors, si $x = (x_1, ..., x_n) \in \mathbb{R}^n$,

$$A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1 A_1 + \cdots + x_n A_n.$$

Exercice. — Calculer

$$\begin{pmatrix} 1 & 2 \\ 4 & 5 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 0 & -1 & 2 \\ 3 & 4 & -6 \end{pmatrix}.$$

Solution. —

$$\begin{pmatrix} 1 & 2 \\ 4 & 5 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 0 & -1 & 2 \\ 3 & 4 & -6 \end{pmatrix} = \begin{pmatrix} 6 & 7 & -10 \\ 15 & 16 & -22 \\ 12 & 11 & -18 \end{pmatrix}.$$

Systèmes linéaires. — Les matrices ont été introduites au XIXe siècle pour noter les systèmes d'équations linéaires de façon abrégée. En effet, le système

$$\left\{ \begin{array}{ccccccc} A_{11}x_1 & + & \cdots & + & A_{1n}x_n & = & B_1 \\ A_{21}x_1 & + & \cdots & + & A_{2n}x_n & = & B_2 \\ \vdots & & & & \vdots & \vdots & \\ A_{m1}x_1 & + & \cdots & + & A_{mn}x_n & = & B_m \end{array} \right.$$

peut être écrit en termes de matrices :

$$AX = B,$$

avec

$$A = (A_{ij}), \quad X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} B_1 \\ \vdots \\ B_n \end{pmatrix}.$$

Par exemple, l'équation matricielle

$$\begin{pmatrix} 0 & -1 & 2 \\ 3 & 4 & -6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

représente le système suivant, à trois inconnues et deux équations :

$$\left\{ \begin{array}{ccccccc} & & -x_2 & + & 2x_3 & = & 2 \\ 3x_1 & + & 4x_2 & - & 6x_3 & = & 1, \end{array} \right.$$

dont l'équation (*) ci-dessus exhibe une solution $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 3 \end{pmatrix}$.

Propriétés des opérations. — L'addition et la multiplication par les scalaires jouissent de propriétés analogues à celles des opérations correspondantes pour les vecteurs de $\mathbb{R}^3$. L'opération nouvelle est la multiplication, qui vérifie notamment :

– Distributivité : $A(B+B') = AB + AB'$ et $(A+A')B = AB + A'B$.

– Associativité : $A(BC) = (AB)C$ (cette propriété sera évidente plus tard, mais il est toujours possible de la vérifier directement)

– Attention, la multiplication n'est pas commutative, puisque par exemple

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Ainsi, on ne peut pas déduire que $AB = CA$ à partir du fait que $B = C$!

Définitions. — Les matrices carrées ($m = n$) sont particulièrement importantes et on note $M_n(\mathbb{R}) = M_{nn}(\mathbb{R})$. De plus,

$$0 = 0_{M_n(\mathbb{R})} = (0), \quad I = I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \in M_n(\mathbb{R}).$$

La matrice I est l'élément neutre de la multiplication de $M_n(\mathbb{R})$: $AI = IA = A$ quelle que soit $A \in M_n(\mathbb{R})$.

Une *inverse* de $A \in M_n\mathbb{R}$ est une matrice $B \in M_n(\mathbb{R})$ telle que

$$AB = BA = I_n.$$

Si une telle matrice B existe (ce qui n'est pas toujours le cas), elle est unique (si B' est aussi un inverse de A, en multipliant l'égalité $AB = I$ à gauche par B' on obtient $B = B'$) ; on la note A^{-1} et on dit que A est *inversible*.

Par exemple, on peut vérifier que

$$\begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix}^{-1} = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}.$$

Nous verrons plus tard que s'il existe une matrice B telle que $AB = I$ *ou* $BA = I$, alors A est inversible et $B = A^{-1}$. Mais comme la multiplication matricielle n'est pas commutative, ceci n'est pas évident, et devient même faux pour les matrices rectangulaires.

L'ensemble des matrices $n \times n$ inversibles s'appelle le *groupe linéaire* de rang n et se note $GL_n(\mathbb{R})$.

Proposition. — Soient $A, B \in M_n(\mathbb{R})$. Si A et B sont inversibles, il en est de même de AB et $(AB)^{-1} = B^{-1}A^{-1}$.

Plus généralement, si $A_1, ..., A_m \in M_n(\mathbb{R})$ sont inversibles, il en est de même de $A_1 \cdots A_m$ et son inverse est ${A_m}^{-1} \cdots {A_1}^{-1}$.

Démonstration – En effet, grâce à l'associativité de la multiplication,

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = I$$

et de même

$$(B^{-1}A^{-1})(AB) = B^{-1}(A^{-1}A)B = I.$$

Démontrons la seconde assertion par récurrence sur m. L'affirmation est vraie pour $m = 1$, par définition de l'inversibilité. Supposons maintenant que l'affirmation est vraie pour $m = k$, et déduisons-la pour $m = k + 1$. Soient $A_1, ..., A_{k+1}$ des matrices inversibles. Notons $P = A_1 \cdots A_k$ le produit des k premières matrices. D'après l'hypothèse de récurrence, P est inversible et $P^{-1} = {A_k}^{-1} \cdots {A_1}^{-1}$. Par ailleurs, $A_1 \cdots A_{k+1} = PA_{k+1}$ est le produit de deux matrices inversibles, donc, d'après la première assertion de la proposition, $A_1 \cdots A_{k+1}$ est inversible d'inverse

$$ {A_{k+1}}^{-1} P^{-1} = {A_{k+1}}^{-1} {A_k}^{-1} \cdots {A_1}^{-1}, $$

et l'affirmation est vraie au rang $m = k + 1$. □

3.2. Classification des systèmes linéaires

Théorème de réduction des systèmes linéaires. — Le système linéaire du § 3.1 est équivalent à un unique système linéaire $A'X = B'$ (au sens que ces systèmes ont le même ensemble de solutions), où A' est de la forme

$$ A' = \begin{pmatrix} 1 & * & * & 0 & * & * & 0 & ... \\ & & & 1 & * & * & 0 & ... \\ & & & & & & 1 & ... \\ & & & & & & & ... \end{pmatrix}, $$

(les coefficients non écrits sont nuls et les étoiles sont des coefficients inconnus), i.e.

(1) sur chaque ligne le premier coefficient non nul vaut 1 (il est appelé un *pivot*) ;

(2) le pivot d'une ligne est strictement à droite du pivot des lignes précédentes ;

(3) les coefficients au-dessus d'un pivot sont nuls.

La matrice A' s'appelle la *forme réduite* de A.

Une inconnue est *essentielle* si sa colonne possède un pivot, et *libre* sinon.

Exemple. — Dans un système réduit d'inconnues (x, y, z) dont le membre de gauche est défini par la matrice

$$ \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix}, $$

x et y sont essentielles et z est libre. En effet, z peut être choisi librement, tandis que les deux premières équations permettent d'exprimer x et y en fonction de z. Si par exemple le membre de droite du système est $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, on obtient

$$ x = 1 - z \quad \text{et} \quad y = -2z. $$

Si le membre de droite était $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, le système serait incompatible et n'aurait aucune solution.

Démonstration du théorème. — Nous démontrerons l'existence, mais pas l'unicité. L'existence se démontre avec l'algorithme fang-cheng (ou algorithme du pivot de Gauss). Nous avons déjà rencontré cet algorithme dans le calcul des déterminants en petite dimension § 2.5.

Notons

$$A = \begin{pmatrix} A_{11} & \dots & A_{1n} \\ \vdots & & \vdots \\ A_{m1} & \dots & A_{mn} \end{pmatrix}$$

la matrice du système (nous n'utiliserons pas explicitement le membre de droite). On peut supposer cette matrice non nulle.

(1) *Choix du pivot.* Si tous les coefficients de la première colonne sont nuls, considérer la sous-matrice obtenue en supprimant la première colonne, et recommencer jusqu'à l'obtention d'une première colonne non nulle.

Ensuite, quitte à permuter une ligne avec la première on peut supposer que $A_{11} \neq 0$, et même, quitte à diviser la première ligne par A_{11}, que $A_{11} = 1$: la matrice devient

$$\begin{pmatrix} 1 & \dots & A_{1n} \\ A_{21} & \dots & A_{2n} \\ \vdots & & \vdots \\ A_{m1} & \dots & A_{mn} \end{pmatrix}.$$

(2) *Élimination des coefficients sous le pivot.* Poser $\ell_i' = \ell_i - A_{i1}\ell_1$ pour tout $i = 2...m$ ramène à une matrice de la forme

$$\begin{pmatrix} 1 & \dots & A_{1n} \\ 0 & \dots & A_{2n} \\ \vdots & & \vdots \\ 0 & \dots & A_{mn} \end{pmatrix}.$$

(3) *Récurrence.* Considérons la sous-matrice

$$\begin{pmatrix} A_{22} & \dots & A_{2n} \\ \vdots & & \vdots \\ A_{m2} & \dots & A_{mn} \end{pmatrix}.$$

obtenue en négligeant la première ligne et la première colonne. Si elle est de dimension non nulle (i.e. elle possède au moins une ligne et une colonne), on recommence à la première étape avec cette sous-matrice. Sinon on continue.

Ainsi, à chaque fois qu'on itère les étapes 1-3, la matrice perd une ligne et une colonne. Au bout d'un nombre fini d'étape (au plus $\min(m, n)$), on sort de la boucle.

(4) *Élimination des coefficients au-dessus du pivot.* Pour chaque pivot de position (i_0, j_0) (considérés par ordre croissant de i_0) on pose $\ell_i' = \ell_i - A_{ij_0}\ell_{i_0}'$ $(i = 1, ..., i_0 - 1)$.

Exercice. — Soit a un nombre réel. Résoudre le système suivant en fonction de a :

$$\begin{cases} ax + y + z + t = 0 \\ x + 2y + z + t = 0 \\ x + y + z + t = 0. \end{cases}$$

Solution. — Notons E_a l'ensemble des solutions du système. Le membre de gauche du système est associé à la matrice

$$A = \begin{pmatrix} a & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

(1) On cherche la forme réduite de A.

$$A \sim \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ a & 1 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1-a & 1-a & 1-a \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1-a & 1-a \end{pmatrix}$$

(où le signe $\sim$ est utilisé entre deux matrices obtenues l'une à partir de l'autre par une ou plusieurs opérations élémentaires).

– Pour $a = 1$, on a

$$A \sim \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

On en déduit l'existence de deux variables libres (et deux variables essentielles correspondant aux positions de pivot) ; géométriquement, l'ensemble des solutions est un plan vectoriel de $\mathbb{R}^4$.

– Pour $a \neq 1$, on a

$$A \sim \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Il y a cette fois trois variables essentielles et une variable libre ; géométriquement, l'ensemble des solutions est une droite vectorielle.

(2) Pour $a = 1$:

$$E_1 = \{(-z-t, 0, z, t) \in \mathbb{R}^4, \quad z \in \mathbb{R}, t \in \mathbb{R}\}$$

et, pour $a \neq 1$,

$$E_a = \{(0, 0, -t, t) \in \mathbb{R}^4, t \in \mathbb{R}\}.$$

Exercice. — Montrer que l'ensemble des solutions d'un système linéaire à cinq inconnues complexes ne peut pas avoir exactement 7 éléments.

Solution. — Le nombre N de solutions d'un système se voit facilement sur sa forme réduite (que les coefficients soient réels ou complexes). Notons A la matrice du membre de gauche et B le vecteur colonne du membre de droite, après réduction.

– Nous avons la trichotomie suivante :

Supposons d'abord que A possède une ligne nulle et que le coefficient de B correspondant à cette ligne soit non nul. On est dans la situation ou l'une des équations est, par exemple, signifie qu'on a une équation comme

$$0x + 0y + 0z = 36.$$

Cette équation n'a pas de solution, donc le système non plus : $N = 0$.

Subdivisons le cas contraire en deux sous-cas :

– S'il existe un pivot par colonne, toutes les inconnues sont essentielles et il existe une solution unique : $N = 1$.

– Sinon, certines inconnues sont libres et il existe autant d'infinités de solutions : $N = \infty$.
Donc un système linéaire possède 0, 1 ou une infinité de solutions ; il est impossible que $N = 7$.

Corollaire. — Soit $A \in M_n(\mathbb{R})$ une matrice carrée. Les propriétés suivantes sont équivalentes :

(1) La matrice réduite de A est l'identité

(2) Quel que soit $B \in \mathbb{R}^n$, l'équation $Ax = B$ possède une solution unique

(3) A est inversible

(4) La matrice réduite de A possède un pivot par ligne

(5) La matrice réduite de A possède un pivot par colonne

(6) $\det A \neq 0$.

Nous laissons au lecteur le plaisir de définir le déterminant d'une matrice carrée A de dimension quelconque, sur le modèle de ce qui a été fait précédemment en dimension 2 ou 3.

Démonstration – Notons A' la matrice réduite de A.

(1) $\Leftrightarrow$ (2) : Montrons d'abord l'*implication contraposée* non (2) $\Rightarrow$ non (1) de l'implication (1) $\Rightarrow$ (2) (ces deux implications sont équivalentes). Si A' n'est pas l'identité, A' ne possède pas un pivot par colonne. Donc le système $Ax = B$ possède au moins une inconnnue libre, et l'ensemble de ses solutions est donc 0 ou l'infini.

Réciproquement, montrons l'implication non (1) $\Rightarrow$ non (2). Comme A' n'est pas l'identité, et comme A' est carrée, A' possède au moins une colonne sans pivot. Donc l'ensemble des solutions de $Ax = B$ est nul ou infini.

(1) $\Leftrightarrow$ (3) : Si $A' = I$, l'algorithme d'inversion de A (§ 3.3) aboutit, donc A est inversible.

Réciproquement, si A est inversible, par des opérations sur les lignes on peut ramener A à l'identité ($A' = I$).

(1) $\Leftrightarrow$ (4) : Si $A' = I$, A' possède évidemment un pivot par ligne.

Réciproquement, supposons que A' possède un pivot par ligne. En général A' possède au plus un pivot par ligne. Donc, ici, A' possède exactement un pivot par ligne, donc n pivots. Comme A' est carrée et comme A' possède au plus un pivot par colonne, ici A' possède exactement un pivot par colonne. Donc A' possède un pivot par ligne et un pivot par colonne : $A' = I$.

(1) $\Leftrightarrow$ (5) : Cette équivalence découle du même raisonnement en intervertissant le rôle des lignes et des colonnes.

(1) $\Leftrightarrow$ (6) : Les calculs de A' et de $\det A$ se font tous deux par des opérations élémentaires inversible sur des lignes, qui ramènent à A' et à $\det A'$, d'où l'équivalence voulue.

□

Exercice. — Montrer que, par trois points du plan non alignés, il passe un unique cercle.

Solution. — L'équation du cercle de centre $I = (a, b)$ et de rayon $r > 0$ est

$$(x - a)^2 + (y - b)^2 = r^2,$$

soit

$$x^2 + y^2 - 2ax - 2by + a^2 + b^2 - r^2 = 0.$$

Inversement, l'équation

$$x^2 + y^2 - 2ax - 2by + c = 0$$

est l'équation du cercle de centre $I = (a, b)$ et de rayon $r = \sqrt{a^2 + b^2 - c}$ si $a^2+b^2-c \geq 0$; si $a^2 + b^2 - c < 0$, aucun point ne satisfait cette équation.

Il s'agit de trouver a, b et r pour que trois points donnés $M_i = (x_i, y_i)$, $i = 1, 2, 3$, satisfassent une équation

$$x_i^2 + y_i^2 - 2ax_i - 2by_i + c = 0,$$

avec $a^2+b^2-c \geq 0$. Il s'agit de trois équations linéaires à trois inconnues. Le déterminant de ce système vaut

$$\begin{vmatrix} 2x_1 & 2y_1 & -1 \\ 2x_2 & 2y_2 & -1 \\ 2x_3 & 2y_3 & -1 \end{vmatrix} = -4 \begin{vmatrix} x_1 - x_3 & y_1 - y_3 \\ x_2 - x_3 & y_2 - y_3 \end{vmatrix}.$$

Si les points M_i ne sont pas alignés, cette quantité est non nulle et, d'après le corollaire § 3.2, il existe une unique solution (a, b, c) au système. De plus, comme l'équation du cercle admet alors trois solutions (une seule suffirait, ici), on a bien $a^2 + b^2 - c \geq 0$.

3.3. Opérations élémentaires sur les lignes et les colonnes d'une matrice

Nous allons formaliser les trois opérations élémentaires introduites dans l'algorithme fangcheng (§§ 2.5 et 3.2), en terme de multiplication matricielle. Cela nous permettra de résoudre des problèmes importants : l'inversion des matrices carrées, la multiplicativité du déterminant, et la mise sous forme normale d'une matrice rectangulaire.

Définitions. — Notons E_{ij} la matrice carrée dont tous les coefficients sont nuls, sauf le coefficient (i, j) qui vaut 1 (attention E_{ij} n'est donc pas le coefficient (i, j) d'une matrice E !) :

$$E_{ij} = (\delta_{ki}\delta_{lj})_{kl},$$

où

$$\delta_{ki} = \begin{cases} 1 & \text{si } k = i \\ 0 & \text{sinon} \end{cases}$$

(*symbole de Kronecker*, du nom du mathématicien et logicien allemand L. KRONECKER, 1823–1891).

Nous définirons trois matrices élémentaires, correspondant aux trois opérations élémentaires du § 2.5 (les coefficients non spécifiés sont nuls, les coefficients signalés par des points sont égaux à 1) :

– *Matrice de permutation-transposition* :

$$S_{ij} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} = I - E_{ii} - E_{jj} + E_{ij} + E_{ji}$$

– *Matrice de dilatation* :

$$D_i(a) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & a & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} = I + (a-1)E_{ii}, \quad a \neq 1$$

– *Matrice de transvection* : $T_{ij}(a) = \begin{pmatrix} 1 & & & \\ & \ddots & a & \\ & & \ddots & \\ & & & 1 \end{pmatrix} = I + aE_{ij}, \quad i \neq j.$

Lemme. — Soit A une matrice rectangulaire dont le nombre de lignes vaut m. Si E est l'une des trois matrices élémentaires de taille $m \times m$ ci-dessus, la matrice EA s'obtient à partir de A respectivement en :

– échangeant les lignes i et j ($\ell_i' = \ell_j$ et $\ell_j' = \ell_i$) pour une matrice de permutation
– multipliant scalairement la ligne i par a ($\ell_i' = a\ell_i$) pour une matrice de dilatation
– ajoutant à la ligne i la ligne j multipliée par a ($\ell_i' = \ell_i + a\ell_j$) pour une matrice de transvection.

Si le nombre de colonnes de A vaut m, la matrice AE s'obtient par les opérations analogue sur les colonnes de A.

Démonstration – Calculons par exemple l'effet d'une matrice de transvection

$$T_{ij}(a) = (\delta_{kl} + a\delta_{ki}\delta_{lj})_{kl},$$

où l et m sont respectivement les indices muets de ligne et de colonne dans $T_{ij}(a)$, sur la matrice $A = (A_{lm})_{lm}$. Par définition du produit matriciel,

$$T_{ij}(a)A = (\delta_{kl} + a\delta_{ki}\delta_{lj})_{kl}(A_{lm})_{lm} = \left(\sum_l (\delta_{kl} + a\delta_{ki}\delta_{lj})A_{lm}\right)_{km}.$$

Par définition du symbole de Kronecker (cf. § 3.3), le coefficient d'indice (k, m) de $AT_{ij}(a)$ vaut donc

$$(AT_{ij}(a))_{km} = A_{km} + \begin{cases} aA_{jm} & \text{si } k = i \\ 0 & \text{sinon.} \end{cases}$$

Les lignes $k \neq i$ sont inchangées, tandis que la ligne i de $AT_{ij}(a)$ s'obtient en ajoutant, à la ligne i de A, la ligne j de A multipliée par a.

L'effet des autres types de matrices élémentaires se calcule de façon analogue. □

Corollaire. — Toute matrice élémentaire est inversible et son inverse est une matrice élémentaire :

$$S_{ij}{}^{-1} = S_{ij} = S_{ji}, \quad D_i(a)^{-1} = D_i\left(a^{-1}\right) \quad \text{et} \quad T_{ij}(a)^{-1} = T_{ij}(-a).$$

Démonstration – Par exemple, une fois qu'on a changé ℓ_i en $\ell_i' = \ell_i + a\ell_j$, pour revenir à la matrice de départ il suffit de poser $\ell_i'' = \ell_i' - a\ell_j$. Ceci montre que $T_{ij}(a)^{-1} = T_{ij}(-a)$, et ainsi de suite. □

Exercice. — La matrice
$$M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$
est-elle inversible ?

Solution. — Si $N = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$, le produit
$$NM = \begin{pmatrix} 0 & a \\ 0 & b \end{pmatrix}$$
ne peut pas être égal à I (parce que le coefficient en haut à gauche est nul, donc différent de 1, quel que soit le choix de N). Donc M n'est pas inversible.

Application 1 : algorithme d'inversion d'une matrice. — Comment inverser une matrice inversible A ? Une stratégie possible est de se ramener, par une série d'opérations élémentaires sur ses lignes, à la matrice I. Comme chaque opération se traduit par la multiplication par une matrice élémentaire, on obtient une égalité
$$I = E_k \cdots E_2 E_1 A,$$
et ce qui précède montre que $A^{-1} = E_k \cdots E_1$, soit, trivialement,
$$A^{-1} = E_k \cdots E_1 I.$$
Autrement dit, il suffit de réaliser en parallèle les mêmes opérations élémentaires sur A et sur I, pour transformer A en la matrice identité ; la matrice ainsi obtenue en partant de I sera précisément A^{-1}.

Remarquons qu'on pourrait tout aussi bien opérer sur les colonnes, et donc construire une suite finie de matrices élémentaires $F_1, ..., F_l$ telles que
$$I = AF_1...F_l,$$
i.e.
$$A^{-1} = F_1...F_l.$$
La seule chose importante est de ne pas mélanger les opérations sur les lignes et sur les colonnes !

Exercice. — La matrice
$$M = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix}$$
est-elle inversible et, si oui, quel est son inverse ?

Solution. — Suivons la méthode d'inversion proposée ; si l'on arrive au bout (i.e. si, par une suite d'opérations élémentaires, on se ramène à l'identité), cela signifiera que M est inversible :
$$M = \begin{pmatrix} \boxed{1} & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix} \quad \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} = I$$
$$L_1 M = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 5 \\ 0 & 5 & 7 \end{pmatrix} \quad \begin{matrix} \\ 2\ell_1 - \ell_2 \\ 3\ell_1 - \ell 3 \end{matrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ 3 & 0 & -1 \end{pmatrix} = L_1$$

$$L_2L_1M = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix} \qquad \tfrac{5\ell_2-\ell_3}{18} \qquad \frac{1}{18}\begin{pmatrix} 18 & 0 & 0 \\ 36 & -18 & 0 \\ 7 & -5 & 1 \end{pmatrix} = L_2L_1$$

$$L_3L_2L_1M = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \begin{matrix} \ell_1 - 3\ell_3 \\ \ell_2 - 5\ell_3 \end{matrix} \quad \frac{1}{18}\begin{pmatrix} -1 & 5 & -1 \\ 1 & 7 & -5 \\ 7 & -5 & 1 \end{pmatrix} = L_3L_2L_1$$

$$I = L_4L_3L_2L_1M = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 1 \end{pmatrix} \quad \ell_1 - 2\ell_2 \quad \frac{1}{18}\begin{pmatrix} -5 & 1 & 7 \\ 1 & 7 & -5 \\ 7 & -5 & 1 \end{pmatrix} = L_4L_3L_2L_1 = M^{-1}.$$

Exercice. — Calculer la matrice inverse, quand elle existe, de la matrice

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R}).$$

Solution. — Par la même méthode, on trouve que A est inversible si et seulement si $ad - bc \neq 0$, auquel cas

$$A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Dans la quantité qui apparait au dénominateur, on reconnaît le déterminant de A (§ 2.5).

Application 2 : forme normale d'une matrice. — Soit $A \in M_{mn}(\mathbb{R})$. Il existe un entier $r \in \{0, ..., \min(m,n)\}$ et deux matrices inversibles $P \in GL_m(\mathbb{R})$ et $Q \in GL_n(\mathbb{R})$ telles que

$$PAQ = J_r, \quad J_r := \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix},$$

où la matrice J_r a tous ses coefficients nuls, sauf les r premiers coefficients sur la diagonale, qui valent 1.

Si A est carrée ($m = n$), elle est inversible si et seulement si $r = n$.

Démonstration – On déduit de l'algorithme fang-cheng du § 3.2 (auquel il faut ajouter les étapes consistant à éliminer les coefficients non nuls au-dessus des pivots, par des opérations sur le colonnes) et de l'interprétation matricielle des opérations élémentaires sur les lignes et les colonnes, qu'il existe des matrices d'opérations élémentaires $E_1, ...E_k$ et $F_1, ..., F_l$ telles que

$$E_k...E_1AF_1...F_l = J_r.$$

Il ne reste qu'à poser $P = E_k...E_1$ et $Q = F_1...F_l$.

Si $m = n$, comme P et Q sont inversibles, A est inversible si et seulement si J_r est inversible, c'est-à-dire $r = n$. $\square$

Application 3 : multiplicativité du déterminant. — Le déterminant est multiplicatif :

$$\det(AB) = \det A \det B, \quad A, B \in M_3(\mathbb{R}).$$

Le lecteur avide de généralisation aura deviné que cette formule vaut pour des matrices carrées de dimension quelconque. La démonstration ci-dessous peut elle aussi être lue en pensant à des matrices 3×3 ou $n \times n$ au choix.

Démonstration – Les règles de calcul du déterminant du § 2.5 montrent que

$$\det S_{ij} = -1, \quad \det D_i(a) = a \quad \text{et} \quad \det T_{ij}(a) = 1.$$

On a vu dans le chapitre précédent comment les opérations élémentaires sur les colonnes modifient le déterminant d'une matrice carrée. On voit donc que si M une matrice carrée et si E est une matrice élémentaire,

$$\det(ME) = \det M \det E.$$

Prenons maintenant deux matrices carrées A et B de même dimension. D'après le § 3.3, on peut décomposer A et B sous la forme

$$A = E_1 ... E_k J_r F_1 ... F_l \quad \text{et} \quad B = E'_1 ... E'_k J_{r'} F'_1 ... F'_l,$$

où k, l, k', l' sont des entiers et où les matrices E_i, F_i, E'_i et F'_i sont des matrices d'opérations élémentaires.

Supposons pour commencer que A et B sont inversibles. Alors $J_r = J_{r'} = I$, donc

$$A = E_1 ... E_k \, F_1 ... F_l \quad \text{et} \quad B = E'_1 ... E'_k \, F'_1 ... F'_l.$$

Mais d'après ce qui précède, par récurrence on voit que

$$\begin{aligned}
\det(AB) &= \det(E_1 ... E_k \, F_1 ... F_l \, E'1 ... E'_k \, F'_1 ... F'_l) \\
&= (\det E_1) ... (\det E_k) \, (\det F_1) ... (\det F_l) \, (\det E'_1) ... (\det E'_{k'}) \, (\det F'_1) ... (\det F'_{l'}) \\
&= \det(E_1 ... E_k \, F_1 ... F_l) \, \det(E'1 ... E'_k \, F'_1 ... F'_l) \\
&= \det A \, \det B.
\end{aligned}$$

Si A ou B n'est pas inversible, d'après le corollaire § 3.2 on a $\det A$ ou $\det B = 0$, donc $\det A \det B = 0$. Il ne reste qu'à montrer qu'alors $\det(AB) = 0$. Reprenons le calcul de ci-dessus dans ce cas. Supposons par exemple que $J_r \neq I$ et $J_{r'} = I$ (les autres cas étant analogues). On a

$$\begin{aligned}
\det(AB) &= \det(E_1 ... E_k J_r F_1 ... F_l \, E'1 ... E'_k \, F'_1 ... F'_l) \\
&= (\det E_1 ... E_k J_r) \, (\det F_1) ... (\det F_l) \, (\det E'_1) ... (\det E'_{k'}) \, (\det F'_1) ... (\det F'_{l'}).
\end{aligned}$$

Or les matrices $E_1, ..., E_k$ sont inversibles et la matrice J_r ne l'est pas, donc $E_1 ... E_k J_r$ n'est pas inversible et son déterminant est nul. Donc $\det AB = 0$. $\square$

Application 4 : Invariance du déterminant par transposition. — Soit $A = (A_{ij})_{(i,j)} \in M_n(\mathbb{R})$. La *matrice transposée* de A est la matrice ${}^tA = (A_{ji})_{(i,j)}$ obtenue en écrivant les colonnes en ligne et vice-versa.

Par exemple,

$$t \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}.$$

Aussi, si $u, v \in \mathbb{R}^n$, un calcul direct montre qu'on a

$$(Au) \cdot v = u \cdot ({}^t A v)$$

pour tous $u, v \in \mathbb{R}^3$, et, plus généralement, si $A, B \in M_n(\mathbb{R})$,

$${}^t(AB) = {}^t B {}^t A.$$

L'interprétation géométrique de la transposition des matrices, qui dépasse le cadre de ce cours, et sera expliquée dans le cours d'algèbre linéaire avec le concept de dualité vectorielle.

Le fait est que le déterminant invariant par transposition :

$$\det A = \det {}^t A, \quad A \in M_n(\mathbb{R}).$$

Démonstration – Nous allons déduire l'invariance voulue du fait vérifiable directement que, si E est une matrice élémentaire,

$$\det E = \det {}^t E.$$

(Si E est une matrice de permutation-transposition ou une matrice de dilatation, c'est immédiat puisque dans ces cas $E = {}^t E$. Si E est une matrice de transvection, on voit que $\det E = 1 = \det {}^t E$.)

D'après § 3.3, il existe un entier r et des matrices élémentaires $E_1, ..., E_k$ et $F_1, ..., F_l$ tels que

$$A = E_1 ... E_k \, J_r \, F_1 ... F_l;$$

alors, en utilisant le fait que ${}^t J_r = J_r$, par récurrence on voit que

$${}^t A = {}^t F_l ... {}^t F_1 \, J_r \, {}^t E_k ... {}^t E_1.$$

Donc

$$\begin{aligned}
\det A &= \det(E_1 ... E_k \, J_r \, F_1 ... F_l) \\
&= (\det E_1) ... (\det E_k)(\det J_r)(\det F_1) ... (\det F_l) \quad \text{(multiplicativité, § 3.3)} \\
&= (\det {}^t F_l) ... (\det {}^t F_1)(\det J_r)(\det {}^t E_k) ... (\det {}^t E_1) \\
&= \det({}^t F_l ... {}^t F_1 \, J_r \, {}^t E_k ... {}^t E_1) \quad \text{(multiplicativité, § 3.3)} \\
&= \det {}^t A.
\end{aligned}$$

□

Compléments

Structure d'anneau. — Un *anneau* est un ensemble R muni de deux opérations $+$ et $\times$ telles que

- $(R, +)$ est un groupe commutatif, dont on note 0 l'élément neutre
- $\times$ est associative et possède un élément neutre, noté 1
- L'addition est distributive sur la multiplication.

Nous venons de voir que $(M_n(\mathbb{R}), +, \times)$ est un anneau et que $(GL_n(\mathbb{R}), \times)$ est un groupe non commutatif.

Volume k-dimensionnel dans $\mathbb{R}^n$. — On peut montrer que le volume k-dimensionnel du parallélépipède engendré par k vecteurs $v_1, ..., v_k$ dans $\mathbb{R}^n$ est

$$\sqrt{\det(V^t V)},$$

où V est la matrice à n lignes dont les k colonnes sont les v_i, et où ${}^tV$ désigne la matrice transposée de V (cf. § 3.3) ; la démonstration utilise les concepts de dualité et d'algèbre bilinéaire, donc dépasse le cadre de ce cours (voir par exemple [**KP08**, Proposition 1.4.3]).

CHAPITRE 4

TRANSFORMATIONS LINÉAIRES DE L'ESPACE

Mots-clef du chapitre. — Transformation linéaire, matrice d'une transformation, inversibilité, base, composante, conjugaison, opérateur orthogonal, vecteur propre

4.1. Matrices et transformations linéaires

Définition. — Une *transformation linéaire* de $\mathbb{R}^3$ est une transformation de $\mathbb{R}^3$, c'est-à-dire une application $A : \mathbb{R}^3 \to \mathbb{R}^3$, qui est *linéaire* au sens déjà rencontré que, pour tous $v, w \in \mathbb{R}^3$ et tous $a, b \in \mathbb{R}$, on a

$$A(av + bw) = aA(v) + bA(w).$$

Remarque. — Si A est linéaire, forcément $A(0) = 0$. En effet, d'après la définition on a en particulier

$$A(0) = A(1.0_{\mathbb{R}^3}) + A(1.0_{\mathbb{R}^3}) = 1\,A(0_{\mathbb{R}^3}) + 1\,A(0_{\mathbb{R}^3}) = 2\,A(0_{\mathbb{R}^3}).$$

Exemples. — L'application $A : \mathbb{R} \to \mathbb{R}$, $x \mapsto x + 1$ n'est pas linéaire parce que $A(0) = 1 \neq 0$.

L'application sin n'est pas linéaire, bien que $\sin 0 = 0$. En effet, on sait par exemple que

$$\sin(\pi/2 + \pi/2) = \sin \pi = 0$$

tandis que

$$\sin \pi/2 + \sin \pi/2 = 2.$$

Exemple. — Si A est une matrice 3×3, l'application encore notée A,

$$\mathbb{R}^3 \to \mathbb{R}^3, \quad x \mapsto A(x) := Ax$$

est linéaire.

Notons A_1, A_2 et A_3 les vecteurs-colonne de A. Au § 3.1, on avait vu que

$$\boxed{Ax = x_1A_1 + x_2A_2 + x_3A_3.}$$

En appliquant cette formule successivement à $x = i$, j et k, on voit que le n-ième vecteur colonne de A est l'image par l'application linéaire A du n-ième vecteur de la base canonique.

Proposition. — Réciproquement, toute application linéaire est de cette forme.

En effet, une application linéaire $\tilde{A}$ étant donnée, il suffit de définir la matrice A comme la matrice $(\tilde{A}(i), \tilde{A}(j), \tilde{A}(k))$. Alors

$$\tilde{A}(x) = x_1 A_1 + x_2 A_2 + x_3 A_3 = Ax.$$

Autrement dit,

> Il est équivalent de se donner une transformation linéaire de $\mathbb{R}^3$ ou une matrice réelle 3×3. Nous identifierons systématiquement l'une à l'autre.

Proposition. — La composition des applications et la multiplication matricielle coïncident :

$$(AB)x = A(B(x))$$

pour toutes matrices $A, B \in M_3(\mathbb{R})$ et tout vecteur $x \in \mathbb{R}^3$. En particulier, A est inversible comme application si et seulement si elle est inversible comme matrice, auquel cas les deux inverses coïncident.

Il est conseillé de retrouver la démonstration suivante plutôt que de la lire...

Démonstration – Notons $A = (A_{mn})_{1\leq m,n\leq 3}$, $B = (B_{np})_{1\leq n,p\leq 3}$ et $x = (x_p)_{1\leq p\leq 3}$. Alors l'application linéaire associée au produit matriciel AB :

$$\begin{aligned}
(AB)x &= \left(\sum_n A_{mn}B_{np}\right)_{mp} (x_p)_p \\
&= \left(\sum_n \sum_p A_{mn}B_{np}x_p\right)_m \\
&= \sum_n \sum_p A_n B_{np} x_p \\
&= A(Bx),
\end{aligned}$$

coïncide bien avec la composée des applications linéaires associées à A et B.

Si A est inversible comme application et si l'on note A^{-1} sa bijection réciproque, pour tout x on a

$$x = A^{-1}(Ax) = (A^{-1}A)x,$$

donc $A^{-1}A = I$ et de façon similaire $AA^{-1} = I$, donc A est inversible aussi comme matrice et sa matrice inverse est la matrice de sa bijection réciproque.

La réciproque est analogue. □

Exercice. — Montrer que le produit matriciel est associatif :

$$A(BC) = (AB)C.$$

Solution. — On a

$$((AB)C)(x) = (AB)(Cx) = A(B(Cx)) = A((BC)x) = (A(BC))x$$

pour tout x, d'où l'associativité.

Corollaire immédiat. — Soit A une transformation linéaire inversible. Alors son inverse A^{-1} est linéaire.

Exercice. — Montrer le corollaire directement, sans passer par les matrices.

Solution. — Notons $y = Ax$ et $y' = Ax'$. On veut montrer que

$$A^{-1}(ay + a'y') = aA^{-1}(y) + a'A^{-1}(y').$$

Comme A est bijective, cette égalité est équivalente à la suivante :

$$ay + a'y' = A(ax + a'x'),$$

qui est satisfaite par hypothèse.

4.2. Bases

Définition. — Une *base* de $\mathbb{R}^3$ est une famille $F = (u, v, w)$ de trois vecteurs de $\mathbb{R}^3$ telle que l'application linéaire F correspondante soit inversible, i.e. $\det F \neq 0$ (cf. § 3.2).

Comme

$$F \begin{pmatrix} a \\ b \\ c \end{pmatrix} = au + bv + cw,$$

le fait que F est une base équivaut à celui que our tout vecteur $x \in \mathbb{R}^3$ il existe d'uniques $a, b, c \in \mathbb{R}$ tels que

$$au + bv + cw = x.$$

Les coefficients a, b et c sont les *composantes* de x dans la base F, et sont donnés par la formule

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = F^{-1}x.$$

Proposition. — F est une base si et seulement si F est libre : $\forall a, b, c \in \mathbb{R} \quad (au + bv + cw) = 0 \Rightarrow a = b = c = 0$

Démonstration – Si F est une base, F est trivialement libre (prendre $x = 0$).

Réciproquement, supposons F libre. D'après le théorème de réduction du § 3.2, la matrice réduite de F est l'identité. Donc, si $x \in \mathbb{R}^3$, le système linéaire

$$au + bv + cw = x$$

d'inconnues $a, b, c \in \mathbb{R}$ possède une solution unique. □

Exercice. — Les trois vecteurs

$$u = i + j - k, \quad v = i - j + k, \quad w = -i + j + k$$

forment-ils une base de $\mathbb{R}^3$?

Solution. — On a

$$(u, v, w) = \begin{pmatrix} \boxed{1} & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -1 \\ 0 & \boxed{1} & -1 \\ 0 & 1 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \sim I_3,$$

donc la famille (u, v, w) est libre ; c'est une base de $\mathbb{R}^3$.

4.3. Conjugaison

Définition. — Deux transformations linéaires $A, B \in M_3(\mathbb{R})$ sont *conjuguées* s'il existe une transformation linéaire inversible F de $\mathbb{R}^3$, appelée une *conjugaison*, telle que le diagramme d'applications suivant commute :

$$\begin{tikzcd} \mathbb{R}^3 \arrow[r, "A"] & \mathbb{R}^3 \\ \mathbb{R}^3 \arrow[u, "F"] \arrow[r, "B"] & \mathbb{R}^3, \arrow[u, "F"'] \end{tikzcd}$$

c'est-à-dire

$$AF = FB.$$

(La conjugaison dans ce sens ne doit pas être confondue avec la conjugaison des nombres complexes. Par ailleurs, dans les ouvrages français, on dit parfois *semblable* à la place de *conjugué.*)

Proposition. — Si A et B sont conjuguées, $\det A = \det B$.

Démonstration – En effet, $B = F^{-1}AF$ donc, d'après le caractère multiplicatif du déterminant (§ 3.3),

$$\det B = \det(F^{-1}AF) = \det(F^{-1}) \det A \det F = \det A \det F^{-1}F = \det A.$$

□

Interprétation géométrique. — Le rapport de volume entre un parallélépipède dont un sommet est l'origine, et son image par la transformation A, ne dépend pas du parallélépipède.

Démonstration – Pour comparer ce rapport pour deux parallélépipèdes différents, il suffit d'appliquer la formule ci-dessus avec la conjugaison F donnée par la transformation qui envoie le premier parallélépipède sur le second. □

Cette interprétation géométrique, à son tour, rend naturelle la multiplicativité du déterminant. En effet, le déterminant apparaît comme le facteur par lequel la transformation linéaire A multiplie les volumes. Donc le déterminant de AB est le produit des facteurs $\det A$ et $\det B$. (Notons cependant que nous avons utilisé la multiplicativité du déterminant pour interpréter ainsi le déterminant. Cet argument ne démontre donc pas la multiplicativité.)

Exercice. — Montrer qu'une transformation linéaire de $\mathbb{R}^2$ (définie de façon analogue aux transformations linéaires de $\mathbb{R}^3$) de la forme

$$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}, \quad a, b, d \in \mathbb{R},$$

est conjuguée à une matrice diagonale. En déduire une méthode pour calculer A^{100}.

Solution. — Supposons par analyse que A soit conjuguée à une matrice diagonale

$$B = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

par une conjugaison $F = (F_1, F_2)$. Notons e_1 et e_2 les vecteurs de la base canonique de $\mathbb{R}^2$. Dans le carré commutatif d'applications caractérisant la conjugaison de A et B (§ 4.3), regardons les images des vecteurs e_i :

$$\begin{tikzcd} F_i \arrow[r, "A"] & A(F_i) = \\ e_i \arrow[u, "F"] \arrow[r, "B"] & \lambda_i e_i, \arrow[u, "F"'] \end{tikzcd}$$

Donc

$$A(F_i) = \lambda_i F_i.$$

Les vecteurs F_i vérifient cette propriété très particulière que leur image par A est proportionnelle à eux-même. On dit que les F_i sont des *vecteurs propres* de A.

On voit qu'il suffit de trouver deux tels vecteurs libres F_1 et F_2. Ces vecteurs sont deux solutions libres (donc, en particulier, non nulles) de l'équation

$$Af = \lambda f$$

d'inconnue $f = (x, y)$ pour un certain réel $\lambda \in \mathbb{R}$. Cette équation se ramène au système

$$\begin{cases} (a-\lambda)x + by = 0 \\ bx + (d-\lambda)y = 0, \end{cases}$$

soit

$$\begin{cases} (a-\lambda)x + by = 0 \\ ((d-\lambda)(a-\lambda) - b^2)x = 0. \end{cases}$$

Pour que ce système possède une solution non triviale, il faut et il suffit que

$$\lambda^2 - (a+d)\lambda + ad - b^2 = 0.$$

Le discriminant de cette équation de degré deux est $\Delta = (a-d)^2 + 4b^2 \geq 0$. Si $\Delta = 0$, $b = 0$ et A est déjà diagonale. Supposons donc que $b \neq 0$. Alors l'équation possède deux racines réelles distinctes λ_1 et λ_2. Quand $\lambda = \lambda_i$, le système ci-dessus se réduit à une équation scalaire, dans laquelle x est essentielle et y libre. Dans chacun des deux cas, en fixant $y = 1$ et on obtient l'unique valeur possible de x. Notons F_1 et F_2 les deux vecteurs ainsi obtenus. Ils forment une base de $\mathbb{R}^2$, parce que s'ils étaient liés, on aurait $\lambda = \beta$, ce qui a été exclu. D'après le raisonnement ci-dessus, on a ainsi construit une conjugaison $F = (F_1, F_2)$ entre A et B : $A = FBF^{-1}$.

Comment calculer simplement A^{100} ? Remarquons déjà que

$$A^2 = FBF^{-1}\,FBF^{-1} = FB^2F^{-1},$$

où le carré de B est immédiat à calculer puisque B est diagonale :

$$B^2 = \begin{pmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{pmatrix}.$$

Par récurrence on voit que, pour tout $n \geq 0$,

$$A^n = FB^nF^{-1}, \quad B^n = \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix}.$$

Donc en particulier le calcul de A^{100} se réduit au calcul du produit de trois matrices :

$$A^{100} = F \begin{pmatrix} \lambda_1^{100} & 0 \\ 0 & \lambda_2^{100} \end{pmatrix} F^{-1}.$$

Nous allons voir comment certains autres problème sont simplifiés par le choix d'une conjugaison adaptée.

4.4. Exemple : Opérateurs orthogonaux

Proposition et définition. — Soit A une transformation linéaire de $\mathbb{R}^3$. Les propriétés suivantes sont équivalentes :

(1) A conserve le produit scalaire : $(Au)\cdot(Av) = u\cdot v$

(2) ${}^tAA = I$

(3) Les vecteurs-colonne A_1, A_2 et A_3 de A forment une *base orthonormale*, i.e. ils forment une base telle que

$$A_1{}^2 = A_2{}^2 = A_3{}^2 = 1, \quad \text{et} \quad A_1\cdot A_2 = A_1\cdot A_3 = A_2\cdot A_3 = 0$$

($A_1{}^2 = A_1\cdot A_1$, etc.).

Si elles sont vérifiées, A est qualifiée d'*orthogonale.* On note $O_3(\mathbb{R})$ et on appelle *groupe orthogonal de la dimension* 3 l'ensemble des transformations orthogonales de $\mathbb{R}^3$.

Démonstration – Comme $(Au)\cdot(Av) = ({}^tAAu)\cdot v$, les propriétés (1) et (2) sont équivalentes. Pour déduire (3) de (1), il suffit de prendre pour u et v les vecteurs de la base canonique. Pour déduire (1) de (3), il suffit d'utiliser la bilinéarité du produit scalaire. □

Dans la suite on suppose que $A \in O_3(\mathbb{R})$.

Lemme. — Il existe un vecteur $w \in \mathbb{R}^3$ unitaire ($\|w\| = 1$) tel que $Aw = \pm w$.

Démonstration – Commençons par chercher plus généralement un vecteur $w \neq 0$ tel que $Aw = \lambda w$, avec $\lambda \in \mathbb{R}$, soit

$$(A - \lambda I)w = 0\ ;$$

un tel vecteur, envoyé par A sur un multiple de lui-même, est ce qui s'appelle un *vecteur propre* de A. Pour cela, il faut et il suffit que $\det(A-\lambda I) = 0$. Cette dernière équation est polynomiale de degré 3 en λ, à coefficients réels. Elle possède donc au moins une racine réelle (exercice du § 1.3), que nous noterons λ. Notons w une solution de norme unité.

Il ne reste qu'à montrer que $\lambda = \pm 1$. Par bilinéarité du produit scalaire on a

$$Aw\cdot Aw = \lambda^2 w^2.$$

Or, par définition d'une transformation orthogonale,

$$Aw\cdot Aw = w\cdot w.$$

Donc $\lambda^2 = 1$ comme voulu. □

Lemme. — Le déterminant de A vaut ± 1.

Démonstration – Par définition d'un opérateur orthogonal, ${}^tAA = I$ et, d'après le caractère multiplicatif du déterminant (§ 3.3),

$$\det {}^tAA = (\det A)^2 = \det I = 1,$$

donc $\det A = \pm 1$. □

Lemme. — Rappelons que le vecteur w a été défini au § 4.4. Le plan vectoriel orthogonal à w, $P = \{v \in \mathbb{R}^3 \ ; \ v \cdot w = 0\}$, est stable par A :

$$v \cdot w = 0 \Rightarrow (Av) \cdot w = 0.$$

Démonstration – Soit $v \in \mathbb{R}^3$ orthogonal à w. On a

$$(Av) \cdot w = (Av) \cdot (\pm Aw) = \pm v \cdot w = 0.$$

□

Théorème. — $A : \mathbb{R}^3 \to \mathbb{R}^3$ est conjuguée à

$$B = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & \pm 1 \end{pmatrix}.$$

La droite vectorielle $\mathbb{R}w$ est fixée par A et s'appelle un *axe* de A ; l'angle θ est un *angle* de A.

Démonstration – Soient u et v deux vecteurs libres de P, supposés orthonormés ($u \cdot v = 0$, $\|u\| = \|v\| = 1$), $F = (u, v, w) \in M_3(\mathbb{R})$ et $B = F^{-1}AF$:

$$\begin{tikzcd} \mathbb{R}^3 \arrow[r, "A"] & \mathbb{R}^3 \\ \mathbb{R}^3 \arrow[u, "F"] \arrow[r, "B"] & \mathbb{R}^3. \arrow[u, "F"'] \end{tikzcd}$$

D'après les lemmes §§ 4.4–4.4, B est de la forme

$$B = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & \pm 1 \end{pmatrix}.$$

Comme les transformations A et F sont orthogonales, $B = F^{-1}AF$ l'est aussi (d'après la caractérisation (1) du § 4.4, le produit de deux transformations orthogonales est une transformation orthogonale).

Les deux vecteurs (a, c) et (b, d) forment une base orthonormée de $\mathbb{R}^2$:

$$a^2 + c^2 = b^2 + d^2 = 1, \quad ab + cd = 0.$$

Quitte à changer w en $-w$, on peut d'ailleurs changer le signe du déterminant

$$\begin{vmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & \pm 1 \end{vmatrix} = \pm \begin{vmatrix} a & b \\ c & d \end{vmatrix},$$

et supposer par exemple que $\begin{vmatrix} a & b \\ c & d \end{vmatrix} > 0$.

Notons θ et ϕ les angles tels que

$$e^{i\theta} = a + ic \quad \text{et} \quad e^{i\phi} = b + id.$$

On a

$$ab + cd = \cos(\theta - \phi) = 0,$$

donc $\theta - \phi = \pi/2 \pmod \pi$. Donc il existe $\epsilon = \pm 1$ tel que $b = -\epsilon \sin\theta$ et $d = \epsilon\cos\theta$. Comme

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc = \sin(\theta - \phi) = \epsilon,$$

on a en réalité $\epsilon = 1$. Donc B est de la forme souhaitée. □

Remarquons que $\mathbb{R}w$ est unique (comme droite vectorielle invariante) si et seulement si $\theta \neq 0 \pmod{2\pi}$. Par ailleurs, l'angle θ de la rotation n'est lui-même déterminé modulo 2π qu'au signe près, comme on le voit en changeant w en $-w$.

Définition. — Une *rotation* est une transformation orthogonale dont le déterminant vaut 1.

D'après le théorème § 4.4, une telle rotation est conjuguée à une matrice

$$B = \begin{pmatrix} \cos\theta & sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1. \end{pmatrix}$$

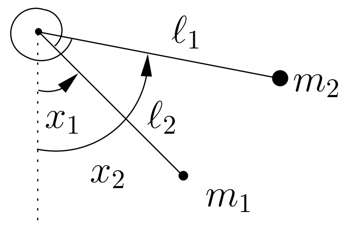
Exercice. — Montrer que la rotation d'angle θ et dont l'axe est dirigé par le vecteur unitaire w est donnée par

$$R_{w,\theta}(x) = \cos\theta\, x + \sin\theta\,(w \wedge x) + (1 - \cos\theta)\,(x.w)\,w.$$

Solution. — Il suffit de vérifier la formule sur trois vecteurs formant une base de $\mathbb{R}^3$, par exemple sur un vecteur $x \neq 0$ quelconque orthogonal à w, sur w et sur $x \wedge w$.

4.5. Exemple : Petites oscillations de deux pendules couplés

Notations. — Considérons deux pendules (masses attachées au bout d'une tige dont l'autre extrémité est fixée de façon à laisser la tige osciller librement dans un plan fixe) oscillant dans des plans parallèles, soumis à la pesanteur, et couplés par un ressort spiral.



On note x_1 et x_2 les angles que ces pendules font avec la verticale. Ce sont des fonctions du temps t.

En Mécanique, on montre que, sous des hypothèses et dans des unités convenables,

$$\begin{cases} \ddot{x}_1 = -\omega_1 \sin x_1 + (x_2 - x_1) \\ \ddot{x}_2 = -\omega_2 \sin x_2 + (x_1 - x_2) \end{cases}$$

(c'est par exemple une conséquence du théorème du moment cinétique) ; $\ddot{x}_j$ est l'accélération angulaire (dérivée seconde de x_j par rapport à t) du pendule j, $-\omega_j \sin x_j$ est le terme de pesanteur tandis que $x_k - x_j$ est le terme de couplage dû au ressort.

Nous allons déterminer le mouvement approché de ce système au voisinage des points d'équilibre $x_1 = x_2 = 0$. Rappelons que $\sin x \simeq x$ et que cette approximation est d'autant meilleure que x est petit, ce qui permet de remplacer avec une bonne précision le système exact par sa linéarisation :

$$\begin{cases} \ddot{x}_1 = -\omega_1 x_1 + (x_2 - x_1) \\ \ddot{x}_2 = -\omega_2 x_2 + (x_1 - x_2), \end{cases}$$

soit

$$\ddot{x} = Ax, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad A = \begin{pmatrix} -(\omega_1 + 1) & 1 \\ 1 & -(\omega_2 + 1) \end{pmatrix}.$$

Pour simplifier encore, nous allons faire les calculs en supposant les deux pendules identiques : $\omega_1 = \omega_2 =: \omega$.

Exercice : Faire l'étude avec deux pendules différents.

Sans le ressort, la matrice du système serait diagonale

$$A = \begin{pmatrix} -\omega & 0 \\ 0 & -\omega \end{pmatrix},$$

et l'équation différentielle vectorielle ci-dessus se ramène à la résolution de deux équations différentielles scalaires indépendantes. La Théorie des Équations différentielles scalaires montre alors que les solutions seraient les fonctions de la forme

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} a_1 \cos(\omega t + \varphi_1) \\ a_2 \cos(\omega t + \varphi_2) \end{pmatrix}, \quad a_1, a_2, \varphi_1, \varphi_2 \in \mathbb{R}$$

(c'est un calcul immédiat de vérifier que la fonction $x_j : t \mapsto a_j \cos(\omega t + \varphi_j)$ est solution de $\ddot{x}_j = -\omega x_j$; le fait important est qu'on obtient bien toutes les solutions ainsi).

Quand le ressort est là, la matrice A n'est plus diagonale mais on va voir qu'elle est conjuguée à une matrice diagonale.

Cherchons les vecteurs v tels que Av soit proportionnel à v (tandis que $Av = v_1 A_1 + ... + v_n A_n$ est généralement une combinaison linéaires des autres vecteurs de la base) : $Av = \lambda v$, où λ est un scalaire.

Définition. — Un *vecteur propre* de A est un vecteur v non nul tel qu'il existe $\lambda \in \mathbb{C}$ tel que $Av = \lambda v$. Le scalaire λ est une *valeur propre* de A.

Lemme. — Les valeurs propres de A sont les racines du *polynôme caractéristique* de A : $\det(A - \lambda I) = 0$.

Démonstration – Une valeur propre est en effet un nombre complexe λ tel que l'équation $Av = \lambda v$ admette au moins une solution v non triviale ($v = 0$ est toujours solution, mais ne nous aide pas à construire une base adaptée). C'est dire que $A - \lambda I$ n'est pas inversible, d'où l'affirmation. □

Exercice : Montrer que $\det(A - \lambda I)$ est un polynôme de degré n si $A \in M_n(\mathbb{R})$.

Rappelons qu'un polynôme de degré n possède au plus n zéros complexes distincts (cf. § 1.3). Il est important de considérer des valeurs propres complexes même si A est un endomorphisme de $\mathbb{R}^n$: si A possède des valeurs propres imaginaires, ses valeurs propres réelles et les vecteurs propres associés, en nombre $< n$, ne peuvent en effet suffire pour construire une base. Même en complexifiant le problème on ne peut pas toujours diagonaliser A.

. — Mais revenons à la transformation

$$A = \begin{pmatrix} -\alpha & 1 \\ 1 & -\alpha \end{pmatrix}, \quad \alpha = 1 + \omega$$

de $\mathbb{R}^2$ dont dépend le mouvement de nos deux pendules couplés. Quels en sont les vecteurs propres ?

L'équation $Av = \lambda v$ est équivalente ($\ell_1' = -\ell_1 - (\alpha+\lambda)\ell_2$) à

$$\begin{pmatrix} 0 & (\alpha+\lambda)(\alpha+\lambda)-1 \\ 1 & -\alpha-\lambda \end{pmatrix} v = 0.$$

Ce système n'a de solution v non triviale que si le paramètre λ est lui-même racine de

$$P(\lambda) = \lambda^2 + 2\alpha\lambda + \alpha^2 - 1$$

(voilà le polynôme caractéristique de P, mais en plus de le calculer on a rendu triangulaire le système d'équations de v), i.e. λ prend l'une des deux valeurs

$$\lambda_\pm := -\alpha \pm 1.$$

Un vecteur propre associé à $\lambda_\pm$ est

$$v = b_\pm := \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}$$

(on pourrait prendre n'importe quel autre vecteur proportionnel non nul).
Considérons donc la base

$$B = (b_+, b_-) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

faite des deux vecteurs propres trouvés et la conjugaison induite par la matrice B. Notons

$$A_B = B^{-1}AB = \begin{pmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{pmatrix}$$

la matrice conjuguée à A (la forme particulière de A_B est évidente sans calcul, mais elle peut être vérifiée en effectuant le produit des trois matrices).
Comme la conjugaison B est indépendante du temps, l'équation différentielle satisfaite par $y = B^{-1}x$ est

$$\ddot{y} = B^{-1}\ddot{x} = B^{-1}Ax = B^{-1}ABy = A_B y.$$

C'est dire que les composantes de $y = (y_+, y_-)$ satisfont des équations différentielles scalaires indépendantes l'une de l'autre :

$$\ddot{y}_+ = \lambda_+ y_+, \quad \ddot{y}_- = \lambda_- y_-.$$

Remarquons que $\lambda_\pm < 0$ et notons $\lambda_\pm = -\omega_\pm^2$. Alors $y_\pm(t)$ est de la forme

$$y_+ = Y_+ \cos(\omega_+ t + \varphi_+), \quad y_- = Y_- \cos(\omega_- t + \varphi_-), \quad Y_\pm, \varphi_\pm \in \mathbb{R}.$$

On récupère enfin $x(t)$ par une simple multiplication par la conjugaison B :

$$x(t) = B\begin{pmatrix} y_+(t) \\ y_+(t) \end{pmatrix} = Y_+ \begin{pmatrix} \cos(\omega_+ t + \varphi_+) \\ \cos(\omega_+ t + \varphi_+) \end{pmatrix} + Y_- \begin{pmatrix} \cos(\omega_- t + \varphi_-) \\ \cos(\omega_- t + \varphi_-) \end{pmatrix}.$$

On voit que x est une combinaison linéaire de deux solutions particulières, l'une pour laquelle les deux pendules sont en phase (coefficient Y_+), et l'autre pour laquelle les deux pendules sont en opposition de phase.

Exercice. — Montrer que si $|\omega_+ - \omega_-| \ll \omega_+$ alors l'amplitude d'oscillation de chacun des deux pendules peut moduler, de sorte que périodiquement un seul des deux pendules semble osciller.

Compléments

Exercice. — L'exponentielle d'une matrice M est la matrice notée e^M définie par la somme infinie

$$e^M = I + M + \frac{M}{2} + \frac{M}{3!} + \cdots = \sum_{n\geq 0} \frac{M^n}{n!}\ ;$$

le cours d'analyse montre que

- la suite $I + M + \cdots + \frac{M^n}{n!}$ converge (chaque coefficient converge vers un réel),
- $M\, e^M = e^M M$,
- et, si $M = M_t$ dépend différentiablement d'un paramètre $t \in \mathbb{R}$, la dérivée de e^{M_t} est

$$\frac{de^{M_t}}{dt} = e^{M_t}\frac{dM_t}{dt}.$$

(1) Calculer e^M quand M est diagonale.

(2) Trouver l'équation différentielle satisfaite par le vecteur mobile $x(t) := e^{tM}x_0$, où $A \in M_n(\mathbb{R})$ et $x_0 \in \mathbb{R}^n$.

Solution. —

(1) Si $M = \mathrm{Diag}\,(\lambda_1, ..., \lambda_n)$, $e^M = \mathrm{Diag}\,(e^{\lambda_1}, ..., e^{\lambda_n})$.

(2) On a

$$\frac{dx}{dt}(t) = e^{tM}Mx_0 = Me^{tM}Mx_0,$$

donc $x(t)$ satisfait l'équation différentielle

$$\frac{dx}{dt}(t) = Mx(t).$$

Exercice. — Soit $f : M_n(\mathbb{R}) \to \mathbb{R}$ une fonction continue. Montrer que f est invariante par conjugaison ($f(F^{-1}AF) = f(A)$) si et seulement si elle est invariante par commutation ($f(AB) = f(BA)$).

Solution. — Si f est invariante par conjugaison,

$$\begin{aligned} f(BA) &= f((A^{-1}A)(BA)) \quad \text{parce que } A^{-1}A = I \\ &= f(A^{-1}(AB)A) \quad \text{par associativité du produit matriciel} \\ &= f(AB) \quad \text{par hypothèse.} \end{aligned}$$

Notons qu'ici on a tacitement supposé A inversible. Si ce n'est pas le cas, on peut trouver une suite $(A_n)_{n\geq 0}$ de matrices inversibles qui converge vers A (pourquoi ?) ; ce qui précède montre que $f(A_nB) = f(BA_n)$, et on obtient l'égalité voulue en passant à la limite.

La réciproque est évidente.

Exercice. — Pour toute matrice carrée $M \in M_3(\mathbb{R})$, on définit la *trace* de M comme le nombre

$$\mathrm{tr}(M) = M_{11} + M_{22} + M_{33}.$$

Le lecteur admettra ou vérifiera que, si N est une autre matrice 3×3,

$$\mathrm{tr}(MN) = \mathrm{tr}(NM).$$

(1) Soit $M = \mathrm{Diag}(\lambda_1, \lambda_2, \lambda_3)$ une matrice diagonale. Montrer que

$$\left.\frac{d}{dt}\right|_{t=0} \det(I + tM) = \mathrm{tr}\, M.$$

(2) En déduire cette formule quand M est diagonalisable. On admettra que la même formule reste valide en général ; l'interpréter géométriquement.

(3) Considérons l'équation différentielle linéaire $\dot{M}x$. D'après l'exercice § 4.5, une solution en est $x(t) = e^{tM}x(0)$. L'application linéaire $x(0) \mapsto x(t)$ multiplie donc les volumes par le facteur $\det e^{tM}$. Quel est la dérivée de ce rapport de volume, par rapport à t en $t = 0$ (formule de J. Liouville, mathématicien français, 1809–1882) ?

Solution. —

(1) On a

$$\begin{aligned}\det(I+tM) &= \det\begin{pmatrix}1+t\lambda_1 & 0 & 0\\ 0 & 1+t\lambda_2 & 0\\ 0 & 0 & 1+t\lambda_3\end{pmatrix}\\ &= (1+t\lambda_1)(1+t\lambda_2)(1+t\lambda_3)\\ &= 1+t(\lambda_1+\lambda_2+\lambda_3)+o(t),\end{aligned}$$

donc

$$\left.\frac{d}{dt}\right|_{t=0}\det(I+tM) = \operatorname{tr} M.$$

(2) Si M est conjuguée à une matrice diagonale : $M = FDF^{-1}$ avec D diagonale, la multiplicativité du déterminant montre que

$$\det(I+tM) = \det(F(F^{-1}F+tD)F^{-1}) = \det(FF^{-1})\det(I+tD) = \det(I+tD).$$

D'après la formule obtenue à la première question, maintenant appliquée à D,

$$\left.\frac{d}{dt}\right|_{t=0}\det(I+tM) = \left.\frac{d}{dt}\right|_{t=0}\det(I+tD) = \operatorname{tr} D.$$

Or

$$\operatorname{tr} D = \operatorname{tr}(F^{-1}FM) = \operatorname{tr}(F^{-1}MF) = \operatorname{tr} M.$$

Finalement, on a la même formule que dans le premier cas.

$$\left.\frac{d}{dt}\right|_{t=0}\det(I+tM) = \operatorname{tr} M.$$

Géométriquement, cette formule signifie que, quand on fait subir des variations infinitésimales aux vecteurs de la base canonique, seules les variations de chaque vecteur dans sa propre direction a un effet sur le volume du parallélépipède engendré.

(3) D'après ce qui précède, pour l'équation différentielle $\dot{x} = Mx$, le taux d'accroissement du volume des solutions par rapport à t en $t = 0$ est

$$\left.\frac{d}{dt}\right|_{t=0}\det e^{tA} = \operatorname{tr} A.$$

BIBLIOGRAPHIE

[Art91] M. Artin – *Algebra*, Prentice Hall Inc., Englewood Cliffs, NJ, 1991.

[Car61] H. Cartan – *Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes*, Avec le concours de Reiji Takahashi, Enseignement des Sciences. Hermann, Paris, 1961.

[Gab01] P. Gabriel – *Matrices, Géométrie, Algèbre linéaire*, Nouvelle Bibliothèque mathématique, Cassini, 2001.

[Gra97] A. Gramain – *Géométrie élémentaire*, Hermann, 1997.

[KP08] S. Krantz & H. Parks – *Geometric integration theory*, Birkhauser, 2008.

[Kri07] J.-L. Krivine – *Théorie des ensembles*, 2e éd., Nouvelle bibliothèque mathématique, no. 5, Cassini, 2007.

[PD96] F. Pham & H. Dillinger – *Algèbre linéaire*, Collection Bibliothèque des sciences, Diderot, 1996.

[Sha05] I. R. Shafarevich – *Basic notions of algebra*, Encyclopaedia of Mathematical Sciences, vol. 11, Springer-Verlag, Berlin, 2005, Translated from the 1986 Russian original by Miles Reid, Reprint of the 1997 English translation [MR1634541], Algebra, I.

Guide. —
- **[PD96]** version détaillée et avancée de ce cours
- **[Gab01]** livre riche, avec de nombreuses explications historiques
- **[Art91]** excellent livre d'algèbre pour aller beaucoup plus loin
- **[Sha05]** panorama magistral de l'algèbre moderne.

INDEX