Université Paris Sud

Séries de Fourier et Géométrie euclidienne

M254

Année universitaire 2016 - 2017
(version avec corrigés)

La dernière version de ce polycopié est disponible en ligne à l'adresse :
`http://www.math.u-psud.fr/~auvray/polyM254.pdf`.
Il est par ailleurs certain que de nombreux points peuvent encore être améliorés ; n'hésitez pas à me faire part de vos suggestions en m'écrivant à l'adresse
`hugues.auvray"at"math.u-psud.fr`.

Avant-propos

Les *séries de Fourier* abordées dans ce cours constituent une théorie mathématique élémentaire des phénomènes vibratoires et ondulatoires, et font en tant que telles appel à des notions récurrentes en sciences physiques, telles que fréquence et spectre. Ces objets furent introduits par J. Fourier comme outil d'analyse dans sa modélisation de la diffusion de la chaleur *via* l'équation du même nom [1]. Outre des ramifications théoriques profondes, les séries de Fourier ont depuis trouvé nombre d'applications industrielles, notamment en synthèse sonore et traitement d'images (mp3 et RMN). De manière plus proche de la chimie, ajoutons que l'on peut aussi voir les séries de Fourier comme une version discrète de la *transformation de Fourier*, d'utilité notable en cristallographie.

La seconde partie du cours est plus géométrique. On y introduit la notion de *produit scalaire*, qui correspond à une quantification de l'idée d'orthogonalité – et permet de corriger un énoncé tel que le théorème de Pythagore pour les triangles non rectangles : c'est al-Kashi. Liées au produit scalaire sont ensuite étudiées les *matrices symétriques*, version matricielle du concept-clé d'opérateur auto-adjoint, que l'on rencontre par exemple en mécanique quantique. La diagonalisation étant toujours possible, elles offrent un cadre simple d'application de la théorie spectrale en algèbre linéaire. On termine par les *matrices orthogonales*, dont l'étude permet une compréhension élémentaire des transformations rigides dans le plan et l'espace.

Prérequis

Il sera appréciable d'aborder ce cours en étant familier avec :

- les ensembles de nombres usuels ($\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$), et les bases du langage formel ($\forall$, $\in$, $\Rightarrow$, *etc* ; en particulier, dès qu'un simple « donc » fera l'affaire, on évitera d'utiliser un $\Rightarrow$ ou autre $\Leftrightarrow$, qui serait mal à propos) ;
- les notions de fonction, graphe, continuité, dérivabilité, primitive, intégrale (bien distinguer ces deux dernières) ; les méthodes élémentaires de calcul d'intégrales ; les fonctions usuelles : fonctions polynomiales, logarithme, exponentielles, **fonctions trigonométriques** ; quelques formules de trigonométrie (telles que $\cos(a+b) = \cos a \cos b - \sin a \sin b$, $\sin(a+b) = \sin a \cos b + \sin b \cos a$, et formules inverses) ;
- le calcul sur les nombres complexes ;
- la manipulation des matrices : somme, produit, calcul de déterminants, applications aux systèmes linéaires ; et éventuellement, un soupçon d'algèbre linéaire : notion d'espace vectoriel et de sous-espace, bases, dimension, et produit vectoriel en dimension 3.

1. *Théorie analytique de la chaleur*, 1822.

Mode d'emploi

Les ☛ en marge indiquent les notions indispensables, autrement dit celles sans lesquelles il n'est guère utile de continuer. Les ATTENTION , accompagnés de ⚠ en marge, indiquent des points parfois contre-intuitifs qu'il est sage de prendre le temps de méditer.

Les exercices présents au fil du texte le sont surtout à titre illustratif et « digestif ». Ils ne sont donc pas nécessairement représentatifs du niveau d'un examen ; on se reportera pour cela plutôt vers les TD et les sujets de devoir. Les exercices marqués d'un * demandent quant à eux un peu de ténacité.

Bien que les démonstrations soient une partie inhérente aux mathématiques, on a privilégié ici une exposition mettant en avant quelques énoncés assez simples, charnières de ce cours. Des preuves significatives, d'aspect plus technique mais de ce fait souvent instructives, sont mises à disposition du lecteur en fin de chapitre sous forme d'appendice.

1 Rappels sur les suites (numériques)

La première moitié de ce cours a pour but l'introduction des séries de Fourier. La notion de *série* étant dérivée de celle de *suite*, commençons par quelques rappels sur cette dernière.

1.1 Définition, notation

Les *suites*, objet d'attention de ce chapitre, sont en quelque sorte un analogue discret des fonctions, où les entiers remplacent la variable continue. On peut présenter les suites comme collections infinies de nombres :

☛ **Définition 1.1 (Suite numérique)** *On appelle* ***suite à valeurs réelles/ complexes*** *une famille (infinie) de nombres réels/complexes indexée (numérotée) par les entiers naturels.*

Étant donnés des nombres réels ou complexes s_0, s_1, etc., on note

$$(s_n)_{n\geqslant 0}, \qquad \text{ou éventuellement} \quad (s_n)_{n\in\mathbb{N}}, \tag{1}$$

la suite qu'ils constituent.

Si $m \in \mathbb{N}$, on dit que s_m est le ***terme de rang*** *m de $(s_n)_{n\geqslant 0}$.*

Exemple 1.2 *On peut voir les entiers naturels, rangés dans l'ordre habituel 0, 1, 2, etc., comme une suite : c'est la suite $(n)_{n\geqslant 0}$. On peut faire de même, par exemple, avec les carrés et les puissances de 2 successives ; on considère alors respectivement les suites $(n^2)_{n\geqslant 0}$ et $(2^n)_{n\geqslant 0}$. On appelle suite nulle la famille constituée par 0 répété indéfiniment, i.e. la suite $(0)_n$ dont tout les termes sont nuls.*

Remarque 1.3

- *Dans la notation $(s_n)_{n\geqslant 0}$, le choix de la lettre n en indice est arbitraire ; on aurait tout aussi bien pu noter $(s_k)_{k\geqslant 0}$, $(s_\ell)_{\ell\geqslant 0}$, ou même $(s_{\text{blabla}})_{\text{blabla}\geqslant 0}$. Tout ce qui compte, c'est que l'indice ait le même nom à l'intérieur et à l'extérieur de la parenthèse (tant que ce nom n'est pas déjà pris) ; on évitera donc de rebaptiser $(s_n)_{n\geqslant 0}$ par « $(s_n)_{k\geqslant 0}$ » ou « $(s_s)_{s\geqslant 0}$ ».*
- *Si l'on considère la suite des inverses des entiers naturels, on est contraint d'éviter le terme de rang 0. Plus généralement, on peut vouloir parler d'une famille de nombres s_{1000}, s_{1001}, etc., sans faire appel à d'éventuels termes de rang 0, 1, ..., 999 – soit que de tels termes ne soient pas définis, soit qu'on ne s'y intéresse pas. Deux solutions sont possibles :*
  1. *soit on reste fidèle à la notation (1) en « décalant l'indice », et l'on considère la suite $(s_{n+1000})_{n\geqslant 0}$;*
  2. *soit on assouplit la notation, et l'on s'autorise à noter $(s_n)_{n\geqslant 1000}$ la suite des s_n « commençant au rang 1000 ».*

On peut bien sûr remplacer 1000 par n'importe quel entier n_0 pertinent, et se ramener à $(s_{n+n_0})_{n\geqslant 0}$ ou $(s_n)_{n\geqslant n_0}$; par exemple, la suite des entiers strictement positifs est indifféremment désignée par $(n+1)_{n\geqslant 0}$ ou $(n)_{n\geqslant 1}$.

Exercice 1.4 *Quel est le terme de rang 5 de la suite $(3^n)_{n\geqslant 0}$? Le terme de rang 10 de la suite $(2^{-k})_{k\geqslant 0}$? Le terme de rang 100 de la suite $\left(\frac{2(\ell+1)^2}{1+\sqrt{\ell}}\right)_{\ell\geqslant 0}$?*

1.2 Définir des suites

La méthode la plus immédiate pour définir une suite $(s_n)_{n\geqslant 0}$ *ex nihilo* est de le faire *via* des *formules* : on voit par exemple aisément quel est l'objet noté $(n^3)_{n\geqslant 0}$. On peut utiliser une formule terme à terme ; considérer la suite $\left(\frac{n^n\sqrt{2\pi n}}{e^n}\right)_{n\geqslant 1}$, revient à parler de la suite $(s_n)_{n\geqslant 1}$ définie par : $s_n = \frac{n^n\sqrt{2\pi n}}{e^n}$ pour tout $n \geqslant 1$. Cet artifice peut s'avérer utile lorsqu'une formule globale ne suffit pas : considérer l'exemple de la suite $(t_n)_n \geqslant 0$ définie par $t_n = 2^n$ si $n \leqslant 99$, et $t_n = 3^n$ si $n \geqslant 100$.

Par ailleurs, si l'on dispose déjà d'une ou de plusieurs suites, leur appliquer des fonctions, ou les opérations usuelles, permet de définir de nouvelles suites. On peut ainsi passer de la suite $(n)_{n\geqslant 0}$ à la suite $(e^n)_{n\geqslant 0}$ puis à la suite $\big(\sin(e^n)\big)_{n\geqslant 0}$, ou, étant données une suite $(s_n)_{n\geqslant 0}$ et une fonction f telle que $f(s_n)$ **soit défini pour tout** $n \geqslant 0$ – ATTENTION donc si $f = \ln$ ou autre ! –, passer à la suite $\big(f(s_n)\big)_{n\geqslant 0}$. Étant données deux suites $(s_n)_{n\geqslant 0}$ et $(t_n)_{n\geqslant 0}$, on peut considérer la suite somme $(s_n + t_n)_{n\geqslant 0}$, la suite produit $(s_n t_n)_{n\geqslant 0}$, ainsi que la suite quotient $(s_n/t_n)_{n\geqslant 0}$, si $t_n \neq 0$ pour tout n. Même si $(t_n)_{n\geqslant 0}$ est constante, non nulle dans le cas quotient, on a bien un nouvel objet en général.

Mentionnons encore la *définition récursive* : on définit le terme d'ordre 0, puis pour tout $n \geqslant 0$, le terme d'ordre $n+1$ à partir du terme d'ordre n ou d'autres termes de rang inférieur ; la suite de Fibonacci en est un célèbre exemple. Enfin, on peut aussi envisager des *définitions implicites*, où par exemple le terme d'ordre $n \geqslant 0$ est défini comme solution d'une équation de paramètre n ; par exemple : « soit $(s_n)_{n\geqslant 1}$ telle que pour tout $n \geqslant 1$, s_n est la solution de l'équation $s^n + s - 1 = 0$ dans $[0, 1]$ ».

1.3 Convergence, divergence, limite

1.3.1 Définitions. — Du fait que $\mathbb{N}$ est ordonné, voir une famille de nombres comme une suite induit une certaine dynamique : étant donnée une suite $(s_n)_{n\geqslant 0}$, on a tendance à se dire que « s_q vient après s_p » dès que $q \geqslant p$.

Il est de ce point de vue naturel de se demander comment se comporte une suite au fur et à mesure que l'on parcourt ses termes dans l'ordre, ce que l'on peut reformuler pour une suite $(s_n)_{n\geqslant 0}$ en : que peut-on dire sur s_n lorsque n

devient grand ? Un premier type de réponse, qualitatif, peut s'illustrer par les exemples suivants. On considère les suites $(s_n)_{n\geqslant 0} = \big((-1)^n\big)_{n\geqslant 0}$, $(t_n)_{n\geqslant 0} = \big(n\big)_{n\geqslant 0}$, et $(u_n)_{n\geqslant 0} = \big(2^{-n}\big)_{n\geqslant 0}$; il est alors raisonnable de dire que lorsque n est grand, « s_n oscille indéfiniment (entre 1 et -1) », « t_n devient aussi grand qu'on veut », et « u_n devient aussi petit qu'on veut ».

Ces expressions restant un peu floues, on propose de préciser le langage comme suit :

☛ **Définition 1.5 (Convergence)** *Soit $(s_n)_{n\geqslant 0}$ une suite. On dit que **la suite** (s_n) **converge, ou tend, vers 0** si s_n est arbitrairement petit (en valeur absolue) dès que n est assez grand, soit :*

$$\forall \varepsilon > 0,\ \exists N = N(\varepsilon),\ \forall n \geqslant N,\ |s_n| \leqslant \varepsilon. \tag{2}$$

*Plus généralement, si $\ell \in \mathbb{R}$ (cas réel) ou $\ell \in \mathbb{C}$ (cas complexe), on dit que **la suite** (s_n) **converge, ou tend, vers** ℓ si l'écart entre s_n et ℓ est arbitrairement petit (en valeur absolue) dès que n est assez grand, soit :*

$$\forall \varepsilon > 0,\ \exists N = N(\varepsilon),\ \forall n \geqslant N,\ |s_n - \ell| \leqslant \varepsilon. \tag{3}$$

*Lorsqu'un tel ℓ existe, on dit que $(s_n)_{n\geqslant 0}$ **est convergente, de limite** ℓ, et on note : $\ell = \lim\limits_{n\to+\infty} s_n$.*

*Dans le cas contraire, c'est-à-dire lorsqu'aucun ℓ ne satisfait (3), on dit que $(s_n)_{n\geqslant 0}$ **est divergente**.*

Remarque 1.6 ATTENTION, *il existe des suites divergentes ! C'est le cas par exemple de $\big((-1)^n\big)_{n\geqslant 0}$ et de $(n)_{n\geqslant 0}$ (pourquoi ?). On s'assurera donc, avant de parler de la limite d'une suite donnée, d'avoir mentionné, sinon prouvé, son existence !*

Exemple 1.7 *La suite (2^{-n}) tend vers 0 ; si en effet je fixe un nombre $\varepsilon > 0$ (petit, c'est ce qui importe !), disons $\varepsilon < 1$, alors*[2] *pour **tout** $n \geqslant N := \lceil \frac{\ln 2}{\ln \varepsilon} \rceil$, $|2^{-n}| = 2^{-n} \leqslant \varepsilon$. Bien sûr, un tel N (qui dépend de ε !) n'est pas choisi au hasard ; il s'est agi pour le trouver de résoudre au brouillon l'inéquation $2^{-n} \leqslant \varepsilon$, via notamment l'utilisation de* ln.

Exercice 1.8 *Montrer que $\big(\frac{n}{n+1}\big)_{n\geqslant 0}$ converge ; quelle est sa limite ?*

Exercice 1.9 *Soit $(s_n)_{n\geqslant n_0}$ une suite réelle positive, i.e. pour tout $n \geqslant n_0$, $s_n \geqslant 0$. On suppose que (s_n) est convergente, de limite ℓ ; montrer que $\ell \geqslant 0$.*

2. $\lceil\cdot\rceil$ représente l'arrondi vers le haut : $\lceil 7,28\rceil = 8$, $\lceil\pi\rceil = 4$, $\lceil -1,513\rceil = -1$, *etc.*

La définition 1.5 permet une première distinction entre des suites telles que (2^{-n}) d'une part, et $\big((-1)^n\big)$ et (n) d'autre part. Les comportements de $\big((-1)^n\big)$ et (n) étant sensiblement distincts – l'une reste « confinée » tandis que l'autre « explose » –, on affine l'analyse, dans le cas réel, avec la notion de *limite infinie* :

☛ **Définition 1.10 (Limite infinie)** *Soit (s_n) une suite réelle. On dit que (s_n)* ***diverge vers*** *$+\infty$ si s_n est arbitrairement grand dès que n est assez grand, soit :*

$$\forall A \in \mathbb{R},\ \exists N = N(A),\ \forall n \geqslant N,\ s_n \geqslant A\,; \tag{4}$$

on note alors : $\lim\limits_{n\to+\infty} s_n = +\infty$.

De même, on dit que (s_n) ***diverge vers*** *$-\infty$ si s_n est arbitrairement négativement grand dès que n est assez grand, soit :*

$$\forall A \in \mathbb{R},\ \exists N = N(A),\ \forall n \geqslant N,\ s_n \leqslant A\,;$$

on note alors : $\lim\limits_{n\to+\infty} s_n = -\infty$. *Ceci équivaut à :* $\lim\limits_{n\to+\infty}(-s_n) = +\infty$

Manifestement, on a donc : $\lim\limits_{n\to+\infty} n = +\infty$, mais ni : $\lim\limits_{n\to+\infty}(-1)^n = -\infty$, ni : $\lim\limits_{n\to+\infty}(-1)^n = -\infty$; $\big((-1)^n\big)$ n'a pas de limite, finie ou infinie. Ceci est le cas de la plupart des suites ! D'où une nouvelle injonction à justifier l'existence d'une limite avant de la manipuler.

Exercice 1.11 (« Croissances comparées 1 ») *Soit $\alpha \in \mathbb{R}$ fixé. En revenant à la définition, déterminer l'éventuelle limite de $(n^\alpha)_{n\geqslant 1}$; démontrer que $(n^\alpha \ln n)_{n\geqslant 1}$ tend vers 0 si $\alpha < 0$ et vers $+\infty$ si $\alpha \geqslant 0$; démontrer que $(n^\alpha \exp n)_{n\geqslant 1}$ tend toujours vers $+\infty$.*

1.3.2 Établir des convergences. — Établir des convergences « à la main » n'est pas ce qu'il y a de plus aisé, ni de plus réjouissant. Voyons comment combiner des résultats de convergence déjà démontrés, pour en déduire de nouveaux, de manière moins laborieuse :

Proposition 1.12 (Opérations sur les limites 1) *Soient (s_n) et (t_n) deux suites réelles convergentes de limites respectives ℓ et m, et soit $\lambda \in \mathbb{R}$. Alors les suites $(s_n + t_n)$, (λs_n) et $(s_n t_n)$ sont convergentes, de limites respectives $\ell + m$, $\lambda\ell$ et ℓm. Si de plus $m \neq 0$, alors $t_n \neq 0$ à partir d'un certain rang, et alors $\left(\frac{s_n}{t_n}\right)$ est convergente, de limite $\frac{\ell}{m}$.*

On a des résultats analogues dans le cas complexe.

La démonstration du cas de la somme $(s_n + t_n)$ est instructive ; on la traite en fin de chapitre (§1.4).

On peut aussi parfois combiner des limites infinies, ou limite finie et limite infinie :

Proposition 1.13 (Opérations sur les limites 2) *Soient (s_n) et (t_n) deux suites réelles, avec (s_n) convergente de limite ℓ, et (t_n) divergente, de limite $+\infty$ / $-\infty$. Alors la suite $(s_n + t_n)$ diverge, vers $+\infty$ / $-\infty$ et la suite $\left(\frac{s_n}{t_n}\right)$, bien définie à partir d'un certain rang, converge vers 0.*

Si $\ell > 0$, alors $(s_n t_n)$ diverge vers $+\infty$ / $-\infty$, tandis que si $\ell < 0$, $(s_n t_n)$ diverge vers $-\infty$ / $+\infty$.

Ces résultats s'étendent au cas « $\ell = +\infty$ », considéré > 0 ; on a de même une extension, en-dehors de la somme, au cas « $\ell = -\infty$ », considéré < 0.

Exemple 1.14 *Dans le cas $\ell = 0$, on ne peut rien dire en général de la suite $(s_n t_n)$: prendre par exemple pour (s_n) la suite $\left(\frac{(-1)^n}{n}\right)_{n\geqslant 1}$, et pour (t_n) successivement $\left(\sqrt{n}\right)_{n\geqslant 1}$, $(n)_{n\geqslant 1}$ et $(n^2)_{n\geqslant 1}$.*

Dans le cas $\ell = -\infty$, on ne peut rien dire en général de la suite $(s_n + t_n)$: prendre par exemple pour (s_n) la suite $(-n)_{n\geqslant 0}$, et pour (t_n) successivement $\left(\frac{n}{2}\right)_{n\geqslant 0}$, $(n)_{n\geqslant 0}$, $(n + 376)_{n\geqslant 0}$ et $(2n)_{n\geqslant 0}$.

Enfin, au prix de la continuité, on peut combiner limites de suite et fonction :

Proposition 1.15 (Composition des limites) *Soit (s_n) une suite réelle convergente de limite $\ell \in \mathbb{R}$, et soit f une fonction définie au voisinage de ℓ, continue en ℓ. Alors la suite $\big(f(s_n)\big)$ est bien définie à partir d'un certain rang, et converge, vers $f(\ell)$.*

On a un résultat analogue si f tend vers $+\infty$ ou $-\infty$ en ℓ, ou si (s_n) diverge vers $+\infty$ / $-\infty$, et f admet une limite en $+\infty$ / $-\infty$, finie ou infinie.

La proposition 1.15 permet par exemple de lever certaines indéterminations apparues au moment d'énoncer la proposition 1.13.

Exemple 1.16 (« Croissances comparées 2.0 ») *Soit $\alpha < 0$ et $a > 1$; alors la suite $(a^n n^\alpha)_{n\geqslant 1}$ tend vers $+\infty$:*

- *étape 1 : on « fait descendre les exposants » avec le logarithme, en posant, pour tout $n \geqslant 1$, $s_n = \ln(a^n n^\alpha)$, de sorte que[3] : $a^n n^\alpha = \exp(s_n)$.*
- *étape 2 : pour $n \geqslant 1$, on a : $s_n = n \ln a + \alpha \ln n = n\big(\ln a + \alpha\frac{\ln n}{n}\big)$; d'après l'exercice 1.11, $\frac{\ln n}{n} \xrightarrow{n\to+\infty} 0$, donc d'après les propositions 1.12 puis 1.13, $\ln a + \alpha\frac{\ln n}{n} \xrightarrow{n\to+\infty} \ln a > 0$ (car $a > 1$) puis $s_n = n\big(\ln a + \alpha\frac{\ln n}{n}\big) \xrightarrow{n\to+\infty} +\infty$;*
- *étape 3 : comme $a^n n^\alpha = \exp(s_n)$ pour tout n avec (s_n) divergeant vers $+\infty$ et comme $\lim_{+\infty} \exp = +\infty$, on a $\lim_{n\to+\infty} a^n n^\alpha = +\infty$ d'après la proposition 1.15.*

3. Qu'a-t-on gagné ? On connaît les propriétés de continuité de l'exponentielle ainsi que ses limites en $+\infty$ et $-\infty$, ce qui permettra de conclure si l'on détermine le comportement de (s_n) lorsque $n \to +\infty$, ce qui est un peu plus facile.

Exercice 1.17 (« Croissance comparée 2.1 ») *En s'inspirant de l'exemple 1.16, étudier le comportement à l'infini de* $(a^n n^\alpha)_{n\geqslant 1}$ *selon les valeurs possibles pour* $a > 0$ *et pour* $\alpha \in \mathbb{R}$.

On considère désormais les limites des exercices 1.11 et 1.17 comme standard ; on s'en servira donc sans les rejustifier.

1.3.3 Monotonie et convergence. — On voit dans le paragraphe suivant un critère de convergence d'une nature différente de ceux des propositions 1.12, 1.13 et 1.15. On va avoir besoin de la notion de *croissance*, spécifique des suites réelles :

Définition 1.18 (Croissance, monotonie) *Soit* $(s_n)_{n\geqslant n_0}$ *une suite à valeurs réelles. On dit que* (s_n) *est* ***croissante*** *si, pour tout* $n \geqslant n_0$, $s_{n+1} \geqslant s_n$. *On dit que* (s_n) *est* ***décroissante*** *si, pour tout* $n \geqslant n_0$, $s_{n+1} \leqslant s_n$, *i.e. si* $(-s_n)$ *est croissante. Si* (s_n) *est soit croissante, soit décroissante, on dit qu'elle est* ***monotone***.

Exemple 1.19 *La suite* (2^n) *est croissante, la suite* (2^{-n}) *décroissante. De même, si* $a > 0$, *alors* (a^n) *est croissante si* $a \geqslant 1$, *et décroissante si* $a \leqslant 1$.

Exercice 1.20 (*) *Soit* $(s_n)_{n\geqslant 0}$ *une suite croissante convergente, de limite* $\ell \in \mathbb{R}$. *Montrer que pour tout* n, $s_n \leqslant \ell$.

Comme le suggère la terminologie, une suite croissante « grandit ». Supposons que l'on sache qu'une telle suite ne peut dépasser un certain seuil ; intuitivement, on se dit alors que cette suite va progressivement « s'arrêter de grandir »pour tendre vers une sorte de « taille maximale ». On peut justifier formellement cette idée, ce qui donne :

Proposition 1.21 (Convergence des suites croissantes majorées) *Soit une suite* <u>*croissante*</u> $(s_n)_{n\geqslant n_0}$ *; on suppose qu'il existe un* $M \in \mathbb{R}$ *tel que* (s_n) *reste à valeurs plus petites que* M, *soit :*

$$\forall n \geqslant n_0,\ s_n \leqslant M. \tag{5}$$

Alors (s_n) *est convergente ; de plus, sa limite est inférieure à* M.

Exemple 1.22 *Un être humain donné grandit. Or on sait (c'est génétique) qu'il ne dépassera pas 2,80m. La proposition nous dit alors qu'il tendra vers une taille adulte* ℓ *– et peut-être l'atteindra –, et que* $\ell \leqslant$*2,80m.*

Définition 1.23 *Soient* $(s_n)_{n\geqslant n_0}$ *une suite réelle et* $M \in \mathbb{R}$ *vérifiant* (5) *ci-dessus ; on dit alors que* s_n ***est majorée par*** M.

Plus généralement, on dit d'une suite réelle (s_n) *qu'elle est* ***majorée*** *(tout court) si elle est majorée par* M *pour un certain* $M \in \mathbb{R}$.

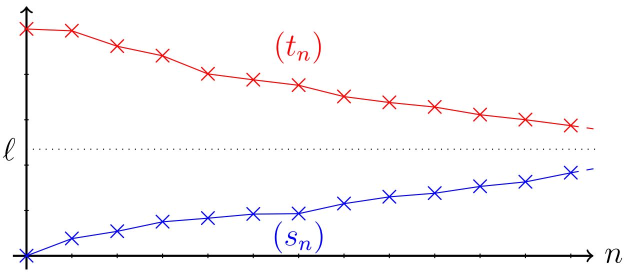
Corollaire 1.24 *Soient $(s_n)_{n\geqslant n_0}$ et $(t_n)_{n\geqslant n_0}$ deux suites réelles, telles que :*

- *(s_n) est croissante et (t_n) est décroissante ;*
- *$(t_n - s_n)$ tend vers 0.*

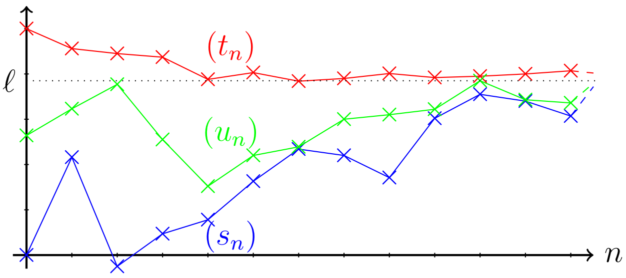
Alors (s_n) et (t_n) sont convergentes, et ont même limite.

Exemple 1.25 *Tout nombre réel admet un développement décimal, éventuellement infini ; à proprement parler, pour tout nombre réel, on peut mettre en évidence deux suites, les développements décimaux inférieur et supérieur, dont la limite commune est le réel de départ. Par exemple, $0,333 \leqslant \frac{1}{3} \leqslant 0,334$ et $0,334 - 0,333 = 10^{-3}$, $0,333333 \leqslant \frac{1}{3} \leqslant 0,333334$ et $0,333334 - 0,333333 = 10^{-6}$, et ainsi de suite ; les suites $(0,\underbrace{3\cdots3}_{n \text{ « 3 »}})_{n\geqslant1}$ et $(0,\underbrace{3\cdots3}_{(n-1) \text{ « 3 »}}4)_{n\geqslant1}$ convergent, vers le réel $\frac{1}{3}$.*

Le corollaire 1.24 peut s'illustrer par un diagramme du type :



Par ailleurs, un diagramme voisin tel que :



permet de rendre compte du fameux :

Théorème 1.26 (Lemme des gendarmes) *Soient (s_n), (t_n) et (u_n) trois suites réelles telles que :*

- *(s_n) et (t_n) sont convergentes, et ont même limite ℓ ;*
- *(u_n) est « coincée » entre (s_n) et (t_n) : pour tout n, $s_n \leqslant u_n \leqslant t_n$;*

Alors (u_n) converge, vers ℓ.

Noter que l'on a retiré les hypothèses de monotonie, mais rajouté des hypothèses de convergence.

1.4 Appendice 1 : démonstration de la proposition 1.12 (cas d'une somme)

Soient (s_n) et (t_n) deux suites, supposées réelles pour le moment, de limites respectives ℓ et m ; on veut voir que $(s_n + t_n)$ prend des valeurs arbitrairement proches de $L := \ell + m$. On commence par montrer que cette suite somme prend, pour n assez grand, des valeurs différant de L d'au plus 1.

Remarquons que pour tout n,

$$\begin{aligned} |(s_n + t_n) - L| &= |(s_n + t_n) - (\ell + m)| = |(s_n + t_n) - (\ell + m)| \\ &= |(s_n - \ell) + (t_n - m)| \\ &\leqslant |s_n - \ell| + |t_n - m| \end{aligned}$$

par inégalité triangulaire. Mutualisons les efforts, et demandons à $|s_n - \ell|$ et $|t_n - m|$ d'être chacun plus petit que $\frac{1}{2}$, de sorte que $|(s_n + t_n) - L|$ soit plus petit que 1. Ce souhait est réaliste, à condition de prendre $n \geqslant N := N\big(\frac{1}{2}\big)$ [4] pour s_n, et disons $n \geqslant N' := N'\big(\frac{1}{2}\big)$ pour t_n. En prenant (n'importe quel) $n \geqslant \max\big(N, N'\big)$, on a donc : $|(s_n + t_n) - L| \leqslant 1$.

La valeur 1 ne jouant *a priori* pas de rôle particulier ici, on voit que l'on aurait pu remplacer 1 par 0,1, ou par 0,001, et ainsi de suite, à condition d'ajuster à chaque fois N et N' ; dans tous les cas, on obtient bien qu'*à condition de prendre n assez grand, $|(s_n + t_n) - L|$ est plus petit que n'importe quel nombre strictement positif fixé au préalable*, soit, formellement :

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N},\ \forall n \geqslant N,\ |(s_n + t_n) - L| \leqslant \varepsilon.$$

En d'autres termes, $(s_n + t_n)$ converge, vers $L = \ell + m$.

Dans le cas complexe, la preuve reste valide, en interprétant les $|\cdot|$ comme des modules. □

Exercice 1.27 *Traiter de même le cas produit « $s_n t_n \xrightarrow{n \to +\infty} \ell m$ » dans la proposition 1.12, en remarquant que toute suite convergente est bornée, i.e. que la suite de ses valeurs absolues/de ses modules est majorée (cf. définition 1.23).*

1.5 Appendice 2 : le symbole de sommation $\sum$

1.5.1 Définition, utilisation. — Pour des raisons de concision et de netteté, il est parfois utile d'adopter une notation synthétique pour les sommes comptant un grand nombre de termes ; par exemple, des expressions telles que

$$1 + 2 + \cdots + 100, \qquad \text{ou} \qquad 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{100},$$

4. En suivant les notations de (3)

ne sont pas très commodes à manipuler, en plus de faire implicitement appel à la sagesse du lecteur pour donner un sens au « $\cdots$ » central. Ces défauts sont encore accentués lorsque les bornes de sommation (dans notre exemple, 1 et 100) sont des paramètres pouvant varier, et non plus des nombres fixes, comme dans les expressions

$$1 + 2 + \cdots + n \qquad \text{et} \qquad 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} \tag{6}$$

pour un entier $n \geqslant 1$. On s'accordera donc sur le fait que des notations telles que

$$\sum_{k=1}^{n} k, \qquad \text{ou} \qquad \sum_{k=1}^{n} \frac{1}{k}, \tag{7}$$

se lisant « somme des k / des $\frac{1}{k}$ pour k allant de 1 à n », remplacent avantageusement celles utilisées en (6) ; ainsi,

$$\sum_{k=1}^{100} k = 1 + 2 + \cdots + 100, \qquad \text{et} \qquad \sum_{k=1}^{100} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{100}.$$

Définition 1.28 (Symbole de sommation $\sum$) *Soit n_0 un entier, et soit $(s_n)_{n \geqslant n_0}$ une suite. Pour $n \geqslant n_0$, on note $\sum_{k=n_0}^{n} s_k$ la somme des termes de rangs n_0 à n :*

$$\sum_{k=n_0}^{n} s_k = s_{n_0} + s_{n_0+1} + \cdots + s_{n-1} + s_n.$$

Remarque 1.29 *Le choix de la lettre k comme indice de sommation est arbitraire ; on aurait tout aussi bien pu noter $\sum_{\ell=n_0}^{n} s_\ell$. On prendra seulement soin de choisir un nom différent de celui des bornes de sommation, ou de toute autre quantité, fixé en-dehors de la somme.*

Si la quantité sous le signe somme ne dépend pas de l'indice, autrement dit, si $s_{n_0} = \cdots = s_n = c$ pour une certaine constante c, alors :

$$\sum_{k=n_0}^{n} s_k = s_{n_0} + \cdots + s_n = \underbrace{c + \cdots + c}_{n-n_0+1 \text{ fois}} = (n - n_0 + 1)c \tag{8}$$

ATTENTION *aussi aux occurrences de l'indice sous le signe somme ; par exemple, à n fixé, $\sum_{k=n_0}^{n} s_n$ est a priori distinct de $\sum_{k=n_0}^{n} s_k$, puisqu'il désigne la somme des s_n, et non des s_k, pour k allant de n_0 à n. Comme s_n ne dépend pas de k (n est fixé hors de la somme, tandis que k n'a de sens qu'à l'intérieur de la somme !), et joue donc le rôle d'une constante pour cet indice, on a fatalement*

$$\sum_{k=n_0}^{n} s_n = (n - n_0 + 1)s_n.$$

Exercice 1.30 *Soit $n \geqslant 1$. Que vaut $\sum_{k=1}^{n} 1$? Montrer que $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$.*

Exercice 1.31 (Relation de Chasles) *Soit $(s_n)_{n \geqslant 0}$ une suite, réelle ou complexe. Soient n et m deux entiers tels que $0 \leqslant m < n$. Montrer que :*

$$\sum_{k=0}^{n} s_k = \sum_{k=0}^{m} s_k + \sum_{k=m+1}^{n} s_k.$$

Exercice 1.32 *Soit $(s_n)_{n \geqslant 0}$ une suite, réelle ou complexe. On définit une suite $(u_n)_{n \geqslant 0}$ par $u_0 = s_0$, et, pour tout $n \geqslant 1$, $u_n = s_n - s_0$. Montrer que la suite $\big(\sum_{k=0}^{n} u_k\big)_{n \geqslant 0}$ n'est autre que (s_n).*

On suppose (s_n) réelle ; (u_n) l'est donc également, et on suppose de plus que pour tout $n \geqslant 1$, $u_n \geqslant 0$. Montrer que (s_n) est croissante.

Exercice 1.33 (Somme des termes d'une suite géométrique) *Soit $a \in \mathbb{C}$, $a \neq 1$. Montrer que pour tout $n \geqslant 0$,*

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

(à retenir !), et plus généralement que $\displaystyle\sum_{k=m}^{n} a^k = \frac{a^m - a^{n+1}}{1-a}$ *dès que $0 \leqslant m \leqslant n$.*

Que dire si $a = 1$?

1.5.2 Linéarité de $\sum$. — On termine par des propriétés immédiates du symbole de sommation $\sum$, directement déduites de la commutativité de l'addition (« $a+b = b+a$ »), et de la distributivité de la multiplication pour l'addition (« $a(b+c) = ab+bc$ »).

Proposition 1.34 (Linéarité de $\sum$) *Soient $(s_n)_{n \geqslant n_0}$ et $(t_n)_{n \geqslant n_0}$ deux suites, et soit $\lambda \in \mathbb{C}$. Alors $\sum_{k=n_0}^{n}(s_k + t_k) = \big(\sum_{k=n_0}^{n} s_k\big) + \big(\sum_{k=n_0}^{n} t_k\big)$, et $\sum_{k=n_0}^{n}(\lambda s_k) = \lambda\big(\sum_{k=n_0}^{n} s_k\big)$.*

Remarque 1.35 ATTENTION*, de même que l'identité $ac+bd = (a+b)(c+d)$ est généralement fausse, il est en général faux que $\sum_{k=1}^{n} s_k t_k = \big(\sum_{k=1}^{n} s_k\big)\big(\sum_{k=1}^{n} t_k\big)$: prendre $n = 2$, et $s_k = t_k = 1$ pour $k = 1, 2$.*

1.6 Corrigés des exercices du chapitre 1

1.6.1 Exercice 1.4. — Ce sont respectivement $3^5 = 243$, $2^{-10} = \frac{1}{1024}$, et $\frac{2\cdot 101^2}{1+10} = \frac{20402}{11}$.

1.6.2 Exercice 1.8. — Pour $n \geqslant 0$, $\frac{n}{n+1} = 1 - \frac{1}{n+1}$; comme $\frac{1}{n+1}$ devient petit lorsque n devient grand, on estime que $\left(\frac{n}{n+1}\right)$ tend vers 1. Justification : soit $\varepsilon > 0$; alors $\left|\frac{n}{n+1} - 1\right| = \left|\frac{1}{n+1}\right| \leqslant \varepsilon$ dès que $n \geqslant N(\varepsilon) = \lceil \frac{1}{\varepsilon} \rceil - 1$. Ceci valant pour tout $\varepsilon > 0$, on a : $\left(\frac{n}{n+1}\right)$ converge, vers 1.

1.6.3 Exercice 1.9. — Soit $\varepsilon > 0$, par exemple $\varepsilon = 0,1$. Pour (au moins un) n assez grand, $s_n - \ell \leqslant \varepsilon$, donc $\ell \geqslant s_n - \varepsilon$; or $s_n \geqslant 0$ et $\varepsilon = 0,1$, donc $\ell \geqslant -0,1$. De même $\ell \geqslant -0,001$, $-0,000001$, ou encore $-10^{-(\text{nombre d'atomes dans l'univers})}$; ℓ est plus grand que tout nombre strictement négatif, aussi proche de 0 soit-il, et donc $\ell \geqslant 0$ – sinon, $\ell < 0$ serait plus grand que $\frac{\ell}{2}$ (qui est < 0), ce qui est intenable pour un nombre strictement négatif.

1.6.4 Exercice 1.11. — Étudions $(n^\alpha)_{n\geqslant 0}$ selon la valeur de α. Si $\alpha = 0$, on a la suite constante égale à 1, qui tend clairement vers 1. Si $\alpha > 0$, on montre que $\lim_{n\to+\infty} n^\alpha = +\infty$: soit $A \in \mathbb{R}$; on peut supposer $A > 0$, et pour tout $n \geqslant N(A) := \lceil A^{1/\alpha} \rceil$, on a $n^\alpha \geqslant A$; comme ceci vaut pour tout A, on a bien la limite annoncée. Si $\alpha < 0$, on montre que $\lim_{n\to+\infty} n^\alpha = 0$: soit $\varepsilon > 0$; alors pour tout $n \geqslant N(\varepsilon) := \lceil \varepsilon^{1/\alpha} \rceil$, $|n^\alpha| = n^\alpha \leqslant \varepsilon$; comme ceci vaut pour tout $\varepsilon > 0$, on a bien la limite annoncée.

Étudions le comportement de $(n^\alpha \ln n)$ à l'infini. On commence par le cas $\alpha \geqslant 0$, plus facile : soit $A > 0$; pour tout $n \geqslant 1$, $n^\alpha \geqslant 1$ tandis que pour tout $n \geqslant N(A) := \lceil e^A \rceil$, $\ln n \geqslant A$, donc $n^\alpha \ln n \geqslant A$ dès que $n \geqslant N(A)$. Ceci valant pour tout A, $(n^\alpha \ln n)$ diverge, vers $+\infty$. Supposons à présent $\alpha < 0$. Soit $n \geqslant 2$; alors $n^{-\alpha} = 1 + (-\alpha)\int_1^n t^{-\alpha-1}\,dt \geqslant (-\alpha)\int_{\sqrt{n}}^n t^{-\alpha}\frac{dt}{t} \geqslant (-\alpha)n^{-\alpha/2}\int_{\sqrt{n}}^n \frac{dt}{t} = \frac{(-\alpha)}{2}n^{-\alpha/2}\ln n$, et donc $n^\alpha \ln n = \frac{\ln n}{n^{-\alpha}} \leqslant \frac{2}{\alpha}n^{\alpha/2}$. Ainsi, si $\varepsilon > 0$, dès que $n \geqslant N(\varepsilon) = \lceil \left(\frac{\varepsilon(-\alpha)}{2}\right)^{2/\alpha} \rceil + 1$, $|n^\alpha \ln n| = n^\alpha \ln n \leqslant \frac{2}{\alpha}n^{\alpha/2} \leqslant \varepsilon$, et ce pour tout $\varepsilon > 0$, d'où : $(\frac{2}{\alpha}n^{\alpha/2})$ converge, vers 0.

On procède de manière analogue pour les suites $(n^\alpha e^n)$; pour le cas $\alpha < 0$, on utilise par exemple que $e^x = 1 + \int_0^x e^t\,dt \geqslant x$ pour $x \geqslant 0$, ce qui appliqué à $x = \frac{n}{2(-\alpha)}$ donne $e^{\frac{n}{2(-\alpha)}} \geqslant \frac{n}{2(-\alpha)}$, puis $e^n \geqslant \frac{n^{2(-\alpha)}}{(2(-\alpha))^{\frac{n}{2(-\alpha)}}}$ pour tout $n \geqslant 1$...

1.6.5 Exercice 1.17. — On a fait le cas $\alpha < 0$, $a > 1$ dans l'exemple 1.16. Si $\alpha = 0$ et $a > 1$, alors $(n^\alpha a^n) = (a^n)$, qui diverge vers $+\infty$; si $\alpha > 0$ et $a > 1$, alors (n^α) et (a^n) divergent vers $+\infty$, de même donc que leur produit $(n^\alpha a^n)$, ce qui conclut l'étude du cas $a > 1$.

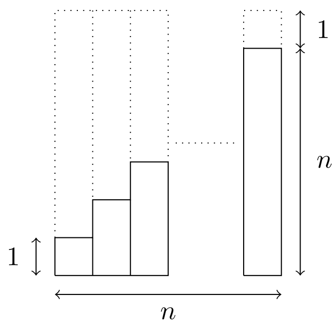
Si $a = 1$, on est ramené à l'étude de (n^α) ; selon que $\alpha < 0$, $\alpha = 0$ ou $\alpha > 0$, on sait que cette suite converge vers 0, converge vers 1, ou diverge vers $+\infty$, respectivement.

Reste le cas $a < 1$; on a alors $a^n \xrightarrow{n\to+\infty} 0$, de sorte que si $\alpha < 0$ (et donc $n^\alpha \xrightarrow{n\to+\infty} 0$) ou $\alpha = 0$ (et donc $n^\alpha \xrightarrow{n\to+\infty} 1$), par produit, on a : $(a^n n^\alpha)$ converge, vers 0. Finalement, on étudie le dernier sous-cas $a < 1$, $\alpha > 0$, où la proposition 1.13 ne permet pas de conclure ; on se ramène comme dans l'exemple 1.16 à l'étude de $(n \ln a + \alpha \ln n)$ *via* le logarithme. Pour tout $n \geqslant 1$, $n \ln a + \alpha \ln n = n\big(\ln a + \alpha \frac{\ln n}{n}\big)$, et la quantité entre parenthèses converge vers $\ln a$, qui est < 0 car $a < 1$. Ainsi $n \ln a + \alpha \ln n \xrightarrow{n\to+\infty} -\infty$, et comme $\lim_{-\infty} \exp = 0$, on a : $(n^\alpha a^n) = \big(\exp(n \ln a + \alpha \ln n)\big)$ converge, vers 0.

1.6.6 Exercice 1.20. — On procède par l'absurde, en supposant que l'assertion « pour tout $n \geqslant 0$, $s_n \leqslant \ell$ » est fausse, c'est-à-dire : il existe un certain n_0 tel que $s_{n_0} > \ell$. Comme (s_n) est croissante, pour tout $n \geqslant n_0$, on a $s_n \geqslant s_{n_0}$, soit : $s_n - s_{n_0} \geqslant 0$, ou : $(s_n - s_{n_0})_{n \geqslant n_0}$ est positive. Or $(s_n - s_{n_0})_{n \geqslant n_0}$ est convergente, de limite $\ell - s_{n_0}$; d'après l'exercice 1.9, on a donc : $\ell - s_{n_0} \geqslant 0$, *i.e.* $\ell \geqslant s_{n_0}$, ce qui contredit l'inégalité $s_{n_0} > \ell$ vue plus haut.

Conclusion : pour tout $n \geqslant 0$, $s_n \leqslant \ell$.

1.6.7 Exercice 1.30. — La somme $\sum_{k=1}^n 1$ n'est autre qu'une somme de n termes valant tous 1, et vaut donc n. Pour démontrer la formule sur $\sum_{k=1}^n k$, on peut procéder graphiquement, avec un dessin du style :



(à méditer). On peut aussi procéder par récurrence [5] . La formule est clairement vraie pour $n = 1$, car $\frac{1\cdot(1+1)}{2} = 1$. Supposons qu'elle soit vraie pour un certain $n \geqslant 1$

5. *Cf. Bases du Raisonnement*, §2.6.

(hypothèse de récurrence) ; alors $\sum_{k=1}^{n+1} k = \underbrace{n+1}_{\text{terme d'ordre } n+1} + \underbrace{\sum_{k=1}^{n} k}_{\text{somme jusqu'à l'ordre } n} = n+1+\frac{n(n+1)}{2}$ par hypothèse de récurrence et donc $\sum_{k=1}^{n+1} k = \frac{n(n+1)}{2}$: la formule est vraie au rang $n+1$. Ceci achève la preuve.

1.6.8 Exercice 1.31. — Il s'agit d'un jeu d'écriture : en revenant à la définition et en mettant en valeurs les termes intermédiaires s_m et s_{m+1}, on a :

$$\sum_{k=0}^{n} s_k = s_0 + \cdots + s_m + s_{m+1} + \cdots + s_n;$$

or on reconnaît dans le membre de droite $\sum_{k=0}^{m} s_k$, sous la forme $s_0 + \cdots + s_m$, suivi de $\sum_{k=m+1}^{n} s_k$, sous la forme $s_{m+1} + \cdots + s_n$, d'où le résultat.

1.6.9 Exercice 1.32. — Dire que $\left(\sum_{k=0}^{n} u_k\right)_{n\geqslant 0}$ et $(s_n)_{n\geqslant 0}$ sont une même suite revient exactement à dire que : pour tout $n \geqslant 0$, $\sum_{k=0}^{n} u_k = s_n$, ce que nous allons vérifier par récurrence. Cette égalité vaut pour $n = 0$, car $\sum_{k=0}^{0} u_k$ est réduite à u_0, qui vaut s_0 par définition. Soit $n \geqslant 0$, et supposons l'égalité vraie au rang n ; alors $\sum_{k=0}^{n+1} u_k = u_{n+1} + \sum_{k=0}^{n} u_k = (s_{n+1} - s_n) + s_n$, en utilisant l'hypothèse de récurrence et la définition de u_{n+1}, et donc $\sum_{k=0}^{n+1} u_k = s_{n+1}$: ceci achève notre démonstration par récurrence.

Dans le cas où les u_k sont des réels positifs, la croissance de (s_n) découle des définitions, puisque pour tout $n \geqslant 0$, on a alors : $s_{n+1} = s_n + u_n \geqslant s_n$ (car $u_n \geqslant 0$).

1.6.10 Exercice 1.33. — On pourrait une fois encore procéder par récurrence. De manière alternative, si $n \geqslant 0$:

$$\begin{aligned}(1-a)\sum_{k=0}^{n} a^k &= \sum_{k=0}^{n} a^k - \sum_{k=0}^{n} a \cdot a^k \\ &= 1 + a + \cdots + a^{n-1} + a^n - (a + \cdots + a^n + a^{n+1}) = 1 - a^{n+1}\end{aligned}$$

par télescopage. On conclut en divisant par $1-a$ qui est non nul (car $a \neq 1$).

La même méthode fonctionne en remplaçant 0 par m comme borne inférieure de sommation. Bien sûr, dans le cas $a = 1$, on récupère directement $\sum_{k=m}^{n} 1 = n - m + 1$.

2 Séries numériques

On continue sur les suites, mais on change de point de vue. Au lieu de voir une suite donnée sous l'angle de la succession de ses termes, on va dans ce chapitre s'intéresser à la suite de *ses variations*. De même que lorsque l'on regarde une fonction, regarder sa dérivée peut être instructif, on va voir que ce changement de point de vue permet de dégager de nouveaux résultats, notamment en termes de *critères de convergence*.

2.1 Séries : définitions

Étant donnée une suite, on peut en constituer une nouvelle, en prenant comme terme d'ordre n la somme des $(n+1)$ premiers termes de la suite initiale :

☛ **Définition 2.1 (Séries numériques)** *Soit $(u_n)_{n\geqslant 0}$ une suite numérique. La* ***série de terme général*** *u_n, $n \geqslant 0$, est la suite $(s_n)_{n\geqslant 0}$ définie par :*

$$\text{(9)} \qquad \textit{pour tout } n \geqslant 0, \quad s_n = u_0 + \cdots + u_n = \sum_{k=0}^{n} u_k.$$

On note cette série $\left(\sum_{k\geqslant 0} u_k\right)$.

Remarque 2.2 *Comme pour les suites, on peut faire varier le point de départ, et démarrer de n'importe quel entier n_0 lorsque cela a un sens ; on notera $\left(\sum_{k\geqslant n_0} u_k\right)$ le cas échéant.*

⚠ **Remarque 2.3** Attention *ici au vocabulaire ; si l'on considère une <u>série</u> $\left(\sum_{k\geqslant 0} u_k\right)$ en tant que telle, son terme (général) sera donc u_n (ou u_k) ; si néanmoins on voit cette série comme une <u>suite</u>, son terme de rang $n \geqslant 0$ sera $s_n = \sum_{k=0}^{n} u_k$.*

La définition 2.5 ci-dessous a pour but de clarifier (un peu) ce point.

Exemple 2.4 (La série harmonique) *On appelle série harmonique la série de terme général $\frac{1}{n}$, $n \geqslant 1$, i.e. la série $\left(\sum_{k\geqslant 1} \frac{1}{k}\right)$. C'est donc la suite indexée par les entiers strictement positifs, dont le terme de rang n est $1 + \cdots + \frac{1}{n}$, pour tout $n \geqslant 1$.*

Définition 2.5 (Sommes partielles) *On reprend les notations de la définition 2.1. Soit $n \geqslant 0$; la* ***somme partielle d'ordre*** *n de la série $\left(\sum_{k\geqslant 0} u_k\right)$ est la somme jusqu'au terme de rang n des u_k, c'est-à-dire $\sum_{k=0}^{n} u_k$ (ce qui n'est autre que le terme de rang n, s_n, de la suite $(s_k)_{k\geqslant 0}$ associée à $\left(\sum_{k\geqslant 0} u_k\right)$.)*

Exercice 2.6 *Soit $n \geqslant 1$; quelle est la somme partielle d'ordre n de la série harmonique (exemple 2.4) ? Quelle est la somme partielle d'ordre n de la série $\left(\sum_{k\geqslant 0} (-1)^k\right)$?*

Exercice 2.7 *Soit $n \geqslant 1$; quelle est la somme partielle d'ordre n de la série $\left(\sum_{k \geqslant 0}(-1)^k k\right)$? En utilisant le résultat de l'exercice 1.30, calculer cette quantité.*

Avec ce vocabulaire, on a que la suite associée à une série n'est autre que la suite des sommes partielles successives de cette série. L'opération réciproque, soit la passage du point de vue des suites au point de vue des séries pour un même objet, a en fait déjà été vu : ce n'est autre que l'exercice 1.32.

2.2 Convergence, somme d'une série

Comme les séries sont des suites (certes vues d'un angle différent), les notions développées pour les suites s'adaptent directement à ce nouveau cadre :

Définition 2.8 (Convergence, somme d'une série) *Soit $\left(\sum_{k \geqslant n_0} u_k\right)$ une série, et $(s_n)_{n \geqslant n_0}$ la suite associée (pour tout $n \geqslant n_0$, $s_n = \sum_{k=n_0}^{n} u_k$).*

*On dit que **la série** $\left(\sum_{k \geqslant n_0} u_k\right)$ **est convergente** si la suite (s_n) l'est.*

*La limite de (s_n) est alors appelée **somme** de la série $\left(\sum_{k \geqslant n_0} u_k\right)$, notée $\sum_{k=n_0}^{+\infty} u_k$.*

Remarque 2.9 ATTENTION *aux différentes utilisation du symbole $\sum$ ici ! La notation $\left(\sum_{k \geqslant n_0} u_k\right)$ (parenthèses et borne inférieure uniquement) désigne une série, i.e. un objet « dynamique » ; la notation $\sum_{k=n_0}^{n} u_k$ (pas de parenthèses, bornes inférieure et supérieure) désigne une somme partielle, i.e. un nombre, dépendant directement des bornes en jeu ; la notation $\sum_{k=n_0}^{+\infty} u_k$ (pas de parenthèses, borne inférieure finie, borne supérieure infinie) désigne une limite, i.e. un nombre fixe. Ainsi, $\left(\sum_{n \geqslant 1} \frac{1}{n}\right)$ est toujours bien définie, $\sum_{k=1}^{n} \frac{1}{k}$ n'est bien défini que si la borne n est précisée, et la bonne définition de $\sum_{k=1}^{+\infty} \frac{1}{k}$ reste pour l'heure en suspens, faute d'avoir tranché le problème de la convergence de la série en jeu (voir l'exercice 2.15 ci-dessous).*

Remarque 2.10 ATTENTION *aussi aux utilisations abusives de $\sum_{k=n_0}^{+\infty}$: comme pour toute limite, on s'assurera d'avoir établi l'existence d'une quantité telle que $\sum_{k=n_0}^{+\infty} u_k$ avant de la manipuler... Sans ces précautions élémentaires, on « démontre » facilement que $\sum_{k=0}^{+\infty}(-1)^k$ vaut simultanément (!) 0, 1, et $\frac{1}{2}$...*

Exemple 2.11 *Comme toute suite peut s'écrire sous la forme d'une série, toute suite convergente peut être vue comme série convergente ! De manière plus naturelle, soit $a \in \mathbb{C}$, et considérons la série $\left(\sum_{k \geqslant 0} a^k\right)$; alors cette série est convergente ssi $|a| < 1$ (voir l'exercice 2.12 ci-dessous).*

Exercice 2.12 (Achille et la tortue) *Démontrer que la série $\left(\sum_{n \geqslant 0} \frac{1}{2^n}\right)$ converge, et calculer sa somme. (*INDICATION *: utiliser le résultat de l'exercice 1.33). Déterminer de même la nature de la série $\left(\sum_{n \geqslant 0} a^n\right)$ pour $a \in \mathbb{C}$ quelconque.*

Définition 2.13 (Divergence) *On dit d'une série non-convergente qu'elle est **divergente**.*

*Si la suite des sommes partielles d'une série réelle diverge vers $+\infty$ (resp. $-\infty$), on dit que la série en jeu **diverge vers** $+\infty$ (resp. $-\infty$).*

Intuitivement, une série évolue au fur et à mesure qu'on lui ajoute des termes, que l'on peut dès lors voir comme des variations. Il semble donc que si une série converge, et donc tend à ne plus évoluer, alors ses variations tendent à devenir nulles. Une première condition nécessaire (mais loin d'être suffisante !) à la convergence d'une série est en effet la suivante :

Proposition 2.14 *Soit $\left(\sum_{n\geqslant n_0} u_n\right)$ une série numérique. On suppose que cette série est convergente. Alors la suite $(u_n)_{n\geqslant n_0}$ converge vers 0.*

Dans l'exemple 2.11, convergence de la série $\left(\sum_{n\geqslant 0} a^n\right)$ et convergence de la suite $(a^n)_{n\geqslant 0}$ vers 0 coïncident ; ce n'est pas toujours le cas, comme l'illustre l'exercice suivant :

Exercice 2.15 *On considère la série harmonique $\left(\sum_{n\geqslant 1} \frac{1}{n}\right)$. Soit $m \geqslant 0$; en utilisant éventuellement la relation de Chasles (exercice 1.31), montrer que :*

$$\sum_{n=1}^{2^m} \frac{1}{n} = 1 + \frac{1}{2} + \sum_{n=3}^{4} \frac{1}{n} + \cdots + \sum_{n=2^{m-1}+1}^{2^m} \frac{1}{n}, \tag{10}$$

puis que $\sum_{n=1}^{2^m} \frac{1}{n} \geqslant 1 + \frac{m}{2}$. En déduire que la série harmonique diverge vers $+\infty$.

2.3 Séries (réelles) à termes positifs

2.3.1 Cas des séries majorées. — Bien souvent, lorsque l'on manipule des nombres, il est plus facile de travailler avec des quantités positives. Ceci s'observe aussi dans le cas des séries ; par exemple, d'après l'exercice 1.32, une série réelle à termes positifs, vue comme suite, est croissante. Le critère sur les suites croissantes majorées (proposition 1.21) devient naturellement dans ce contexte :

Proposition 2.16 (Convergence des séries à termes positifs et à sommes partielles majorées) *Soit $\left(\sum_{n\geqslant n_0} u_n\right)$ une série réelle à termes positifs dont la suite des sommes partielles est majorée, i.e. pour une certain $M \in \mathbb{R}$, et pour tout $n \geqslant n_0$, on a : $\sum_{k=n_0}^{n} u_k \leqslant M$. Alors $\left(\sum_{n\geqslant n_0} u_n\right)$ est convergente, de somme inférieure à M.*

Remarque 2.17 *Comme dans le cas des suites croissantes majorées, ce résultat ne donne qu'une information a priori très partielle sur la somme/limite.*

Exemple 2.18 *On a vu que la série harmonique divergeait. Que se passe-t-il lorsque l'on regarde une série dont le terme général tend sensiblement plus vite vers 0, par exemple* $\left(\sum_{k\geqslant 1}\frac{1}{k^2}\right)$ *? On peut commencer par remarquer que pour tout* $k \geqslant 2$, $\frac{1}{k^2} \leqslant \frac{1}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k}$. *Ceci nous permet de majorer les sommes partielles de notre série : pour tout* $n \geqslant 2$,

$$\sum_{k=1}^{n}\frac{1}{k^2} = 1 + \sum_{k=2}^{n}\frac{1}{k^2} \leqslant 1 + \sum_{k=2}^{n}\left(\frac{1}{k-1} - \frac{1}{k}\right) = 1 + \sum_{k=2}^{n}\frac{1}{k-1} - \sum_{k=2}^{n}\frac{1}{k}$$
$$= 1 + \left(1 + \cdots + \frac{1}{n-1}\right) - \left(\frac{1}{2} + \cdots + \frac{1}{n}\right) = 1 + 1 - \frac{1}{n} \leqslant 2$$

(ce qui est encore vrai pour $n = 1$*). D'après la proposition, la série* $\left(\sum_{k\geqslant 1}\frac{1}{k^2}\right)$ *est donc convergente. Mais tout ce que l'on sait de sa somme pour l'heure, c'est qu'elle est inférieure à 2.*

Exercice 2.19 *Soit* $\alpha > 1$*. En établissant que pour tout* $k \geqslant 2$*, on a :* $\frac{1}{k^\alpha} \leqslant \int_{k-1}^{k}\frac{dt}{t^\alpha}$*, montrer que la série* $\left(\sum_{k\geqslant 1}\frac{1}{k^\alpha}\right)$ *est convergente.*

2.3.2 Un nouveau critère de convergence. — On a vu que, de manière informelle, le terme général d'une série convergente variait de moins en moins. Intuitivement, *une série dont le terme général varie encore moins que celui d'une série convergente devrait donc en être d'autant plus convergente* ; c'est cette idée que la proposition suivante concrétise dans le cas des séries à termes positifs :

Proposition 2.20 (Comparaison de séries à termes positifs) *Soient* $(u_n)_{n\geqslant n_0}$ *et* $(v_n)_{n\geqslant n_0}$ *deux suites à termes positifs. On suppose que la série* $\left(\sum_{n\geqslant n_0} v_n\right)$ *converge. Alors* $\left(\sum_{n\geqslant n_0} u_n\right)$ *converge, et on a :* $\sum_{n=n_0}^{+\infty} u_n \leqslant \sum_{n=n_0}^{+\infty} v_n$.

Grâce aux résultats vus plus haut, la démonstration de cette proposition est relativement simple : comme $\left(\sum_{n\geqslant n_0} v_n\right)$ converge, la suite de ses sommes partielles, croissante, est majorée par sa limite $\ell := \sum_{n=n_0}^{+\infty} v_n$ (proposition 1.21) ; de plus, pour tout $n \geqslant n_0$, $\sum_{n=n_0}^{n} u_k \leqslant \sum_{n=n_0}^{n} v_k$, puisque $u_k \leqslant v_k$ pour tout $k = n_0, \ldots, n$; ainsi, pour tout $n \geqslant n_0$, $\sum_{n=n_0}^{n} u_k \leqslant \ell$, et on conclut par la proposition 2.16.

Exemple 2.21 *On pose* $0! = 1$ *et pour tout* $n \geqslant 1$, $n! = 1 \times 2 \times \cdots \times n$. *Comme* $0 \leqslant \frac{1}{n!} \leqslant \frac{5}{(n+1)^2}$ *pour tout* $n \geqslant 0$*, la série* $\left(\sum_{n\geqslant 0}\frac{1}{n!}\right)$ *converge.*

2.4 Convergence absolue

2.4.1 Définition. — Une série réelle n'a en général pas de raison d'être à termes positifs ; *a fortiori*, une série complexe n'a en général pas de raison d'être à termes réels positifs. On aimerait cependant pouvoir employer le critère de la proposition

2.20 dans un cadre moins restrictif que celui des séries à termes réels positifs. Notons néanmoins qu'étant donnée une série numérique, on peut toujours construire une série à termes réels positifs, en substituant au terme général son module/sa valeur absolue :

$$\text{on remplace} \quad \Big(\sum_{n \geqslant n_0} u_n\Big) \quad \text{par} \quad \Big(\sum_{n \geqslant n_0} |u_n|\Big);$$

par exemple, à $\big(\sum_{n\geqslant 1} \frac{(-1)^{n-1}}{n}\big)$ on substitue la *nouvelle* série $\big(\sum_{n\geqslant 1} \frac{1}{n}\big)$. La notion de *convergence absolue* est utile pour analyser cette transition.

☛ **Définition 2.22 (Convergence absolue)** *Soit* $\big(\sum_{n\geqslant n_0} u_n\big)$ *une série numérique. On dit que* $\big(\sum_{n\geqslant n_0} u_n\big)$ ***converge absolument*** *si la série des modules/valeurs absolues associée* $\big(\sum_{n\geqslant n_0} |u_n|\big)$ *converge.*

Exemple 2.23

1. *On a vu que la série à termes positifs* $\big(\sum_{n\geqslant 1} \frac{1}{n^2}\big)$ *est convergente (exemple 2.18) ; ainsi, la série* $\big(\sum_{n\geqslant 1} \frac{(-1)^n}{n^2}\big)$ *est absolument convergente, de même que toute série du type* $\big(\sum_{n\geqslant 1} \frac{a_n}{n^2}\big)$ *avec* $|a_n| = 1$ *pour tout* $n \geqslant 1$*.*
2. *Puisque la série harmonique diverge, la série* $\big(\sum_{n\geqslant 1} \frac{(-1)^{n-1}}{n}\big)$ *n'est pas absolument convergente.*

2.4.2 Utilisation de la convergence absolue. — Si l'on suit la terminologie, la convergence absolue doit être plus forte que la convergence (tout court) ; en termes de logique, il serait raisonnable que toute série numérique absolument convergente soit convergente, *i.e.* que pour les séries numériques, la convergence absolue implique la convergence. C'est ce résultat fondamental que nous énonçons à présent :

Théorème 2.24 (La convergence absolue implique la convergence) *Toute série numérique absolument convergente est convergente. En d'autres termes, pour qu'une série soit convergente, <u>il suffit</u> qu'elle soit absolument convergente.*

On renvoie à l'appendice en fin de chapitre pour une preuve de ce théorème.

Remarque 2.25 *Malgré les apparences, ce théorème n'a rien d'un simple jeu de langage ! Il dit en substance que l'on peut parfois décider de la convergence d'une série donnée en travaillant sur une autre série – à savoir la série des modules associée. Par exemple, on sait que* $\big(\sum_{n\geqslant 1} \frac{(-1)^n}{n^2}\big)$ *est absolument convergente, ce qui veut exactement dire que* $\big(\sum_{n\geqslant 1} \frac{1}{n^2}\big)$ *est convergente ; c'est le théorème 2.24 qui permet alors bien d'en déduire la convergence de* $\big(\sum_{n\geqslant 1} \frac{(-1)^n}{n^2}\big)$*.*

Exemple 2.26 *La série* $\left(\sum_{n\geqslant 0}\frac{(-1)^n}{n!}\right)$ *est absolument convergente (cf. exemple 2.21), donc convergente.*

Exercice 2.27 *Soit* $x\in\mathbb{R}$*, fixé. En combinant la proposition 2.20 avec le théorème 2.24, démontrer que la série* $\left(\sum_{n\geqslant 1}\frac{\cos(nx)}{n^2}\right)$ *converge.*

Remarque 2.28 ATTENTION*, l'implication (convergence absolue* $\Rightarrow$ *convergence) est à sens unique, dans la mesure où il existe des séries convergentes, mais non absolument convergentes (de même qu'il existe des rectangles qui ne sont pas des carrés). Par exemple,* $\left(\sum_{n\geqslant 1}\frac{(-1)^{n-1}}{n}\right)$ *n'est pas absolument convergente ; on démontre néanmoins « à la main » qu'elle est convergente. En appelant* $(\sigma_n)_{n\geqslant 1}$ *la suite des sommes partielles de cette série, on a, pour tout* $n\geqslant 1$ *:* $s_n\leqslant\sigma_n\leqslant t_n$*, où* $s_n := 1-\frac{1}{2}+\cdots+\frac{1}{2k-1}-\frac{1}{2k}$*, et* $t_n := 1-\frac{1}{2}+\cdots+\frac{1}{2k+1}$ *avec* k *tel que* $n=2k$ *ou* $2k+1$*. On vérifie que* (s_n) *est croissante,* (t_n) *décroissante, et comme*[6] $(t_n-s_n)=\left(\frac{1}{2\lfloor n/2\rfloor+1}\right)$ *tend vers 0, la convergence de* $(\sigma_n)_{n\geqslant 1}$ *s'en suit par le corollaire 1.24 combiné aux gendarmes.*

Exercice 2.29 (Critère de Leibniz, $*$) *En s'inspirant de l'exemple ci-dessus, démontrer que pour toute suite (positive) décroissante* $(v_n)_{n\geqslant n_0}$ *convergeant vers 0, la série* $\left(\sum_{n\geqslant n_0}(-1)^n v_n\right)$ *converge.*

2.5 Appendice : démonstration du théorème 2.24

Soit $\left(\sum_{n\geqslant n_0}u_n\right)$ une série absolument convergente. Pour fixer les idées, on suppose que $n_0=0$; on suppose aussi pour l'instant qu'il s'agit d'une série réelle. On définit deux suites $(a_n)_{n\geqslant 0}$ et $(b_n)_{n\geqslant 0}$ à partir de $(u_n)_{n\geqslant 0}$ comme suit : pour chaque $n\geqslant 0$, on pose $a_n=u_n$ si $u_n\geqslant 0$ et $a_n=0$ si $u_n\leqslant 0$ (cohérent pour $u_n=0$), et $b_n=0$ si $u_n\geqslant 0$ et $b_n=u_n$ si $u_n\leqslant 0$; on a alors (toujours) $u_n=a_n+b_n$ (si par exemple $u_n\geqslant 0$, $a_n+b_n=u_n+0=u_n$, et de même si $u_n\leqslant 0$).

Si l'on démontre que les séries $\left(\sum_{n\geqslant 0}a_n\right)$ et $\left(\sum_{n\geqslant 0}b_n\right)$ sont convergentes, on pourra en déduire par somme que $\left(\sum_{n\geqslant 0}(a_n+b_n)\right)=\left(\sum_{n\geqslant 0}u_n\right)$ l'est aussi. Or par construction, pour tout $n\geqslant 0$, $a_n\geqslant 0$; si l'on sait majorer a_n par le terme général d'une série convergente, alors on pourra en déduire la convergence de $\left(\sum_{n\geqslant 0}a_n\right)$ par la proposition 2.20. La seule série convergente dont on dispose pour l'instant est $\left(\sum_{n\geqslant 0}|u_n|\right)$; or il se trouve que pour tout n, $a_n\leqslant|u_n|$: c'est automatique si $u_n\leqslant 0$, auquel cas $a_n=0$, et si $u_n\geqslant 0$, alors $a_n=u_n=|u_n|$.

Ainsi, $\left(\sum_{n\geqslant 0}a_n\right)$ est convergente ; par symétrie, $\left(\sum_{n\geqslant 0}b_n\right)$ l'est aussi – travailler sur $\left(\sum_{n\geqslant 0}(-b_n)\right)$, et voir que pour tout $n\geqslant 0$, $0\leqslant -b_n\leqslant|u_n|$. On a donc que $\left(\sum_{n\geqslant 0}u_n\right)=\left(\sum_{n\geqslant 0}(a_n+b_n)\right)$ est convergente.

6. $\lfloor\cdot\rfloor$ représente l'arrondi vers le bas, appelé « partie entière ».

Dans le cas d'une série complexe absolument convergente, $\left(\sum_{n\geqslant n_0} w_n\right)$ disons, on se ramène au cas précédent en posant $u_n = \mathfrak{Re}(w_n)$ et $v_n = \mathfrak{Im}(w_n)$ pour tout n. On a alors $0 \leqslant |u_n|, |v_n| \leqslant |w_n|$ pour tout n, et comme $\left(\sum_{n\geqslant n_0} |w_n|\right)$ est convergente, par la proposition 2.20, les séries *réelles* $\left(\sum_{n\geqslant n_0} |u_n|\right)$ et $\left(\sum_{n\geqslant n_0} |v_n|\right)$ convergent. Le cas traité plus haut nous dit que $\left(\sum_{n\geqslant n_0} u_n\right)$ et $\left(\sum_{n\geqslant n_0} v_n\right)$ sont bien convergentes, et donc finalement que $\left(\sum_{n\geqslant n_0}(u_n + iv_n)\right) = \left(\sum_{n\geqslant n_0} w_n\right)$ est convergente. □

2.6 Corrigés des exercices du chapitre 2

2.6.1 Exercice 2.6. — La somme partielle d'ordre n de la série harmonique est le nombre $\sum_{k=1}^{n} \frac{1}{k} = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$. Dans le cas de la série $\left(\sum_{k\geqslant 0}(-1)^k\right)$, la somme partielle d'ordre n est $\left(\sum_{k=0}^{n}(-1)^k\right)$, mais on peut ici être plus explicite : en effet, $\sum_{k=0}^{n}(-1)^k = 1 - 1 + 1 - 1 + \cdots + (-1)^n$. En se débarrassant de tous les $1 - 1$ de cette somme, on obtient donc 0 si la somme se termine par -1, *i.e.* si n est impair, et 1 si la somme se termine par $+1$, *i.e.* si n est pair.

2.6.2 Exercice 2.12. — D'après l'exercice 1.33, on a, pour tout $n \geqslant 0$, $\sum_{k=0}^{n} \frac{1}{2^k} = \frac{1-\frac{1}{2^{n+1}}}{1-\frac{1}{2}} = 2 - \frac{1}{2^n}$; comme $\left(\frac{1}{2^n}\right)_{n\geqslant 0}$ converge (vers 0), la suite des sommes partielles de $\left(\sum_{n\geqslant 0} \frac{1}{2^n}\right)$ converge (vers $2 - 0 = 2$) ; autrement dit $\left(\sum_{n\geqslant 0} \frac{1}{2^n}\right)$ converge et sa somme vaut 2.

Dans le cas général de la série $\left(\sum_{n\geqslant 0} a^n\right)$, si $a \neq 1$, la somme partielle d'ordre n vaut $\frac{1-a^{n+1}}{1-a}$. La convergence (de la suite des sommes partielles donc) de la série est donc équivalente à celle de $(a^n)_{n\geqslant 0}$, *i.e.* à $|a| < 1$; la somme $\sum_{n=0}^{+\infty} a^n$ vaut alors $\frac{1}{1-a}$. Si enfin $a = 1$, la suite des sommes partielles n'est autre que $(n+1)_{n\geqslant 0}$, d'où une divergence avérée !

2.6.3 Exercice 2.7. — La somme partielle d'ordre n de $\left(\sum_{k\geqslant 0}(-1)^k k\right)$ est $\sum_{k=0}^{n}(-1)^k k$. Pour expliciter cette quantité grâce à la formule connue sur $\sum_{k=0}^{n} k$, on additionne ces deux sommes, de sorte que grâce à l'alternance des signes (les $(-1)^k$), on obtienne une simplification dans les calculs :

$$\sum_{k=0}^{n}(-1)^k k + \sum_{k=0}^{n} k = \sum_{k=0}^{n} \underbrace{\left(1 + (-1)^k\right)}_{\substack{=2 \text{ si } k \text{ pair,} \\ 0 \text{ si } k \text{ impair}}} k = 2 \sum_{\substack{0\leqslant k\leqslant n, \\ k \text{ pair}}} k.$$

Or un entier naturel k est pair et $\leqslant n$ si et seulement s'il s'écrit 2ℓ avec $0 \leqslant \ell \leqslant \frac{n}{2}$, et donc $0 \leqslant \ell \leqslant \frac{\lfloor n \rfloor}{2}$, d'où :

$$\sum_{k=0}^{n}(-1)^k k + \sum_{k=0}^{n} k = 2\sum_{\ell=0}^{\lfloor n/2 \rfloor} 2\ell = 4\sum_{\ell=0}^{\lfloor n/2 \rfloor} \ell.$$

On peut à présent appliquer le résultat l'exercice 1.30, avec $\lfloor \frac{n}{2} \rfloor$ et n comme bornes supérieures, ce qui donne :

$$\sum_{k=0}^{n}(-1)^k k = 4\sum_{\ell=0}^{\lfloor n/2 \rfloor} \ell - \sum_{k=0}^{n} k = 4\frac{\lfloor \frac{n}{2} \rfloor(\lfloor \frac{n}{2} \rfloor + 1)}{2} - \frac{n(n+1)}{2}.$$

Si $n = 2m$ est pair, cette expression devient : $2m(m+1) - m(2m+1) = m = \frac{n}{2}$; si $n = 2m+1$ est impair, on obtient : $2m(m+1) - (2m+1)(m+1) = -m-1 = -\frac{n+1}{2}$ (ce dont on pouvait rapidement avoir l'intuition en regardant les $\sum_{k=0}^{n}(-1)^k k$ pour de petits n).

2.6.4 Exercice 2.15. — La première question est un problème simple de réécriture, dont la solution consiste à regrouper les $\frac{1}{n}$, $n \in \{2, \ldots, 2^m\}$, en « paquets » successifs de longueur 2^q, $q = 0, \cdots, m-1$:

$$\sum_{n=1}^{2^m} \frac{1}{n} = 1 + \underbrace{\frac{1}{2}}_{1\,(=2^0)\text{ terme}} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{2\,(=2^1)\text{ termes}} + \underbrace{\frac{1}{5} + \cdots + \frac{1}{8}}_{4\,(=2^2)\text{ termes}} + \cdots + \underbrace{\frac{1}{2^{m-1}+1} + \cdots + \frac{1}{2^m}}_{2^{m-1}\text{ termes}},$$

d'où (10), en remplaçant $\frac{1}{3} + \frac{1}{4}$ par $\sum_{n=3}^{4} \frac{1}{n}$, $\frac{1}{5} + \cdots + \frac{1}{8}$ par $\sum_{n=5}^{8} \frac{1}{n}$, et ainsi de suite jusqu'à $\frac{1}{2^{m-1}+1} + \cdots + \frac{1}{2^m}$, remplacé par $\sum_{n=2^{m-1}+1}^{2^m} \frac{1}{n}$. Notons que l'on aurait pu également procéder à rebours, en appliquant successivement m fois la relation de Chasles au membre de droite de (10) pour en déduire le membre de gauche. Pour obtenir la minoration, on minore chacun des m paquets ci-dessus par sa longueur fois son plus petit terme, en l'occurence le dernier ; ainsi $\frac{1}{3} + \frac{1}{4} \geqslant 2 \times \frac{1}{4} = \frac{1}{2}$, $\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} \geqslant 4 \times \frac{1}{8} = \frac{1}{2}$, et ainsi de suite jusqu'à $\frac{1}{2^{m-1}+1} + \cdots + \frac{1}{2^m} \geqslant 2^{m-1} \times \frac{1}{2^m} = \frac{1}{2}$. On a ainsi que $\sum_{n=1}^{2^m} \frac{1}{n}$, qui s'écrit comme 1 plus une somme de m paquets plus grands que $\frac{1}{2}$, est minorée par $1 + \frac{m}{2}$.

La divergence de la série harmonique est alors claire : si $A \geqslant 0$, pour tout $n \geqslant 2^{2\lfloor A \rfloor}$, on a $\sum_{k=1}^{n} \frac{1}{k} \geqslant \sum_{k=1}^{2^{2\lfloor A \rfloor}} \frac{1}{k} \geqslant 1 + \frac{2\lfloor A \rfloor}{2} \geqslant A$.

2.6.5 Exercice 2.19. — Soit $k \geqslant 2$. Pour tout $t \in [k-1, k]$, $t^\alpha \leqslant k^\alpha$ car $\alpha > 0$, donc $\frac{1}{k^\alpha} \leqslant \frac{1}{t^\alpha}$. En intégrant en t sur $[k-1, k]$, il vient donc : $\frac{1}{k^\alpha} = \int_{k-1}^{k} \frac{dt}{k^\alpha} \leqslant \int_{k-1}^{k} \frac{dt}{t^\alpha}$, et ceci pour tout $k \geqslant 2$.

La série $\left(\sum_{k\geqslant 1}\frac{1}{k^\alpha}\right)$ étant à termes positifs, il suffit d'en établir une majoration pour en déduire sa convergence. Soit donc $n \geqslant 2$; on a :

$$\sum_{k=1}^{n}\frac{1}{k^\alpha} = 1 + \sum_{k=2}^{n}\underbrace{\frac{1}{k^\alpha}}_{\leqslant\int_{k-1}^{k}\frac{dt}{t^\alpha}} \leqslant 1 + \sum_{k=2}^{n}\int_{k-1}^{k}\frac{dt}{t^\alpha} = 1 + \int_1^n \frac{dt}{t^\alpha},$$

d'après la relation de Chasles intégrale [7] appliquée successivement $(n-2)$ fois. Or $\int_1^n \frac{dt}{t^\alpha} = \left[-\frac{1}{\alpha-1}t^{1-\alpha}\right]_1^n = \frac{1}{\alpha-1}(1-n^{1-\alpha}) \leqslant \frac{1}{\alpha-1}$ car $1-\alpha<0$, donc $\sum_{k=1}^{n}\frac{1}{k^\alpha} \leqslant 1+\frac{1}{\alpha-1}$, et ceci pour tout $n \geqslant 2$ (ainsi que pour $n=1$ de manière immédiate). On conclut par la proposition 2.16.

2.6.6 Exercice 2.27. — Pour tout $n \geqslant 1$, $0 \leqslant \left|\frac{\cos(nx)}{n^2}\right| \leqslant \frac{1}{n^2}$. Comme $\left(\sum_{n\geqslant 1}\frac{1}{n^2}\right)$ est convergente (exemple 2.18), d'après la proposition 2.20, la série $\left(\sum_{n\geqslant 1}\left|\frac{\cos(nx)}{n^2}\right|\right)$ est convergente. En d'autres termes, $\left(\sum_{n\geqslant 1}\frac{\cos(nx)}{n^2}\right)$ est absolument convergente, donc convergente, par le théorème 2.24.

2.6.7 Exercice 2.29. — Suivant l'indication, on va intercaler la suite (σ_n) des sommes partielles de $\left(\sum_{n\geqslant n_0}(-1)^n v_n\right)$ entre une suite croissante (s_n) et une suite décroissante (t_n), dont la différence tend vers 0. Pour $n \geqslant n_0$, on pose donc $\sigma_n = \sum_{j=n_0}^{n}(-1)^j v_j$, et on appelle également s_n la somme partielle jusqu'à l'entier impair égal ou immédiatement supérieur à n, et t_n la somme partielle jusqu'à l'entier pair égal ou immédiatement inférieur à n (et $t_{n_0}=0$ si n_0 est impair). Si $n=2k$ est pair, s_n est la somme jusqu'à $2k+1$ et t_n jusqu'à $2k$, et de même si $n=2k+1$ est impair, s_n est la somme jusqu'à $2k+1$ et t_n jusqu'à $2k$. Ainsi, pour n s'écrivant $2k$ ou $2k+1$, en raison de l'alternance des signes,

$$s_n = (-1)^{n_0}v_{n_0} + \cdots + v_{2k} - \underbrace{v_{2k+1}}_{\leqslant 0} \leqslant (-1)^{n_0}v_{n_0} + \cdots + v_{2k} = t_n.$$

De plus, selon que $n=2k$ ou $2k+1$, on a $\sigma_n = s_n$ ou $\sigma_n = t_n$, et l'inégalité ci-dessus peut être raffinée en la chaîne $s_n \leqslant \sigma_n \leqslant t_n$, valable pour tout $n \geqslant n_0$

Par ailleurs, en revenant aux définitions, pour n pair, $s_{n+1} = s_n$, tandis que pour n impair, $s_{n+1} = s_n + v_{n+1} - v_n \leqslant s_n$ car (v_j) décroît ; on en déduit que $(s_n)_{n\geqslant n_0}$ est décroissante. On démontre de même que (t_n) est croissante.

Enfin, suivant que $n=2k$ ou $2k+1$, $0 \leqslant s_n - t_n = v_{2k+1} \leqslant v_n$; comme $v_n \xrightarrow{n\to+\infty} 0$, la suite $(s_n - t_n)$ converge vers 0. Le corollaire 1.24 nous dit donc que (s_n) et (t_n) sont convergentes ; le lemme des gendarmes nous alors que (σ_n) converge, *i.e.* que la série $\left(\sum_{n\geqslant n_0}(-1)^n v_n\right)$ converge.

7. $\int_a^b f(t)\,dt + \int_b^c f(t)\,dt = \int_a^c f(t)\,dt$; ici donc, $\int_1^2 \frac{dt}{t^\alpha} + \int_2^3 \frac{dt}{t^\alpha} + \cdots + \int_{n-1}^{n}\frac{dt}{t^\alpha} = \int_1^n \frac{dt}{t^\alpha}$.

3 Séries de Fourier

Nous abordons maintenant le cœur de la partie d'Analyse de ce cours : les séries de Fourier. Leur théorie est basée sur un *principe de décomposition et de transformation*. Illustrons ce principe par un exemple simple : à un nombre réel a, on peut associer la suite de ses décimales, (a_n) disons, suite à partir de laquelle on peut constituer une série $\big(\sum_{n\geqslant 0} 10^{-n}a_n\big)$, dont la somme [8] est le nombre de départ a ; en résumé :

$$\text{nombre réel } a \rightsquigarrow \begin{array}{c}\text{suite } (a_n)\\ \text{des décimales de } a\end{array} \rightsquigarrow \begin{array}{c}\text{série}\\ \big(\sum\limits_{n\geqslant 0} 10^{-n}a_n\big)\end{array} \rightsquigarrow \begin{array}{c}\text{passage à la}\\ \text{somme } \sum\limits_{n=0}^{+\infty} 10^{-n}a_n = a.\end{array}$$

On retombe donc sur a... mais en exploitant la notion de convergence, on constate que la troncature de l'écriture décimale infinie d'un nombre permet de se faire une idée assez précise de ce nombre : ainsi $0,333333333$ est très proche de $\frac{1}{3}$, l'erreur étant inférieure à 10^{-8}, et cette précision peut être améliorée à l'envi.

Les objets de départ en théorie des séries de Fourier ne sont plus des nombres, mais des *fonctions* ; on considère plus particulièrement des fonctions périodiques, très présentes en sciences physiques dans l'analyse de phénomènes ondulatoires/vibratoires. Le schéma sera alors le suivant :

$$\begin{array}{c}\text{fonction}\\ \text{périodique } f\end{array} \rightsquigarrow \begin{array}{c}\text{suites } \big(a_n(f)\big) \text{ et } \big(b_n(f)\big)\\ \text{(coefficients de Fourier de } f\text{)}\end{array} \rightsquigarrow$$

$$\rightsquigarrow \begin{array}{c}\text{série de Fourier de } f,\\ \text{constituée \textit{via} les}\\ \text{coefficients de Fourier}\end{array} \rightsquigarrow \begin{array}{c}\text{convergence de la}\\ \text{série de Fourier de } f\\ \text{vers } f \text{ elle-même}\end{array}$$

De même que dans la situation précédente, le développement tronqué d'une fonction périodique, modélisant par exemple un signal ou une note de violon, permettra de cerner certaines de ses propriétés. On verra de plus certains cas où le développement de Fourier, utilisé pleinement, permet d'aboutir à des relations numériques non triviales, ce qui ne manque pas de sel, si l'on se rappelle l'origine thermodynamique de cette théorie.

3.1 Coefficients de Fourier

3.1.1 Fonctions 2π-périodiques. — On commence par définir la notion fondamentale de ce chapitre, qui est celle de *fonction 2π-périodique*. Un phénomène

8. Par exemple, à π on associe la suite $(\pi_n)_{n\geqslant 0} := (3,1,4,1,5,9,2,\ldots)$, puis la série $\big(\sum_{n\geqslant 0} 10^{-n}\pi_n\big)$, convergente, dont la somme n'est autre que π.

périodique est un phénomène qui se répète à intervalles (de temps) réguliers, indéfiniment ; pour simplifier les définitions, on demande seulement à ce que le phénomène, modélisé par une fonction d'une variable (que l'on pourra assimiler au temps), soit exactement identique à lui-même à 2π d'intervalle – puisque alors le phénomène sera encore inchangé 2π encore plus tard, soit à 4π d'intervalle de la première mesure, et ainsi de suite.

☛ **Définition 3.1 (Fonction 2π-périodique)** *Une fonction d'une variable à valeurs réelles ou complexes $f : \mathbb{R} \to \mathbb{R}$ ou $\mathbb{C}$ est dite 2π-**périodique** si elle reste inchangée lorsque l'on décale la variable de 2π, c'est-à-dire :*

$$\forall t \in \mathbb{R},\ f(t+2\pi) = f(t). \tag{11}$$

Exemple 3.2 *Les fonctions constantes sont 2π-périodiques ! De manière moins triviale, les fonctions* $\cos$ *et* $\sin$ *sont 2π-périodiques, ainsi que leur combinaison* $\cos + i \sin = \left\{t \mapsto e^{it}\right\}$ *– ce qui motive 2π comme choix de période pour ce chapitre.*

Exercice 3.3 *Soit f une fonction 2π-périodique. Montrer que pour tout $m \in \mathbb{Z}$, la fonction $t \mapsto f(mt)$ est 2π-périodique.*

⚠ ATTENTION, il est important que le paramètre m soit ici entier ; par exemple, $t \to \sin(t/2)$ n'est pas 2π-périodique (pourquoi ?), bien que sin le soit. En revanche, on dispose de plus de souplesse lorsque l'on *translate* la variable :

Exercice 3.4 *Soit f une fonction 2π-périodique. Montrer que pour tout $a \in \mathbb{R}$, la fonction $t \mapsto f(t+a)$ est 2π-périodique.*

Enfin, comme vu avec sin et cos plus haut, on peut combiner entre elles des fonctions périodiques :

Exercice 3.5 *Soient f et g une fonction 2π-périodiques. Montrer que pour tous λ, $\mu \in \mathbb{C}$, la fonctions $\lambda f + \mu g$ est 2π-périodique ; montrer que fg est également 2π-périodique.*

Ces exercices nous donnent déjà de quoi construire de nombreux exemples de fonctions 2π-périodiques ; un exemple fondamental est celui des *polynômes trigonométriques*, qui ne sont autres que des polynômes, auxquels on substitué la fonction trigonométrique élémentaire $t \mapsto e^{it}$, ou son inverse $t \mapsto e^{-it}$, à la variable :

Définition 3.6 *(Polynômes trigonométriques) Un **polynôme trigonométrique** est une fonction – nécessairement 2π-périodique – de la forme*

$$t \longmapsto \sum_{k=-N}^{N} p_k e^{ikt}$$

où $N \in \mathbb{N}$, et les p_k, $k = -N, \ldots, N$, sont dans $\mathbb{C}$.

3.1.2 Définition des coefficients de Fourier. —

☛ **Définition 3.7 (Coefficients de Fourier)** *Soit f une fonction 2π-périodique. On appelle* ***coefficients de Fourier de*** f *les nombres $a_n(f)$, $b_n(f)$ ($n \in \mathbb{N}^*$) et $c_m(f)$ ($m \in \mathbb{Z}$) définis*[9] *par :*

$$a_n(f) = \frac{1}{\pi}\int_0^{2\pi} f(t)\cos(nt)\,dt, \qquad b_n(f) = \frac{1}{\pi}\int_0^{2\pi} f(t)\sin(nt)\,dt,$$
$$\text{et} \quad c_m(f) = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-imt}\,dt. \tag{12}$$

⚠ *(Attention : prendre garde à la présence du signe $-$ dans le e^{-imt} de la dernière intégrale, ainsi qu'au coefficient $\frac{1}{2\pi}$ au lieu de $\frac{1}{\pi}$).*

Exemple 3.8 *(à connaître !) Soit $k \in \mathbb{Z}$ fixé ; on note $\mathfrak{e}_k$ la fonction $t \mapsto e^{ikt}$. Alors pour tout $m \in \mathbb{Z}$,*

$$c_m(\mathfrak{e}_k) = \frac{1}{2\pi}\int_0^{2\pi} \mathfrak{e}_k(t)e^{-imt}\,dt = \frac{1}{2\pi}\int_0^{2\pi} e^{ikt}e^{-imt}\,dt = \frac{1}{2\pi}\int_0^{2\pi} e^{i(k-m)t}\,dt.$$

Or si $k \neq m$, alors $k-m \neq 0$, donc $t \mapsto e^{i(k-m)t}$ admet pour primitive $\frac{1}{i(k-m)}e^{i(k-m)t}$, et ainsi $c_m(\mathfrak{e}_k) = \frac{1}{2\pi}\Big[\frac{1}{i(k-m)}e^{i(k-m)t}\Big]_0^{2\pi} = \frac{1}{2i\pi(k-m)}(1-1) = 0$. Si en revanche $m = k$, i.e. si $k - m = 0$, alors $t \mapsto e^{i(k-m)t}$ est constante égale à 1, et par suite $c_m(\mathfrak{e}_k) = \frac{1}{2\pi}\int_0^{2\pi} 1\,dt = 1$. En résumé :

$$c_m(\mathfrak{e}_k) = \begin{cases} 1, & \text{si } k = m, \\ 0, & \text{si } k \neq m. \end{cases}$$

Remarque 3.9 *Soit f une fonction 2π-périodique. Supposons $m > 0$; alors pour tout $t \in [0, 2\pi]$, $e^{-imt} = \cos(mt) - i\sin(mt)$, donc par la formule (12), $c_m(f) = \frac{1}{2}[a_m(f) - ib_m(f)]$. De même, si $m < 0$, alors $c_m(f) = \frac{1}{2}[a_{-m}(f) + ib_{-m}(f)]$.*

On peut aussi inverser ces relations, en disant que pour tout $n > 0$, $a_n(f) = c_n(f) + c_{-n}(f)$, et $b_n(f) = i\big(c_n(f) - c_{-n}(f)\big)$.

Remarque 3.10 (Linéarité) *Les coefficients de Fourier sont linéaires ! En d'autres termes, si f et g sont 2π-périodiques et si $\lambda, \mu \in \mathbb{C}$, alors $a_n(\lambda f + \mu g) = \lambda a_n(f) + \mu a_n(g)$, $b_n(\lambda f + \mu g) = \lambda b_n(f) + \mu b_n(g)$ pour tout $n \geqslant 1$, et $c_m(\lambda f + \mu g) = \lambda c_m(f) + \mu c_m(g)$ pour tout $m \in \mathbb{Z}$.*

Par contre, <u>on n'a pas ce genre de propriété pour la multiplication</u> ; en général, $a_n(fg)$ et $a_n(f)a_n(g)$ n'ont rien à voir, et de même pour les coefficients b_n et c_m.

9. Pour peu que les intégrales aient un sens, ce qui dans le cadre de ce cours sera toujours le cas !

Exercice 3.11 *En utilisant la remarque 3.9, calculer, pour $k \in \mathbb{Z}$ fixé et pour tout $n > 0$, les coefficients $a_n(\mathfrak{e}_k)$ et $b_n(\mathfrak{e}_k)$.*

D'après la définition 3.7, les coefficients de Fourier a_n, b_n et c_m sont calculés grâce à une intégrale sur le segment $[0, 2\pi]$. Le résultat qui suit nous donne un peu de souplesse quant à cet intervalle de calcul :

Proposition 3.12 *Soit f une fonction 2π-périodique. Alors les $\int_0^{2\pi}$ de la définition 3.7 peuvent être remplacés par $\int_A^{2\pi+A}$; autrement dit, on a*

$$(13) \qquad \begin{aligned} a_n(f) &= \frac{1}{\pi}\int_A^{2\pi+A} f(t)\cos(nt)\,dt, \qquad b_n(f) = \frac{1}{\pi}\int_A^{2\pi+A} f(t)\sin(nt)\,dt, \\ &\text{et} \quad c_m(f) = \frac{1}{2\pi}\int_A^{2\pi+A} f(t)e^{-imt}\,dt, \end{aligned}$$

avec $A \in \mathbb{R}$ quelconque.

Par exemple, on prendra souvent $A = -\pi$, pour se ramener à un calcul sur $[-\pi, \pi]$.

À ce stade, nous avons décrit la partie

$$\begin{array}{c} \text{fonction} \\ \text{périodique } f \end{array} \rightsquigarrow \begin{array}{c} \text{suites } \big(a_n(f)\big) \text{ et } \big(b_n(f)\big) \\ \text{(coefficients de Fourier de } f) \end{array}$$

du diagramme en introduction de ce chapitre. Passons à la partie suivante, qui explique comment recombiner les coefficients de Fourier d'une fonction donnée pour compléter le diagramme de l'introduction.

3.2 Série de Fourier d'une fonction 2π-périodique

☛ **Définition 3.13** *Soit f une fonction 2π-périodique. Soit $t \in \mathbb{R}$. On appelle **série de Fourier de f au point** t la série*

$$(14) \qquad \Big(c_0(f) + \sum_{n\geqslant 1} \big[a_n(f)\cos(nt) + b_n(f)\sin(nt)\big]\Big),$$

ou, de manière équivalente, la série

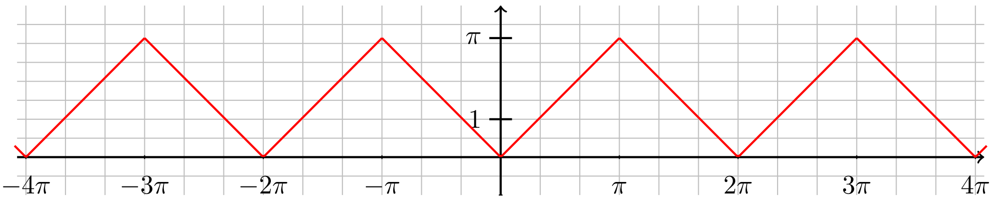
$$(15) \qquad \Big(c_0(f) + \sum_{n\geqslant 1} \big[c_n(f)e^{int} + c_{-n}(f)e^{-int}\big]\Big).$$

⚠ Attention, la série d'une fonction en un point t dépend de ce point, puisqu'en changeant t, on change les séries écrites en (14) et (15) !

Exercice 3.14 *Justifier l'équivalence énoncée dans la définition 3.13 : étant donnée une fonction 2π-périodique f, démontrer que $a_n(f)\cos(nt) + b_n(f)\sin(nt) = c_n(f)e^{int} + c_{-n}(f)e^{-int}$ pour tout $t \in \mathbb{R}$ et tout $n \geqslant 1$.*

Dans de la plupart des exercices sur les séries de Fourier, on demande de calculer la série de Fourier d'une fonction donnée. D'après la définition 3.13, cela revient donc à calculer les coefficients de Fourier, avant de les recombiner sous forme de série selon la formule (14) ou (15) !

Exemple 3.15 *(Il existe plusieurs variantes de cet exemple). On considère f la fonction 2π-périodique (sur $\mathbb{R}$), donnée par $t \mapsto |t|$ **sur** $[-\pi, \pi]$; le graphe de cette fonction sera donc une répétition tous les 2π de la ligne brisée joignant les points $(-\pi, \pi)$, $(0, 0)$ et (π, π) (notons qu'il n'y a pas de problème de définition car $|-\pi| = \pi = |\pi|$) :*



On commence par calculer le coefficient $c_0(f)$, en prenant $[-\pi, \pi]$ comme intervalle d'intégration : $c_0(f) = \frac{1}{2\pi}\int_{-\pi}^{\pi} |t|\, dt = \frac{1}{2\pi}\left(\int_{-\pi}^{0}(-t)\, dt + \int_0^{\pi} t\, dt\right) = \frac{1}{2\pi}\left([-t^2/2]_{-\pi}^{0} + [t^2/2]_0^{\pi}\right)$, soit : $c_0(f) = \frac{\pi}{2}$. Remarquons que les deux demi-intégrales ($\int_{-\pi}^{0}$ et $\int_0^{\pi}$) contribuent à parts égales au résultat : ceci est dû à ce que l'intégrande, $t \mapsto |t|$, est pair sur le segment $[-\pi, \pi]$ symétrique par rapport à 0. De même – cf. exercice ci-dessous – on a, pour tout $n \geqslant 1$, $b_n(f) = 0$, tandis que $a_n(f) = 2 \cdot \frac{1}{\pi}\int_0^{\pi} t\cos(nt)\, dt$; ce dernier calcul requiert une intégration par parties : $\int_0^{\pi} t\cos(nt)\, dt = \left[\frac{t\sin(nt)}{n}\right]_0^{\pi} - \int_0^{\pi}\frac{\sin(nt)}{n} dt = 0 + \left[\frac{\cos(nt)}{n^2}\right]_0^{\pi} = \frac{(-1)^n - 1}{n^2}$. Au final, on a donc $a_n(f) = 0$ si n est pair, et $a_n(f) = -\frac{4}{\pi n^2}$ si n est impair. On peut donc écrire la série de Fourier de f en $t \in \mathbb{R}$ sous la forme :

$$\left(\frac{\pi}{2} - \sum_{\substack{n>0 \\ n \text{ impair}}} \frac{4}{\pi n^2}\cos(nt)\right). \tag{16}$$

Exercice 3.16 (Utilisation des symétries) *Montrer que pour une fonction 2π-périodique paire f, on a $a_n(f) = \frac{2}{\pi}\int_0^{\pi} f(t)\cos(nt)\, dt$ et $b_n(f) = 0$ pour tout $n \geqslant 1$. Montrer que de même pour une fonction 2π-périodique impaire g, on a $c_0(g) = 0$, $a_n(g) = 0$ et $b_n(f) = \frac{2}{\pi}\int_0^{\pi} g(t)\sin(nt)\, dt$ pour tout $n \geqslant 1$.*

En utilisant la convergence de la série $\left(\sum_{n\geqslant 1}\frac{1}{n^2}\right)$ avec les critères de convergence du chapitre 2, notamment la proposition 2.20 et le théorème 2.24, on obtient que pour tout $t \in \mathbb{R}$, la série (16) est convergente. Il est donc naturel de se poser la question suivante : que vaut sa somme – qui *a priori* dépend de t, comme la série elle-même ?

Réponse à la prochaine section...

3.3 Convergence des séries de Fourier

Pour répondre à la question de la somme éventuelle de la série de Fourier d'une fonction 2π-périodique, penchons-nous sur l'exemple élémentaire d'un polynôme trigonométrique, $P : t \mapsto \sum_{k=-N}^{N} p_k e^{ikt}$ disons ; on utilise les coefficients c_m. Soit $m \in \mathbb{Z}$; par linéarité, $c_m(P) = c_m(p_{-N}\mathfrak{e}_{-N} + \cdots + p_N\mathfrak{e}_N) = p_{-N}c_m(\mathfrak{e}_{-N}) + \cdots + p_N c_m(\mathfrak{e}_N)$, *i.e.* $c_m(P) = \sum_{k=-N}^{N} p_k c_m(\mathfrak{e}_k)$. Or on connaît les $c_m(\mathfrak{e}_k)$ (exemple 2.21) ! $c_m(\mathfrak{e}_k) = 1$ si $k = m$, et 0 sinon ; par suite, $c_m(P) = p_m$ si $-N \leqslant m \leqslant N$, et $c_m(P) = 0$ si $|m| > N$. En particulier, seul un nombre fini de $c_m(P)$ sont non-nuls ; étant donné $t \in \mathbb{R}$, la série de Fourier $\big(p_0 + \sum_{n\geqslant 1}[p_n e^{int} + p_{-n}(f)e^{-int}]\big)$ a donc ses termes nuls, hormis un nombre fini, et est donc convergente, de somme $p_0 + \sum_{n=1}^{N}[p_n e^{int} + p_{-n}e^{-int}]$, puisque $p_n e^{int} + p_{-n}e^{-int} = 0$ si $|n| > N$.

Remarquons que cette somme peut être réécrite sous la forme $\sum_{n=-N}^{N} p_n e^{int}$, qui n'est autre que $P(t)$, et ceci pour tout $t \in \mathbb{R}$! On a donc démontré que *dans le cas d'un polynôme trigonométrique, la série de Fourier converge en tout point, vers la valeur du polynôme en ce point.*

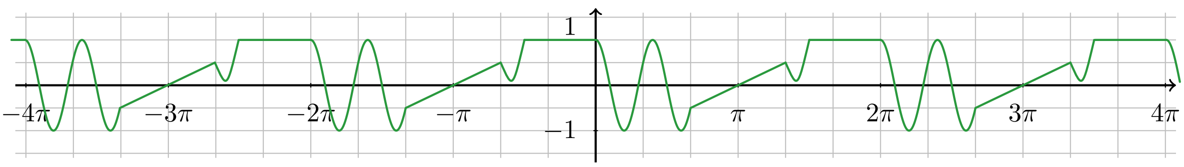
Le résultat principal de ce chapitre nous dit que ceci reste vrai de manière plus générale, à condition de demander un peu de régularité sur la fonction étudiée :

☛ **Théorème 3.17 (Théorème de Dirichlet C^0)** *Soit f une fonction 2π-périodique, « continue et C^1 par morceaux ». Alors pour tout $t \in \mathbb{R}$, la série de Fourier de f au point t est convergente, et sa somme, notée*

$$c_0(f) + \sum_{n=1}^{+\infty} \big[a_n(f)\cos(nt) + b_n(f)\sin(nt)\big], \quad \text{ou bien} \quad \sum_{m=-\infty}^{+\infty} c_m(f)e^{imt},$$

vaut précisément $f(t)$.

Plutôt qu'une définition abstraite, donnons un exemple typique de graphe de fonction continue et C^1 par morceaux :



(on a donc une fonction *continue*, qui peut *présenter des angles* – ne pas être dérivable – en certains points [10], mais admet *partout dérivées à droite et à gauche*).

10. Les points de non-dérivabilité sont en nombre fini sur une période.

En se rappelant de la définition de la convergence, on peut interpréter ce théorème en disant qu'en un certain sens, les sommes partielles (finies, donc) de la série de Fourier d'une fonction approchent de manière arbitraire la fonction elle-même. À une erreur aussi petite que voulue près, on peut donc ramener la donnée d'un signal *continu*, périodique, à la donnée d'un nombre *fini* de ses coefficients de Fourier, suffisants pour reconstituer une somme partielle de série de Fourier, version tronquée du signal de départ.

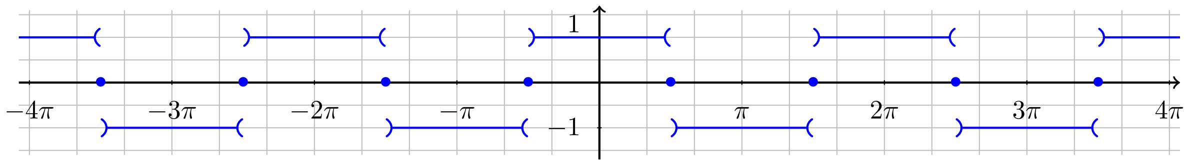
Exemple 3.18 *La fonction « en dents de scie » f de l'exemple 3.15 est continue et C^1 par morceaux ; le théorème 3.17 nous dit alors que sa série de Fourier converge en tout point (ce que nous savions certes déjà, mais il y a des situations moins explicites), et surtout qu'en tout point, somme et fonction coïncident, i.e.*

$$f(t) = \frac{\pi}{2} - \sum_{\substack{n=1 \\ n \text{ impair}}}^{+\infty} \frac{4}{\pi n^2} \cos(nt),$$

pour tout $t \in \mathbb{R}$. En faisant par exemple $t = 0$ dans cette relation, on obtient $\sum_{n=1,\, n \text{ impair}}^{+\infty} \frac{4}{\pi n^2} = \frac{\pi}{2}$.

Exercice 3.19 (Réponse au problème de Bâle) *À partir de la dernière égalité ci-dessus, montrer que $\sum_{n=1}^{+\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.*

On aurait aussi pu considérer une fonction périodique simple, comme la fonction créneau :



qui présente des discontinuités. Comme on peut encore calculer ses coefficients de Fourier, et donc constituer ses séries de Fourier, on peut à nouveau s'interroger sur la convergence de celles-ci. Le résultat suivant fournit la réponse :

Théorème 3.20 (Théorème de Dirichlet général) *Soit f une fonction 2π-périodique, « C^1 par morceaux ». Alors pour tout $t \in \mathbb{R}$, la série de Fourier de f au point t est convergente, et sa somme vaut précisément $f(t)$ si f est continue au point t, ou $\frac{1}{2}\big(f_{\mathrm{g}}(t) + f_{\mathrm{d}}(t)\big)$ sinon.*

Dans cet énoncé, $f_{\mathrm{g}}(t)$ désigne la limite à gauche de f en t (souvent notée $f(t^-)$), $f_{\mathrm{d}}(t)$ la limite à droite (souvent notée $f(t^+)$). Une *fonction C^1 par morceaux* est la même chose qu'une fonction continue C^1 par morceaux, à ceci près qu'elle peut admettre des points de discontinuité (en nombre fini sur une période).

3.4 Égalité de Parseval

On voit dans cette partie conclusive une manière alternative de recombiner les coefficients de Fourier d'une fonction 2π-périodique, de sorte à retomber non plus sur la fonction elle-même, mais sur une quantité intégrale calculée à partir de cette fonction.

Observons la propriété suivante des polynômes trigonométriques : on choisit $P : t \mapsto \sum_{k=-N}^{N} p_k e^{ikt}$, et on calcule son « énergie » $\frac{1}{2\pi}\int_0^{2\pi} |P(t)|^2\,dt$:

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi} |P(t)|^2\,dt &= \frac{1}{2\pi}\int_0^{2\pi} P(t)\overline{P(t)}\,dt = \frac{1}{2\pi}\int_0^{2\pi}\Big(\sum_{k=-N}^{N} p_k e^{ikt}\Big)\Big(\sum_{\ell=-N}^{N} \overline{p_\ell} e^{-i\ell t}\Big)\,dt\\ &= \frac{1}{2\pi}\sum_{k,\ell=-N}^{N} p_k\overline{p_\ell}\int_0^{2\pi} e^{i(k-\ell)t}\,dt.\end{aligned}$$

Or on a vu que $\int_0^{2\pi} e^{i(k-\ell)t}\,dt$ est nul si $k \neq \ell$, et vaut 2π si $k = \ell$. La somme sur k et ℓ ci-dessus peut être remplacée par $\sum_{k,\ell=-N,k=\ell}^{N}$, soit $\sum_{m=-N}^{N}$, et $\frac{1}{2\pi}\int_0^{2\pi} |P(t)|^2\,dt = \sum_{m=-N}^{N} |p_m|^2$. Cette quantité peut encore s'écrire $\sum_{m=-\infty}^{+\infty} |c_m(P)|^2$, en se souvenant que $c_m(P) = p_m$ si $|m| \leqslant N$, et $c_m(P) = 0$ sinon. En résumé,

$$\frac{1}{2\pi}\int_0^{2\pi} |P(t)|^2\,dt = \sum_{m=-\infty}^{+\infty} |c_m(P)|^2,$$

dès que P est un polynôme trigonométrique, et ce type d'égalité se généralise à des fonctions 2π-périodiques bien plus générales :

Théorème 3.21 (Égalité de Parseval) *Soit f une fonction 2π-périodique « de carré intégrable sur une période » (par exemple, continue par morceaux). Alors la série $\big(|c_0(f)|^2 + \frac{1}{2}\sum_{n\geqslant 1}[|a_n(f)|^2 + |b_n(f)|^2]\big)$, qui s'écrit également $\big(|c_0(f)|^2 + \sum_{n\geqslant 1}[|c_n(f)|^2 + |c_{-n}(f)|^2]\big)$, est convergente, de somme $\frac{1}{2\pi}\int_0^{2\pi} |f(t)|^2\,dt$, soit :*

$$(17)\qquad |c_0(f)|^2 + \frac{1}{2}\sum_{n=1}^{+\infty}[|a_n(f)|^2 + |b_n(f)|^2] = \sum_{m=-\infty}^{+\infty} |c_m(f)|^2 = \frac{1}{2\pi}\int_0^{2\pi} |f(t)|^2\,dt.$$

À nouveau, on a un énoncé en deux temps : on établit la convergence d'une série, puis on en précise la somme. Noter également qu'ici, la série ne dépend plus d'un point ; par conséquent, sa somme non plus.

Exercice 3.22 *Que donne l'égalité de Parseval appliquée à la fonction en dents de scie des exemples 3.15 et 3.18 ?*

Exercice 3.23 *Soit f une fonction 2π-périodique de classe C^1 ; calculer les $c_m(f')$ en fonction des $c_m(f)$. Utiliser l'égalité de Parseval pour démontrer que :*

$$\frac{1}{2\pi}\int_0^{2\pi}|f|^2 \leqslant |c_0(f)|^2 + \frac{1}{2\pi}\int_0^{2\pi}|f'|^2.$$

3.5 Appendice : démonstration de la proposition 3.12

On va démontrer plus généralement que pour toute fonction 2π-périodique $F : \mathbb{R} \to \mathbb{C}$ et tout $A \in \mathbb{R}$, on a $\int_0^{2\pi} F(t)\,dt = \int_A^{2\pi+A} F(t)\,dt$; prendre $F : t \mapsto f(t)\cos(nt)$, $f(t)\sin(nt)$ ou $f(t)e^{-imt}$ comme dans l'énoncé suffira alors à conclure.

On écrit tout d'abord $A = 2k\pi + a$, $k = \lfloor \frac{A}{2\pi} \rfloor \in \mathbb{Z}$, $a \in [0, 2\pi[$, de sorte que par relation de Chasles,

$$\text{(18)} \qquad \int_A^{2\pi+A} F(t)\,dt = \int_{2k\pi+a}^{2k\pi} F(t)\,dt + \int_{2k\pi}^{2(k+1)\pi} F(t)\,dt + \int_{2(k+1)\pi}^{2(k+1)\pi+a} F(t)\,dt.$$

Le changement de variable $u = t - 2k\pi$ dans l'intégrale centrale $\int_{2k\pi}^{2(k+1)\pi} F(t)\,dt$, la transforme en $\int_0^{2\pi} F(u + 2k\pi)\,du$; comme F est 2π-périodique, on a $F(u + 2k\pi) = F\big(u + 2(k \pm 1)\pi\big) = \cdots = F(u)$ pour tout tout $u \in [0, 2\pi]$, et l'intégrale est donc égale à $\int_0^{2\pi} F(u)\,du$.

Traitons l'intégrale $\int_{2k\pi+a}^{2k\pi} F(t)\,dt$, réécrite $-\int_{2k\pi}^{2k\pi+a} F(t)\,dt$; on procède cette fois au changement de variable $v = t+2\pi$, qui donne $-\int_{2k\pi}^{2k\pi+a} F(t)\,dt = -\int_{2(k+1)\pi}^{2(k+1)\pi+a} F(v-2\pi)\,dv = -\int_{2(k+1)\pi}^{2(k+1)\pi+a} F(v)\,dv$, toujours par 2π-périodicité.

La variable d'intégration étant muette, cette intégrale est donc l'opposée de l'intégrale $\int_{2(k+1)\pi}^{2(k+1)\pi+a} F(t)\,dt$; on a donc au final $\int_A^{2\pi+A} F(t)\,dt = \int_0^{2\pi} F(u)\,du = \int_0^{2\pi} F(t)\,dt$ (à nouveau, la variable d'intégration est muette !). □

3.6 Corrigés des exercices du chapitre 3

3.6.1 Exercice 3.3. — Fixons un entier m. Si $m = 0$, la fonction $t \mapsto f(mt)$ est la fonction constante égale à $f(0)$, et est donc 2π-périodique. Supposons à présent $m > 0$; si $t \in \mathbb{R}$, alors $f\big(m(t + 2\pi)\big) = f(mt + 2m\pi) = f\big(mt + 2(m-1)\pi\big)$ car f est 2π-périodique, et de même $f\big(mt + 2(m-1)\pi\big) = f\big(mt + 2(m-2)\pi\big) = \cdots = f(mt + 2\pi)$, *i.e.* $f\big(m(t + 2\pi)\big) = f(mt + 2\pi)$, et ce pour tout $t \in \mathbb{R}$: la fonction $t \mapsto f(mt)$ est bien 2π-périodique.

Enfin, si $m < 0$, on procède de la même manière, en ajoutant autant de fois que nécessaire 2π à l'argument [11] de $f\big(m(t+2\pi)\big)$ pour retomber sur $f(t)$, et conclure à l'identique.

3.6.2 Exercice 3.4. — Soit $t \in \mathbb{R}$. Alors $f\big((t+2\pi)+a\big) = f(t+2\pi+a) = f(t+a+2\pi) = f\big((t+a)+2\pi\big) = f(t+a)$, par 2π-périodicité de f. Ceci vaut pour tout réel t ; la fonction $t \mapsto f(t+a)$ est donc 2π-périodique.

3.6.3 Exercice 3.5. — Soit $t \in \mathbb{R}$. Alors par définition, $(\lambda f + \mu g)(t+2\pi) = \lambda \cdot f(t+2\pi) + \mu \cdot g(t+2\pi)$; or par 2π-périodicité, $f(t+2\pi) = f(t)$ et $g(t+2\pi) = g(t)$ donc $(\lambda f + \mu g)(t+2\pi) = \lambda \cdot f(t) + \mu \cdot g(t) = (\lambda f + \mu g)(t)$. Ceci est vrai pour tout $t \in \mathbb{R}$; Par conséquent, $\lambda f + \mu g$ est 2π-périodique.

On procède exactement de la même manière (on « écrit les choses ») pour démontrer que fg est 2π-périodique.

3.6.4 Exercice 3.11. — Soit $n > 0$. On commence par le calcul de $a_n(\mathfrak{e}_k)$. D'après la remarque 3.9, $a_n(\mathfrak{e}_k) = \frac{1}{2}\big(c_n(\mathfrak{e}_k) + c_{-n}(\mathfrak{e}_k)\big)$. Si $k = \pm n$ (soit $n = \pm k$), alors $a_n(\mathfrak{e}_k) = 1$ – sachant qu'on n'a pas $k = n$ et $k = -n$ simultanément, puisque $n \neq 0$; si $k \neq \pm n$ (soit $n = \pm k$), alors $a_n(\mathfrak{e}_k) = 0$.

Calculons à présent $b_n(\mathfrak{e}_k)$; toujours par la remarque 3.9, $b_n(\mathfrak{e}_k) = \frac{1}{2i}\big(c_{-n}(\mathfrak{e}_k) - c_n(\mathfrak{e}_k)\big)$. En conséquence, si $k = n$, $b_n(\mathfrak{e}_k) = \frac{-1}{2i}$; si $k = -n$, $b_n(\mathfrak{e}_k) = \frac{1}{2i}$; et si $k \neq \pm n$, $b_n(\mathfrak{e}_k) = 0$.

3.6.5 Exercice 3.14. — Fixons f et t comme dans l'énoncé. On utilise la relation mentionnée dans la remarque 3.9 : pour tout $n \geqslant 1$, $a_n(f) = c_n(f) + c_{-n}(f)$, et $b_n(f) = i\big(c_n(f) - c_{-n}(f)\big)$. Ceci donne, en utilisant les formules de de Moivre :

$$\begin{aligned}
&a_n(f)\cos(nt) + b_n(f)\sin(nt)\\
&\qquad = \big(c_n(f) + c_{-n}(f)\big)\frac{e^{int}+e^{-int}}{2} + i\big(c_n(f) - c_{-n}(f)\big)\frac{e^{int}-e^{-int}}{2i}\\
&\qquad = c_n(f)\Big(\frac{e^{int}+e^{-int}}{2} + i\frac{e^{int}-e^{-int}}{2i}\Big) + c_{-n}(f)\Big(\frac{e^{int}+e^{-int}}{2} - i\frac{e^{int}-e^{-int}}{2i}\Big)\\
&\qquad = c_n(f)e^{int} + c_{-n}(f)e^{-int}.
\end{aligned}$$

3.6.6 Exercice 3.16. — Traitons le cas pair ; on fixe f une fonction paire 2π-périodique, et $n \geqslant 1$. Comme la fonction $t \mapsto \cos(nt)$ est paire puisque cos l'est, le produit $F : t \mapsto f(t)\cos(nt)$ l'est également. De plus, en utilisant la proposition 3.12 avec $A = -\pi$, $a_n(f) = \frac{1}{\pi}\int_{-\pi}^{\pi} F(t)\,dt = \frac{1}{\pi}\big(\int_{-\pi}^{0} F(t)\,dt + \int_0^{\pi} F(t)\,dt\big)$.

11. Dans une expression de la forme $g(x)$ où g est une fonction, l'*argument* est x ; par exemple, dans « $\sin(2\pi/3)$ », l'argument est $2\pi/3$.

Pour conclure, il s'agit de voir que $\int_{-\pi}^{0} F(t)\,dt = \int_0^\pi F(t)\,dt$. Procédons au *changement de variable* $u = -t$ dans la première intégrale, qui donne $\int_{-\pi}^{0} F(t)\,dt = \int_\pi^0 F(-u)\,(-du)$; en échangeant les bornes et en remarquant que $F(-u) = F(u)$ pour tout $u \in [0,\pi]$, on a bien $\int_{-\pi}^{0} F(t)\,dt = \int_0^\pi F(u)\,du$, qui n'est autre que $\int_0^\pi F(t)\,dt$, la variable d'intégration étant muette (de même que $\sum_{k=1}^N v_k = \sum_{n=1}^N v_n$), d'où le résultat : $a_n(f) = 2 \times \frac{1}{\pi}\int_0^\pi F(t)\,dt = \frac{2}{\pi}\int_0^\pi f(t)\cos(nt)\,dt$.

Montrons que $b_n(f) = 0$. Ici, le produit $t \mapsto f(t)\sin(nt)$ est impair, car sin est impair. Comme $b_n(f) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin(nt)\,dt = \frac{1}{\pi}\big(\int_{-\pi}^{0} f(t)\sin(nt)\,dt + \int_0^\pi f(t)\sin(nt)\,dt\big)$ (proposition 3.12), il suffit pour conclure de voir que les deux demies intégrales se compensent, soit : $\int_{-\pi}^{0} f(t)\sin(nt)\,dt = -\int_0^\pi f(t)\sin(nt)\,dt$. On le fait exactement comme dans le cas précédent *via* le changement de variable $u = -t$, en remarquant que $f(-u)\sin\big(n(-u)\big) = -f(u)\sin(nu)$ pour tout $u \in [0,\pi]$.

On traite le cas impair de manière identique.

3.6.7 Exercice 3.19. — Remarquons tout d'abord que puisque les entiers $\geqslant 1$ se partagent (strictement) entre entiers $\geqslant 1$ pairs et entiers $\geqslant 1$ impairs, on a

$$\sum_{n=1}^{+\infty} \frac{1}{n^2} = \sum_{\substack{n=1 \\ n \text{ pair}}}^{+\infty} \frac{1}{n^2} + \sum_{\substack{n=1 \\ n \text{ impair}}}^{+\infty} \frac{1}{n^2},$$

les trois séries en jeu étant convergentes. De plus, les n pairs strictement positifs sont exactement les $2k$, avec $k \geqslant 1$; ainsi,

$$\sum_{\substack{n=1 \\ n \text{ pair}}}^{+\infty} \frac{1}{n^2} = \sum_{k=1}^{+\infty} \frac{1}{(2k)^2} = \frac{1}{4}\sum_{k=1}^{+\infty} \frac{1}{k^2},$$

et ceci n'est autre que $\frac{1}{4}\sum_{n=1}^{+\infty}\frac{1}{n^2}$, la variable de sommation étant muette. De l'équation précédente, on déduit que $\sum_{n=1}^{+\infty}\frac{1}{n^2} = \frac{1}{4}\sum_{n=1}^{+\infty}\frac{1}{n^2} + \sum_{n=1,\, n \text{ impair}}^{+\infty}\frac{1}{n^2}$, soit $\sum_{n=1,\, n \text{ impair}}^{+\infty}\frac{1}{n^2} = \frac{3}{4}\sum_{n=1}^{+\infty}\frac{1}{n^2}$. En injectant dans l'égalité $\sum_{n=1,\, n \text{ impair}}^{\infty}\frac{4}{\pi n^2} = \frac{\pi}{2}$ de l'exemple 3.18, il vient $\frac{3}{\pi}\sum_{n=1}^{+\infty}\frac{1}{n^2} = \frac{\pi}{2}$, soit $\sum_{n=1}^{+\infty}\frac{1}{n^2} = \frac{\pi^2}{6}$.

3.6.8 Exercice 3.22. — Appelons f la fonction en jeu. Comme on connaît ses coefficients de Fourier, on peut calculer son « énergie » $\frac{1}{2\pi}\int_0^{2\pi}|f(t)|^2\,dt$, qui est égale à $\frac{1}{2\pi}\int_{-\pi}^{\pi}|f(t)|^2\,dt$ par 2π-périodicité de $|f|^2 : t \mapsto |f(t)|^2$. Cette quantité est est encore égale à $\frac{1}{\pi}\int_0^\pi |f(t)|^2\,dt$ car $|f|^2$ est paire. Or $\int_0^\pi |f(t)|^2\,dt = \int_0^\pi t^2\,dt = \frac{\pi^3}{3}$, d'où une énergie égale à $\frac{2\pi^2}{3}$.

L'égalité de Parseval (théorème 3.21) nous dit donc :

$$\frac{\pi^2}{3} = |c_0(f)|^2 + \frac{1}{2}\sum_{n=1}^{+\infty}[|a_n(f)|^2 + |b_n(f)|^2] = \frac{\pi^2}{4} + \frac{1}{2}\sum_{\substack{n=1 \\ n \text{ impair}}}^{+\infty} \frac{16}{\pi^2 n^4}$$

(avec convergence des séries en jeu), d'où l'on tire l'identité $\sum_{n=1}^{+\infty}\frac{1}{n^4} = \frac{\pi^4}{90}$, de la même manière que dans l'exercice 3.19.

3.6.9 Exercice 3.23. — Fixons $m \in \mathbb{Z}$. Pour ramener le calcul de $c_m(f)$ à celui de $c_m(f)$, on utilise une *intégration par parties*, légitime car f et $t \mapsto e^{-imt}$ sont de classe C^1 :

$$c_m(f) = \frac{1}{2\pi}\int_0^{2\pi} f'(t)e^{-imt}\,dt = \left[f(t)e^{-imt}\right]_0^{2\pi} - \frac{1}{2\pi}\int_0^{2\pi} f(t)(-im)e^{-imt}\,dt.$$

Or $[f(t)e^{-imt}]_0^{2\pi} = 0$ car $t \mapsto f(t)e^{-imt}$ est 2π-périodique, d'où : $c_m(f') = imc_m(f)$; en particulier, $c_0(f')$ est nul, et pour tout $m \neq 0$, $|c_m(f')|^2 = m^2|c_m(f)|^2 \geqslant |c_m(f)|^2$! Par conséquent, en utilisant Parseval pour justifier la convergence des séries en jeu,

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi}|f(t)|^2\,dt &= \sum_{m=-\infty}^{+\infty}|c_m(f)|^2 = |c_0(f)|^2 + \sum_{\substack{m=-\infty \\ m\neq 0}}^{+\infty}|c_m(f)|^2 \\ &\leqslant |c_0(f)|^2 + \sum_{\substack{m=-\infty \\ m\neq 0}}^{+\infty}|c_m(f')|^2 = |c_0(f)|^2 + \frac{1}{2\pi}\int_0^{2\pi}|f'(t)|^2\,dt.\end{aligned}$$

4 Produit scalaire

On passe maintenant à la partie Géométrie de ce cours, avec l'introduction du *produit scalaire*, pilier de la géométrie euclidienne – qui gouverne cet univers aux échelles intermédiaires. Une motivation auxiliaire à ce chapitre réside dans le calcul des coordonnées dans une base donnée, qui implique la résolution d'un système/inversion d'une matrice en général, ce qui est coûteux en temps de calcul. On verra que le calcul est direct si l'on a une base adaptée à un produit scalaire que l'on s'est fixé.

On donne, par souci de généralité, les définitions en dimension finie arbitraire $n \geqslant 1$; on se contentera souvent des dimensions 2 et 3 pour les exemples et exercices.

On note $(e_1, \ldots, e_n)$ la *base canonique* de $\mathbb{R}^n$; e_i est donc le vecteur $(0, \ldots, 0, \overset{i}{\widehat{1}}, 0, \ldots, 0)$ pour tout $i = 1, \ldots, n$ – ou bien le vecteur transposé lors de calculs matriciels... Notions et résultats (plus avancés) spécifiques aux *sous-espaces vectoriels* apparaissant au fil du texte sont indiqués par un liseré bleu en marge.

4.1 Produit scalaire de deux vecteurs

En dimension deux ou trois, le produit scalaire de deux vecteurs – qui comme son nom l'indique, produit un *scalaire*, *i.e.* un *nombre*, à partir de vecteurs – est simplement donné par la somme des produits coordonnée par coordonnée. On peut généraliser ceci immédiatement en tout dimension :

☛ **Définition 4.1 (Produit scalaire, norme)** *Soient $x = (x_1, \ldots, x_n)$ et $y = (y_1, \ldots, y_n)$ deux vecteurs de $\mathbb{R}^n$. Le* ***produit scalaire (standard) de** x **et** y* *est la quantité*

$$x_1 y_1 + \cdots + x_n y_n = \sum_{j=1}^{n} x_j y_j, \tag{19}$$

notée $\langle x, y \rangle$.

Lorsque $y = x$, on a $\langle x, x \rangle \geqslant 0$; on appelle alors ***norme de** x* *le réel $\sqrt{\langle x, x \rangle}$, que l'on note $\|x\|$; autrement dit, $\|x\| \geqslant 0$ et $\|x\|^2 = \langle x, x \rangle$.*

Remarque 4.2 *Pour tout $x \in \mathbb{R}^n$, on a $\|x\| \geqslant 0$. Or si $\|x\| = 0$, alors $\|x\|^2 = 0$, i.e. $x_1^2 + \cdots + x_n^2 = 0$, et ceci n'est possible que si $x_1 = \cdots = x_n = 0$. Ainsi, $\|x\| = 0$ ssi x est le vecteur nul, et dans le cas contraire, $\|x\| > 0$.*

Exercice 4.3 *Calculer les produits scalaires deux à deux des vecteurs $x = (0, 1)$, $y = (2, 1)$, et $z = (3, -1)$. Calculer également la norme de ces vecteurs.*

Même chose dimension 3, avec les vecteurs $u = (1, 1, 1)$, $v = (1, 2, 3)$, $w = (2, 0, -1)$.

Remarque 4.4 *Soit* $x = (x_1, \dots, x_n) \in \mathbb{R}^n$*. Soit* $i \in \{1, \dots, n\}$ *; comme seule la* i*-ème coordonnée de* e_i *est non nulle, et vaut 1, on a :*

$$\langle x, e_i \rangle = x_1 \cdot 0 + \cdots + x_{i-1} \cdot 0 + x_i \cdot 1 + x_{i+1} \cdot 0 + \cdots + x_n \cdot 0 = x_i,$$

et ceci pour tout $i = 1, \dots, n$*. Comme de plus* $x = x_1 e_1 + \cdots + x_n e_n$*, on a la décomposition :*

$$x = \langle x, e_1 \rangle e_1 + \cdots + \langle x, e_n \rangle e_n = \sum_{i=1}^{n} \langle x, e_i \rangle e_i, \tag{20}$$

et ceci pour tout $x \in \mathbb{R}^n$*.*

Cette identité semble assez tautologique ; nous verrons néanmoins dans la section 4.3 que l'on peut la généraliser d'autres bases que $(e_1, \dots, e_n)$*.*

De même qu'en dimensions 2 et 3, on décrète que deux vecteurs sont perpendiculaires, ou « orthogonaux », si leur produit scalaire est nul :

Définition 4.5 (Orthogonalité) *Deux vecteurs* x *et* y *de* $\mathbb{R}^n$ *sont dits **orthogonaux** si* $\langle x, y \rangle = 0$*. On écrit alors :* $x \perp y$*.*

Exemple 4.6 *Le vecteur nul* $(0, \dots, 0)$ *est orthogonal à tout autre vecteur ; immédiat, mais fondamental ! Autre exemple :* $(1, 0, 1)$ *est orthogonal à* $(1, 0, -1)$ *et à* $(0, 1, 0)$*.*

Exercice 4.7 *En dimension 3 : démontrer que* $(1, 0, 0)$*,* $(0, 1, 0)$ *et* $(0, 0, 1)$ *sont deux à deux orthogonaux.*

D'un point de vue syntaxique, l'expression « x et y sont orthogonaux » laisse penser que l'on pourrait aussi bien dire « y et x sont orthogonaux », autrement dit que x et y jouent un rôle symétrique dans cette affaire... Éclaircissons ce point – entre autres – dès à présent :

Proposition 4.8 *Le produit scalaire est :*

- ***bilinéaire** : pour tous* x*,* x'*,* y*,* $y' \in \mathbb{R}^n$*,* λ*,* $\mu \in \mathbb{R}$*,*

$$\langle x + \lambda x', y + \mu y' \rangle = \langle x, y \rangle + \lambda \langle x', y \rangle + \mu \langle x, y' \rangle + \lambda\mu \langle x', y' \rangle ;$$

- ***symétrique** : pour tous* x*,* $y \in \mathbb{R}^n$*,*

$$\langle x, y \rangle = \langle y, x \rangle.$$

Exercice 4.9 *Montrer, pour tout* x *de* $\mathbb{R}^n$ *et tout* λ *de* $\mathbb{R}$*, que* $\|\lambda x\| = |\lambda| \cdot \|x\|$*. En déduire, pour* $x \neq 0$*, la valeur de la norme de* $\frac{1}{\|x\|} x$

Exercice 4.10 *Soient y et $z \in \mathbb{R}^n$ fixés tels que pour tout $x \in \mathbb{R}^n$, on ait $\langle x, y\rangle = \langle x, z\rangle$. Montrer que $y = z$.*

De ces propriétés élémentaires du produit scalaire découlent les fameux :

Proposition 4.11 (Théorème de Pythagore, et sa réciproque) *Pour tous x, $y \in \mathbb{R}^n$, on a l'équivalence : $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ ssi $x \perp y$.*

Bien noter le caractère logique de cet énoncé : l'égalité a lieu si on a orthogonalité des vecteurs, et seulement dans ce cas.

On peut encore exploiter bilinéarité et symétrie pour démontrer :

Proposition 4.12 (Inégalité de Schwarz) *Pour tous x, $y \in \mathbb{R}^n$, on a :*

$$|\langle x, y\rangle| \leqslant \|x\| \cdot \|y\|.$$

⚠ ATTENTION à l'orthographe : *Schwarz* s'écrit avec un w, et sans t devant le z...

4.2 Produit scalaire, famille de vecteurs – et sous-espaces vectoriels

4.2.1 Notion d'orthogonal. —

Définition 4.13 *Soit S un sous-ensemble (pas nécessairement linéaire) de $\mathbb{R}^n$. On appelle **orthogonal de** S, noté $S^\perp$, le sous-ensemble de $\mathbb{R}^n$ donné par :*

$$S^\perp = \{x \in \mathbb{R}^n \mid \forall y \in S, \langle x, y\rangle = 0\}.$$

*En d'autres termes, $S^\perp$ est le sous-ensemble de $\mathbb{R}^n$ constitué des vecteurs de $\mathbb{R}^n$ qui sont simultanément orthogonaux à **tous** les vecteurs de S.*

Exemple 4.14 *Fixons S au hasard. Alors sûrement, $0 \in S^\perp$! En effet, 0 est orthogonal à tout le monde dans $\mathbb{R}^n$, donc a fortiori dans S, et entre donc dans la constitution de $S^\perp$.*

Par ailleurs, dans $\mathbb{R}^n$, l'orthogonal de $\{0\}$ est $\mathbb{R}^n$ (tout le monde est orthogonal à 0), et l'orthogonal de $\mathbb{R}^n$ est $\{0\}$ (il n'y a que 0 qui soit orthogonal à tout le monde).

Exercice 4.15 *Montrer que le passage à l'orthogonal est « décroissant », c'est-à-dire : si $S_1 \subset S_2 \subset \mathbb{R}^n$, alors $S_2^\perp \subset S_1^\perp$.*

Exercice 4.16 *Soit S une partie quelconque de $\mathbb{R}^n$. Soient x et x' deux vecteurs de $S^\perp$, et soit $\lambda \in \mathbb{R}$. Que dire de $x + x'$? De λx ?*

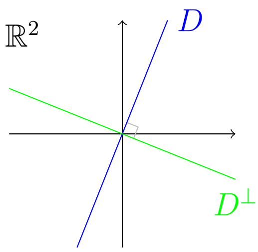
Exercice 4.17 *Quel est l'orthogonal de l'axe des abscisses dans $\mathbb{R}^2$? Dans $\mathbb{R}^3$?*

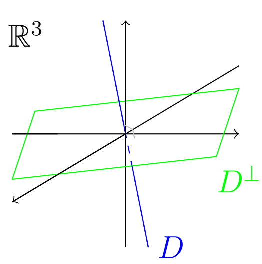
Proposition 4.18 *Dans $\mathbb{R}^2$, l'orthogonal d'une droite est une droite.*

Dans $\mathbb{R}^3$, l'orthogonal d'une droite est un plan ; l'orthogonal d'un plan est une droite.

Plus généralement, l'orthogonal d'un sous-espace vectoriel F de $\mathbb{R}^n$ est un sous-espace vectoriel de $\mathbb{R}^n$, de dimension complémentaire, c'est-à-dire : $\dim_{\mathbb{R}}(F) + \dim_{\mathbb{R}}(F^{\perp}) = n$.

Illustrons ce résultat en petites dimensions :





Exercice 4.19 *Soit $S \subset \mathbb{R}^n$. Montrer que si $0 \in S$, alors $S \cap S^{\perp} = \{0\}$, et que si $0 \notin S$, alors $S \cap S^{\perp} = \varnothing$.*

De l'exercice 4.19 et de la proposition 4.18 découle qu'*un sous-espace et son orthogonal sont supplémentaires* : étant donné F un sous-espace de $\mathbb{R}^n$, tout $x \in \mathbb{R}^n$ s'écrit de manière unique $x_F + x_{F^{\perp}}$, avec $x_F \in F$ et $x_{F^{\perp}} \in F^{\perp}$. D'où :

Définition 4.20 (Projection orthogonale) *Soit F un sous-espace vectoriel de $\mathbb{R}^n$; pour tout $x \in \mathbb{R}^n$, on appelle* ***projection orthogonale de*** *x* ***sur*** *F le vecteur x_F, où $x_F + x_{F^{\perp}}$ est la décomposition de x selon F et $F^{\perp}$.*

On note : $\wp_F(x) = x_F$.

*4.2.2 Ortho*NORM*alité. —*

☛ **Définition 4.21 (Orthonormalité)** *On dit qu'une famille de vecteurs $(v_i)_{i\in I}$ est* ***orthonormée****, ou* ***orthonormale****, si :*

- *ses éléments sont deux à deux orthogonaux : pour tous $i, j \in I$, $i \neq j$, $\langle v_i, v_j \rangle = 0$;*
- *ses éléments sont de norme 1 : pour tout $i \in I$, $\|v_i\| = \sqrt{\langle v_i, v_i \rangle} = 1$.*

Exercice 4.22 *Montrer que toute sous-famille d'une famille orthonormale est encore orthonormale : si $(v_i)_{i\in I}$ orthonormale et si $J \subset I$, alors $(v_i)_{i\in J}$ est orthonormale.*

On admet que dans $\mathbb{R}^n$, une famille orthonormale se compose d'au plus n éléments (*cf.* exercice suivant).

Exercice 4.23 **(∗)** *Démontrer que dans $\mathbb{R}^n$, une famille orthonormale se compose d'au plus n éléments ; on pourra, partant d'une famille orthonormale $(v_i)_{1\leqslant i\leqslant n+1}$, aboutir à une contradiction en utilisant la matrice $V = (\langle v_i, e_j\rangle)_{1\leqslant i,j\leqslant n}$, qui est inversible, d'inverse sa transposée (le démontrer !).*

Définition 4.24 **(Base orthonormée)** *Une famille de vecteurs de $\mathbb{R}^n$ est appelée **base orthonormée** (ou **orthonormale**) si elle est composée de n vecteurs.*

Remarque 4.25 *La terminologie est cohérente : une base orthonormée est bien une base, dans laquelle on peut donc parler de coordonnées pour n'importe quel vecteur. On voit dans la section suivante comment calculer ces coordonnées.*

Cette définition se généralise aux sous-espaces vectoriels de $\mathbb{R}^n$, en disant qu'une base orthonormée d'un sous-espace de dimension p est une famille orthonormée à p éléments de ce sous-espace.

4.3 Utilisation et construction de bases orthonormées

4.3.1 Calculs de coordonnées. — Étant donnée une base $\mathcal{B}$ de $\mathbb{R}^n$, *a priori* différente de la base canonique, et un vecteur x, il est en général fastidieux de *calculer les coordonnées de x dans $\mathcal{B}$*, car cela revient à résoudre un système, ou encore à inverser une matrice, ce qui demande un nombre d'opérations de l'ordre de n^3 si l'on utilise le pivot de Gauss.

Dans le cas néanmoins où l'on dispose d'une base orthonormée, on peut procéder de manière plus directe :

Proposition 4.26 *Soit $(v_1, \ldots, v_n)$ une base orthonormée de $\mathbb{R}^n$. Alors pour tout vecteur $x \in \mathbb{R}^n$ on a :*

$$x = \langle x, v_1\rangle v_1 + \cdots + \langle x, v_n\rangle v_n = \sum_{k=1}^{n} \langle x, v_k\rangle v_k.$$

Plus généralement, si F est un sous-espace vectoriel de $\mathbb{R}^n$ et $(v_1, \ldots, v_p)$ en est une base orthonormée, alors pour tout x de F, on a : $x = \langle x, v_1\rangle v_1 + \cdots + \langle x, v_p\rangle v_p = \sum_{k=1}^{p} \langle x, v_k\rangle v_k$.

Si l'on dispose d'une base orthonormée pour $\mathbb{R}^n$, le nombre d'opérations dans un calcul de coordonnées est ainsi réduit à un ordre de n^2.

Remarquer, dans la proposition 4.26, que la formule donnée dans le cas « sous-espace vectoriel » n'est valide que pour les vecteurs x du sous-espace F considéré.

Si l'on garde le sous-espace, mais que l'on s'autorise des vecteurs plus généraux, on obtient alors une expression simple pour le calcul des projections orthogonales :

Corollaire 4.27 *Soit F un sous-espace vectoriel de $\mathbb{R}^n$, admettant une base orthonormée $(v_1, \ldots, v_p)$. Alors pour tout x de $\mathbb{R}^n$, on a :*

$$\wp_F(x) = \sum_{k=1}^{p} \langle x, v_k \rangle v_k.$$

4.3.2 Construction de bases orthonormées. — Une base prise au hasard – ou apparaissant au cours d'un problème, telles les « bases de solutions » – ont malheureusement fort peu de chances d'être orthonormées.

Le résultat suivant nous dit que l'on peut toutefois « corriger » toute base en une base orthonormée :

Théorème 4.28 (Théorème d'orthonormalisation de Schmidt) *Soit $(f_1, \ldots, f_n)$ une base de $\mathbb{R}^n$. Alors il existe une base orthonormée $(v_1, \ldots, v_n)$ de $\mathbb{R}^n$, qui vérifie de plus : pour tout $j = 1, \ldots, n$,*

$$\mathrm{Vect}(v_1, \ldots, v_j) = \mathrm{Vect}(f_1, \ldots, f_j). \tag{21}$$

Le résultat analogue en partant de la base d'un sous-espace vectoriel de $\mathbb{R}^n$ reste vrai.

Remarque 4.29 *La propriété définie par l'équation* (21) *se traduit simplement en termes matriciels, en disant que la matrice de passage de $(v_1, \ldots, v_n)$ à $(f_1, \cdots, f_n)$ (et réciproquement, par symétrie) est triangulaire supérieure, i.e. de la forme*

$$\begin{pmatrix} \times & \ldots & \ldots & \times \\ 0 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & \times \end{pmatrix}.$$

On retiendra surtout de cet énoncé que l'on peut modifier toute base « intelligemment » pour en tirer une base orthonormée ; on renvoie aux TD pour la méthode à appliquer (et qui consiste à appliquer « concrètement » la preuve ci-dessous).

4.4 Appendice : démonstration du théorème 4.28

On démontre le résultat général, et on travaille par récurrence sur la dimension p d'un sous-espace F de $\mathbb{R}^n$ fixé. Si $p = 1$, alors on prend $v_1 = \frac{1}{\|f_1\|} f_1$.

Si $p > 1$, on pose $F' = \mathrm{Vect}(f_1, \ldots, f_{p-1})$; alors $\dim(F') = p - 1$. Par hypothèse de récurrence, il existe une base orthonormée $(v_1, \ldots, v_{p-1})$ de F' telle que $\mathrm{Vect}(v_1, \ldots, v_j) = \mathrm{Vect}(f_1, \ldots, f_j)$ pour tout $j = 1, \ldots, p-1$. Il s'agit à présent de construire le vecteur v_p de sorte que $\mathrm{Vect}(v_1, \ldots, v_p) = \mathrm{Vect}(f_1, \ldots, f_p) = F$, et que $(v_1, \ldots, v_p)$ soit orthonormée. Soit

$$w = f_p - \wp_{F'}(f_p)\,;$$

alors w est orthogonal à $F' = \mathrm{Vect}(v_1, \ldots, v_j)$, donc aux v_j, $j = 1, \ldots, p-1$, et donc, pour peu que $w \neq 0$, $\big(v_1, \ldots, v_{p-1}, \frac{1}{\|w\|} w\big)$ est orthonormée (donc libre).

Pour voir que $w \neq 0$, il suffit de voir que $\mathrm{Vect}(v_1, \ldots, v_{p-1}, w) = F$, car $\mathrm{Vect}(v_1, \ldots, v_{p-1}, 0) = F' \subsetneq F$, ce qui permettra de conclure. Or

$$\begin{aligned}\mathrm{Vect}(v_1, \ldots, v_{p-1}, w) &= \mathrm{Vect}\,\big(v_1, \ldots, v_{p-1}, f_p - \wp_{F'}(f_p)\big)\\ &= \mathrm{Vect}(v_1, \ldots, v_{p-1}, f_p)\end{aligned}$$

car $\wp_{F'}(f_p) \in F' = \mathrm{Vect}(v_1, \ldots, v_{p-1})$. Comme $\mathrm{Vect}(v_1, \ldots, v_{p-1}) = \mathrm{Vect}(f_1, \ldots, f_{p-1})$, on a finalement bien $\mathrm{Vect}(v_1, \ldots, v_{p-1}, w) = \mathrm{Vect}(f_1, \ldots, f_{p-1}, f_p) = F$.

Pour terminer la preuve, on pose $v_p = \frac{1}{\|w\|} w$, de sorte que $\mathrm{Vect}(v_1, \ldots, v_{p-1}, v_p) = \mathrm{Vect}(v_1, \ldots, v_{p-1}, w) = F$. □

4.5 Corrigés des exercices du chapitre 4

4.5.1 Exercice 4.3. — On a :

$$\begin{aligned}\langle x, y\rangle &= 0 \cdot 2 + 1 \cdot 1 = 1,\\ \langle x, z\rangle &= 0 \cdot 3 + 1 \cdot (-1) = -1,\\ \langle y, z\rangle &= 2 \cdot 3 + 1 \cdot (-1) = 5,\end{aligned}$$

tandis que

$$\begin{aligned}\langle x, x\rangle &= 0^2 + 1^2 = 1,\\ \langle y, y\rangle &= 2^2 + 1^2 = 5,\\ \langle z, z\rangle &= 3^2 + (-1)^2 = 10,\end{aligned}$$

d'où $\|x\| = 1$, $\|y\| = \sqrt{5}$ et $\|z\| = \sqrt{10}$ (ne pas oublier la racine carrée !).

De même, $\langle u, v\rangle = 1 \cdot 1 + 1 \cdot 2 + 1 \cdot 3 = 6$, $\langle u, w\rangle = 1 \cdot 2 + 1 \cdot 0 + 1 \cdot (-1) = 1$, $\langle v, w\rangle = 1{\cdot}2 + 2{\cdot}0 + 3{\cdot}(-1) = -1$ et $\langle u, u\rangle = 1^2+1^2+1^2 = 3$, $\langle v, v\rangle = 1^2+2^2+3^2 = 14$ $\langle w, w\rangle = 2^2 + 0^2 + (-1)^2 = 5$, d'où $\|u\| = 1$, $\|v\| = \sqrt{14}$ et $\|w\| = \sqrt{5}$.

4.5.2 Exercice 4.7. — On a $\langle(1,0,0),(0,1,0)\rangle = 1\cdot 0+0\cdot 1+0\cdot 0 = 0$, donc $(1,0,0)\perp(0,1,0)$. De même $(1,0,0)\perp(0,0,1)$ et $(0,1,0)\perp(0,1,0)$ par permutation circulaire des coordonnées.

4.5.3 Exercice 4.9. — Soient $x\in\mathbb{R}^n$ et $\lambda\in\mathbb{R}$. Alors $\|\lambda x\|^2 = \langle\lambda x,\lambda x\rangle = \lambda^2\langle x,x\rangle = \lambda^2\|x\|^2$. En prenant les racines carrées – et en prenant garde au fait que $\sqrt{\lambda^2} = |\lambda|$! – on a bien $\|\lambda x\| = |\lambda|\cdot\|x\|$.

Par conséquent, si $x\neq 0$ – de sorte que $\|x\|>0$ –, on aura $\left\|\frac{1}{\|x\|}x\right\| = \frac{1}{\|x\|}\|x\| = 1$.

4.5.4 Exercice 4.10. — Soit $x\in\mathbb{R}^n$. Alors $\langle x,y-z\rangle = \langle x,y\rangle-\langle x,z\rangle = 0$, car $\langle x,y\rangle = \langle x,z\rangle$. Ceci vaut pour tout $x\in\mathbb{R}^n$; en particulier, en faisant $x=y-z$, on obtient : $\|y-z\|^2 = \langle y-z,y-z\rangle = 0$, et donc $y-z=0$ (par la remarque 4.2), *i.e.* $y=z$.

4.5.5 Exercice 4.16. — Montrons que $x+x'$ est encore dans $S^\perp$. Soit $y\in S$. Alors $\langle x+x',y\rangle = \langle x,y\rangle+\langle x',y\rangle = 0+0 = 0$, car $\langle x,y\rangle = \langle x',y\rangle = 0$, puisque x et $x'\in S^\perp$. L'égalité $\langle x+x',y\rangle = 0$ valant ainsi pour tout y de S, on a : $x+x'\in S^\perp$.

Montrons de même que $\lambda x\in S^\perp$. Pour tout y, $\langle\lambda x,y\rangle = \lambda\langle x,y\rangle = \lambda\cdot 0 = 0$, car $x\in S^\perp$. Ainsi, x est orthogonal à tout $y\in S$, *i.e.* : $x\in S^\perp$.

4.5.6 Exercice 4.17. — L'habitude nous fait deviner que dans le plan ($\mathbb{R}^2$), l'orthogonal de l'axe des abscisses est l'axe des ordonnées. Vérifions-le : soit x dans l'axe des ordonnées, s'écrivant donc $t\binom{0}{1}$, $t\in\mathbb{R}$; alors pour tout y de l'axe des abscisses, donc de la forme $s\binom{1}{0}$, $s\in\mathbb{R}$, on a :

$$\langle x,y\rangle = \left\langle t\tbinom{0}{1}, s\tbinom{1}{0}\right\rangle = ts\left\langle \tbinom{0}{1},\tbinom{1}{0}\right\rangle = 0,$$

car $\left\langle\binom{0}{1},\binom{1}{0}\right\rangle = 0$. Ainsi, x est dans l'orthogonal de l'axe des abscisses, et ce dès que x est un vecteur de l'axe des ordonnées : l'axe des ordonnées est donc inclus dans l'orthogonal de l'axe des abscisses.

Pour conclure, il s'agit de vérifier l'inclusion réciproque. Soit donc $x=\binom{s}{t}$ dans l'orthogonal de l'axe des abscisses ; en particulier, x est orthogonal au vecteur $\binom{1}{0}$, *i.e.* $\left\langle x,\binom{1}{0}\right\rangle = 0$. Or $\left\langle x,\binom{1}{0}\right\rangle = \left\langle\binom{s}{t},\binom{1}{0}\right\rangle = s$, d'où $s=0$, puis $x=\binom{0}{t}$, soit : x est un vecteur de l'axe des ordonnées. L'inclusion réciproque voulue est démontrée ; l'orthogonal de l'axe des abscisses dans $\mathbb{R}^2$ est l'axe des ordonnées.

On démontre de la même manière que dans $\mathbb{R}^3$, l'orthogonal de l'axe des abscisses est le plan vertical Oyz.

4.5.7 Exercice 4.19. — Soit $x \in S \cap S^\perp$. Alors x, en tant qu'élément de $S^\perp$, est orthogonal à tous les éléments de S, donc à lui-même, puisqu'on a aussi $x \in S$! Autrement dit, $0 = \langle x, x\rangle = \|x\|^2$, *i.e.* $\|x\| = 0$. D'après la remarque 4.2, on a donc $x = 0$.

Or si $0 \notin S$, un tel x, dont la seule particularité est d'être élément de $S \cap S^\perp$, ne saurait exister, et donc $S \cap S^\perp = \varnothing$. Si en revanche $0 \in S$, alors comme $0 \in S^\perp$ (exemple 4.14), on a $0 \in S \cap S^\perp$, et par suite $S \cap S^\perp = \{0\}$.

4.5.8 Exercice 4.22. — Il s'agit de remarquer que l'on ne rend pas les conditions d'orthonormalité plus exigeantes en supprimant des membres de la famille, ce qui semble cohérent avec le caractère « direct » de ces conditions.

Concrètement, soient i, $j \in J$, $i \neq j$; alors en particulier i, $j \in I$, et donc $\langle v_i, v_j\rangle = 0$. De même, $\|v_i\| = 1$ dès que $i \in J$, puisque alors $i \in I$. Ceci valant en toute généralité, $(v_i)_{i\in J}$ est une famille orthonormée.

4.5.9 Exercice 4.23. — Suivant l'énoncé, on se donne une famille orthonormale $(v_i)_{1\leqslant i\leqslant n+1}$ à $n+1$ éléments de $\mathbb{R}^n$, et on cherche une contradiction. On décompose v_{n+1} suivant la base canonique $(e_j)_{1\leqslant j\leqslant n}$ de $\mathbb{R}^n$:

$$v_{n+1} = \lambda_1 e_1 + \cdots + \lambda_n e_n.$$

Le produit scalaire avec v_1 – rappelons que $v_1 \perp v_{n+1}$ – nous donne :

$$0 = \langle v_1, v_{n+1}\rangle = \lambda_1\langle v_1, e_1\rangle + \cdots + \lambda_n\langle v_1, e_n\rangle$$

de même, pour tout $i = 2, \ldots, n$, on obtient en prenant le produit scalaire avec v_i :

$$0 = \langle v_i, v_{n+1}\rangle = \lambda_1\langle v_i, e_1\rangle + \cdots + \lambda_n\langle v_i, e_n\rangle.$$

On peut résumer ces n équations en : $V \cdot v_{n+1} = 0$, où $V = (\langle v_i, e_j\rangle)_{1\leqslant i,j\leqslant n}$.

Si l'on démontre que V est inversible, on aura terminé, car $V \cdot v_{n+1} = 0$ nous dira alors que $v_{n+1} = 0$, ce qui contredira $\|v_{n+1}\| = 1$. Suivant l'indication de l'énoncé, montrons que $V \cdot {}^tV = I_n$, avec ${}^tV = (\langle v_j, e_i\rangle)_{1\leqslant i,j\leqslant n} = (\langle e_i, v_j\rangle)_{1\leqslant i,j\leqslant n}$. Soient k, $l \in \{1, \ldots, n\}$; les règles du calcul matriciel nous disent que le coefficient d'indice (k, l) de $V \cdot {}^tV$ est égal à :

$$\langle v_k, e_1\rangle\langle e_1, v_l\rangle + \cdots + \langle v_k, e_n\rangle\langle e_n, v_l\rangle.$$

En posant $v_k = x_1 e_1 + \cdots + x_n e_n$ et $v_l = y_1 e_1 + \cdots + y_n e_n$, de sorte que $\langle v_k, e_i \rangle = x_i$ et $\langle v_l, e_i \rangle = y_i$ pour tout i, ce coefficient se réécrit :

$$x_1 y_1 + \cdots + x_n y_n, \quad \text{soit} \quad \langle v_k, v_l \rangle.$$

Or $\langle v_k, v_l \rangle = 1$ si $k = l$, 0 sinon ; par conséquent, $V \cdot {}^tV = I_n$ est la matrice avec des 1 sur la diagonale principale et des 0 ailleurs : on a bien $V \cdot {}^tV = I_n$!

En conclusion, il ne saurait exister dans $\mathbb{R}^n$ de famille orthonormale à $n + 1$ éléments, et *a fortiori* à plus de $n + 1$ éléments. Autrement dit, dans $\mathbb{R}^n$, une famille orthonormale admet au plus n éléments.

5 Matrices symétriques

On rappelle la notation $\mathcal{M}_n(\mathbb{R})$ pour *l'ensemble des matrices carrées de taille n à coefficients réels*. L'objet de ce chapitre est l'étude d'une première famille de matrices ayant un comportement « sympathique » vis-à-vis du produit scalaire.

5.1 Symétrie et stabilité

Étant donnée une matrice carrée A de taille n, on peut « tordre » le produit scalaire à l'aide de A, en la faisant agir sur l'un de ses arguments ; cela revient à considérer

$$\langle Ax, y\rangle \quad \text{ou} \quad \langle x, Ay\rangle,$$

au lieu de $\langle x, y\rangle$ (x, $y \in \mathbb{R}^n$).

Exercice 5.1 *On prend $A = \left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)$. Soient $x = \left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right)$ et $y = \left(\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix}\right) \in \mathbb{R}^2$. Montrer que*

$$\langle Ax, y\rangle = x_1y_1 + 2x_2y_1 + x_2y_2 \quad \text{et} \quad \langle x, Ay\rangle = x_1y_1 + 2x_1y_2 + x_2y_2.$$

On rappelle que ${}^tA = \left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right)$. Montrer de même que

$$\langle {}^tAx, y\rangle = x_1y_1 + 2x_1y_2 + x_2y_2 \quad \text{et} \quad \langle x, {}^tAy\rangle = x_1y_1 + 2x_2y_1 + x_2y_2.$$

Que se passe-t-il avec $B = \left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)$?

L'observation faite ici est tout à fait générale :

Proposition 5.2 *Soit $A \in \mathcal{M}_n(\mathbb{R})$. Alors pour tous x, $y \in \mathbb{R}^n$, on a : $\langle Ax, y\rangle = \langle x, {}^tAy\rangle$ (et $\langle x, Ay\rangle = \langle {}^tAx, y\rangle$, par symétrie du produit scalaire).*

La preuve découle d'une écriture explicite des quantités en jeu en termes de coefficients, élémentaire mais légèrement fastidieux. Un peu dans le même style :

Exercice 5.3 *Calculer les produits $\left(\begin{smallmatrix} 1 & 3 \\ 5 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 2 & -1 \\ 4 & 1 \end{smallmatrix}\right)$ et $\left(\begin{smallmatrix} 2 & 4 \\ -1 & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 & 5 \\ 3 & 0 \end{smallmatrix}\right)$. Une observation ?*

Plus généralement, *la transposée d'un produit de matrices est égale au produit des transposées des facteurs <u>dans l'ordre inverse</u>* [12] : ${}^t(M_1 \cdots M_k) = ({}^tM_k) \cdots ({}^tM_1)$.

La proposition 5.2 motive la définition suivante :

☛ **Définition 5.4 (Matrice symétrique)** *Une matrice $A = (a_{ij})_{1 \leqslant i,j \leqslant n} \in \mathcal{M}_n(\mathbb{R})$ est dite **symétrique** si*

$$a_{ji} = a_{ij}$$

pour tous i, $j \in \{1, \ldots, n\}$, c'est-à-dire si : ${}^tA = A$.

⚠ 12. Attention : rappelons que la multiplication des matrices n'est *pas commutative* dès que $n \geqslant 2$.

Exemple 5.5 *Toute matrice diagonale, c'est-à-dire de la forme*

$$\begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix},$$

avec $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$*, est symétrique.*

Exercice 5.6 *Écrire une matrice (carrée) symétrique de taille 3 – si possible non triviale.*

Les matrices symétriques se comportent agréablement quant au produit scalaire. En effet, d'après la proposition 5.2, on a la caractérisation suivante :

Proposition 5.7 *Soit* $A \in \mathcal{M}_n(\mathbb{R})$*. On a alors l'équivalence suivante :* $\langle Ax, y\rangle = \langle x, Ay\rangle$ *pour tous* x*,* $y \in \mathbb{R}^n$ *si et seulement si* A *est symétrique.*

On étudie plus en profondeur les matrices symétriques dans la section suivante. Introduisons la notion, plutôt géométrique, de *stabilité*, qui intervient de manière centrale dans l'analyse à venir :

Définition 5.8 *Soient* $A \in \mathcal{M}_n(\mathbb{R})$*, et* S *un sous-ensemble de* $\mathbb{R}^n$*. On pose*

$$AS = \{Ax \mid x \in S\},$$

☛ *et on dit que* S *est* ***stable par*** A *si* $AS \subset S$*.*

Exemple 5.9 *Bien sûr,* $\{0\}$ *et* $\mathbb{R}^n$ *sont stables pour toute matrice de* $\mathcal{M}_n(\mathbb{R})$*.*

Autre exemple : dans le plan $\mathbb{R}^2$*, la première bissectrice, de vecteur directeur* $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$*, est stable par la matrice* $\left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)$*.*

Exercice 5.10 *Démontrer que la première bissectrice est stable par toute matrice* 2×2 *dont la somme des coefficients par ligne est constante (i.e. toute matrice* $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ *telle que* $a + b = c + d$*).*

Les notions de stabilité, d'orthogonal, et de matrices transposées s'articulent de la manière suivante :

Proposition 5.11 *Soit* $A \in \mathcal{M}_n(\mathbb{R})$*, et soit* S *un sous-ensemble de* $\mathbb{R}^n$*, stable par* A*. Alors* $S^\perp$ *est stable par* ${}^tA$*.*

En particulier, si A *est symétrique, alors* $S^\perp$ *est également stable par* A*.*

Exercice 5.12 *Déduire de l'exemple 5.9 que la droite de vecteur directeur* $\left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right)$ *est stable par la matrice* $\left(\begin{smallmatrix}1&2\\2&1\end{smallmatrix}\right)$.

On peut exploiter l'observation faite dans cet exercice, à savoir l'existence de droites stables pour une matrice donnée, comme suit. Étant donné un vecteur $x = \left(\begin{smallmatrix}s\\t\end{smallmatrix}\right)$ de $\mathbb{R}^2$, l'« action » de $\left(\begin{smallmatrix}1&2\\2&1\end{smallmatrix}\right)$ sur x est relativement compliquée, puisque le calcul des coordonnées de $\left(\begin{smallmatrix}1&2\\2&1\end{smallmatrix}\right)x$ « mélange » s et t.

On peut en revanche décomposer x selon les droites stables mises en évidence, c'est-à-dire écrire

$$x = \frac{s+t}{2}\begin{pmatrix}1\\1\end{pmatrix} + \frac{s-t}{2}\begin{pmatrix}1\\-1\end{pmatrix} = \frac{s+t}{2}u_1 + \frac{s-t}{2}u_2,$$

où l'on a posé $u_1 = \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$ et $u_2 = \left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right)$. L'intérêt de cette écriture réside en ce que l'action de $\left(\begin{smallmatrix}1&2\\2&1\end{smallmatrix}\right)$ sur u_1 et u_2 est simple, puisque

$$\begin{pmatrix}1&2\\2&1\end{pmatrix}u_1 = 3u_1 \quad \text{et} \quad \begin{pmatrix}1&2\\2&1\end{pmatrix}u_2 = -u_2\,;$$

par suite, on a donc :

$$\begin{pmatrix}1&2\\2&1\end{pmatrix}x = \frac{s+t}{2}\begin{pmatrix}1&2\\2&1\end{pmatrix}u_1 + \frac{s-t}{2}\begin{pmatrix}1&2\\2&1\end{pmatrix}u_2 = 3\frac{(s+t)}{2}u_1 - \frac{s-t}{2}u_2.$$

En d'autres termes, on a simplement multiplié la première coordonnée de x selon (u_1, u_2) par 3, et la seconde par -1, au lieu de les mélanger comme c'était le cas en calculant selon la base canonique.

On voit dans la section suivante que cette simplification, répondant au doux nom de « *diagonalisation* », concerne en réalité toutes les matrices symétriques.

Exercice 5.13 *Soit* $x \in \mathbb{R}^2$. *Exploiter ce qui précède pour calculer* $\left(\begin{smallmatrix}1&2\\2&1\end{smallmatrix}\right)^2 x$, $\left(\begin{smallmatrix}1&2\\2&1\end{smallmatrix}\right)^3 x$, *puis* $\left(\begin{smallmatrix}1&2\\2&1\end{smallmatrix}\right)^k x$ *pour un entier* $k \geqslant 1$ *quelconque.*

5.2 Diagonalisation

5.2.1 Énoncé du théorème. — Comme annoncé, le résultat principal de cette section est une généralisation du constat effectué plus haut sur la matrice $\left(\begin{smallmatrix}1&2\\2&1\end{smallmatrix}\right)$:

☛ **Théorème 5.14 (Réduction des matrices symétriques)** *Soit* $A \in \mathcal{M}_n(\mathbb{R})$. *Alors il existe des nombres réels* $\lambda_1, \dots, \lambda_n$ *(non nécessairement tous distincts) et une base orthonormée* $(v_1, \dots, v_n)$ *de* $\mathbb{R}^n$ *tels que :*

$$Av_1 = \lambda_1 v_1, \quad Av_2 = \lambda_2 v_2, \quad \dots \quad , \quad Av_n = \lambda_n v_n.$$

Définition 5.15 (Base de réduction, vecteurs propres, valeurs propres) *Une telle famille $(v_1, \dots, v_n)$ est appelée* ***base de réduction*** *pour A ; ses éléments sont appelés* ***vecteurs propres*** *de A, et pour $j \in \{1, \dots, n\}$, λ_j est appelée* ***valeur propre*** *associée à v_j.*

Exemple 5.16 *Reprenons la matrice $\left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)$ de la section précédente. On sait déjà que $\left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)u_1 = 3u_1$ et $\left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)u_2 = -u_2$, et que $u_1 \perp u_2$. En « normalisant » u_1 et u_2, c'est-à-dire en les divisant par leur norme, on obtient le couple $(v_1, v_2) = \left(\frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right), \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)\right)$; ce couple est orthonormé[13], et vérifie $\left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)v_1 = 3v_1$ et $\left(\begin{smallmatrix} 1 & 2 \\ 2 & 1 \end{smallmatrix}\right)v_2 = -v_2$. Comme on est en dimension 2, on a bien une base de réduction pour la matrice considérée ; les valeurs propres associées sont 3 et -1.*

⚠ Attention, l'hypothèse de symétrie du théorème 5.14 est fondamentale ! On ne peut par exemple pas trouver de base de réduction pour une matrice telle que $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$.

Par ailleurs, ce théorème est plutôt abstrait : il nous donne l'existence des λ_j et des v_j sans nous dire comment les calculer... Il existe pour cela des méthodes, notamment en dimensions 2 et 3, qui seront détaillées en TD.

Remarque 5.17 *Ne pas hésiter à exploiter le caractère orthonormé de la base de réduction donné par le théorème 5.14 ! En reprenant les notations de l'énoncé, et si $x \in \mathbb{R}^n$, la proposition 4.26 nous dit que*

$$x = \langle x, v_1 \rangle v_1 + \cdots + \langle x, v_n \rangle v_n.$$

Par conséquent, on a simplement :

$$Ax = \langle x, v_1 \rangle Av_1 + \cdots + \langle x, v_n \rangle Av_n = \lambda_1 \langle x, v_1 \rangle v_1 + \cdots + \lambda_n \langle x, v_n \rangle v_n.$$

5.2.2 Explication du terme « diagonalisation ». — Examinons au préalable l'effet d'une matrice diagonale sur une autre matrice, par multiplication à droite.

Exercice 5.18 *Soit $M = (m_{ij})_{1 \leqslant i,j \leqslant 3} \in \mathcal{M}_3(\mathbb{R})$, et soit $D = \left(\begin{smallmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{smallmatrix}\right)$. Calculer le produit MD.*

---

13. On a tout fait pour : $\langle v_1, v_2 \rangle = \frac{1}{\|u_1\| \cdot \|u_2\|} \langle u_1, u_2 \rangle = 0$, $\|v_1\| = \frac{1}{\|u_1\|}\|u_1\| = 1$, et de même, $\|u_2\| = 1$.

Au vu du résultat de cet exercice, on se convainc que plus généralement, si $M \in \mathcal{M}_3(\mathbb{R})$ et si D est diagonale de taille n, alors MD est la matrice M dont on a multiplié la première colonne par la première valeur diagonale de D, la deuxième colonne par la deuxième valeur diagonale de D, et ainsi de suite jusqu'à la n-ième colonne.

On reprend maintenant les notations du théorème 5.14. On voit v_1 comme le vecteur-colonne de ses coordonnées, et de même pour v_2, *etc.* Si l'on met côte à côte ces n vecteurs de hauteur n, c'est-à-dire si l'on constitue la matrice

$$V = \big(v_1 \,\big|\, v_2 \,\big|\, \cdots \,\big|\, v_n\big),$$

on obtient alors une matrice carrée de taille n, dite *matrice de passage* (de la base canonique) vers $(v_1, \ldots, v_n)$. Les règles de calcul matriciel nous disent à présent que

$$AV = \big(Av_1 \,\big|\, Av_2 \,\big|\, \cdots \,\big|\, Av_n\big)\,;$$

or $Av_j = \lambda_j v_j$ pour tout $j \in \{1, \ldots, n\}$, d'où :

$$AV = \big(\lambda_1 v_1 \,\big|\, \lambda_2 v_2 \,\big|\, \cdots \,\big|\, \lambda_n v_n\big).$$

Il s'avère que la matrice obtenue, c'est-à-dire la matrice V dont on a multiplié la première colonne par λ_1, la deuxième par λ_2, et ainsi de suite jusqu'à la n-ème, est aussi égale au produit $V \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ d'après notre préambule, soit :

$$AV = V \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}. \tag{22}$$

Rappelons pour conclure que V est inversible, et que son inverse n'est autre que sa transposée ${}^tV$ (*cf.* exercice 4.23) ! En multipliant l'égalité (22) à gauche par ${}^tV = V^{-1}$, on conclut [14] que :

$${}^tVAV = V^{-1}AV = V^{-1}V \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

En multipliant la matrice (symétrique) A de part et d'autre par une matrice de passage et son inverse, on a donc transformé A en une matrice *diagonale*, d'où le terme « diagonalisation ».

---

14. La multiplication des matrices est en revanche *associative*.

5.3 Appendice : démonstration du théorème 5.14

On admet que tout polynôme à coefficients complexes non constant admet au moins une racine dans $\mathbb{C}$.

On commence par chercher *une* valeur λ et *un* vecteur v non nul tels que $Av = \lambda v$; autrement dit, on veut un réel λ tel que l'équation

$$(A - \lambda I_n)v = 0$$

admette une solution v non nulle, c'est-à-dire un λ tel que la matrice $(A - \lambda I_n)$ soit de déterminant nul (*cf.* cours du premier semestre).

Notons $\chi_A(\lambda)$ ce déterminant, qui dépend de λ (χ, « chi », prononcé *ki*, comme *caractéristique*). Les règles de calcul de déterminant nous disent que

$$\chi_A(\lambda) = \det(A - \lambda I_n) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n_1,n} \\ a_{n1} & \dots & a_{n,n-1} & a_{nn} - \lambda \end{vmatrix}$$

est un *polynôme* en λ, de degré $n \geqslant 1$. Il existe donc $\lambda \in \mathbb{C}$ tel que $\chi_A(\lambda) = 0$. Pour montrer que $\lambda \in \mathbb{R}$ (ce que l'on voulait à l'origine), on revient à l'équation $(A - \lambda I_n)v = 0$, dont on sait désormais qu'elle admet une solution non nulle, *à coefficients complexes* ; notons z ce vecteur de $\mathbb{C}^n$, vérifiant donc : $Az = \lambda z$, $z \neq 0$. En multipliant cette relation à gauche par ${}^t\bar{z} = (\overline{z_1}, \dots, \overline{z_n})$, on obtient, tous calculs faits :

$$\sum_{i,j=1}^{n} a_{ij}\overline{z_i}z_j = {}^t\bar{z}Az = \lambda {}^t\bar{z} \cdot z = \lambda(|z_1|^2 + \dots + |z_n|^2).$$

Comme A est symétrique (et réelle !), la quantité $\sum_{i,j=1}^{n} a_{ij}\overline{z_i}z_j$ est égale à sa conjuguée, donc réelle. En outre, comme $z \neq 0$, $|z_1|^2 + \dots + |z_n|^2 > 0$. Par conséquent, $\lambda = \left(\sum_{i,j=1}^{n} a_{ij}\overline{z_i}z_j\right)/(|z_1|^2 + \dots + |z_n|^2)$ est bien réel ! La matrice *réelle* $A - \lambda I_n$ étant de déterminant nul, il existe donc $v \in \mathbb{R}^n$ non nul tel que $(A - \lambda I_n)v = 0$, *i.e.* $Av = \lambda v$.

Quitte à remplacer v par $\frac{1}{\|v\|}v$, on peut supposer que $\|v\| = 1$. On complète v en une base orthonormée $(v, w_2, \dots, w_n)$ en choisissant w_2 normalisé dans $\{v\}^\perp$, puis w_3 dans $\{v, w_2\}^\perp$, *etc.* [15]. Posons $W = (v|w_2|\cdots|w_n)$; alors (par calcul matriciel

15. Cet algorithme produisant une famille orthonormale, on est sûr qu'il s'arrête au bout d'au plus $(n-1)$ étapes. Pour vérifier qu'il ne s'arrête pas avant, supposons que $\{v, w_2, \dots, w_k\}^\perp = \{0\}$ $(k \leqslant n)$; autrement dit, la seule solution du système d'équations $v_1x_1 + \dots + v_nx_n = (w_2)_1x_1 + \dots + (w_2)_nx_n = \dots = (w_k)_1x_1 + \dots + (w_k)_nx_n = 0$ est 0. Or ceci implique que la matrice associée (en remplissant les dernières lignes de 0) soit de déterminant non nul, (*cf.* cours du premier semestre), et donc qu'il n'y a pas de ligne de 0, *i.e.* que $k = n$.

direct)

$$ {}^tWAW = \left(\begin{array}{c|c} {}^tvAv & ({}^tvAw_i)_{2\leqslant i\leqslant n} \\ \hline ({}^tw_jAv)_{2\leqslant j\leqslant n} & ({}^tw_jAw_i)_{2\leqslant i,j\leqslant n} \end{array}\right). $$

Or $Av = \lambda v$ donc ${}^tvAv = \lambda {}^tvv = \lambda\|v\|^2 = \lambda$, ${}^tvAw_i = \langle v, Aw_i\rangle = 0$ car $w_i \in \{v\}^\perp$ qui est stable par A, pour $i = 2,\ldots,n$; de plus ${}^tw_jAv = 0$ car la matrice-produit ${}^tWAW$ est symétrique : ${}^t({}^tWAW) = {}^tW{}^tAW = {}^tWAW$. En posant $B = ({}^tw_jAw_i)_{2\leqslant i,j\leqslant n} \in \mathcal{M}_{n-1}(\mathbb{R})$, on a donc :

$$ {}^tWAW = \left(\begin{array}{c|c} \lambda & 0 \\ \hline 0 & B \end{array}\right), $$

avec B symétrique, toujours par symétrie de ${}^tWAW$! En admettant le théorème en dimension $n-1$ (ce qui est rigoureux dans le cadre d'un raisonnement par récurrence), on sait qu'il existe une matrice $V_1 \in \mathcal{M}_{n-1}(\mathbb{R})$, inversible, d'inverse sa transposée ${}^tV_1$, telle que :

$$ {}^tV_1BV_1 = \begin{pmatrix} \lambda_2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}, $$

avec $\lambda_2,\ldots,\lambda_n \in \mathbb{R}$. En posant $V_2 = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & V_1 \end{array}\right) \in \mathcal{M}_n(\mathbb{R})$, on a ${}^tV_2V_2 = I_n$, et

$$ \begin{aligned} {}^t(WV_2)AWV_2 = {}^tV_2{}^tWAWV_2 &= \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & {}^tV_1 \end{array}\right)\left(\begin{array}{c|c} \lambda & 0 \\ \hline 0 & B \end{array}\right)\left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & V_1 \end{array}\right) = \left(\begin{array}{c|c} \lambda & 0 \\ \hline 0 & {}^tV_1BV_1 \end{array}\right) \\ &= \left(\begin{array}{c|ccc} \lambda & 0 & \cdots & 0 \\ \hline 0 & \lambda_2 & 0 & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{array}\right). \end{aligned} $$

Autrement dit, en posant $V = WV_2 = (v_1|\cdots|v_n)$, on a que $(v_1,\ldots,v_n)$ est une base orthonormée de $\mathbb{R}^n$ car ${}^tVV = {}^t(WV_2)WV_2 = {}^tV_2{}^tWWV_2 = I_n$, et ${}^tVAV$ est diagonale : $(v_1,\ldots,v_n)$ est une base de réduction pour A. □

Exercice 5.19 *Normaliser le vecteur* $\left(\begin{smallmatrix}1\\2\end{smallmatrix}\right) \in \mathbb{R}^2$*, puis le compléter en une base orthonormée de* $\mathbb{R}^2$*. Même question avec le vecteur* $\left(\begin{smallmatrix}1\\2\\1\end{smallmatrix}\right) \in \mathbb{R}^3$*.*

5.4 Corrigés des exercices du chapitre 5

5.4.1 Exercice 5.1. — On a $Ax = \left(\begin{smallmatrix}x_1+2x_2\\x_2\end{smallmatrix}\right)$, et donc $\langle Ax, y\rangle = (x_1+2x_2)y_1 + x_2y_2 = x_1y_1 + 2x_2y_1 + x_2y_2$.

5.4.2 Exercice 5.3. — En appliquant les règles de calcul sur les matrices, on obtient : $\left(\begin{smallmatrix}1&3\\5&0\end{smallmatrix}\right)\left(\begin{smallmatrix}2&-1\\4&1\end{smallmatrix}\right) = \left(\begin{smallmatrix}14&2\\10&-5\end{smallmatrix}\right)$, et $\left(\begin{smallmatrix}2&4\\-1&1\end{smallmatrix}\right)\left(\begin{smallmatrix}1&5\\3&0\end{smallmatrix}\right) = \left(\begin{smallmatrix}14&10\\2&-5\end{smallmatrix}\right)$. Sur cet exemple, on a donc que transposer les facteurs d'un produit et en inverser l'ordre produit une transposition sur le résultat.

5.4.3 Exercice 5.6. — On a bien sûr la matrice nulle, et la matrice identité

$$I_3 = \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix},$$

ainsi que n'importe quelle matrice diagonale. Mais on a également

$$\begin{pmatrix}1&2&3\\2&4&5\\3&5&6\end{pmatrix},$$

entre autres...

5.4.4 Exercice 5.10. — Soit $\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)$ comme dans l'énoncé, *i.e.* telle que $a + b = c + d := \alpha$. Soit x un vecteur de la première bissectrice ; x s'écrit donc $\left(\begin{smallmatrix}t\\t\end{smallmatrix}\right)$ pour un $t \in \mathbb{R}$. On a donc $\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)\left(\begin{smallmatrix}t\\t\end{smallmatrix}\right) = \left(\begin{smallmatrix}(a+b)t\\(c+d)t\end{smallmatrix}\right) = \alpha t\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$, qui est bien un vecteur de la première bissectrice. Ceci valant pour tout vecteur de cette droite, on a bien que la première bissectrice est stable par notre matrice $\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)$.

5.4.5 Exercice 5.12. — D'après la proposition 5.11, comme $\left(\begin{smallmatrix}1&2\\2&1\end{smallmatrix}\right)$ est symétrique, cette matrice stabilise l'orthogonal de ses parties stables. Or la première bissectrice est stable pour cette matrice (exemple 5.9) ; son orthogonal l'est donc également. Mais cet orthogonal n'est autre que la droite de vecteur directeur $\left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right)$ – on sait que c'est une droite par la proposition 4.18, qui contient $\left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right)$ car ce vecteur est orthogonal à tous les $\left(\begin{smallmatrix}t\\t\end{smallmatrix}\right)$ –, d'où le résultat.

5.4.6 Exercice 5.13. — Suivant le principe de l'exemple 5.9, posons $x = \left(\begin{smallmatrix}s\\t\end{smallmatrix}\right)$; alors $x = \frac{s+t}{2}u_1 + \frac{s-t}{2}u_2$, où $u_1 = \left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$ et $u_2 = \left(\begin{smallmatrix}1\\-1\end{smallmatrix}\right)$. Comme $\left(\begin{smallmatrix}1&2\\2&1\end{smallmatrix}\right)$ multiplie u_1 par 3 et u_2 par -1, on a

$$\begin{pmatrix}1&2\\2&1\end{pmatrix}x = 3\frac{(s+t)}{2}u_1 - \frac{s-t}{2}u_2 = \begin{pmatrix}s\\2t\end{pmatrix}$$

(comme déjà vu), puis

$$\begin{pmatrix}1&2\\2&1\end{pmatrix}^2 x = 3\frac{(s+t)}{2}\begin{pmatrix}1&2\\2&1\end{pmatrix}u_1 - \frac{s-t}{2}\begin{pmatrix}1&2\\2&1\end{pmatrix}u_2 = 9\frac{(s+t)}{2}u_1 + \frac{s-t}{2}u_2 = \begin{pmatrix}5s\\4t\end{pmatrix},$$

et de même,

$$\begin{pmatrix}1&2\\2&1\end{pmatrix}^3 x = 9\frac{(s+t)}{2}\begin{pmatrix}1&2\\2&1\end{pmatrix}u_1+\frac{s-t}{2}\begin{pmatrix}1&2\\2&1\end{pmatrix}u_2 = 27\frac{(s+t)}{2}u_1-\frac{s-t}{2}u_2 = \begin{pmatrix}13s\\14t\end{pmatrix}.$$

En poursuivant de la sorte, on a ainsi

$$\begin{pmatrix}1&2\\2&1\end{pmatrix}^k x = 3^k\frac{(s+t)}{2}u_1 + (-1)^k\frac{s-t}{2}u_2 = \begin{pmatrix}\frac{3^k+(-1)^k}{2}s\\ \frac{3^k-(-1)^k}{2}t\end{pmatrix}$$

pour tout $k \geqslant 1$.

5.4.7 Exercice 5.18. — En appliquant les règles du calcul matriciel, on trouve :

$$MD = \begin{pmatrix}m_{11}&m_{12}&m_{13}\\m_{21}&m_{22}&m_{23}\\m_{31}&m_{32}&m_{33}\end{pmatrix}\begin{pmatrix}1&0&0\\0&2&0\\0&0&3\end{pmatrix} = \begin{pmatrix}m_{11}&2m_{12}&3m_{13}\\m_{21}&2m_{22}&3m_{23}\\m_{31}&2m_{32}&3m_{33}\end{pmatrix}.$$

Autrement dit, MD est la matrice M dont on a multiplié la première colonne par la première valeur diagonale de D, la deuxième colonne par la deuxième valeur diagonale de D, et la troisième colonne par la troisième valeur diagonale de D.

5.4.8 Exercice 5.19. — Le vecteur $\binom{1}{2} \in \mathbb{R}^2$ est de norme $\sqrt{1^2+2^2} = \sqrt{5}$; on a donc comme premier vecteur normalisé $v = \frac{1}{\sqrt{5}}\binom{1}{2} \in \mathbb{R}^2$. Pour trouver un vecteur orthogonal à v qui soit non trivial, on applique le principe suivant (valable en dimension 2) : $\binom{-b}{a}$ est orthogonal à $\binom{a}{b}$, car leur produit scalaire vaut $-b \cdot a + a \cdot b = 0$. On trouve donc $w = \frac{1}{\sqrt{5}}\binom{-2}{1}$ (qui est aussi de norme 1) comme vecteur orthogonal à v. Par suite, (v,w) est une base orthonormée de $\mathbb{R}^2$.

Le principe évoqué ne fonctionne plus en dimension 3 ; pour trouver un vecteur orthogonal au vecteur normalisé $\frac{1}{\sqrt{6}}\begin{pmatrix}1\\2\\1\end{pmatrix}$, on procède donc « à la main »[16], en cherchant <u>une</u> solution non triviale à l'équation

$$x_1 + 2x_2 + x_3 = 0.$$

On peut par exemple prendre $x_1 = 2$, $x_2 = -1$ et $x_3 = 0$, ce qui mène au vecteur normalisé $\frac{1}{\sqrt{5}}\begin{pmatrix}2\\-1\\0\end{pmatrix}$. En dimension 3, on dispose en revanche du *produit vectoriel*, qui permet de conclure en produisant le dernier vecteur de la base orthonormée voulue :

$$\frac{1}{\sqrt{6}}\begin{pmatrix}1\\2\\1\end{pmatrix} \wedge \frac{1}{\sqrt{5}}\begin{pmatrix}2\\-1\\0\end{pmatrix} = \frac{1}{\sqrt{30}}\begin{pmatrix}1\\2\\-5\end{pmatrix}.$$

16. On peut aussi utiliser le produit vectoriel, en multipliant le vecteur de départ avec un vecteur qui ne lui est pas proportionnel, puis en normalisant le résultat.

6 Matrices orthogonales en dimensions 2 et 3

6.1 Matrices orthogonales

On a vu dans les deux chapitres précédents intervenir un certain type de matrices, notées V, et vérifiant l'identité

$$V^tV = I_n.$$

Pour mémoire, cette relation signifie que V est inversible, et que $V^{-1} = {}^tV$; elle entraîne donc également l'identité

$${}^tVV = I_n,$$

propriété [17] que nous avons utilisée discrètement dans la discussion du paragraphe explicatif 5.2.2. Ces matrices ont été constituées comme *matrices de bases orthonormées*, c'est-à-dire de la forme

$$V = \big(v_1 \big| \cdots \big| v_n\big) \tag{23}$$

avec $(v_1, \ldots, v_n)$ une base orthonormée de $\mathbb{R}^n$, et sont intervenues dans la *diagonalisation des matrices symétriques* : le théorème 5.14 nous dit que l'on peut, à toute matrice symétrique A, associer (au moins !) une base orthonormée $(v_1, \ldots, v_n)$ de $\mathbb{R}^n$, et donc une matrice V par (23), telles que ${}^tVAV$ soit une matrice diagonale, constituée des valeurs propres de A.

On va s'attacher dans ce dernier chapitre à analyser d'un peu plus près ces matrices V, et l'ensemble qu'elles constituent. On commence par fixer la terminologie :

☛ **Définition 6.1** *On appelle* ***matrice orthogonale*** *une matrice $V \in \mathcal{M}_n(\mathbb{R})$ vérifiant*

$$V^tV = I_n \qquad ou \qquad {}^tVV = I_n$$

(qui sont équivalentes). On appelle ***groupe orthogonal****, et l'on note $\mathcal{O}_n(\mathbb{R})$, l'ensemble de ces matrices, soit :*

$$\mathcal{O}_n(\mathbb{R}) = \big\{V \in \mathcal{M}_n(\mathbb{R}) \big| \, V^tV = I_n \text{ ou } {}^tVV = I_n\big\}. \tag{24}$$

Certes, les éléments de $\mathcal{O}_n(\mathbb{R})$ sont « orthogonaux », mais pourquoi parle-t-on de « groupe » ? Ceci tient à sa *structure*, comme le décrit la proposition suivante :

Proposition 6.2 *L'ensemble $\mathcal{O}_n(\mathbb{R})$ est un groupe pour la multiplication des matrices :*

17. Propriété assez spécifique aux matrices, et que l'on démontre soit par calcul brut, soit en utilisant un chouïa d'algèbre linéaire.

- *il contient un élément neutre, I_n ;*
- *si V_1, $V_2 \in \mathcal{O}_n(\mathbb{R})$, alors $V_1V_2 \in \mathcal{O}_n(\mathbb{R})$;*
- *si $V \in \mathcal{O}_n(\mathbb{R})$, alors V admet dans $\mathcal{O}_n(\mathbb{R})$ un inverse : ${}^tV$.*

Ces assertions se vérifient facilement : on a bien $I_n \in \mathcal{O}_n(\mathbb{R})$, car ${}^tI_n = I_n$, et donc ${}^tI_nI_n = I_n^2 = I_n$; si V_1, $V_2 \in \mathcal{O}_n(\mathbb{R})$, alors ${}^t(V_1V_2)V_1V_2 = {}^tV_2{}^tV_1V_1V_2 = I_n$; si $V \in \mathcal{O}_n(\mathbb{R})$, alors ${}^tV{}^t({}^tV) = {}^tVV = I_n$, et on sait que ${}^tVV = V{}^tV = I_n$.

Exercice 6.3 *Soient $V_1 = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$ et $V_2 = \frac{1}{5}\left(\begin{smallmatrix} 3 & 4 \\ -4 & 3 \end{smallmatrix}\right)$; montrer que V_1 et V_2 sont dans $\mathcal{O}_2(\mathbb{R})$. Même question en dimension 3 avec*

$$W_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad et \qquad W_2 = \frac{1}{2}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & \sqrt{3} \\ 0 & -\sqrt{3} & 1 \end{pmatrix}.$$

Exercice 6.4 *Calculer les déterminants des matrices V_1, V_2, W_1 et W_2 de l'exercice 6.3.*

Le résultat de ce dernier exercice participe d'un phénomène général :

Proposition 6.5 *Soit $V \in \mathcal{O}_n(\mathbb{R})$. Alors* $\det(V) = \pm 1$.

La preuve tient en quelques mots : étant donnée $V \in \mathcal{O}_n(\mathbb{R})$, en appliquant le déterminant à l'identité $V{}^tV = I_n$, on obtient : $(\det V)^2 = \det(V{}^tV) = \det(I_n) = 1$ – car « déterminant d'un produit = produit des déterminants » et « déterminant de la transposée = déterminant de la matrice initiale ». La proposition s'en déduit alors, car les seuls réels de carré 1 sont 1 et -1.

Ceci nous permet de constituer un nouveau sous-ensemble de matrices :

☛ **Proposition-Définition 6.6 (Groupe spécial orthogonal)** *On appelle **groupe spécial orthogonal**, et l'on note $\mathcal{SO}_n(\mathbb{R})$, l'ensemble des matrices orthogonales de déterminant 1, soit :*

$$\mathcal{SO}_n(\mathbb{R}) = \left\{V \in \mathcal{M}_n(\mathbb{R}) \middle| {}^tVV = I_n \text{ et } \det(V) = 1\right\}.$$

Comme $\mathcal{O}_n(\mathbb{R})$, c'est un groupe pour la multiplication des matrices.

On pourra parler de *matrices spéciales orthogonales*, ou encore *matrices orthogonales directes*, pour les éléments de $\mathcal{SO}_n(\mathbb{R})$, soit les matrices orthogonales de déterminant 1.

6.2 Description des matrices (spéciales) orthogonales en dimensions 2 et 3

L'objet de cette section est de décrire, à l'aide de « matrices standard », les matrices orthogonales introduites dans la section précédente, en basses dimensions (2 et 3). Les matrices standard étant attachées à des *transformations géométriques simples*, on exploitera cette description pour décrire géométriquement les matrices orthogonales générales.

6.2.1 La dimension 2. — On commence par le cas spécial orthogonal : soit

$$V = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathcal{SO}_2(\mathbb{R})\,;$$

on a donc comme contraintes sur les coefficients les équations $a^2+c^2 = b^2+d^2 = 1$, $ab+cd = 0$, et $ad-bc = 1$, que l'on peut réécrire sous la forme :

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} a \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} d \\ -b \end{pmatrix},$$

c'est-à-dire :

$${}^tV\begin{pmatrix} a \\ c \end{pmatrix} = {}^tV\begin{pmatrix} d \\ -b \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

La matrice ${}^tV$ étant inversible, on en déduit que les vecteurs $\left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right)$ et $\left(\begin{smallmatrix} d \\ -b \end{smallmatrix}\right)$ coïncident, soit :

$$d = a \quad \text{et} \quad c = -b,$$

avec de plus $a^2+b^2 = 1$. Notons que cette dernière équation entraîne que $|a| \leqslant 1$; on peut donc écrire $a = \cos\theta$ pour un certain $\theta \in \mathbb{R}$, déterminé au signe et à 2π près. On a alors $b^2 = 1-(\cos\theta)^2 = (\sin\theta)^2$, soit $b = \pm\sin\theta$; quitte à remplacer θ par $-\theta$ (ce qui ne change pas $\cos\theta$), on a donc $b = -\sin\theta$. L'« angle » θ est alors uniquement déterminé à 2π près, d'où :

☛ **Théorème 6.7 (Matrices spéciales orthogonales de taille 2)** *Les éléments de $\mathcal{SO}_2(\mathbb{R})$ sont les matrices de la forme*

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad \theta \in \mathbb{R}. \tag{25}$$

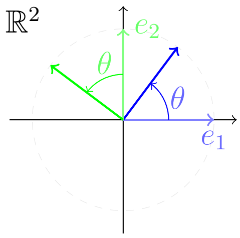
Plus précisément, étant donnée $V \in \mathcal{SO}_2(\mathbb{R})$, il existe $\theta \in \mathbb{R}$, unique modulo 2π[18]*, tel que $V = \left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$; réciproquement, les matrices de cette forme sont dans $\mathcal{SO}_2(\mathbb{R})$.*

18. C'est-à-dire uniquement déterminé à 2π près.

Exercice 6.8 *Vérifier qu'en effet, on a bien* $\left(\begin{smallmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{smallmatrix}\right) \in \mathcal{SO}_2(\mathbb{R})$ *pour tout* $\theta \in \mathbb{R}$.

Exercice 6.9 *Soient* θ, $\varphi \in \mathbb{R}$. *Calculer le produit* $\left(\begin{smallmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{smallmatrix}\right)\left(\begin{smallmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{smallmatrix}\right)$.

Le résultat de l'exercice 6.9 n'est pas dû au hasard ; en effet, $\left(\begin{smallmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{smallmatrix}\right)$ transforme le vecteur $e_1 = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ en $\left(\begin{smallmatrix}\cos\theta\\ \sin\theta\end{smallmatrix}\right)$, et le vecteur $e_2 = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$ en $\left(\begin{smallmatrix}-\sin\theta\\ \cos\theta\end{smallmatrix}\right)$:



On constate donc que lorsqu'elle agit sur le plan $\mathbb{R}^2$, la matrice $\left(\begin{smallmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{smallmatrix}\right)$ n'est autre que *la rotation d'angle* θ !

Par suite, interprété géométriquement, le produit

$$\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{pmatrix}$$

est la rotation d'angle φ suivie de la rotation d'angle θ, soit au final *la rotation d'angle* $\varphi + \theta$, ce qui matriciellement correspond bien à

$$\begin{pmatrix}\cos(\varphi+\theta) & -\sin(\varphi+\theta)\\ \sin(\varphi+\theta) & \cos(\varphi+\theta)\end{pmatrix}.$$

Une fois que l'on a déterminé l'angle de la rotation qui leur est associée (leur « θ »), multiplier entre elles les matrices de $\mathcal{SO}_2(\mathbb{R})$ est donc aisé, puisqu'il s'agit alors simplement d'additionner les angles[19] !

On conclut ce paragraphe par le cas des matrices orthogonales de déterminant -1, telles

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

Exercice 6.10 *En calculant* Se_1 *et* Se_2, *identifier la transformation géométrique correspondant à* S.

19. Et l'on voit en particulier que restreinte à $\mathcal{SO}_2(\mathbb{R})$, la multiplication des matrices redevient commutative.

Noter qu'on a un comportement analogue pour la matrice $\left(\begin{smallmatrix}-1 & 0\\ 0 & 1\end{smallmatrix}\right)$, qui renverse les abscisses sans changer les ordonnées. Cette observation se généralise comme suit :

Proposition 6.11 *Les éléments de $\mathcal{O}_2(\mathbb{R})$ qui ne sont pas dans $\mathcal{SO}(2)$ sont les matrices de la forme*

$$\begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix}, \quad \theta\in\mathbb{R}. \tag{26}$$

Géométriquement, une telle matrice s'interprète comme la symétrie orthogonale, ou réflexion, par rapport à la droite faisant un angle $\theta/2$ avec l'axe des abscisses.

Remarque 6.12 Attention *à la place du signe $-$ dans les expressions* (25) *et* (26)*. Malgré leur air de famille, ce petit changement de place a des effets drastiques – à commencer par le déterminant, qui passe de 1 à -1 !*

On peut exploiter le théorème 6.7 de description de $\mathcal{SO}_2(\mathbb{R})$ pour obtenir un tel résultat. Soit $V\in\mathcal{O}_2(\mathbb{R})$ de déterminant -1 ; alors $VS\in\mathcal{O}_2(\mathbb{R})$, et $\det(VS) = (\det V)(\det S) = (-1)\cdot(-1) = 1$. Par suite, $VS\in\mathcal{SO}_2(\mathbb{R})$, donc VS est de la forme $\left(\begin{smallmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{smallmatrix}\right)$ pour un $\theta\in\mathbb{R}$. Comme $S^2 = I_2$, on en déduit :

$$V = (VS)S = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}S = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} = \begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix},$$

d'où la première partie de la proposition 6.11 (en vérifiant aussi que les matrices de cette forme sont bien orthogonales de déterminant -1, ce qui est élémentaire). Revenons alors à l'identité $V = \left(\begin{smallmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{smallmatrix}\right)S$, que l'on peut réécrire sous la forme $V = \left(\begin{smallmatrix}\cos(\theta/2) & -\sin(\theta/2)\\ \sin(\theta/2) & \cos(\theta/2)\end{smallmatrix}\right)^2 S$, une rotation d'angle θ étant égale à deux rotations successives d'angle $\theta/2$. On utilise alors :

Exercice 6.13 *Soit $\varphi\in\mathbb{R}$. Montrer que :*

$$\begin{pmatrix}\cos\varphi & -\sin\varphi\\ \sin\varphi & \cos\varphi\end{pmatrix}S = S\begin{pmatrix}\cos\varphi & \sin\varphi\\ -\sin\varphi & \cos\varphi\end{pmatrix}.$$

En effet, en posant $\varphi = \theta/2$, on fait passer l'une des rotations d'angle $\theta/2$ de gauche à droite de S, en la changeant en rotation d'angle $-\theta/2$:

$$\begin{pmatrix}\cos(\theta/2) & -\sin(\theta/2)\\ \sin(\theta/2) & \cos(\theta/2)\end{pmatrix}S = S\begin{pmatrix}\cos(\theta/2) & \sin(\theta/2)\\ -\sin(\theta/2) & \cos(\theta/2)\end{pmatrix},$$

d'où : $V = \begin{pmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{pmatrix} S \begin{pmatrix} \cos(\theta/2) & \sin(\theta/2) \\ -\sin(\theta/2) & \cos(\theta/2) \end{pmatrix}$, soit : $V = R_{\theta/2} S R_{-\theta/2}$ avec les abréviations évidentes. Posons alors $v_1 = R_{\theta/2} e_1$ et $v_2 = R_{\theta/2} e_2$, de sorte que $e_1 = R_{\theta/2}^{-1} v_1 = R_{-\theta/2} v_1$, et de même, $e_2 = R_{-\theta/2} v_2$. On a alors (en se souvenant que $Se_1 = e_1$ et $Se_2 = -e_2$) : $Vv_1 = R_{\theta/2} S R_{-\theta/2} v_1 = R_{\theta/2} S e_1 = R_{\theta/2} e_1 = v_1$ et de même, $Vv_2 = R_{\theta/2} S R_{-\theta/2} v_2 = R_{\theta/2} S e_2 = -R_{\theta/2} e_2 = -v_2$.

En conclusion, V laisse inchangé v_1 et donc les vecteurs qui lui sont proportionnels, et qui forment la droite du plan faisant un angle de $\theta/2$ avec l'axe des abscisses ; elle transforme en revanche v_2 et les vecteurs qui lui sont proportionnels, formant l'orthogonal de la droite précédente, en leur opposé. Autrement dit, V est la symétrie orthogonale d'axe la droite faisant un angle de $\theta/2$ avec l'axe des abscisses.

6.2.2 La dimension 3. — Dans la paragraphe précédent, nous avons envisagé la description des matrices (spéciales) orthogonales de taille 2 *via* des matrices standard $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ (si det $= 1$) et $S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ (si det $= -1$), permettant de rendre compte simplement des transformations géométriques associées (rotations d'angle θ et réflexion).

En dimension 3, dans le cas det $= 1$, les matrices standard auxquelles nous allons faire référence sont un analogue des matrices de rotation de la dimension 2 :

Exercice 6.14 *Soit $\theta \in \mathbb{R}$. Montrer que la matrice*

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{27}$$

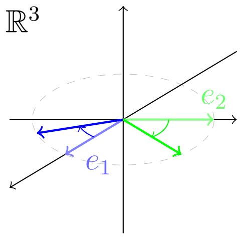
appartient à $\mathcal{SO}_3(\mathbb{R})$.

Exercice 6.15 *Soit une matrice M de la forme*

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

avec $A \in \mathcal{M}_2(\mathbb{R})$. On suppose que $M \in \mathcal{SO}_3(\mathbb{R})$; montrer que $A \in \mathcal{SO}_2(\mathbb{R})$.

La matrice de l'exercice 6.14 s'interprète géométriquement en décomposant $\mathbb{R}^3$ selon le plan horizontal $\mathbb{R}^2$ « des x et des y », et l'axe vertical « des z ». La matrice agit alors comme $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ sur le plan, en laissant inchangé l'axe Oz ; autrement dit, il s'agit de la *rotation d'axe Oz et d'angle θ* (« vue du dessus ») :



Toutefois, contrairement à ce qui était décrit en dimension 2 par le théorème 6.7, toutes les matrices de $\mathcal{SO}_3(\mathbb{R})$ de sont pas immédiatement de la forme (27) ; en témoigne la matrice

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

On peut néanmoins s'y ramener *via* un « changement de base orthonormale » :

Théorème 6.16 (Matrices spéciales orthogonales de taille 3) *Les matrices de $\mathcal{SO}_3(\mathbb{R})$ sont les matrices de la forme*

$$V \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} {}^tV, \tag{28}$$

avec $V \in \mathcal{O}(3)$. Géométriquement, une telle matrice s'interprète comme une[20] *rotation d'angle θ et d'axe $\mathbb{R}w$, où $V = (u|v|w)$.*

Remarque 6.17 *Ce résultat peut sembler décevant en termes de classification, puisque l'on y décrit des matrices de $\mathcal{SO}_3(\mathbb{R})$ – donc en particulier de $\mathcal{O}_3(\mathbb{R})$ – à l'aide de matrices de rotation standard, et de... matrices de $\mathcal{O}_3(\mathbb{R})$! Il semble donc plus pertinent d'en retenir l'interprétation géométrique, qui en substance nous dit qu'*une transformation rigide de l'espace (préservant un point) est toujours une rotation par rapport à un certain axe de l'espace (passant par ce point).

On se focalisera ainsi dans la plupart des exercices sur la détermination des éléments géométriques constitutifs d'une telle rotation, à savoir axe *et* angle.

Exercice 6.18 *Pour $M = (m_{ij})_{1 \leqslant i,j \leqslant 3} \in \mathcal{M}_3(\mathbb{R})$, on appelle trace de M la somme de ces coefficients diagonaux $m_{11} + m_{22} + m_{33}$. En supposant que $M \in \mathcal{SO}_3(\mathbb{R})$, montrer que sa trace est comprise entre -1 et 3.*

(Indication : écrire la trace de M sous la forme $\langle Me_1, e_1\rangle + \langle Me_2, e_2\rangle + \langle Me_3, e_3\rangle$, puis remplacer (e_1, e_2, e_3) par une base orthonormée bien choisie).

20. Sachant qu'il y a deux rotations de la sorte, celle « vue du dessus », et celle « vue du dessous ». On peut lever cette ambiguïté, une fois fixé w, en imposant $V \in \mathcal{SO}_3(\mathbb{R})$.

Pour conclure ce chapitre, on déduit du théorème 6.16 l'énoncé suivant sur les matrices de $\mathcal{O}_3(\mathbb{R})$ qui ne sont pas dans $\mathcal{SO}_3(\mathbb{R})$:

Proposition 6.19 *Les matrices de $\mathcal{O}_3(\mathbb{R})$ qui ne sont pas dans $\mathcal{SO}_3(\mathbb{R})$ sont les matrices de la forme*

$$V \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix} {}^tV, \tag{29}$$

avec $V \in \mathcal{O}(3)$. Une telle matrice s'interprète géométriquement comme une rotation d'angle θ et d'axe $\mathbb{R}w$, où $V = (u|v|w)$, suivie d'une symétrie orthogonale par rapport au plan $(\mathbb{R}w)^\perp$.

En effet, étant donnée $M \in \mathcal{O}_3(\mathbb{R})$, $M \notin \mathcal{SO}_3(\mathbb{R})$, on applique le théorème 6.16 à $(-M)$, qui nous donne l'existence de $V = (u|v|w) \in \mathcal{O}_3(\mathbb{R})$ et de $\varphi \in \mathbb{R}$ (on garde θ pour plus tard) tels que

$$-M = V \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} {}^tV, \quad i.e. \quad M = V \begin{pmatrix} -\cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & -\cos\varphi & 0 \\ 0 & 0 & -1 \end{pmatrix} {}^tV.$$

En posant $\theta = \varphi + \pi$, de sorte que $\cos\theta = -\cos\varphi$ et $\sin\theta = -\sin\varphi$, on a donc bien :

$$M = V \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix} {}^tV.$$

Exercice 6.20 (*) *Soit $\varphi \in \mathbb{R}$. Quel est l'angle (le « θ », dans la forme (29)) associé à une matrice de la forme*

$$V \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ \sin\varphi & -\cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} {}^tV,$$

où $V \in \mathcal{O}_3(\mathbb{R})$?

6.3 Appendice : démonstration du théorème 6.16

Soit $M \in \mathcal{SO}_3(\mathbb{R})$. On commence par chercher un vecteur $w \in \mathbb{R}^3$ jouant le rôle de la dernière colonne de la matrice V de l'énoncé, c'est-à-dire de vecteur directeur de l'axe de M. Par définition, l'axe de M étant inchangé par M, on peut chercher w comme solution non nulle de l'équation

$$Mw = w.$$

Une première étape consiste donc à s'assurer qu'une telle solution existe, ce qui revient à vérifier que $\det(M - I_3) = 0$, ou encore que

$$\chi_M(1) = 0,$$

avec $\chi_M(\lambda) = \det(M - \lambda I_3)$, *cf.* appendice 5.3.

Or les règles de calcul de déterminant nous disent que $\chi_M(\lambda)$ est un polynôme en λ de degré 3, de la forme $-\lambda^3 + \alpha\lambda^2 + \beta\lambda + \gamma$, avec α, β, $\gamma \in \mathbb{R}$; par suite, $\lim_{\lambda\to+\infty} \chi_M(\lambda) = -\infty$, tandis que $\chi_M(0) = \det M = 1$. Le graphe de χ_M sur $\mathbb{R}^+$, continu, intersecte donc au moins une fois l'axe des abscisses, ce qui signifie qu'il existe au moins un $\lambda \in \mathbb{R}^+$ tel que $\chi_M(\lambda) = 0$. En se rappelant que $\chi_M(\lambda) = \det(M - \lambda I_3)$, on sait donc qu'il existe $w \in \mathbb{R}^3$ non nul tel que

$$Mw = \lambda w.$$

Ainsi,

$$\langle Mw, Mw\rangle = \langle \lambda w, \lambda w\rangle = \lambda^2\|w\|^2.$$

Or $\langle Mw, Mw\rangle = \langle w, {}^tMMw\rangle = \langle w, w\rangle = \|w\|^2 > 0$, et donc $\lambda^2 = 1$, soit : $\lambda = 1$, puisque $\lambda \geqslant 0$.

On a donc trouvé un vecteur $w \neq 0$, que l'on peut supposer normalisé, tel que $Mw = w$, ce qui implique en particulier $w = {}^tMMw = {}^tMw$. On complète w en une base orthonormée (u, v, w) de $\mathbb{R}^3$ selon le même algorithme que dans l'appendice 5.3. Alors $\langle Mu, w\rangle = \langle u, {}^tMw\rangle = \langle u, w\rangle = 0$, et de même $\langle v, w\rangle = 0$. Cela signifie, en posant $V = (u|v|w) \in \mathcal{O}_3(\mathbb{R})$, que :

$$M \text{ est de la forme } V \left(\begin{array}{c|c} A & 0 \\ \hline 0 & 1 \end{array}\right) {}^tV,$$

avec $A \in \mathcal{M}_2(\mathbb{R})$. Plus précisément, comme alors $\left(\begin{smallmatrix} A & 0 \\ 0 & 1 \end{smallmatrix}\right) = {}^tVMV$, on a que $\left(\begin{smallmatrix} A & 0 \\ 0 & 1 \end{smallmatrix}\right)$ est dans $\mathcal{O}_3(\mathbb{R})$, et a pour déterminant $\det({}^tV)\det M \det V = \det M = 1$, car $\det({}^tV)\det V = \det({}^tVV) = 1$. Par conséquent, $\left(\begin{smallmatrix} A & 0 \\ 0 & 1 \end{smallmatrix}\right) \in \mathcal{SO}_3(\mathbb{R})$, donc $A \in \mathcal{SO}_2(\mathbb{R})$, *cf.* exercice 6.15. D'après le théorème 6.7, A est de la forme $\left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$, $\theta \in \mathbb{R}$, et donc :

$$M = V \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} {}^tV,$$

avec $V \in \mathcal{O}_3(\mathbb{R})$. $\square$

6.4 Corrigés des exercices du chapitre 6

6.4.1 Exercice 6.3. — Comme V_1 est symétrique, $V_1{}^tV_1 = V_1^2 = \frac{1}{2}\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)^2 = \frac{1}{2}\left(\begin{smallmatrix} 2 & 0 \\ 0 & 2 \end{smallmatrix}\right) = I_2$. Pour V_2 on a : $V_2{}^tV_2 = \frac{1}{25}\left(\begin{smallmatrix} 3 & 4 \\ -4 & 3 \end{smallmatrix}\right)\left(\begin{smallmatrix} 3 & -4 \\ 4 & 3 \end{smallmatrix}\right) = \frac{1}{2}\left(\begin{smallmatrix} 3^2+4^2 & 3\cdot(-4)+4\cdot 3 \\ -4\cdot 3+3\cdot 4 & 3^2+(-4)^2 \end{smallmatrix}\right) = I_2$. Conclusion : V_1 et V_2 sont dans $\mathcal{O}_2(\mathbb{R})$.

Considérons W_1 ; il est clair que ses colonnes forment une base orthonormée de $\mathbb{R}^3$, puisque, à l'ordre près, il s'agit de la base canonique ! Quant à W_2, on peut la mettre sous la forme $\left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & A \end{array}\right)$, avec $A = \frac{1}{2}\left(\begin{smallmatrix} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{smallmatrix}\right)$ que sorte que $W_2 = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & {}^tA \end{array}\right)$; ainsi, $W_2{}^tW_2 = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & A \end{array}\right)\left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & {}^tA \end{array}\right) = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & A{}^tA \end{array}\right)$, et on vérifie que $A^tA = I_2$, d'où il résulte que $W_2{}^tW_2 = I_3$. Conclusion : W_1 et W_2 sont dans $\mathcal{O}_3(\mathbb{R})$

6.4.2 Exercice 6.4. — On a : $\det(V_1) = \frac{1}{2}\big(1\cdot(-1) - 1\cdot 1\big) = 1$, $\det(V_2) = \frac{1}{25}\big(3\cdot 3 - (-4)\cdot 4\big) = \frac{9+16}{25} = 1$, et $\det(W_1) = (-1)\cdot\left|\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right| = -1$ (en développant par rapport à la première colonne), $\det(W_2) = \frac{1}{8}\cdot 2\cdot\left|\begin{smallmatrix} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{smallmatrix}\right| = \frac{1}{4}\big(1\cdot 1 - (-\sqrt{3})\cdot\sqrt{3}\big) = 1$ (en développant par rapport à la première ligne).

6.4.3 Exercice 6.10. — On a $Se_1 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right) = e_1$, et $Se_2 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 \\ -1 \end{smallmatrix}\right) = -e_2$. La matrice S se traduit donc géométriquement par une réflexion d'axe l'axe des abscisses.

6.4.4 Exercice 6.13. — En calculant directement, on constate que l'on obtient

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix}$$

de chaque côté.

6.4.5 Exercice 6.14. — Écrivons la matrice étudiée, M disons, sous la forme $\left(\begin{array}{c|c} A & 0 \\ \hline 0 & 1 \end{array}\right)$, avec donc $A = \left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right) \in \mathcal{SO}_2(\mathbb{R})$ (théorème 6.7).

Ainsi, ${}^tMM = \left(\begin{array}{c|c} {}^tA & 0 \\ \hline 0 & 1 \end{array}\right)\left(\begin{array}{c|c} A & 0 \\ \hline 0 & 1 \end{array}\right) = \left(\begin{array}{c|c} {}^tAA & 0 \\ \hline 0 & 1 \end{array}\right)$, et comme ${}^tAA = I_2$, ${}^tMM = I_3$, c'est-à-dire : $M \in \mathcal{O}_3(\mathbb{R})$. Reste à vérifier que $\det(M) = 1$; ceci résulte simplement de l'égalité $\det(M) = \det(A)$ obtenue en développant $\det(M)$ par rapport à la dernière ligne de M, et du fait que $\det(A) = 1$.

6.4.6 Exercice 6.15. — Comme $M \in \mathcal{SO}_3(\mathbb{R})$, ${}^tMM = I_3$. Or

$${}^tMM = \left(\begin{array}{c|c} {}^tAA & 0 \\ \hline 0 & 1 \end{array}\right),$$

et donc ${}^tAA = I_2$, soit : $A \in \mathcal{O}_2(\mathbb{R})$. De plus, on a en développant par rapport à la dernière ligne $\det M = 1 \cdot \det A$, et comme $\det M = 1$, on a bien $A \in \mathcal{SO}_2(\mathbb{R})$.

6.4.7 Exercice 6.18. — Notons $\operatorname{tr} M$ la trace de M. Suivant l'indication, on a $Me_1 = m_{11}e_1 + m_{12}e_2 + m_{13}e_3$, donc $\langle Me_1, e_1\rangle = m_{11}$; de même, $\langle Me_2, e_2\rangle = m_{22}$ et $\langle Me_3, e_3\rangle = m_{33}$, et donc $\operatorname{tr} M = \langle Me_1, e_1\rangle + \langle Me_2, e_2\rangle + \langle Me_3, e_3\rangle$.

Soit à présent (u, v, w) une base orthonormée, de sorte que M s'écrive sous la forme (28), avec $V = (u|v|w)$; suivant encore l'indication, vérifions que $\operatorname{tr} M = \langle Mu, u\rangle + \langle Mv, v\rangle + \langle Mw, w\rangle$. En remplaçant le premier u par $\langle e_1, u\rangle e_1 + \langle e_2, u\rangle e_2 + \langle e_3, u\rangle e_3 = \sum_{j=1}^{3} \langle e_j, u\rangle e_j$, dans le membre de droite, et de même pour le premier e_2 et le premier e_3, il vient :

$$\begin{aligned}
&\langle Mu, u\rangle + \langle Mv, v\rangle + \langle Mw, w\rangle \\
&\quad = \Big\langle M \sum_{j=1}^{3} \langle e_j, u\rangle e_j, u\Big\rangle + \Big\langle M \sum_{j=1}^{3} \langle e_j, v\rangle e_j, v\Big\rangle + \Big\langle M \sum_{j=1}^{3} \langle e_j, w\rangle e_j, w\Big\rangle \\
&\quad = \Big\langle \sum_{j=1}^{3} \langle e_j, u\rangle Me_j, u\Big\rangle + \Big\langle \sum_{j=1}^{3} \langle e_j, v\rangle Me_j, v\Big\rangle + \Big\langle \sum_{j=1}^{3} \langle e_j, w\rangle Me_j, w\Big\rangle \\
&\quad = \sum_{j=1}^{3} \big\langle \langle e_j, u\rangle Me_j, u\big\rangle + \sum_{j=1}^{3} \big\langle \langle e_j, v\rangle Me_j, v\big\rangle + \sum_{j=1}^{3} \big\langle \langle e_j, w\rangle Me_j, w\big\rangle \\
&\quad = \sum_{j=1}^{3} \big\langle Me_j, \langle e_j, u\rangle u\big\rangle + \sum_{j=1}^{3} \big\langle Me_j, \langle e_j, v\rangle v\big\rangle + \sum_{j=1}^{3} \big\langle Me_j, \langle e_j, w\rangle w\big\rangle \\
&\quad = \sum_{j=1}^{3} \big\langle Me_j, \langle e_j, u\rangle u + \langle e_j, v\rangle v + \langle e_j, w\rangle w\big\rangle
\end{aligned}$$

Or pour $j = 1$, 2 et 3, $\langle e_j, u\rangle u + \langle e_j, v\rangle v + \langle e_j, w\rangle w = e_j$, car (u, v, w) est une base orthonormée de $\mathbb{R}^3$. Par suite,

$$\langle Mu, u\rangle + \langle Mv, v\rangle + \langle Mw, w\rangle = \sum_{j=1}^{3} \langle Me_j, e_j\rangle = \operatorname{tr} M.$$

On conclut alors facilement en remarquant que ${}^tVu = e_1$, de sorte que $Mu = V\big((\cos\theta)e_1 + (\sin\theta)e_2\big) = (\cos\theta)u + (\sin\theta)v$ puis $\langle Mu, u\rangle = \cos\theta$, que de même $Mv = (-\sin\theta)u + (\cos\theta)v$ puis $\langle Mv, v\rangle = \cos\theta$, et enfin que $Mw = w$, donc $\langle Mw, w\rangle = 1$. On a ainsi $\operatorname{tr} M = 1 + 2\cos\theta$, ce qui est bien compris entre -1 et 3, puisque $\cos\theta$ est compris entre -1 et 1.

6.4.8 Exercice 6.20. — Appelons M une matrice de la forme suggérée. Déjà, $M \in \mathcal{O}_3(\mathbb{R})$. De plus, $\det(M) = -1$, *i.e.* $M \notin \mathcal{SO}_3(\mathbb{R})$, et il s'agit donc de voir de quelle manière M s'écrit sous la forme (29) de la proposition 6.19.

On a vu à la fin du paragraphe 6.2.1 l'identité

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{pmatrix} = R_{\varphi/2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} {}^t R_{\varphi/2},$$

avec $R_{\varphi/2} = \begin{pmatrix} \cos(\varphi/2) & -\sin(\varphi/2) \\ \sin(\varphi/2) & \cos(\varphi/2) \end{pmatrix}$. En dimension 3, ceci se traduit par :

$$M = V \left(\begin{array}{c|c} R_{\varphi/2} & 0 \\ \hline 0 & 1 \end{array}\right) \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} {}^t\!\left(\begin{array}{c|c} R_{\varphi/2} & 0 \\ \hline 0 & 1 \end{array}\right) {}^t V.$$

Observons que le matrice centrale n'a pas d'effet selon deux axes (Ox et Oz), tandis qu'elle change la coordonnée selon le troisième axe (Oy) en son opposée ; elle agit donc de manière similaire à

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

qui a pour seul effet de changer la coordonnée selon Oz en son opposée. En d'autres termes, quitte à interchanger les rôles de Oy et Oz, ces deux matrices n'en sont qu'une, ce qui se résume dans la relation :

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = P \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} {}^t P,$$

avec $P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \in \mathcal{O}_3(\mathbb{R})$, qui sert justement à échanger Oy et Oz.

En résumé, on a donc

$$M = W \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} {}^t W,$$

avec $W = V \left(\begin{array}{c|c} R_{\varphi/2} & 0 \\ \hline 0 & 1 \end{array}\right) P \in \mathcal{O}_3(\mathbb{R})$, ce qui correspond à l'écriture (29). L'angle associé à M est donc 0 (modulo 2π).

Une méthode similaire à celle de l'exercice 6.18 nous aurait par ailleurs donné : $\operatorname{tr} M = -1 + 2\cos\theta$ d'une part, et

$$\operatorname{tr} M = \cos\varphi - \cos\varphi + 1 = 1$$

d'autre part. On retrouve bien $\cos\theta = 1$, et donc $\theta = 0$ modulo 2π, même si l'on perd l'information sur la matrice W, non demandée ici.

Notations