



Université Grenoble-Alpes
Centre Drôme-Ardèche

Cours MAT244

**Formes quadratiques, séries
et séries de Fourier pour la physique**

Romain JOLY

Dernière mise à jour : janvier 2016

Table des matières

Chapitre 1 : Formes bilinéaires et quadratiques	1
1 Rappels	1
2 Applications et formes bilinéaires	2
3 Formes quadratiques	5
Chapitre 2 : Espaces euclidiens et préhilbertiens	11
1 Produits scalaires et normes	11
2 Espaces euclidiens et bases orthonormales	13
3 Espaces préhilbertiens et approximations	16
4 Un mot sur les espaces de Hilbert	20
Chapitre 3 : Séries	21
1 Introduction	21
2 Critères de convergence pour les séries de termes positifs	22
3 Séries de fonctions	25
4 Quelques compléments	27
Chapitre 4 : Séries de Fourier	30
1 Introduction	30
2 L'espace des fonctions $\mathcal{C}_{T-per,m}^{k,VP}(\mathbb{R})$	31
3 Séries de Fourier et convergence au sens L^2	34
4 Autres types de convergence	39
5 Série de Fourier en exponentielles complexes	43
Chapitre 5 : Applications aux EDP	46
1 Décompositions de fonctions sur $[0, L]$	46
2 Equation de la chaleur	48
3 Equation de Laplace	52
4 Equation des ondes	54

Chapitre 1 : Formes bilinéaires et quadratiques

Le but de ce chapitre est d'introduire un contexte géométrique qui pourra être utilisé pour décrire les espaces-temps de la relativité, une grande partie des énergies utilisées en physique et les produits scalaires.

1 Rappels

Dans la suite E est un espace vectoriel réel, c'est-à-dire que cet espace est stable par addition (translation) et multiplication par un nombre réel (dilatation) : si x et y sont dans E et si λ et μ sont des réels, alors on peut définir un point $\lambda x + \mu y$ qui est dans E .

Exemples :

- Pour tout $d \geq 1$, \mathbb{R}^d est un espace vectoriel – en particulier, le plan \mathbb{R}^2 , l'espace \mathbb{R}^3 et l'espace-temps \mathbb{R}^4 .
- Si I est un intervalle de \mathbb{R} , l'espace $\mathcal{C}^0(I, \mathbb{R})$ des fonctions à valeurs réelles continues sur I est un espace vectoriel. De même, l'espace $\mathcal{C}^1(I, \mathbb{R})$ des fonctions dérivables et de dérivée continue sur I est un espace vectoriel.
- L'espace des matrices $\mathcal{M}_{m,n}(\mathbb{R})$.

Une application linéaire f entre deux espaces vectoriels E et F est une fonction $f : E \rightarrow F$ qui commute avec les combinaisons linéaires : si x et y sont dans E et si λ et μ sont des réels, alors $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$.

Exemples :

- L'application $x \in \mathbb{R}^3 \mapsto (x_1 + 2x_3, x_2 - x_1) \in \mathbb{R}^2$ est linéaire de \mathbb{R}^3 dans \mathbb{R}^2 .
- La dérivation $f \in \mathcal{C}^1(I, \mathbb{R}) \mapsto f' \in \mathcal{C}^0(I, \mathbb{R})$ est linéaire.

Une forme linéaire φ sur un espace vectoriel E est une application linéaire de E dans \mathbb{R} .

Exemples :

- Si $y \in \mathbb{R}^3$, alors $x \in \mathbb{R}^3 \mapsto \langle x|y \rangle$ est une forme linéaire.
- L'intégrale $f \in \mathcal{C}^0([0, 1], \mathbb{R}) \mapsto \int_0^1 f(x)dx$ est une forme linéaire.

2 Applications et formes bilinéaires

2.1 Définitions

Définitions 1.1. Soient E, E' et F trois espaces vectoriels réels.

- Une application f de $E \times E'$ dans F est dite bilinéaire si pour tout $y \in E', x \mapsto f(x, y)$ est linéaire en x et si pour tout $x \in E, y \mapsto f(x, y)$ est linéaire en y .
- Une forme bilinéaire φ sur E est une application bilinéaire de $E \times E$ dans \mathbb{R} .
- Une application bilinéaire f de $E \times E$ dans F est dite symétrique si : $\forall x, y \in E, f(y, x) = f(x, y)$. Une application bilinéaire f de $E \times E$ dans F est dite anti-symétrique si : $\forall x, y \in E, f(y, x) = -f(x, y)$.

Remarque : Si f est bilinéaire anti-symétrique sur E , alors $f(x, x) = -f(x, x)$ et donc $f(x, x) = 0$.

Exemples :

- Le produit scalaire $(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3 \mapsto \langle x|y \rangle = x_1y_1 + x_2y_2 + x_3y_3$ est une forme bilinéaire symétrique sur \mathbb{R}^3 .
- Le produit coordonnées par coordonnées $(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3 \mapsto x \times y = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1)$ est une application bilinéaire symétrique de $\mathbb{R}^3 \times \mathbb{R}^3$ dans \mathbb{R}^3 .
- Le produit vectoriel $(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3 \mapsto x \wedge y = (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1)$ est une application bilinéaire anti-symétrique de $\mathbb{R}^3 \times \mathbb{R}^3$ dans \mathbb{R}^3 .
- L'application $(f, g) \mapsto \int_0^1 f(x)g(x)dx$ définit une forme bilinéaire symétrique sur $\mathcal{C}^0([0, 1], \mathbb{R})$.
- Le commutateur $(M, N) \in \mathcal{M}_n(\mathbb{R}) \times \mathcal{M}_n(\mathbb{R}) \mapsto [M, N] = MN - NM$ est une application bilinéaire anti-symétrique de $\mathcal{M}_n(\mathbb{R}) \times \mathcal{M}_n(\mathbb{R})$ dans $\mathcal{M}_n(\mathbb{R})$.
- Le déterminant 2×2 $(x, y) \in \mathbb{R}^2 \times \mathbb{R}^2 \mapsto x_1y_2 - x_2y_1$ est une forme bilinéaire anti-symétrique par rapport aux colonnes de la matrice.
- L'application $(x, y) \in \mathbb{R}^2 \times \mathbb{R}^2 \mapsto x_1y_2 - 3x_2y_1 + 2x_2y_2$ est une forme bilinéaire sur \mathbb{R}^2 qui n'est ni symétrique, ni anti-symétrique.

Exercice 1.2. Vérifier les affirmations des exemples ci-dessus.

Définition 1.3. A toute forme bilinéaire φ sur E , on peut définir la forme bilinéaire transposée ${}^t\varphi$ par ${}^t\varphi(x, y) = \varphi(y, x)$.

On note qu'une forme est symétrique si et seulement si ${}^t\varphi = \varphi$ et anti-symétrique si et seulement si ${}^t\varphi = -\varphi$.

Proposition 1.4. Soit E un espace vectoriel et φ une forme bilinéaire. Il existe une unique décomposition $\varphi = \varphi_s + \varphi_a$ avec φ_s une forme bilinéaire symétrique et φ_a une forme bilinéaire anti-symétrique

Démonstration : Pour l'existence, il suffit de poser $\varphi_s = \frac{1}{2}(\varphi + {}^t\varphi)$ et $\varphi_a = \frac{1}{2}(\varphi - {}^t\varphi)$. Par ailleurs, cette décomposition est unique puisque si $\varphi = \varphi_s + \varphi_a$, alors $\varphi + {}^t\varphi = 2\varphi_s$. \square

2.2 Représentation matricielle

Soit f une application bilinéaire de $E \times E'$ dans F . La propriété de bilinéarité montre que, pour tout $x_1, x_2 \in E$, $y_1, y_2 \in E'$ et $\lambda_1, \lambda_2, \mu_1$ et μ_2 réels, on a

$$f(\lambda_1 x_1 + \lambda_2 x_2, \mu_1 y_1 + \mu_2 y_2) = \lambda_1 \mu_1 f(x_1, y_1) + \lambda_1 \mu_2 f(x_1, y_2) + \lambda_2 \mu_1 f(x_2, y_1) + \lambda_2 \mu_2 f(x_2, y_2) .$$

Le même type de développement montre la propriété suivante dans les espaces de dimension finie. Si (e_1, \dots, e_d) est une base de E et $(e'_1, \dots, e'_{d'})$ est une base de E' et si $x = x_1 e_1 + x_2 e_2 + \dots + x_d e_d$ et si $y = y_1 e'_1 + y_2 e'_2 + \dots + y_{d'} e'_{d'}$ alors

$$f(x, y) = \sum_{i=1}^d \sum_{j=1}^{d'} x_i y_j f(e_i, e'_j) .$$

On en déduit les propriétés suivantes.

Proposition 1.5. *On suppose que E et E' sont de dimension finie avec (e_1, \dots, e_d) et $(e'_1, \dots, e'_{d'})$ comme bases respectives.*

- Une application bilinéaire est connue si et seulement si on connaît les images $f(e_i, e'_j)$ des vecteurs de base.
- Une application bilinéaire f de $E \times E$ dans F est symétrique si et seulement si $f(e_i, e_j) = f(e_j, e_i)$ pour tout $i \neq j$.
- Une application bilinéaire f de $E \times E$ dans F est anti-symétrique si et seulement si $f(e_i, e_j) = -f(e_j, e_i)$ pour tout $i \neq j$ et $f(e_i, e_i) = 0$.

En particulier, si φ est une forme bilinéaire sur $E \times E$ de dimension finie, alors

$$\varphi(x_1 e_1 + \dots + x_d e_d, y_1 e_1 + \dots + y_d e_d) = \sum_{i,j=1}^d c_{i,j} x_i y_j ,$$

où le coefficient réel $c_{i,j}$ est défini par

$$c_{i,j} = \varphi(e_i, e_j) .$$

A toute forme bilinéaire, on peut donc associer la matrice $C = (c_{i,j})_{i,j=1,\dots,d}$ qui la définit dans la base (e_1, \dots, e_d) . Si $x = (x_1, \dots, x_d)$ et $y = (y_1, \dots, y_d)$ dans cette base, alors

$$\varphi(x, y) = {}^t x C y = (x_1, \dots, x_d) \begin{pmatrix} c_{1,1} & \cdots & c_{1,d} \\ \vdots & \ddots & \vdots \\ c_{d,1} & \cdots & c_{d,d} \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_d \end{pmatrix} .$$

La forme bilinéaire φ est symétrique si et seulement si C est symétrique. En outre, la transposée tC est la matrice représentant la forme bilinéaire transposée ${}^t\varphi$.

Exemples :

- Le produit scalaire $(x, y) \in \mathbb{R}^3 \times \mathbb{R}^3 \mapsto \langle x|y \rangle = x_1y_1 + x_2y_2 + x_3y_3$ est représenté, dans la base canonique, par la matrice identité $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ qui est symétrique.
- La forme linéaire $(x, y) \in \mathbb{R}^2 \times \mathbb{R}^2 \mapsto x_1y_2 - 3x_2y_1 + 2x_2y_2$ est représentée, dans la base canonique, par la matrice $\begin{pmatrix} 0 & 1 \\ -3 & 2 \end{pmatrix}$.

2.3 Un peu de géométrie

Définitions 1.6. Soit φ une forme bilinéaire symétrique sur E . Deux vecteurs x et y sont dits orthogonaux (pour φ) si $\varphi(x, y) = 0$. On note alors $x \perp y$ (ou $x \perp_\varphi y$ en cas d'ambiguïté). Un vecteur x est perpendiculaire à un sous-espace F de E si $x \perp y$ pour tout $y \in F$. Enfin, on définit l'orthogonal F^\perp d'un sous-espace comme l'ensemble des vecteurs orthogonaux à F .

Proposition 1.7. Si x_1 et x_2 sont orthogonaux à y , alors $\lambda_1x_1 + \lambda_2x_2$ l'est aussi pour tous réels λ_1 et λ_2 . En particulier,

- F^\perp est un sous-espace vectoriel de E (qui contient 0)
- x est orthogonal à F si et seulement s'il est orthogonal à tous les vecteurs d'une base de F .

Exemples :

- Dans une carte qui n'est pas à l'échelle (penser aux points près des pôles dans une carte du globe), les vecteurs (x_1, x_2) sont transformés en $(x'_1, x'_2) = (x_1, 2x_2)$. Pour garder l'ancienne géométrie, le nouveau produit scalaire dans ces coordonnées est $\langle x'|y' \rangle = x'_1y'_1 + \frac{1}{4}x'_2y'_2$. Donc les vecteurs $(1, 2)$ et $(1, -2)$ sont orthogonaux dans les nouvelles coordonnées.
- On considère la forme bilinéaire symétrique

$$\varphi : (x, y) \in \mathbb{R}^3 \times \mathbb{R}^3 \mapsto x_1y_1 + 2x_2y_3 + 2x_3y_2 .$$

Calculons l'orthogonal de $F = \text{vect}((1, 0, 0), (0, 1, 0))$. Soit x orthogonal à F . C'est équivalent à $x \perp (1, 0, 0)$ et $x \perp (0, 1, 0)$, ce qui se traduit par :

$$x_1 \times 1 + 2x_2 \times 0 + 2x_3 \times 0 = 0 \quad \text{et} \quad x_1 \times 0 + 2x_2 \times 0 + 2x_3 \times 1 = 0 .$$

On doit donc avoir $x_1 = 0$ et $x_3 = 0$. Donc $F^\perp = \mathbb{R} \cdot (0, 1, 0)$.

Définition 1.8. Soit φ une forme bilinéaire symétrique sur E . On appelle noyau de φ et on note $\text{Ker}(\varphi)$ l'ensemble des vecteurs orthogonaux à tout E (c'est-à-dire que $\text{Ker}(\varphi) = E^\perp$). On dit que φ est dégénérée si $\text{Ker}(\varphi)$ n'est pas réduit à $\{0\}$ et non-dégénérée sinon.

Proposition 1.9. Soit φ une forme bilinéaire symétrique sur \mathbb{R}^d et C la matrice associée une base quelconque. Alors, φ est dégénérée si et seulement si $\text{Ker}(C)$ n'est pas réduit à 0 , i.e. $\det(C) = 0$. Le rang de φ est défini comme le rang de C et en particulier, $\text{rang}(\varphi) = d - \dim(\text{Ker}(\varphi))$.

Exemples :

- Soit $c > 0$, la forme bilinéaire $\varphi((x, y, z, t), (x', y', z', t')) = xx' + yy' + zz' - c^2tt'$ est non-dégénérée sur l'espace-temps \mathbb{R}^4 .
- La forme bilinéaire symétrique

$$\varphi : (f, g) \in \mathcal{C}^0([0, 1], \mathbb{R}) \times \mathcal{C}^0([0, 1], \mathbb{R}) \longmapsto \int_0^1 f(x)g(x)dx$$

est non-dégénérée.

- La forme bilinéaire symétrique sur \mathbb{R}^2 $\varphi(x, y) = x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2$ est dégénérée avec comme noyau $\text{Ker}(\varphi) = \mathbb{R} \cdot (1, -1)$.

3 Formes quadratiques

3.1 Définition et lien avec les formes bilinéaires

Etant donnée une forme bilinéaire φ sur un espace vectoriel E , on peut considérer la fonction

$$q : x \in E \longmapsto \varphi(x, x).$$

Cette fonction est appelée *forme quadratique* associée à φ . Le terme *quadratique* vient de la propriété $q(\lambda x) = \lambda^2 q(x)$ et du fait que, dans \mathbb{R}^d , la forme q va s'écrire comme un polynôme homogène de degré deux en les coordonnées.

Exemples : On reprend les formes bilinéaires juste ci-dessus.

- La forme quadratique de Lorentz sur l'espace-temps \mathbb{R}^4 est donnée par $q(x, y, z, t) = x^2 + y^2 + z^2 - c^2t^2$, où c est la vitesse de la lumière.
- $q(f) = \int_0^1 |f(x)|^2 dx$ est une forme quadratique sur $\mathcal{C}^0([0, 1], \mathbb{R})$.
- $q(x) = x_1^2 + x_2^2 + 2x_1x_2$ est une forme quadratique sur \mathbb{R}^2 .

Si on dispose d'une forme quadratique, peut-on retrouver la forme bilinéaire associée? Notons tout d'abord que si $\varphi = \varphi_s + \varphi_a$ est la décomposition d'une forme bilinéaire donnée

par la proposition 1.4, alors $q(x) = \varphi(x, x) = \varphi_s(x, x)$. La partie anti-symétrique n'a donc aucun effet sur la forme quadratique. On peut donc se restreindre à chercher une forme bilinéaire symétrique φ_s associée à q (qu'on appelle *forme polaire*). Pour retrouver φ_s à partir de q , on note que

$$q(x + y) = q(x) + 2\varphi_s(x, y) + q(y) .$$

Proposition 1.10. *Soit q une forme quadratique sur E associée à une forme bilinéaire symétrique φ_s . Alors, pour tout x et y dans E , on a :*

– *L'identité du parallélogramme*

$$q(x + y) + q(x - y) = 2(q(x) + q(y)) .$$

– *L'identité de polarisation*

$$\varphi_s(x, y) = \frac{1}{2}(q(x + y) - q(x) - q(y)) .$$

En particulier, pour toute forme quadratique q , il existe une unique forme bilinéaire symétrique φ_s associée et elle est donnée par l'identité de polarisation. Toutes les formes bilinéaires associées à q ne diffèrent que par leur partie anti-symétrique φ_a .

Dorénavant, on associera automatiquement la forme quadratique et la forme bilinéaire symétrique donnée par l'identité de polarisation. On peut définir les formes quadratiques de manière intrinsèque comme suit.

Définition 1.11. *Une forme quadratique sur E est une application $q : E \rightarrow \mathbb{R}$ vérifiant que $(x, y) \mapsto \frac{1}{2}(q(x + y) - q(x) - q(y))$ est une forme bilinéaire sur E .*

Les définitions suivantes sont importantes pour la géométrie associée aux formes quadratiques.

Définitions 1.12. *Soit q une forme quadratique sur E . On dit que q est dégénérée (resp. non-dégénérée) si la forme bilinéaire associée est dégénérée (resp. non-dégénérée). On appelle vecteur isotrope un vecteur x tel que $q(x) = 0$. L'ensemble des vecteurs isotrope forme le cône isotrope de q . On dit que q est définie si q n'a pas de vecteurs isotropes non nuls. On dit que q est positive (resp. négative) si $q(x) \geq 0$ pour tout $x \in E$ (resp. $q(x) \leq 0$).*

Exemple : La forme quadratique de Lorentz $q(x, y, z, t) = x^2 + y^2 + z^2 - c^2t^2$ sur \mathbb{R}^4 est non-dégénérée mais possède un cône isotrope non-trivial appelé cône de lumière (voir exercices).

Proposition 1.13. *Une forme quadratique q qui est définie est soit définie positive, soit définie négative.*

Dans le cas $E = \mathbb{R}^d$, on peut représenter q par la matrice symétrique C représentant φ et on identifie le rang de q à celui de φ . On a

$$q(x) = \varphi(x, x) = {}^t x C x = (x_1, \dots, x_d) \begin{pmatrix} c_{1,1} & \cdots & c_{1,d} \\ \vdots & \ddots & \vdots \\ c_{d,1} & \cdots & c_{d,d} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} .$$

3.2 Réduction et signature

Théorème 1.14. Réduction de Gauss

Soit q une forme quadratique sur \mathbb{R}^d . Il existe k réels non nuls $(\alpha_i)_{i=1\dots k}$ et k formes linéaires indépendantes $(\ell_i)_{i=1\dots k}$ (donc $k \leq d$) tels que

$$\forall x \in \mathbb{R}^d, \quad q(x) = \alpha_1(\ell_1(x))^2 + \alpha_2(\ell_2(x))^2 + \dots + \alpha_k(\ell_k(x))^2 .$$

Les formes ℓ_i sont indépendantes dans le sens que si $c_1\ell_1(x) + c_2\ell_2(x) + \dots + c_k\ell_k(x) = 0$ pour tout x , alors tous les c_i sont nuls. Par exemple, si ℓ_1 dépend de x_1 et pas les autres ℓ_i , puis ℓ_2 dépend de x_2 mais pas les formes suivantes etc., alors les ℓ_i sont indépendantes. De manière générale, on peut mettre les coefficients des formes linéaires dans une matrice $d \times k$ et les formes sont indépendantes si et seulement si on peut trouver k lignes de la matrice telles que le déterminant $k \times k$ correspondant est non nul.

Exemples : On se place sur \mathbb{R}^3 .

- Les formes linéaires $\ell_1(x) = x_1 + x_2 - 2x_3$, $\ell_2(x) = 2x_2 + x_3$ et $\ell_3(x) = x_3$ sont indépendantes.
- Les formes linéaires $\ell_1(x) = x_1 + x_2 + x_3$, $\ell_2(x) = x_1 - x_2 + x_3$ et $\ell_3(x) = x_2$ ne sont pas indépendantes car $\ell_1 - \ell_2 - 2\ell_3 = 0$.
- Les formes linéaires $\ell_1(x) = x_1 + x_2 + x_3$, $\ell_2(x) = x_1 - x_2 + x_3$ et $\ell_3(x) = x_3$ sont indépendantes car la matrice $\begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$ est de déterminant non nul.

La démonstration du théorème 1.14 se fait par récurrence, en construisant petit à petit les termes $\alpha_1(\ell_1(x))^2$. En particulier, la démonstration est constructive et donne une méthode algorithmique de réduction. Pour éviter les notations trop lourdes, nous allons simplement voir comment procéder une étape de réduction sur des exemples.

Supposons que nous avons une forme quadratique

$$q(x) = \sum_{i,j} c_{ij} x_i x_j .$$

1er cas : il existe un terme carré $c_{ii}x_i^2$ avec $c_{ii} \neq 0$.

On regroupe alors tous les termes contenant x_i en les mettant dans un carré du type

$c_{ii}(x_i + \sum_{j \neq i} \alpha_j x_j)^2$ et en soustrayant les termes en trop. Tout ce qui reste en dehors de ce carré ne dépendra pas de x_i .

2ème cas : il n'existe aucun terme carré $c_{ii}x_i^2$ non nul.

On prend un terme type $c_{ij}x_i x_j$ et on regroupe tous les termes comprenant x_i ou x_j dans un produit $c_{ij}(x_i + \sum_{k \neq i,j} \alpha_k x_k)(x_j + \sum_{k \neq i,j} \beta_k x_k)$. Puis on utilise la formule $ab = \frac{1}{4}(a+b)^2 - \frac{1}{4}(a-b)^2$. Les termes venant de $(a+b)$ et $(a-b)$ sont indépendants et tous les autres termes ne feront plus intervenir x_i et x_j .

Exemples :

1. On veut réduire $q(x, y, z) = x^2 + 2xy - yz$. On utilise le terme x^2 comme pivot. On a $q(x, y, z) = (x+y)^2 - y^2 - yz$. Puis on utilise le terme $-y^2$ comme pivot et on obtient $q(x, y, z) = (x+y)^2 - (y+z/2)^2 + z^2/4$.
2. On veut réduire $q(x, y, z) = xy + yz$. On utilise le terme xy comme pivot et on écrit $q(x, y, z) = (x+z)y = \frac{1}{4}(x+z+y)^2 - \frac{1}{4}(x+z-y)^2$.
3. On veut réduire $q(x, y, z) = -x^2 + 2xy - y^2 + 3yz$. On écrit

$$q(x, y, z) = -(x-y)^2 + 3yz = -(x-y)^2 + \frac{3}{4}(y+z)^2 - \frac{3}{4}(y-z)^2 .$$

La réduction de Gauss permet de récupérer beaucoup d'informations sur q .

Théorème 1.15. Soit q une forme quadratique sur \mathbb{R}^d qu'on suppose mise sous forme réduite

$$q(x) = \alpha_1(\ell_1(x))^2 + \dots + \alpha_k(\ell_k(x))^2 .$$

La forme bilinéaire symétrique associée à q est donnée par $\varphi(x, y) = \alpha_1 \ell_1(x) \ell_1(y) + \dots + \alpha_k \ell_k(x) \ell_k(y)$. Elle est dégénérée si et seulement si $k < d$. De plus, q est positive (resp. négative) si et seulement si tous les α_i sont positifs (resp. négatifs) et q est définie positive (resp. définie négative) si et seulement si tous les α_i sont (strictement) positifs (resp. négatifs) et $k = d$.

Théorème 1.16. Principe d'inertie de Sylvester

Soit q une forme quadratique sur \mathbb{R}^d . Quelle que soit la forme réduite $q = \alpha_1(\ell_1)^2 + \dots + \alpha_k(\ell_k)^2$, l'entier k , le nombre k_+ de α_i strictement positifs et le nombre k_- de α_i strictement négatifs sont toujours les mêmes. L'entier k correspond à $d - \dim(\text{Ker}(q))$. Le couple (k_+, k_-) est appelé signature de q .

On note aussi la signature sous la forme $(+, \dots, +, -, \dots, -, 0, \dots)$, avec k_+ signes $+$, k_- signes $-$ et $n-k$ zéros. Ainsi, la forme quadratique de Lorentz $q(x, y, z, t) = x^2 + y^2 + z^2 - c^2 t^2$ est de signature $(3, 1)$ ou $(+, +, +, -)$ sur \mathbb{R}^4 . Une forme quadratique positive n'aura que

des + et des 0 dans sa signature et une forme quadratique définie positive n'aura que des +.

Exemple : Nous allons tenter une autre réduction de la forme quadratique $q(x, y, z) = x^2 + 2xy - yz$ déjà vue ci-dessus. Pour cela, on utilise le terme yz comme pivot et on a

$$q(x, y, z) = x^2 - y(z - 2x) = x^2 - \frac{1}{4}(y + z - 2x)^2 + \frac{1}{4}(y - z + 2x)^2 .$$

On a bien une réduction différente, mais la signature est la même.

3.3 Inégalité de Cauchy-Schwarz

Nous allons démontrer une inégalité qui sera très importante quand la suite du cours quand on considèrera des produits scalaires.

Théorème 1.17. Inégalité de Cauchy-Schwarz

Soit E un espace vectoriel réel et φ une forme bilinéaire symétrique positive sur E associée à la forme quadratique q . Alors,

$$\forall (x, y) \in E \times E , \quad |\varphi(x, y)| \leq \sqrt{q(x)}\sqrt{q(y)} .$$

Si en outre, φ est définie, alors l'inégalité a lieu si et seulement si x et y sont colinéaires.

Démonstration : Soit x et y donnés. On considère pour tout $t \in \mathbb{R}$

$$P(t) = \varphi(tx + y, tx + y) = t^2\varphi(x, x) + 2t\varphi(x, y) + \varphi(y, y) .$$

Comme φ est positive, le polynôme du second degré $P(t)$ doit avoir un discriminant négatif. C'est-à-dire qu'on doit avoir $4\varphi(x, y)^2 - 4\varphi(x, x)\varphi(y, y) \leq 0$, ce qui conduit à l'inégalité voulue. Par ailleurs, s'il y a égalité, alors $P(t)$ a une racine double et donc il existe $t \in \mathbb{R}$ tel que $\varphi(tx + y, tx + y) = 0$. Si φ est définie, cela implique $tx + y = 0$ et donc $y = -tx$. \square

On peut faire une autre démonstration¹, moins astucieuse mais exploitant les invariances du problème. On commence par constater que, pour tout x et y , $\varphi(x - y, x - y) \geq 0$ entraîne que

$$\varphi(x, y) \leq \frac{1}{2}\varphi(x, x) + \frac{1}{2}\varphi(y, y) . \tag{1.1}$$

Comme le membre de gauche de (1.1) est invariant par changement $x \mapsto -x$ alors que le membre de droite est changé en son opposé, on peut dire que

$$|\varphi(x, y)| \leq \frac{1}{2}\varphi(x, x) + \frac{1}{2}\varphi(y, y) . \tag{1.2}$$

1. Que je connais par le blog de Terence Tao. Je ne sais pas si elle est de lui.

Puis, on utilise l'invariance $(x, y) \mapsto (\lambda x, \lambda^{-1}y)$ du membre de gauche pour avoir que, pour tout x et y dans E et pour tout $\lambda \in \mathbb{R}^*$,

$$|\varphi(x, y)| \leq \frac{\lambda^2}{2}\varphi(x, x) + \frac{1}{2\lambda^2}\varphi(y, y) . \quad (1.3)$$

Il ne reste plus qu'à trouver le λ optimal, c'est-à-dire pour lequel le second membre est le plus petit, et on retombe sur l'inégalité de Cauchy-Schwarz.

Chapitre 2 : Espaces euclidiens et préhilbertiens

Dans ce chapitre, on va introduire une notion générale de produit scalaire. On pourra ainsi faire de la géométrie dans des espaces généraux, même de dimension infinie, et utiliser les projections orthogonales pour obtenir des approximations faciles à calculer.

1 Produits scalaires et normes

Définition 2.1. Soit E un espace vectoriel réel. Un **produit scalaire** sur E est une forme bilinéaire symétrique définie positive, c'est-à-dire que l'application $(x, y) \in E^2 \mapsto \langle x|y \rangle \in \mathbb{R}$ doit vérifier :

1. pour tout x, y dans E , $\langle x|y \rangle = \langle y|x \rangle$
2. pour tout x, x', y dans E et pour tout λ et μ réel $\langle \lambda x + \mu x'|y \rangle = \lambda \langle x|y \rangle + \mu \langle x'|y \rangle$.
3. pour tout $x \in E$, $\langle x|x \rangle \geq 0$
4. si $\langle x|x \rangle = 0$ alors $x = 0$.

Exemples :

- Dans \mathbb{R}^d , $(x, y) \mapsto \langle x|y \rangle = x_1y_1 + x_2y_2 + \dots + x_dy_d$ est un produit scalaire.
- Dans \mathbb{R}^{3+1} , $(x, y) \mapsto x_1y_1 + x_2y_2 + x_3y_3 - c^2x_4y_4$ n'est pas un produit scalaire (la forme n'est pas définie positive).
- Dans $\mathcal{C}^0([0, 1], \mathbb{R})$, $(f, g) \mapsto \int_0^1 f(x)g(x)dx$ est un produit scalaire.

Exercice 2.2. Démontrer les assertions ci-dessus.

Définition 2.3. Soit E un espace vectoriel réel. Une **norme** sur E est une application $\| \cdot \|$ de E dans \mathbb{R} qui vérifie, pour tout $x \in E$,

1. $\|x\| \geq 0$
2. $\|x\| = 0$ si et seulement si $x = 0$.
3. pour tout $\lambda \in \mathbb{R}$, $\|\lambda x\| = |\lambda| \|x\|$.
4. pour tout x et y dans E , $\|x + y\| \leq \|x\| + \|y\|$ (inégalité triangulaire).

Exemples :

- Dans \mathbb{R}^d , $x \mapsto |x_1| + |x_2| + \dots + |x_d|$, $x \mapsto \sqrt{|x_1|^2 + |x_2|^2 + \dots + |x_d|^2}$ et $x \mapsto \max_{i=1, \dots, d} |x_i|$ sont trois normes, appelées respectivement norme ℓ^1 , norme ℓ^2 et norme ℓ^∞ .
- Dans $\mathcal{C}^0([0, 1], \mathbb{R})$, $f \mapsto \int_0^1 |f(x)| dx$, $f \mapsto \int_0^1 |f(x)|^2 dx$ et $f \mapsto \max_{x \in [0, 1]} |f(x)|$ sont trois normes, appelées respectivement norme L^1 , norme L^2 et norme L^∞ .

Exercice 2.4. *Démontrer les assertions ci-dessus.*

Une norme sur E définit une distance sur E par $d(x, y) = \|x - y\|$ (d'où le nom d'inégalité triangulaire pour la dernière propriété). Le choix de la norme va donc influencer ce qu'on appelle la topologie de E , par exemple la notion de proximité, de limite... On note aussi qu'une norme est une application continue.

Proposition 2.5. *Si $\|\cdot\|$ est une norme sur E , alors*

$$\forall (x, y) \in E^2, \quad \left| \|x\| - \|y\| \right| \leq \|x - y\|. \quad (2.1)$$

En particulier, la norme est une fonction continue de E (muni de la distance associée à la norme) dans \mathbb{R} .

Démonstration : L'inégalité triangulaire implique que

$$\|x\| \leq \|x - y\| + \|y\| \quad \text{et} \quad \|y\| \leq \|y - x\| + \|x\|.$$

Comme $\|y - x\| = \|(-1)(x - y)\| = |-1| \|x - y\| = \|x - y\|$, la combinaison des deux inégalités donne bien l'inégalité (2.1). La continuité vient que, pour tout $\varepsilon > 0$, si $\|x - y\| = d(x, y) < \varepsilon$, alors (2.1) montre que la distance dans \mathbb{R} entre $\|x\|$ et $\|y\|$ est plus petite que ε . \square

Dans la suite, nous allons en fait nous consacrer uniquement aux normes liées à un produit scalaire de la manière suivante.

Proposition 2.6. *Soit $\langle \cdot | \cdot \rangle$ un produit scalaire sur un espace vectoriel réel E . Alors $\|x\| = \sqrt{\langle x | x \rangle}$ définit une norme sur E .*

Démonstration : Il suffit de vérifier les hypothèses de la définition. Le point le plus délicat est l'inégalité triangulaire. Soit x et y dans E . On a

$$\begin{aligned} \|x + y\|^2 &= \langle x + y | x + y \rangle = \langle x | x \rangle + \langle x | y \rangle + \langle y | x \rangle + \langle y | y \rangle = \|x\|^2 + \|y\|^2 + 2\langle x | y \rangle \\ &\leq \|x\|^2 + \|y\|^2 + 2\|x\| \cdot \|y\| = (\|x\| + \|y\|)^2, \end{aligned}$$

où on a utilisé l'inégalité de Cauchy-Schwarz (théorème 1.17). \square

On note que la norme associée à un produit scalaire est simplement la racine carrée de la forme quadratique. On peut donc retrouver le produit scalaire à partir de la norme par la formule de polarisation.

Résumé : en pratique, nous allons voir dans ce cours principalement deux produits scalaires et leur normes associées :

- le produit scalaire usuel $\langle x|y \rangle = x_1y_1 + x_2y_2 + \dots + x_dy_d$ sur \mathbb{R}^d . On pourra se raccrocher à ce cas pour mieux « voir » ce qu'il se passe.
- le produit scalaire L^2 $\langle f|g \rangle = \int_0^1 f(x)g(x)dx$ sur les espaces de fonctions. Ce produit scalaire est parfaitement adapté aux approximations de fonctions, en particulier pour les séries de Fourier.

En récapitulant, outre les propriétés du produit scalaire et de sa norme venant de leur définitions, on a la liste de propriétés suivantes. Soit E un espace vectoriel muni d'un produit scalaire $\langle \cdot | \cdot \rangle$ et soit $\| \cdot \|$ la norme associée. On a :

1. l'inégalité triangulaire $\|x + y\| \leq \|x\| + \|y\|$,
2. l'inégalité de Cauchy-Schwarz $|\langle x|y \rangle| \leq \|x\| \cdot \|y\|$,
3. le théorème de Pythagore $\|x + y\|^2 = \|x\|^2 + \|y\|^2$ si $x \perp y$,
4. l'identité du parallélogramme $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$.

On voit aussi que la notion de produit scalaire nous donne une notion d'angle en posant que l'angle (non signé) θ entre x et y est défini par $\cos \theta = \frac{\langle x|y \rangle}{\|x\| \cdot \|y\|}$.

2 Espaces euclidiens et bases orthonormales

Un *espace euclidien* est un espace vectoriel réel E de dimension finie et muni d'un produit scalaire $\langle \cdot | \cdot \rangle$ (et donc de la norme $\|x\| = \sqrt{\langle x|x \rangle}$ associée). Par exemple, \mathbb{R}^3 muni du produit scalaire usuel est un espace euclidien. Mais l'espace $\mathbb{R}_2[X]$ des polynômes de degré 2, muni du produit scalaire $\langle P|Q \rangle = \int_0^1 P(x)Q(x)dx$ est aussi un espace euclidien et

$$\langle aX^2 + bX + c | a'X^2 + b'X + c' \rangle = \frac{1}{5}aa' + \frac{1}{4}(ab' + ba') + \frac{1}{3}(ac' + ca' + bb') + \frac{1}{2}(cb' + bc') + cc'$$

ne correspond pas au produit scalaire canonique de \mathbb{R}^3 dans la base $(1, X, X^2)$.

Définition 2.7. Une base (e_1, \dots, e_d) de E est dite *orthogonale* si $\langle e_i | e_j \rangle = 0$ pour tout $i \neq j$. La base est dite *orthonormale* si en outre $\|e_i\| = 1$ pour tout i .

La base canonique de \mathbb{R}^d est bien orthonormale pour le produit scalaire canonique, mais la base $(1, X, X^2)$ de $\mathbb{R}_2[X]$ n'est même pas orthogonale pour le produit scalaire L^2 car

$$\langle 1|X \rangle = 1/2.$$

Le grand intérêt des bases orthonormales des espaces euclidiens est la conséquence suivante du théorème de Pythagore. Celle-ci dit, dans un certain sens, que tout espace euclidien se ramène à \mathbb{R}^d dans une base bien choisie.

Proposition 2.8. *Soit E un espace euclidien et (e_1, \dots, e_d) une base orthonormale de E . Alors, pour tout x et y dans E qui se décomposent en $x = x_1e_1 + x_2e_2 + \dots + x_de_d$ et $y = y_1e_1 + y_2e_2 + \dots + y_de_d$, on a*

$$\langle x|e_i \rangle = x_i \quad , \quad \langle x|y \rangle = x_1y_1 + \dots + x_dy_d \quad \text{et} \quad \|x\| = \sqrt{x_1^2 + \dots + x_d^2}.$$

On notera aussi le critère pratique suivant.

Proposition 2.9. *Soit E un espace euclidien de dimension d et soit (e_1, \dots, e_d) une famille de d vecteurs non nuls vérifiant $\langle e_i|e_j \rangle = 0$ pour tout $i \neq j$ (on parle de famille orthogonale). Alors, (e_1, \dots, e_d) est une base orthogonale de E .*

Démonstration : On a déjà une famille de d vecteurs pour un espace de dimension d . Il suffit donc de montrer que la famille est libre. Soient λ_i des réels tels que $\lambda_1e_1 + \dots + \lambda_de_d = 0$. En prenant le produit scalaire par e_i , on trouve

$$0 = \langle \lambda_1e_1 + \dots + \lambda_de_d|e_i \rangle = \lambda_1\langle e_1|e_i \rangle + \dots + \lambda_d\langle e_d|e_i \rangle = \lambda_i\|e_i\|^2.$$

Comme e_i est non nul, $\|e_i\| \neq 0$ et λ_i est nul. On peut faire ce calcul pour tout i et donc tous les λ_i sont nuls. \square

L'existence de bases orthonormales est assurée par le procédé constructif suivant.

Théorème 2.10. Procédé d'orthonormalisation de Gram-Schmidt

Soit (e_1, e_2, \dots, e_d) une base d'un espace euclidien E . Alors, il existe une base orthonormale $(\tilde{e}_1, \tilde{e}_2, \dots, \tilde{e}_d)$ de E telle que, pour tout i , $\text{vect}(\tilde{e}_1, \dots, \tilde{e}_i) = \text{vect}(e_1, \dots, e_i)$.

Démonstration : La démonstration se fait par récurrence sur i en construisant petit à petit la base orthonormale par le procédé suivant. On commence par poser $\tilde{e}_1 = e_1/\|e_1\|$. On note que

$$\|\tilde{e}_1\| = \left\| \frac{e_1}{\|e_1\|} \right\| = \frac{1}{\|e_1\|} \|e_1\| = 1.$$

Puis si on suppose $\tilde{e}_1 \dots \tilde{e}_i$ déjà construits, on pose

$$e'_{i+1} = e_{i+1} - \langle e_{i+1}|e_1 \rangle e_1 - \dots - \langle e_{i+1}|e_i \rangle e_i \quad \text{puis} \quad \tilde{e}_{i+1} = e'_{i+1}/\|e'_{i+1}\|.$$

Par construction \tilde{e}_{i+1} est bien de norme 1 et $\text{vect}(\tilde{e}_1, \dots, \tilde{e}_{i+1}) = \text{vect}(e_1, \dots, e_{i+1})$. Il reste à vérifier que e'_{i+1} (et donc aussi \tilde{e}_{i+1}) est bien orthogonal aux \tilde{e}_j avec $j = 1 \dots i$. Pour cela, il suffit de voir que

$$\langle e'_{i+1} | \tilde{e}_j \rangle = \langle e_{i+1} - \langle e_{i+1} | e_1 \rangle e_1 - \dots - \langle e_{i+1} | e_i \rangle e_i | \tilde{e}_j \rangle = \langle e_{i+1} | \tilde{e}_j \rangle - \langle e_{i+1} | \tilde{e}_j \rangle \langle \tilde{e}_j | \tilde{e}_j \rangle = 0$$

car $\langle e_k | e_j \rangle = 0$ si $k \neq j$ et $\langle e_j | e_j \rangle = \|e_j\|^2 = 1$. \square

Le procédé d'orthonormalisation de Gram-Schmidt peut se résumer à trois étapes :

- i) Projeter le vecteur sur l'orthogonal à l'espace engendré par les vecteurs de base déjà traités.
- ii) Normaliser le vecteur obtenu.
- iii) Remplacer l'ancien vecteur de base par le nouveau vecteur orthonormalisé et passer au vecteur suivant.

Pour finir cette partie, nous allons réaliser deux exemples d'orthonormalisation.

Exemples :

- On se place dans \mathbb{R}^3 muni du produit scalaire canonique. On considère la base

$$(e_1 = (2, 0, 0) , e_2 = (3, -3, 4) , e_3 = (-1, 1, -3)) .$$

La normalisation du premier vecteur donne $\tilde{e}_1 = e_1/2 = (1, 0, 0)$. Puis, on rend e_2 orthogonal à \tilde{e}_1 en posant

$$e'_2 = e_2 - \langle e_2 | \tilde{e}_1 \rangle \tilde{e}_1 = (3, -3, 4) - 3(1, 0, 0) = (0, -3, 4) .$$

Le vecteur e'_2 est de norme 5, donc on obtient $\tilde{e}_2 = (0, -3/5, 4/5)$. Enfin, on s'occupe du dernier vecteur :

$$\begin{aligned} e'_3 &= e_3 - \langle e_3 | \tilde{e}_1 \rangle \tilde{e}_1 - \langle e_3 | \tilde{e}_2 \rangle \tilde{e}_2 \\ &= (-1, 1, -3) - (-1)(1, 0, 0) - (-3/5 - 12/5)(0, -3/5, 4/5) \\ &= (0, -4/5, -3/5) \end{aligned}$$

Un peu miraculeusement, on a déjà que $\|e'_3\| = 1$ et donc $\tilde{e}_3 = e'_3$. La base orthonormale obtenue est donc

$$(\tilde{e}_1 = (1, 0, 0) , \tilde{e}_2 = (0, -3/5, 4/5) , \tilde{e}_3 = (0, -4/5, -3/5)) .$$

- On considère l'espace $\mathbb{R}_2[X]$ des polynômes de degré 2, muni du produit scalaire $\langle P | Q \rangle = \int_0^1 P(x)Q(x)dx$. On souhaite orthonormaliser la base $(1, X, X^2)$. Le vecteur $e_1 = 1$ est déjà de norme 1. L'orthogonalisation du vecteur $e_2 = X$ s'écrit

$$e'_2 = e_2 - \langle e_2 | \tilde{e}_1 \rangle \tilde{e}_1 = X - \int_0^1 1x dx = X - \frac{1}{2} .$$

Puis on calcule

$$\|e'_2\|^2 = \int_0^1 \left(x - \frac{1}{2}\right)^2 dx = 1/12 ,$$

et donc $\tilde{e}_2 = 2\sqrt{3}(X - \frac{1}{2})$. Passons à $e_3 = X^2$

$$\begin{aligned} e'_3 &= e_3 - \langle e_3 | \tilde{e}_1 \rangle \tilde{e}_1 - \langle e_3 | \tilde{e}_2 \rangle e_2 \\ &= X^2 - \int_0^1 1x^2 dx - 2\sqrt{3}\left(X - \frac{1}{2}\right) \int_0^1 2\sqrt{3}x^2\left(x - \frac{1}{2}\right) dx \\ &= X^2 - \frac{1}{3} - \left(X - \frac{1}{2}\right) = X^2 - X + \frac{1}{6} . \end{aligned}$$

Le calcul de la norme de e'_3 donne $\|e'_3\|^2 = \int_0^1 \left(x^2 - x + \frac{1}{6}\right)^2 dx = \frac{1}{180}$ et la base orthonormale est donc

$$\left(1 , \sqrt{12}\left(X - \frac{1}{2}\right) , 6\sqrt{5}\left(X^2 - X + \frac{1}{6}\right) \right) .$$

On notera que, par construction, \tilde{e}_3 est orthogonal à tous les polynômes de degré 1, ce que l'on peut vérifier :

$$\begin{aligned} \langle X^2 - X + \frac{1}{6} | aX + b \rangle &= \int_0^1 (ax + b)\left(x^2 - x + \frac{1}{6}\right) dx \\ &= \int_0^1 ax^3 + (b - a)x^2 + \left(\frac{a}{6} - b\right)x + \frac{b}{6} dx \\ &= \frac{a}{4} + \frac{b - a}{3} + \frac{a/6 - b}{2} + b/6 = 0 . \end{aligned}$$

3 Espaces préhilbertiens et approximations

Dans cette partie, nous allons travailler dans un espace préhilbertien.

Définition 2.11. *Un espace préhilbertien est un espace vectoriel réel E muni d'un produit scalaire $\langle \cdot | \cdot \rangle$ et de sa norme associée $\| \cdot \|$.*

La différence avec les espaces euclidiens réside donc seulement dans le fait qu'un espace préhilbertien peut être (mais pas obligatoirement) de dimension infinie. On pourra aussi penser à un espace de dimension finie mais très grande.

Exemples :

- Les espaces euclidiens.
- L'espace des polynômes $\mathbb{R}[X]$ muni du produit scalaire $\langle P | Q \rangle = \int_0^1 P(x)Q(x)dx$.

- L'espace $\mathcal{C}^0([0, 1], \mathbb{R})$ des fonctions continues sur $[0, 1]$ muni du produit scalaire $\langle f|g \rangle = \int_0^1 f(x)g(x)dx$.
- L'espace $\mathcal{C}^1([0, 1], \mathbb{R})$ des fonctions de classe \mathcal{C}^1 sur $[0, 1]$ muni du produit scalaire $\langle f|g \rangle = \int_0^1 f'(x)g'(x) + f(x)g(x)dx$.

Le concept principal de cette partie sera l'approximation par projection orthogonale sur un sous-espace de dimension finie. L'intérêt est de remplacer au mieux une description ayant besoin d'un nombre très grand, voire infini, de dimensions, par une description très correcte en quelques nombres. Il faut aussi remarquer que dans un espace préhilbertien E , la base peut être non dénombrable et donc que le procédé d'orthonormalisation de Gram-Schmidt n'est pas forcément possible. L'idée de travailler dans des sous-espaces de dimension finie permet de réintroduire des bases orthonormales.

Théorème 2.12. Projection orthogonale

Soit E un espace préhilbertien et soit F un sous-espace de dimension finie de E , dont (e_1, \dots, e_d) est une base orthonormale.

Alors, pour tout $x \in E$, les trois définitions suivantes de Px sont équivalentes.

- i) Px est l'unique vecteur de F tel que $x - Px$ est orthogonal à F .
- ii) Px est l'unique vecteur de F tel que $\|x - Px\| = \min_{y \in F} \|x - y\|$.
- iii) $Px = \langle x|e_1 \rangle e_1 + \dots + \langle x|e_d \rangle e_d$.

Démonstration : Il est clair que la dernière définition correspond à un unique vecteur Px dans F . Ensuite, le Px de iii) vérifie $\langle x - Px|e_i \rangle = \langle x|e_i \rangle - \langle x|e_i \rangle \langle e_i|e_i \rangle = 0$ et donc vérifie la description i). Pour conclure que i) et iii) sont équivalentes, il suffit donc de montrer qu'il n'existe pas plus d'un vecteur y de F tel que $x - y$ est orthogonal à F . Imaginons qu'il existe un autre vecteur y' avec cette propriété. Alors $y - y'$ est dans F et est égal à $(x - y') - (x - y)$ donc est orthogonal à F . Ceci n'est possible que si $y - y' = 0$, c'est-à-dire $y = y'$.

Supposons que Px vérifie $\|x - Px\| = \min_{y \in F} \|x - y\|$. Alors, pour tout $y \in F$ et tout $\lambda \in \mathbb{R}$, on a $\|x - (Px + \lambda y)\|^2 = \|x - Px\|^2 + 2\lambda \langle x - Px|y \rangle + |\lambda|^2 \|y\|^2$. Quand λ tend vers 0, on trouve que l'on doit avoir $\lambda \langle x - Px|y \rangle \geq 0$ pour que Px réalise bien le minimum $\min_{y \in F} \|x - y\|$. Comme cela est vrai pour λ positif comme négatif, cela implique que $\langle x - Px|y \rangle = 0$ pour tout $y \in F$ et donc que $x - Px$ est orthogonal à F . Supposons maintenant que $Px \in F$ vérifie que $x - Px$ est orthogonal à F . Soit $y \in F$, on a $\|x - y\|^2 = \|x - Px + Px - y\|^2 = \|x - Px\|^2 + \|Px - y\|^2$ par le théorème de Pythagore et donc $\|x - y\| \geq \|x - Px\|$ ce qui montre ii). Comme l'on sait déjà que Px est unique dans la définition i), l'équivalence ci-dessus justifie l'unicité dans la définition ii). \square

Nous allons montrer deux exemples d'utilisation de cette projection orthogonale.

- **Régression linéaire.** On mesure une variable $x(t)$ sur une suite de temps t_1, \dots, t_n , avec n assez grand, et on trouve des valeurs x_1, \dots, x_n . On sait que, en théorie, $x(t)$ est une droite $at + b$ et on aimerait retrouver a et b à partir des mesures. Comme il y a des erreurs et imprécisions sur les mesures, les points (t_n, x_n) ne sont pas vraiment sous la forme d'une droite. On aimerait trouver « la meilleure droite possible approchant les points (t_n, x_n) », mais il faudrait savoir ce qu'on entend par là : la droite $at + b$ minimisant le plus grand écart entre $at_i + b$ et x_i ? Celle minimisant l'écart moyen entre les $at_i + b$ et x_i ? Ces deux critères correspondent respectivement à minimiser la distance au sens des normes ℓ^∞ et ℓ^1 . Cela pourrait se faire, mais il est bien plus pratique de chercher à minimiser la distance ℓ^2 car elle provient d'un produit scalaire et le théorème de projection orthogonal est à notre disposition. C'est une *méthode des moindres carrées*, puisqu'on cherche à minimiser la somme des carrés des écarts. Prenons le formalisme suivant : $E = \mathbb{R}^n$ est muni du produit scalaire canonique et on pose $x = (x_n)$. On considère le sous-espace F de E engendré par les vecteurs $\mathbb{1} = (1, 1, \dots, 1)$ et $T = (t_1, \dots, t_n)$. On souhaite trouver a et b tel que $\|x - aT + b\mathbb{1}\|$ soit minimum. Il suffit d'orthonormaliser la base $(\mathbb{1}, T)$ en $(\tilde{e}_1, \tilde{e}_2)$ puis de calculer la projection orthogonale de x sur F à l'aide de cette base. En repassant à la base $(\mathbb{1}, T)$, on obtient les coefficients a et b . Notons que tous ces calculs, y compris le procédé d'orthonormalisation de Gram-Schmidt sont parfaitement algorithmiques et programmables. En fait, on peut montrer les formules

$$a = \frac{\sum_i (t_i - \bar{t})(x_i - \bar{x})}{\sum_i (t_i - \bar{t})^2} = \frac{\frac{1}{n} \sum_i t_i x_i - \bar{t} \bar{x}}{\frac{1}{n} \sum_i t_i^2 - \bar{t}^2} = \frac{\text{covariance}(x, t)}{\text{variance}(t)} \quad b = \bar{x} - a\bar{t}$$

où $\bar{t} = \frac{1}{n} \sum_i t_i$ et $\bar{x} = \frac{1}{n} \sum_i x_i$ sont les moyennes des (t_i) et des (x_i) .

La figure 2.1 montre un exemple d'application du procédé avec 100 points.

Un autre outil statistique qui vient de la géométrie des espaces préhilbertiens est la corrélation linéaire. On considère les données (t_i) et (x_i) et leur moyennes \bar{t} et \bar{x} . Il existe une relation linéaire du type $x = at + b$ si et seulement si les vecteurs $(t_i - \bar{t})$ et $(x_i - \bar{x})$ sont colinéaires. Pour tester cette relation, on calcule le cosinus de l'angle entre ces vecteurs qui est donné par

$$\cos \alpha = \frac{\langle x - \bar{x} | t - \bar{t} \rangle}{\|x - \bar{x}\| \|t - \bar{t}\|} = \frac{\sum_i (x_i - \bar{x})(t_i - \bar{t})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (t_i - \bar{t})^2}}.$$

Ce nombre est appelé *corrélation linéaire* et en général noté r . S'il est proche de ± 1 , alors il existe une relation affine entre x et t . S'il est proche de 0, cette relation n'existe (probablement) pas. Ce nombre r permet donc de mesurer la pertinence de la régression linéaire.

- **Approximation polynômiale.** On souhaite approcher de la meilleure façon possible la fonction $x \mapsto e^{-x}$ sur $[0, 1]$ par un polynôme de degré 2. Encore une fois, on pourrait donner plusieurs notions pour « approcher de la meilleure façon possible ». La plus simple à manier est celle de la norme L^2 : on veut trouver a, b et c tels que $\int_0^1 (e^{-x} - (ax^2 + bx + c))^2 dx$ soit minimal. On se place donc dans $\mathcal{C}^0([0, 1], \mathbb{R})$ muni du produit scalaire L^2 et on note F le sous-espace des polynômes de degré deux. On a déjà une base orthonormale de F (voir ci-dessus). Il suffit donc de calculer :

$$\langle e^{-x} | \tilde{e}_1 \rangle = \int_0^1 e^{-x} \cdot 1 dx = 1 - 1/e .$$

$$\langle e^{-x} | \tilde{e}_2 \rangle = \int_0^1 e^{-x} 2\sqrt{3} \left(x - \frac{1}{2}\right) dx = \sqrt{3}(1 - 3/e)$$

$$\langle e^{-x} | \tilde{e}_3 \rangle = \int_0^1 e^{-x} 6\sqrt{5} \left(x^2 - x + \frac{1}{6}\right) dx = \sqrt{5}(7 - 19/e) .$$

On trouve donc l'approximation suivante de l'exponentielle sur $[0, 1]$

$$e^{-x} \simeq (210 - 570/e)x^2 + (552/e - 204)x + (33 - 87/e) .$$

La figure 2.1 montre que cette approximation est très bonne sur $[0, 1]$. En outre, si on travaille avec des calculs approchés, tous les calculs précédents se programment très facilement sur ordinateur. On peut donc remplacer les courbes de fonctions continues par des courbes polynômiales qui ne sont représentées que par un nombre très petit de paramètres. Ainsi, les courbes de Bézier sont des courbes codées par des polynômes et qui servent dans les dessins assistés par ordinateur, pour coder les polices de caractères etc.

Concernant l'approximation par des polynômes, il vient naturellement la question suivante : si P_d est la projection de la fonction $f(x) = e^{-x}$ sur les polynômes de degré d , est-ce que la suite P_d converge vers e^{-x} pour la norme L^2 ? Il est clair que l'erreur $\|f - P_d\|$ diminue quand d augmente, mais tend-elle vers 0 quand d tend vers $+\infty$? C'est le point intéressant et difficile de ces problèmes d'approximations. Pour la culture mathématique, voici un théorème qui répond à cette question.

Théorème 2.13. Weierstraß

Soit f une fonction réelle continue sur un segment $[a, b]$. Alors, pour tout $\varepsilon > 0$, il existe un polynôme P_ε tel que $\sup_{x \in [a, b]} |f(x) - P_\varepsilon(x)| \leq \varepsilon$.

En particulier, le polynôme P_ε du théorème ci-dessus vérifie

$$\|f - P_\varepsilon\| = \sqrt{\int_a^b |f(x) - P_\varepsilon(x)|^2 dx} \leq \sqrt{b-a} \varepsilon ,$$

et donc P_ε tend vers f quand ε tend vers 0.

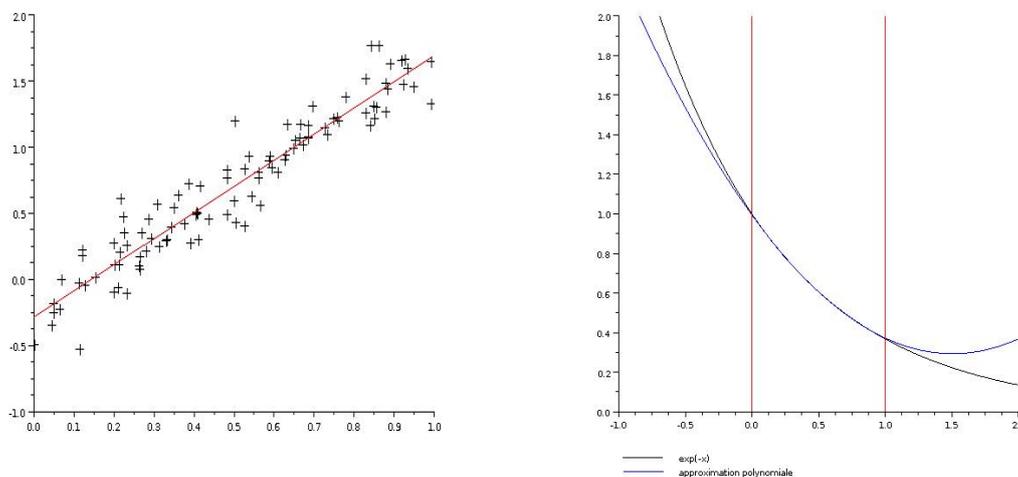


FIGURE 2.1 – A gauche, un exemple de régression linéaire avec 100 points de mesure. A droite, la courbe $x \mapsto e^{-x}$ et son approximation par un polynôme de degré 2 sur $[0, 1]$. On voit bien que l'approximation est excellente sur $[0, 1]$ mais devient mauvaise en dehors.

4 Un mot sur les espaces de Hilbert

Si on parle d'espaces préhilbertiens, c'est qu'il existe des espaces de Hilbert. Ceux-ci sont des espaces préhilbertiens qui ont la propriété supplémentaire d'être *complets*. Nous avons vu plus haut le problème de savoir si une suite d'approximations P_d d'un point f donné tend vers f . En quelque sorte, la complétude signifie que toute suite (P_d) qui est raisonnablement une suite d'approximations (le critère exact est qu'elle est une suite dite *de Cauchy*) doit converger vers quelque chose.

On a par exemple le parallèle suivant. On peut approcher le nombre π par des approximations 3, 3, 1, 3, 14, etc. Mais si on se donne une suite d'approximations 2, 2, 3, 2, 38, 2, 381, etc., cette suite converge-t-elle vers un nombre réel? La réponse est naturellement oui et est liée au fait que \mathbb{R} est complet. Notons que \mathbb{Q} ne l'est pas car la suite d'approximations rationnelles de π ne converge pas vers un rationnel.

Tout espace euclidien comme \mathbb{R}^d est complet et est donc un espace de Hilbert. Mais, nous venons de voir que l'espace des polynômes $\mathbb{R}[X]$ muni du produit scalaire L^2 n'est pas complet car une suite de polynômes peut converger vers e^{-x} qui n'est pas un polynôme. En fait, même $\mathcal{C}^0([0, 1], \mathbb{R})$ muni du produit scalaire L^2 n'est pas complet (son complété est noté $L^2([0, 1], \mathbb{R})$). Quand nous ferons des séries de Fourier, nous serons donc dans un cadre préhilbertien : nous ferons des approximations de fonctions continues par morceaux par des séries de Fourier, mais il existera des séries de Fourier bien définies dont la limite ne sera pas une fonction continue par morceaux.

Chapitre 3 : Séries

Les scientifiques se sont longtemps demandé quel sens donner à une somme infinie. Dans le paradoxe de Zénon d'Elée (V^{ème} avant JC), un caillou est lancé sur un arbre et parcourt la moitié de la distance, puis la moitié de la moitié restante, puis la moitié de ce qui reste. . . Il semble ainsi n'arriver jamais à destination car il lui faut franchir une infinité d'étapes et chacune lui demande un temps non nul. Aujourd'hui, on résoud le paradoxe en écrivant qu'une somme infinie de temps peut très bien donner un temps fini. Ce paradoxe se transcrit exactement par le fait que $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \dots = 1$.

A l'inverse, certaines sommes infinies, même d'éléments de plus en plus petits, peuvent donner un résultat infini. Ainsi, la somme $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \dots$ diverge vers $+\infty$ au fur et à mesure que l'on ajoute des fractions de plus en plus petites.

Le but de ce chapitre est de fournir quelques notions et résultats de base sur les séries, c'est-à-dire les sommes infinies de termes.

1 Introduction

Soit E un espace vectoriel muni d'une norme $\| \cdot \|$. On appelle *série* une suite écrite sous la forme $S_N = \sum_{n=0}^N u_n$. Le terme $u_n \in E$ est appelé le *terme général* de la série et $S_N = \sum_{n=0}^N u_n$ est appelé une *somme partielle*.

Définition 3.1. *On dit que la série $\sum_{n=0}^{\infty} u_n$ est convergente si la suite S_N a une limite $S \in E$, c'est-à-dire que $\lim_{n \rightarrow \infty} \|S_N - S\| = 0$. On appelle alors S la somme de la série et on la note $S = \sum_{n=0}^{+\infty} u_n$. Une série qui ne converge pas dans E est dite divergente.*

Des propriétés élémentaires des limites, on déduit la proposition suivante.

Proposition 3.2. *Soit $\lambda \in \mathbb{R} \setminus \{0\}$, les séries $\sum u_n$ et $\sum \lambda u_n$ ont même nature et si elles convergent alors $\sum_{n \geq 0} \lambda u_n = \lambda \sum_{n \geq 0} u_n$.*

Soient $\sum u_n$ et $\sum v_n$ deux séries convergentes, alors $\sum (u_n + v_n)$ converge et a pour somme $\sum_{n \geq 0} (u_n + v_n) = \sum_{n \geq 0} u_n + \sum_{n \geq 0} v_n$.

Pour tous rangs n_1 et n_2 , les séries $\sum_{n \geq n_1} u_n$ et $\sum_{n \geq n_2} u_n$ ont même nature.

Nous n'allons nous concentrer ici que sur les critères de convergence des séries réelles à termes positifs. En effet, on obtiendra alors la convergence des séries plus générales par le résultat admis suivant.

Proposition 3.3. Soit $\sum_{n \geq 0} u_n$ une série dans \mathbb{R} , resp. \mathbb{C} , resp. un espace euclidien E , telle que $\sum_{n \geq 0} |u_n|$, resp. $\sum_{n \geq 0} |u_n|$, resp. $\sum_{n \geq 0} \|u_n\|$, soit une série convergente dans \mathbb{R} . Alors $\sum_{n \geq 0} u_n$ converge dans \mathbb{R} , \mathbb{C} , resp. E et on dit que $\sum_{n \geq 0} u_n$ converge en valeur absolue, resp. en module, resp. en norme.

On a le critère de divergence suivant.

Proposition 3.4. Soit $\sum_{n \geq 0} u_n$ une série dans E . Si $\sum_{n \geq 0} u_n$ converge alors (u_n) tend vers 0. Autrement dit, si (u_n) ne tend pas vers 0 alors $\sum_{n \geq 0} u_n$ diverge.

Démonstration : Supposons que $S_N = \sum_{n=0}^N u_n$ tende vers une limite S . Alors $u_n = S_n - S_{n-1}$ tend vers $S - S = 0$. \square

Il est important de garder en mémoire que la réciproque n'est pas vraie comme le montre la divergence de $\sum_{n \geq 1} 1/n$ que nous verrons plus loin.

Considérons notre premier exemple typique de séries : les séries géométriques. Soit $a \in \mathbb{C}$ différent de 1, on a que $(a-1)(a^n + a^{n-1} + \dots + a + 1) = a^{n+1} - 1$ et donc

$$S_N = \sum_{n=0}^N a^n = \frac{a^{N+1} - 1}{1 - a}.$$

On obtient donc que $\sum a^n$ converge si et seulement si $|a| < 1$ (pour $a = 1$, il est clair que la série diverge) et la série a alors pour somme

$$\sum_{n \geq 0} a^n = \frac{1}{1 - a}.$$

2 Critères de convergence pour les séries de termes positifs

Comme vu ci-dessus, il fait sens de ne regarder, au moins dans un premier temps, que des séries de réels positifs.

2.1 Critères de comparaison

Notons que si $S_N = \sum_{n=0}^N u_n$ est une série de termes positifs, alors S_N est croissante et seuls deux comportements sont possibles : S_N tend vers l'infini et diverge, ou S_N reste bornée et converge. On en déduit le théorème suivant.

Proposition 3.5. Soient $\sum u_n$ et $\sum v_n$ deux séries de réels positifs tels que $u_n \geq v_n$ à partir d'un certain rang. Si $\sum u_n$ converge, alors $\sum v_n$ converge et si $\sum v_n$ diverge, alors $\sum u_n$ diverge.

Démonstration : Supposons que $\sum u_n$ converge, alors la somme partielle $\sum_{n=0}^N u_n$ est majorée par la somme $S = \sum_{n \geq 0} u_n$. Donc $\sum_{n=0}^N v_n$ est une suite croissante majorée par S et donc convergente. La deuxième partie de la proposition est la contraposée de la première. \square

On a le corollaire suivant.

Corollaire 3.6. Soient $\sum u_n$ et $\sum v_n$ deux séries de réels positifs. Si $u_n \sim v_n$ alors $\sum u_n$ et $\sum v_n$ ont même nature. Si $u_n = o(v_n)$, alors $\sum u_n$ converge si $\sum v_n$ converge et $\sum v_n$ diverge si $\sum u_n$ diverge.

Démonstration : Si $u_n \sim v_n$, alors pour tout $\varepsilon > 0$, il existe un rang à partir duquel $u_n/v_n \in [1 - \varepsilon, 1 + \varepsilon]$ et donc $(1 - \varepsilon)v_n \leq u_n \leq (1 + \varepsilon)v_n$. Il suffit donc de prendre $\varepsilon = 1/2$ et d'utiliser la proposition précédente. De même, si $u_n = o(v_n)$, alors il existe un rang à partir duquel $u_n \leq v_n$. \square

2.2 Critères de D'Alembert et de Cauchy

Les deux critères suivants sont appropriés quand on considère des séries dont le terme général ressemble à un a^n avec $|a| \neq 1$.

Théorème 3.7. Critère de D'Alembert

Soit $\sum_{n \geq 0} u_n$ une série de réels non nuls. Supposons que $|u_{n+1}/u_n|$ ait une limite ℓ quand $n \rightarrow +\infty$. Si $\ell < 1$ alors $\sum_{n \geq 0} u_n$ est absolument convergente et si $\ell > 1$, alors $\sum_{n \geq 0} u_n$ est divergente.

Démonstration : Supposons que $\ell < 1$, il existe un rang n_0 tel que, pour tout $n \geq n_0$, $|u_{n+1}/u_n| \leq a$ avec $a = (\ell + 1)/2 < 1$. On a donc $|u_n| \leq a^{n-n_0}|u_{n_0}|$ pour tout $n \geq n_0$ et donc $|u_n|$ est majoré par le terme d'une série géométrique convergente. On en déduit que $\sum_{n \geq 0} |u_n|$ converge.

Supposons que $\ell > 1$, on a alors $|u_n| \geq a^{n-n_0}|u_{n_0}|$ à partir d'un certain rang, avec $a > 1$. En particulier u_n ne tend pas vers 0 et la série diverge trivialement. \square

Notons bien que l'on ne peut pas conclure quand $\ell = 1$. Le critère de Cauchy se démontre de façon similaire.

Théorème 3.8. Critère de Cauchy

Soit $\sum_{n \geq 0} u_n$ une série de réels non nuls. Supposons que $|u_n|^{1/n}$ ait une limite ℓ quand $n \rightarrow +\infty$. Si $\ell < 1$ alors $\sum_{n \geq 0} u_n$ est absolument convergente et si $\ell > 1$, alors $\sum_{n \geq 0} u_n$ est divergente.

Exemple : La série $\sum_{n \geq 0} n/2^n$ est convergente car

$$\left| \frac{u_{n+1}}{u_n} \right| = \frac{n+1}{2n} \rightarrow \frac{1}{2} < 1.$$

2.3 Séries de Riemann

Dans cette partie, nous allons voir le cas particulier des séries de Riemann qui sont les séries qui s'écrivent

$$S_n = \sum_{n=0}^N \frac{1}{k^\alpha}$$

avec $\alpha > 0$ (notons que si $\alpha \leq 0$, la série diverge trivialement car son terme général ne tend pas vers 0). Le but de ce paragraphe est de montrer le théorème suivant.

Théorème 3.9. *La série de Riemann $\sum_{n=0}^N \frac{1}{k^\alpha}$ converge si et seulement si $\alpha > 1$.*

Pour cela, on utilise le lemme suivant.

Lemme 3.10. *Soit f une fonction continue de \mathbb{R}_+^* dans \mathbb{R}_+ et décroissante. Alors la série $\sum_{n \geq 1} f(n)$ converge si et seulement si la limite $\lim_{n \rightarrow +\infty} \int_1^n f(x) dx$ est finie.*

Démonstration : Comme f est décroissante, on a

$$f(n+1) \leq \int_n^{n+1} f(x) dx \leq f(n)$$

et donc

$$\sum_{n=2}^N f(n) \leq \int_1^N f(x) dx \leq \sum_{n=1}^{N-1} f(n).$$

Quand N tend vers l'infini, on utilise l'une ou l'autre des inégalités pour montrer que soit $\sum_{n=1}^{N-1} f(n)$ tend vers l'infini, soit $\sum_{n=2}^N f(n)$ reste borné est donc converge. \square

On applique alors le lemme précédent à $f(x) = 1/x^\alpha$, ce qui suffit à démontrer le théorème 3.9.

Exemples :

- La série $\sum_{n \geq 1} 1/n^2$ converge.
- La série $\sum_{n \geq 1} 1/n$ diverge vers $+\infty$ à une vitesse logarithmique.
- La série $\sum_{n \geq 0} n/(n^2 + 1)$ diverge car $n/(n^2 + 1) \sim 1/n$ quand n tend vers $+\infty$.
- La série $\sum_{n \geq 0} \sqrt{n}/(n^2 + 1)$ converge car $\sqrt{n}/(n^2 + 1) \sim 1/n^{3/2}$ quand n tend vers $+\infty$.
- La série $\sum_{n \geq 1} \cos(n)/n^2$ converge car $|\cos(n)/n^2| \leq 1/n^2$.

3 Séries de fonctions

Les séries de Fourier sont des sommes de cosinus et sinus et donc des sommes de fonctions. Un autre exemple important de séries de fonctions est le développement en série entière. On sait que pour tout $N \in \mathbb{N}$, $e^x = 1 + x + \dots + x^N/N! + o(x^N)$, mais peut-on écrire carrément que

$$\forall x \in \mathbb{R}, \quad e^x = \sum_{n \geq 0} \frac{x^n}{n!} ? \quad (3.1)$$

Pour écrire (3.1), il y a trois problèmes à considérer :

1. Pour x fixé, la série de (3.1) a-t-elle un sens, c'est-à-dire est-elle convergente ?
2. La série de (3.1) a-t-elle un sens en tant que fonction, par exemple définit-elle bien une fonction continue quand x varie ?
3. La série obtenue est-elle bien égale à l'exponentielle ?

Le premier point peut se traiter par les théorèmes vus plus haut. Le deuxième point est le sujet du résultat suivant.

Théorème 3.11. *Soit I un intervalle de \mathbb{R} et soient (f_n) une suite de fonctions continues sur I . On suppose que $\sum_{n \geq 0} (\sup_{x \in I} |f_n(x)|)$ est une série convergente dans \mathbb{R}_+ . Alors, pour tout $x \in I$, $S(x) = \sum_{n \geq 0} f_n(x)$ est bien définie et la fonction $x \mapsto S(x)$ est continue sur I . En outre, la somme partielle $S_N(x) = \sum_{n=0}^N f_n(x)$ converge uniformément vers $S(x)$ dans le sens où*

$$\sup_{x \in I} \left| S(x) - \sum_{n=0}^N f_n(x) \right| \xrightarrow{N \rightarrow +\infty} 0 .$$

Démonstration : Pour tout $x \in I$, $|f_n(x)| \leq \sup_{y \in I} |f_n(y)|$ et donc, par hypothèse, $\sum_{n \geq 0} f_n(x)$ est une série absolument convergente vers une limite $S(x)$. Montrons que la fonction $x \mapsto S(x)$ est continue. Pour cela, notons que, pour tout $x \in I$ et $N \in \mathbb{N}$,

$$|S(x) - \sum_{n=0}^N f_n(x)| = \left| \sum_{n \geq N+1} f_n(x) \right| \leq \sum_{n \geq N+1} |f_n(x)| \leq \sum_{n \geq N+1} \sup_{y \in I} |f_n(y)| .$$

Comme la série $\sum_{n \geq 0} (\sup_{y \in I} |f_n(y)|)$ converge, on obtient que $|S(x) - \sum_{n=0}^N f_n(x)|$ est borné par une quantité R_N indépendante de x et qui tend vers 0 quand N tend vers $+\infty$. Soit $x_0 \in I$ et $\varepsilon > 0$. On note que $S_N(x) = \sum_{n=0}^N f_n(x)$ est une fonction continue comme somme finie de fonctions continues, il existe donc $\eta > 0$ tel que, pour tout $x \in]x_0 - \eta, x_0 + \eta[$, $|S_N(x_0) - S_N(x)| \leq \varepsilon/2$. On a donc

$$|S(x) - S(x_0)| \leq |S(x) - S_N(x)| + |S_N(x) - S_N(x_0)| + |S_N(x_0) - S(x_0)| \leq 2R_N + \varepsilon/2 .$$

Pour N assez grand, on obtient que $|S(x) - S(x_0)| \leq \varepsilon$, ce qui montre que $S(x)$ est continue.

Il reste à justifier que la convergence est bien uniforme, ce qui vient de

$$\begin{aligned} \sup_{x \in I} \left| S(x) - \sum_{n=0}^N f_n(x) \right| &= \sup_{x \in I} \left| \sum_{n \geq N+1} f_n(x) \right| \\ &\leq \sum_{n \geq N+1} \sup_{x \in I} |f_n(x)| \xrightarrow{N \rightarrow +\infty} 0 . \end{aligned}$$

□

Exemples :

- On prend I de la forme $[-1 + \eta, 1 - \eta]$ avec $\eta \in]0, 1[$. On note que, pour tout $x \in I$, on a $|x^n| \leq (1 - \eta)^n$ et que $\sum_{n \geq 0} (1 - \eta)^n$ est une série géométrique convergente. Donc la fonction $S(x) = \sum_{n \geq 0} x^n$ est bien définie et continue sur $[-1 + \eta, 1 - \eta]$ pour tout $\eta \in]0, 1[$, et donc sur $] - 1, 1[$. Par ailleurs, la somme de la série géométrique nous confirme ceci, puisque $S(x) = \sum_{n \geq 0} x^n = 1/(1 - x)$.
- Sur tout intervalle $I = [-M, M]$, $|x^n/n!|$ est borné par $M^n/n!$. En appliquant le critère de D'Alembert, on obtient que $\sum M^n/n!$ est convergente. On en déduit que $\sum_{n \geq 0} \frac{x^n}{n!}$ définit bien une fonction continue sur tout I , et donc sur \mathbb{R} .

Pour montrer (3.1), il nous reste à prouver que la série converge bien vers l'exponentielle. Pour cela, on utilise la formule de Taylor avec reste intégral.

Proposition 3.12. Soit I un intervalle de \mathbb{R} et f une fonction de classe $\mathcal{C}^{k+1}(I, \mathbb{R})$, alors pour tout $x_0 \in I$,

$$\begin{aligned} \forall x \in I, \quad f(x) &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \dots \\ &\quad + \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + \int_{x_0}^x \frac{f^{(k+1)}(t)}{k!}(x - t)^k dt . \end{aligned}$$

Démonstration : On fait une récurrence sur k en utilisant que

$$\int_{x_0}^x \frac{f^{(k)}(t)}{(k-1)!}(x-t)^{k-1} dt = \left[-\frac{f^{(k)}(t)}{k!}(x-t)^k \right]_{x_0}^x - \int_{x_0}^x -\frac{f^{(k+1)}(t)}{k!}(x-t)^k dt .$$

□

Il s'agit donc de montrer que le reste intégral tend vers 0 dans la formule

$$e^x = \sum_{n=0}^N \frac{x^n}{n!} + \int_0^x \frac{e^t}{N!} (x-t)^N dt .$$

Or, on a

$$\left| \int_0^x \frac{e^t}{N!} (x-t)^N dt \right| \leq |x| \frac{e^{|x|}}{N!} |x|^N \xrightarrow{N \rightarrow +\infty} 0$$

Donc (3.1) est bien vérifiée. Avec les mêmes arguments, on montre que, pour tout $x \in \mathbb{R}$,

$$\cos x = \sum_{n \geq 0} (-1)^n \frac{x^{2n}}{(2n)!} \quad \text{et} \quad \sin x = \sum_{n \geq 0} (-1)^n \frac{x^{2n+1}}{(2n+1)!} .$$

On notera que tout ne se passe pas forcément si bien que pour ces exemples. Ainsi, si la fonction $1/(1-x)$ est définie aussi pour tous les $x \neq 1$, sa série entière $\sum x^n$ ne converge que sur $] -1, 1[$. On peut aussi montrer que $x \mapsto e^{-1/x^2}$, prolongé par 0 en 0, est une fonction infiniment dérivable mais dont le développement de Taylor est toujours nul. Son développement de Taylor converge donc, mais la fonction n'est pas égale à la limite de la somme. On dit que la fonction n'est pas développable en série entière.

4 Quelques compléments

4.1 Séries alternées

On a regardé pour le moment que des séries de termes positifs. Il se trouve que le fait qu'une série change de signe aide à la convergence. Par exemple, la série $\sum_{n \geq 1} (-1)^n/n$ converge, mais elle ne converge pas en valeur absolue. En effet, en regroupant les termes deux à deux, on a que

$$\frac{1}{2n} - \frac{1}{2n+1} = \frac{1}{4n^2 + 2n} \sim \frac{1}{4n^2} .$$

Cette astuce montre qu'en sommant la série astucieusement, on obtient une série de type Riemann qui converge. En fait, on a un principe plus général.

Théorème 3.13. Séries alternées

Soit (v_n) une suite de réels positifs, qui décroît et tend vers 0. Alors la série $\sum_{n \geq 0} (-1)^n v_n$ est convergente.

Démonstration : Il suffit de montrer que $S_N = \sum_{n=0}^{2N} (-1)^n v_n$ et $T_N = \sum_{n=0}^{2N+1} (-1)^n v_n$ sont deux suites adjacentes. □

4.2 Un exemple de somme non-continue

On pourrait croire que l'énoncé du théorème 3.11 est très naturel et qu'il n'y a pas besoin de tant d'hypothèses. En fait, ce n'est pas vrai. Pour tout $x \in [0, 1]$, posons $f_n(x) = x^{n+1} - x^n$ et $S_N(x) = \sum_{n=1}^N f_n(x)$. Par simplification en cascade, on a que $S_N(x) = x^{N+1} - x$ qui est bien une fonction continue. De plus, pour tout $x \in [0, 1]$, la série est convergente et sa limite est

$$S(x) = \begin{cases} -x & \text{si } x \in [0, 1[\\ 0 & \text{si } x = 1 \end{cases}$$

qui ne donne pas une fonction continue sur tout $[0, 1]$.

4.3 Dérivation sous le signe somme

Il n'est pas vrai en général que l'on peut dériver une série de fonction dans le sens que la dérivée de la série serait la série des dérivées. Ceci est toutefois vrai sous de bonnes hypothèses.

Théorème 3.14. *Soit I un intervalle de \mathbb{R} et soient (f_n) une suite de fonctions de classe \mathcal{C}^1 sur I . On suppose que $\sum_{n \geq 0} (\sup_{x \in I} |f_n(x)|)$ et $\sum_{n \geq 0} (\sup_{x \in I} |f'_n(x)|)$ sont des séries convergentes dans \mathbb{R}_+ . Alors, pour tout $x \in I$, $S(x) = \sum_{n \geq 0} f_n(x)$ est bien définie et la fonction $x \mapsto S(x)$ est de classe \mathcal{C}^1 sur I . En outre, on a que*

$$S'(x) = \frac{d}{dx} \left(\sum_{n \geq 0} f_n(x) \right) = \sum_{n \geq 0} f'_n(x) .$$

Démonstration : En utilisant le théorème 3.11, on obtient que les fonctions

$$S(x) = \sum_{n \geq 0} f_n(x) \quad \text{et} \quad T(x) = \sum_{n \geq 0} f'_n(x)$$

sont bien définies et continues sur I . Il reste juste à vérifier que S est dérivable et que T est sa dérivée. Pour cela, on regarde les sommes partielles S_N et T_N et, comme il s'agit de sommes finies, on a que $S'_N(x) = T_N(x)$. Fixant $x_0 \in I$, on écrit donc que, pour tout $x \in I$

$$S_N(x) = S_N(x_0) + \int_{x_0}^x T_N(y) dy .$$

Quand N tend vers $+\infty$, comme $T_N(x)$ converge uniformément vers T , on a

$$\begin{aligned} \left| \int_{x_0}^x T_N(y) dy - \int_{x_0}^x T(y) dy \right| &\leq \left| \int_{x_0}^x T_N(y) - T(y) dy \right| \\ &\leq \int_{x_0}^x |T_N(y) - T(y)| dy \\ &\leq |x_0 - x| \sup_{y \in I} |T_N(y) - T(y)| \longrightarrow 0 . \end{aligned}$$

D'où, à la limite, $S(x) = S(x_0) + \int_{x_0}^x T(y)dy$, ce qui montre que S est une primitive de T et donc que T est la dérivée de S . \square

Chapitre 4 : Séries de Fourier

1 Introduction

Nous n'allons pas trop détailler ici la vie de Joseph Fourier (1768, Auxerre – 1830, Paris), mais celle-ci vaut le coup d'oeil à la fois par sa richesse et par son parcours caractéristique de cette époque : naissance très modeste, pensionnat, élève à l'Ecole Normale de Paris, professeur à l'Ecole Polytechnique (à moins de 30 ans!), conseiller scientifique pendant la campagne d'Egypte, puis préfet de l'Isère et enfin membre de l'Académie de Sciences. On lui doit la création du lycée et de l'université de Grenoble, la route du col de Lautaret ou la découverte de l'effet de serre à l'échelle des planètes. Mais si le nom de Fourier est internationalement connu, c'est pour ses travaux sur l'équation de la chaleur (effectués à Grenoble) et sa résolution par les séries qui portent son nom.

La décomposition d'une fonction f en série de Fourier est l'écriture

$$f(x) = a_0 + \sum_{n \geq 1} a_n \cos\left(\frac{2\pi n}{T}x\right) + b_n \sin\left(\frac{2\pi n}{T}x\right), \quad (4.1)$$

où les a_n et b_n sont des coefficients réels et $T > 0$ une période réelle. On peut remarquer qu'une telle décomposition n'est envisageable que pour les fonctions T -périodiques puisque le membre de droite est T -périodique. Il faudra mathématiquement comprendre dans quel sens la série converge et dans quel sens l'égalité (4.1) a lieu et aussi voir comment calculer en pratique les coefficients a_n et b_n , ce qui sera l'objet de ce chapitre.

Le principe des séries de Fourier est un changement de point de vue. Ainsi, à la place de regarder un signal sonore comme une fonction périodique dépendant du temps, on le regarde comme la superposition d'ondes périodiques de différentes amplitudes. On passe donc d'une vision temporelle à une vision fréquentielle. Si le spectre en fréquence a des propriétés particulières, la vision fréquentielle est bien plus facile à manipuler.

Fourier a utilisé ses séries pour décrire les solutions de l'équation de la chaleur. Mais de nombreuses autres applications des ses séries se retrouvent partout : décomposition et analyse d'un signal sonore, débruitage, compression MP3, compression d'images JPEG...

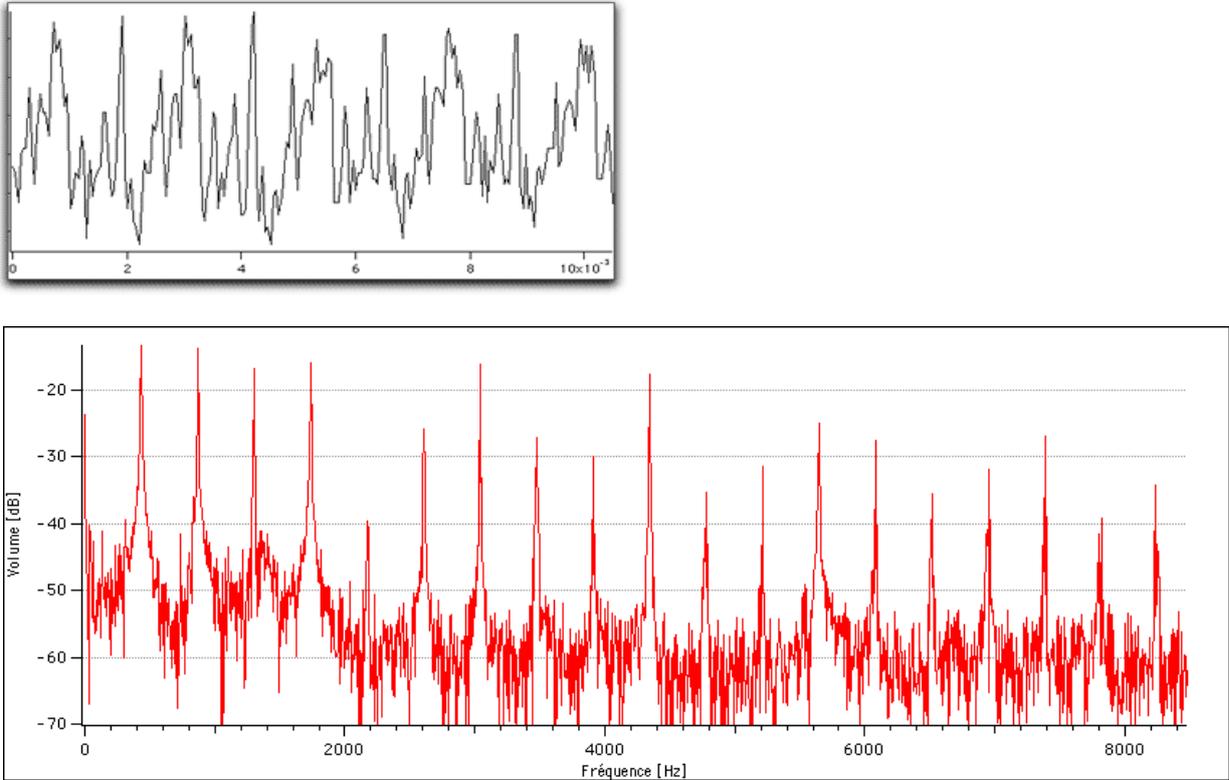


FIGURE 4.1 – Le signal sonore d’un violon et la décomposition spectrale associée.

2 L’espace des fonctions $\mathcal{C}_{T\text{-per},m}^{k,VP}(\mathbb{R})$

On veut travailler avec des fonctions de type signal périodique : son, signal lumineux, signal électronique. . . Soit T la période de ce signal, on peut commencer par considérer les fonctions de classe \mathcal{C}^k qui sont T -périodiques, c’est-à-dire l’espace

$$\mathcal{C}_{T\text{-per}}^k(\mathbb{R}) = \{f \in \mathcal{C}^k(\mathbb{R}) \mid \forall x \in \mathbb{R}, f(x+T) = f(x)\} .$$

En général, les signaux que l’on regardera ne seront pas forcément continus. On va donc regarder des signaux de classe \mathcal{C}^k par morceaux. On définit l’ensemble $\mathcal{C}_m^k(\mathbb{R})$ comme l’ensemble des fonctions telles que, pour tout intervalle $[a, b]$, il existe un nombre fini de points $a = x_1 < x_2 < \dots < x_p = b$ tels que

- i) f soit de classe \mathcal{C}^k sur chaque intervalle $[x_i, x_{i+1}]$,
- ii) f et ses k premières dérivées ont une limite à gauche et à droite en tout point x_i .

Pour avoir des fonctions en outre T -périodiques, on pourra considérer

$$\mathcal{C}_{T\text{-per},m}^k(\mathbb{R}) = \{f \in \mathcal{C}_m^k(\mathbb{R}) \mid \forall x \in \mathbb{R}, f(x+T) = f(x)\} .$$

Le problème des fonctions dans $\mathcal{C}_{T\text{-per},m}^k(\mathbb{R})$, c'est que la valeur aux points de discontinuité

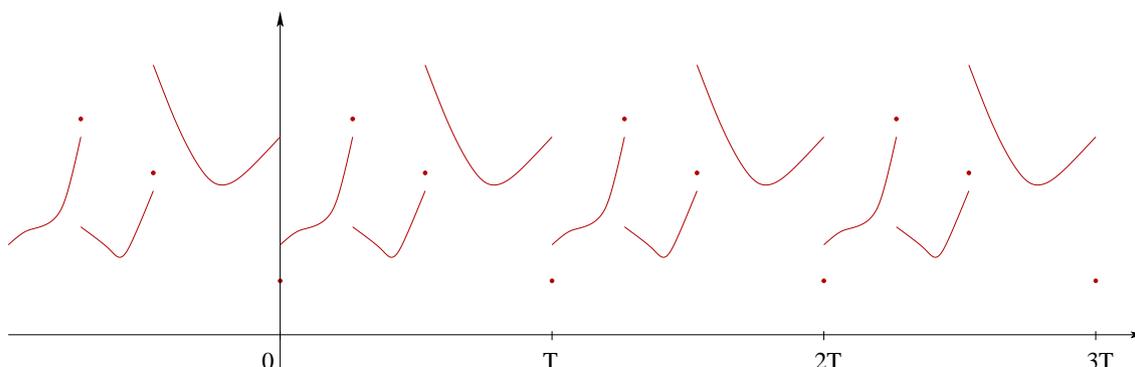


FIGURE 4.2 – Un exemple de fonction dans $\mathcal{C}_{T\text{-per},m}^k(\mathbb{R})$.

peut être arbitraire. On peut par exemple avoir une fonction nulle partout sauf sur un nombre fini de points de $[0, T]$. Or, en physique, la valeur ponctuelle d'une fonction n'a pas vraiment de sens. On s'intéresse plutôt à la valeur moyenne de la fonction sur de petits intervalles de temps où le signal est observé.

On introduit donc la valeur principale $VP(f)$ d'une fonction $f \in \mathcal{C}_m^k(\mathbb{R})$ en posant

$$VP(f)(x_0) = \lim_{h \rightarrow 0} \frac{1}{2h} \int_{x_0-h}^{x_0+h} f(x) dx .$$

Théorème 4.1. Soit $f \in \mathcal{C}_m^k(\mathbb{R})$, alors la valeur principale $VP(f)$ de f est bien définie et vérifie

$$VP(f)(x_0) = \frac{1}{2} \left(\lim_{x \rightarrow x_0^+} f(x) + \lim_{x \rightarrow x_0^-} f(x) \right) .$$

Par conséquent, $VP(f)$ est aussi dans $\mathcal{C}_m^k(\mathbb{R})$ et coïncide avec f sauf aux points de discontinuité.

Démonstration : Comme

$$VP(f)(x_0) = \frac{1}{2} \left(\lim_{h \rightarrow 0} \frac{1}{h} \int_{x_0-h}^{x_0} f(x) dx + \lim_{h \rightarrow 0} \frac{1}{h} \int_{x_0}^{x_0+h} f(x) dx \right) ,$$

il suffit de montrer que

$$\lim_{h \rightarrow 0} \frac{1}{h} \int_{x_0}^{x_0+h} f(x) dx = \lim_{x \rightarrow x_0^+} f(x) ,$$

cette dernière limite, que l'on notera ℓ , existant par définition de $\mathcal{C}_m^k(\mathbb{R})$. Soit $\varepsilon > 0$ et soit $\eta > 0$ tel que $|\ell - f(x)| \leq \varepsilon$ pour $x \in]x_0, x_0 + \eta[$. Pour $h \leq \eta$, on a

$$\begin{aligned} \left| \frac{1}{h} \int_{x_0}^{x_0+h} f(x) dx - \ell \right| &= \left| \frac{1}{h} \int_{x_0}^{x_0+h} (f(x) - \ell) dx \right| \\ &\leq \frac{1}{h} \int_{x_0}^{x_0+h} |f(x) - \ell| dx \\ &\leq \varepsilon . \end{aligned}$$

□

Dans la suite, on ne s'intéressera qu'à la valeur principale des fonctions, on pourra donc ne regarder que les fonctions f qui sont une valeur principale, c'est-à-dire qui vérifie que

$$\forall x_0 \in \mathbb{R}, \quad f(x_0) = \frac{1}{2} \left(\lim_{x \rightarrow x_0^+} f(x) + \lim_{x \rightarrow x_0^-} f(x) \right)$$

ou bien que $VP(f) = f$. On considèrera l'ensemble X_T défini par

$$X_T = \mathcal{C}_{T-per,m}^{k,VP}(\mathbb{R}) = \{f \in \mathcal{C}_{T-per,m}^k(\mathbb{R}) \mid VP(f) = f\} .$$

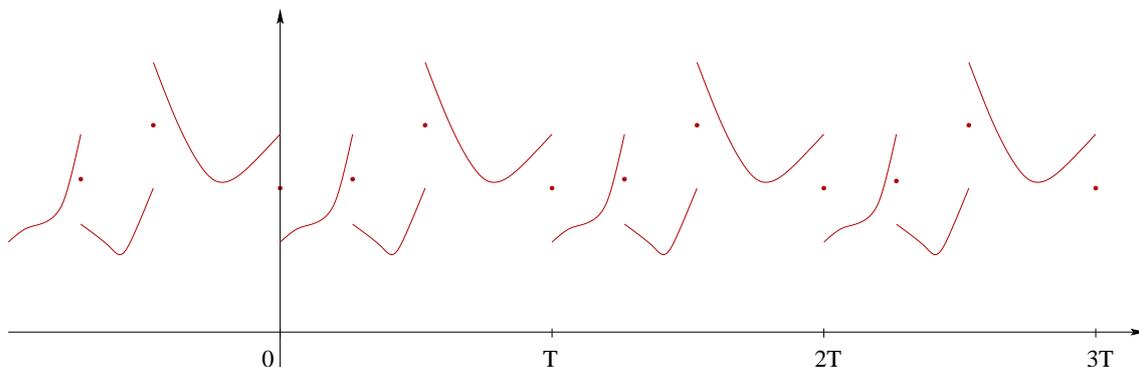


FIGURE 4.3 – Un exemple de fonction dans $\mathcal{C}_{T-per,m}^{k,VP}(\mathbb{R})$: la valeur principale de la fonction de la figure 4.2.

On munira X_T du produit scalaire L^2

$$\langle f|g \rangle = \int_0^T f(x)g(x)dx ,$$

qui fera de lui un espace préhilbertien. Il est important de noter que, sans la condition $VP(f) = f$ dans la définition de X_T , $\langle \cdot | \cdot \rangle$ ne serait pas un produit scalaire car $\|f\| = 0$ pour une fonction f non nulle du type $f(x) = 1$ si $x = 0[T]$ et $f(x) = 0$ sinon.

3 Séries de Fourier et convergence au sens L^2

Dans cette partie, on se placera dans l'espace préhilbertien $X_T = \mathcal{C}_{T-per,m}^{0,VP}(\mathbb{R})$ défini ci-dessus et muni du produit scalaire $\langle f|g \rangle = \int_0^T f(x)g(x)dx$.

3.1 Polynômes trigonométriques

On note $Trig_T^N$ l'espace des polynômes trigonométriques de degré N , c'est-à-dire le sous-espace vectoriel de X_T engendré par la famille

$$1, \cos\left(\frac{2\pi}{T}\cdot\right), \sin\left(\frac{2\pi}{T}\cdot\right), \cos\left(2\frac{2\pi}{T}\cdot\right), \sin\left(2\frac{2\pi}{T}\cdot\right), \dots, \cos\left(N\frac{2\pi}{T}\cdot\right), \sin\left(N\frac{2\pi}{T}\cdot\right).$$

On note que la famille en question est orthogonale. En effet, pour $n \neq m$:

$$\begin{aligned} \int_0^T \cos\left(n\frac{2\pi}{T}x\right) \cos\left(m\frac{2\pi}{T}x\right) dx &= \frac{T}{2\pi} \int_0^{2\pi} \cos(nx) \cos(mx) dx \\ &= \frac{T}{4\pi} \int_0^{2\pi} \cos((n+m)x) + \cos((n-m)x) dx \\ &= 0 \end{aligned}$$

(le calcul est similaire pour les sinus) et, pour $n \geq 0$ et $m \geq 1$,

$$\begin{aligned} \int_0^T \cos\left(n\frac{2\pi}{T}x\right) \sin\left(m\frac{2\pi}{T}x\right) dx &= \frac{T}{2\pi} \int_0^{2\pi} \cos(nx) \sin(mx) dx \\ &= \frac{T}{4\pi} \int_0^{2\pi} \sin((m+n)x) + \sin((m-n)x) dx \\ &= 0 \end{aligned}$$

Pour rendre la base de $Trig_T^N$ orthonormale, il suffit de calculer, pour $n \geq 1$,

$$\int_0^T \cos^2\left(n\frac{2\pi}{T}x\right) dx = \int_0^T \frac{1 + \cos\left(2n\frac{2\pi}{T}x\right)}{2} dx = \frac{T}{2}.$$

Le calcul est similaire pour les sinus et par ailleurs $(\int_0^T 1^2 dx)^{1/2} = \sqrt{T}$. Une base orthonormale de $Trig_T^N$ est donc

$$\frac{1}{\sqrt{T}}, \sqrt{\frac{2}{T}} \cos\left(\frac{2\pi}{T}\cdot\right), \sqrt{\frac{2}{T}} \sin\left(\frac{2\pi}{T}\cdot\right), \sqrt{\frac{2}{T}} \cos\left(2\frac{2\pi}{T}\cdot\right), \sqrt{\frac{2}{T}} \sin\left(2\frac{2\pi}{T}\cdot\right),$$

$$\dots, \sqrt{\frac{2}{T}} \cos\left(N \frac{2\pi}{T} \cdot\right), \sqrt{\frac{2}{T}} \sin\left(N \frac{2\pi}{T} \cdot\right).$$

3.2 Séries de Fourier et théorème de Parseval-Bessel

Soit f une fonction de l'espace préhilbertien X_T . L'ensemble $Trig_T^N$ des polynômes trigonométriques de degré N forme un sous-espace vectoriel de dimension finie dont on connaît une base orthonormale. Si on projète la fonction f sur $Trig_T^N$, on obtient l'expression du projeté $P_N f$ donnée par

$$P_N f(x) = a_0 + \sum_{n=1}^N a_n \cos\left(n \frac{2\pi}{T} x\right) + b_n \sin\left(n \frac{2\pi}{T} x\right)$$

avec

$$\begin{aligned} a_0 &= \frac{1}{T} \int_0^T f(x) dx \\ a_n &= \frac{2}{T} \int_0^T \cos\left(n \frac{2\pi}{T} x\right) f(x) dx \quad \text{pour } n \geq 1 \\ b_n &= \frac{2}{T} \int_0^T \sin\left(n \frac{2\pi}{T} x\right) f(x) dx. \end{aligned}$$

On a vu que les modes en cosinus et sinus sont porteurs d'une énergie $T/2$, c'est-à-dire que $\|\cos(2\pi/T \cdot)\|^2 = T/2$ et que le mode constant est porteur d'une énergie T . Par le théorème de Pythagore, on a que, pour tout N ,

$$|a_0|^2 T + \sum_{n=1}^N |a_n|^2 \frac{T}{2} + \sum_{n=1}^N |b_n|^2 \frac{T}{2} = \|P_N f\|^2 \leq \|f\|^2. \quad (4.2)$$

On en déduit donc que la série $\sum_{n=1}^{+\infty} |a_n|^2 + |b_n|^2$ est convergente. L'inégalité (4.2) est appelée inégalité de Bessel. La convergence de la série de Fourier au sens L^2 s'obtient grâce à l'égalité de Parseval.

Théorème 4.2. Egalité de Parseval-Bessel

Soit f une fonction continue par morceaux et T -périodique sur \mathbb{R} et soient (a_n) et (b_n) ses coefficients de Fourier définis ci-dessus. Alors, la série $\sum_{n=1}^{+\infty} |a_n|^2 + |b_n|^2$ est convergente et

$$|a_0|^2 T + \left(\sum_{n=1}^{+\infty} |a_n|^2 + |b_n|^2 \right) \frac{T}{2} = \|f\|_{L^2}^2 = \int_0^T |f(x)|^2 dx.$$

En outre, f est égale à sa série de Fourier dans le sens où

$$\left\| f - a_0 + \sum_{n \geq 1} a_n \cos\left(n \frac{2\pi}{T} x\right) + b_n \sin\left(n \frac{2\pi}{T} x\right) \right\|_{L^2} \xrightarrow{N \rightarrow +\infty} 0$$

Démonstration : On va tout d'abord justifier que l'on peut se passer de condition du type $f = VP(f)$. En effet, étant donnée une fonction f continue par morceaux et T -périodique sur \mathbb{R} , on peut remplacer f par $VP(f)$ dans l'énoncé. Mais on note que tout ce qui concernera alors $VP(f)$ sera tout autant défini pour f et aura la même valeur, car $VP(f)$ et f ne diffèrent que sur un nombre fini de points. Ainsi, par exemple, $\|f\|_{L^2} = \|VP(f)\|_{L^2}$. Donc, même si seule $VP(f)$ appartient à l'espace préhilbertien X_T , l'énoncé du théorème est aussi valable pour f .

On admet pour le moment la première partie du résultat. La deuxième partie vient alors du théorème de Pythagore puisque $\|f - P_N f\|^2 = \|f\|^2 - \|P_N f\|^2$ et

$$\|P_N f\|^2 = \sum_{n=0}^N |a_n|^2 + \sum_{n=1}^N |b_n|^2 .$$

□

Il faut comprendre le théorème 4.2 comme le fait que f est égale à sa série de Fourier dans le sens de l'espace L^2 qui est l'espace d'énergie. On notera en particulier que l'énergie de f , qui vaut $\int_0^T |f(x)|^2 dx$, est égale à la somme des énergies des modes de vibrations $|a_0|^2 T + (\sum_{n=1}^{+\infty} |a_n|^2 + |b_n|^2) T/2$.

3.3 Exemples

Pour donner quelques exemples, on va prendre le cas classique $T = 2\pi$. Les coefficients de Fourier d'une fonction f continue par morceaux et 2π -périodique sont donnés par :

$$\begin{aligned} a_0(f) &= \frac{1}{2\pi} \int_0^{2\pi} f(x) dx \\ a_n(f) &= \frac{1}{\pi} \int_0^{2\pi} \cos(nx) f(x) dx \quad \text{pour } n \geq 1 \\ b_n(f) &= \frac{1}{\pi} \int_0^{2\pi} \sin(nx) f(x) dx . \end{aligned}$$

Commençons par considérer la fonction en dents de scie définie par $f(x) = x$ sur $[0, 2\pi[$ et par 2π -périodicité. Le calcul donne

$$\begin{aligned} a_0(f) &= \frac{1}{2\pi} \int_0^{2\pi} x dx = \pi \\ a_n(f) &= \frac{1}{\pi} \int_0^{2\pi} x \cos(nx) dx = \frac{1}{\pi} \left[\frac{x}{n} \sin(nx) \right]_0^{2\pi} - \frac{1}{n\pi} \int_0^{2\pi} \sin(nx) dx = 0 \\ b_n(f) &= \frac{1}{\pi} \int_0^{2\pi} x \sin(nx) dx = \frac{1}{\pi} \left[-\frac{x}{n} \cos(nx) \right]_0^{2\pi} + \frac{1}{n\pi} \int_0^{2\pi} \cos(nx) dx = \frac{-2}{n} . \end{aligned}$$

On trouve donc, la série étant à prendre au sens de $L^2([0, 2\pi])$, que

$$f = \pi - \sum_{n \geq 1} \frac{2}{n} \sin(n \cdot) .$$

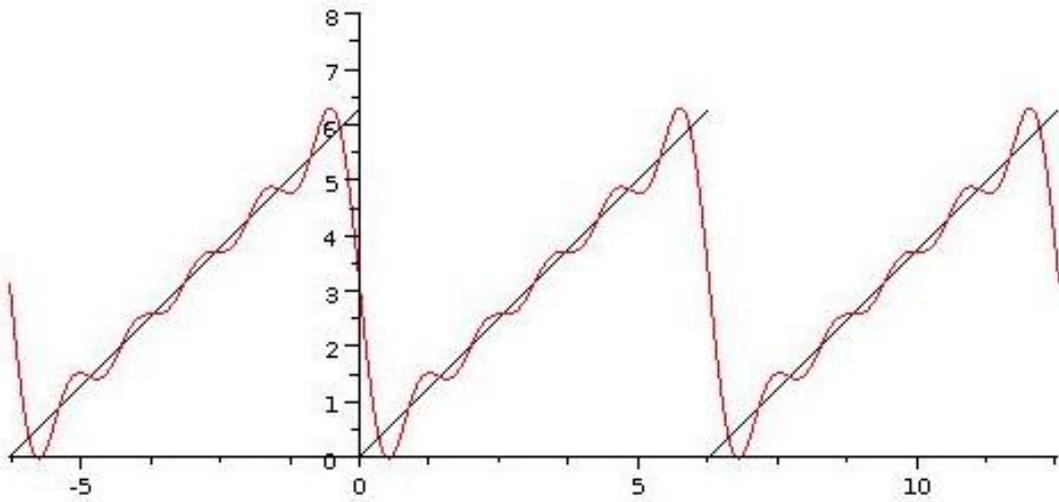


FIGURE 4.4 – La fonction en dents de scie et sa projection sur ses 6 premiers modes de Fourier ($N = 5$).

On peut regarder une fonction créneau définie par $g(x) = 1$ sur $[0, \pi[$, $g(x) = 0$ sur $[\pi, 2\pi[$ et par 2π -périodicité. Le calcul donne

$$\begin{aligned} a_0(g) &= \frac{1}{2\pi} \int_0^\pi dx = \frac{1}{2} \\ a_n(g) &= \frac{1}{\pi} \int_0^\pi \cos(nx) dx = 0 \\ b_n(g) &= \frac{1}{\pi} \int_0^\pi \sin(nx) dx = \frac{1}{\pi} \left[\frac{-1}{n} \cos(nx) \right]_0^\pi = \frac{1 - (-1)^n}{n\pi} . \end{aligned}$$

Pour finir, prenons une fonction continue, par exemple la fonction h_a définie par imparité, 2π -périodicité et

$$h_a = \begin{cases} x/a & \text{si } x \in [0, a] \\ \frac{\pi-x}{\pi-a} & \text{si } x \in [a, \pi] \end{cases}$$

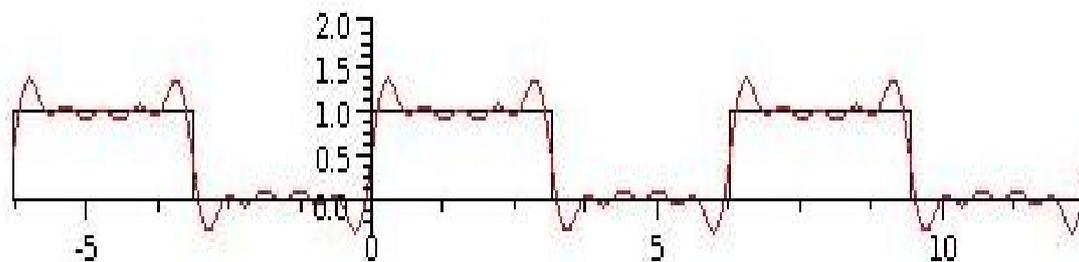


FIGURE 4.5 – La fonction créneau et sa projection sur ses 6 premiers modes de Fourier ($N = 9$).

Le calcul donne

$$\begin{aligned}
 a_n(h_a) &= \frac{1}{\pi} \int_0^{2\pi} h_a(x) \cos(nx) dx = \frac{1}{\pi} \int_0^{\pi} h_a(x) \cos(nx) dx + \frac{1}{\pi} \int_{-\pi}^0 h_a(x) \cos(nx) dx = 0 \\
 b_n(h_a) &= \frac{1}{\pi} \int_0^{2\pi} h_a(x) \sin(nx) dx = \frac{2}{\pi} \int_0^{\pi} h_a(x) \sin(nx) dx \\
 &= \frac{2}{\pi} \int_0^a x \sin(nx) dx + \frac{2}{\pi} \int_a^{\pi} \frac{\pi-x}{\pi-a} \sin(nx) dx \\
 &= \frac{2 \sin(na)}{\pi n^2} \left(\frac{1}{a} + \frac{1}{\pi-a} \right) .
 \end{aligned}$$

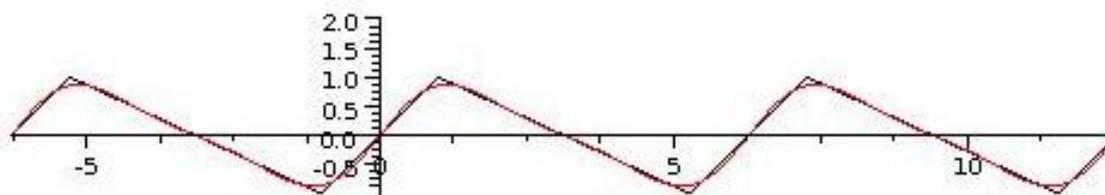


FIGURE 4.6 – La fonction h_1 et sa projection sur ses 3 premiers modes de Fourier.

On notera sur les exemples ci-dessus que :

- la convergence de la série de Fourier vers la fonction est à prendre au sens L^2 , pas au sens de la convergence vers 0 du plus grand écart. Par exemple, dans les figures 4.4 et

4.5, la convergence a l'air mauvaise près des points de discontinuité. Ces oscillations de grandes amplitudes arrivant près des discontinuités forment ce qui est appelé *phénomène de Gibbs*. La figure 4.7 montre un exemple du phénomène dans une image au format jpeg.

- si la fonction est continue et \mathcal{C}^1 par morceaux, comme dans la figure 4.6, la convergence a l'air excellente, même dans le sens du max des écarts. Nous verrons mathématiquement pourquoi dans la partie suivante.
- si une fonction est impaire, les coefficients a_n sont nuls et la série de Fourier ne comprend que des sinus. De même, si une fonction est paire, les coefficients b_n sont nuls et la série de Fourier ne comprend que des cosinus.



FIGURE 4.7 – Une image au format jpeg très compressée (source wikipédia). On remarque le phénomène de Gibbs au niveau de l'interface fond/dessin avec des oscillations fortes des couleurs et une forme qui « bave ».

4 Autres types de convergence

4.1 Convergence ponctuelle

La convergence au sens L^2 ci-dessus est une convergence globale au sens des énergies. Mais cette convergence ne dit rien de ce qui se passe ponctuellement car la valeur d'une fonction en un seul point n'influe pas sur son énergie. On peut espérer toutefois obtenir un résultat si la fonction est continue car si deux fonctions continues ont des valeurs différentes en un point, leur énergies doivent être différentes sur un petit intervalle autour de ce point. Le résultat clef est le suivant.

Théorème 4.3. Théorème de Dirichlet

Soit $f \in \mathcal{C}_{T\text{-per},m}^1$ une fonction \mathcal{C}^1 par morceaux et T -périodique. Alors sa série de Fourier

converge en tout point vers la valeur principale de f au sens où

$$\forall x \in \mathbb{R}, \left| VP(f)(x) - \left(a_0(f) + \sum_{n=1}^N a_n(f) \cos\left(n \frac{2\pi}{T} x\right) + b_n(f) \sin\left(n \frac{2\pi}{T} x\right) \right) \right| \xrightarrow{N \rightarrow +\infty} 0.$$

En particulier, si $f \in \mathcal{C}_{T\text{-per},m}^{1,VP}$, alors f est limite point par point de sa série de Fourier (on parle de limite simple).

On va utiliser le lemme suivant.

Lemme 4.4. Riemann-Lebesgue

Soit $f \in \mathcal{C}_{T\text{-per},m}^0(\mathbb{R})$, alors $\int_0^T f(x) \cos(nx) dx$ comme $\int_0^T f(x) \sin(nx) dx$ tendent vers 0 quand n tend vers $+\infty$.

Démonstration : On remarque tout d'abord, qu'à un facteur près, les intégrales correspondent aux coefficients de Fourier a_n et b_n de f . Comme f est dans $\mathcal{C}_{T\text{-per},m}^0(\mathbb{R})$, on sait par l'inégalité de Bessel que la série $\sum a_n^2 + b_n^2$ converge et donc que a_n et b_n tendent vers 0. \square

Démonstration du théorème 4.3 : Pour simplifier les notations, on va prendre $T = 2\pi$. Calculons la somme partielle de Fourier de f . Pour tout $N \in \mathbb{N}$, on a

$$\begin{aligned} P_N f(x) &= a_0 + \sum_{n=1}^N a_n \cos(nx) + b_n \sin(nx) \\ &= \frac{1}{2\pi} \int_0^{2\pi} f(y) dy + \frac{1}{\pi} \sum_{n=1}^N \cos(nx) \int_0^{2\pi} f(y) \cos(ny) dy + \sin(nx) \int_0^{2\pi} f(y) \sin(ny) dy \\ &= \frac{1}{2\pi} \int_0^{2\pi} f(y) dy + \frac{1}{\pi} \sum_{n=1}^N \int_0^{2\pi} f(y) \cos(n(x-y)) dy \\ &= \int_0^{2\pi} f(y) D_N(x-y) dy \end{aligned}$$

où D_N est le noyau de Dirichlet donné par

$$\begin{aligned} D_N(\xi) &= \frac{1}{2\pi} + \frac{1}{\pi} \sum_{n=1}^N \cos(n\xi) = \frac{1}{2\pi} \sum_{n=-N}^N N e^{in\xi} \\ &= \frac{1}{2\pi} \frac{e^{-iN\xi} - e^{i(N+1)\xi}}{1 - e^{i\xi}} = \frac{1}{2\pi} \frac{e^{-i(N+\frac{1}{2})\xi} - e^{i(N+\frac{1}{2})\xi}}{e^{i\frac{\xi}{2}} - e^{-i\frac{\xi}{2}}} \\ &= \frac{1}{2\pi} \frac{\sin((N+\frac{1}{2})\xi)}{\sin \frac{\xi}{2}}. \end{aligned}$$

On a donc, en utilisant la 2π -périodicité,

$$\begin{aligned} P_N f(x) &= \int_0^{2\pi} f(y) D_N(x-y) dy = \int_0^{2\pi} f(x-y) D_N(y) dy \\ &= \int_{-\pi}^{\pi} f(x-y) D_N(y) dy = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{f(x-y)}{\sin \frac{y}{2}} \sin((N + \frac{1}{2})y) dy \end{aligned}$$

En développant $\sin((N + \frac{1}{2})y)$, on pourrait presque utiliser le lemme 4.4... sauf que $\frac{f(x-y)}{\sin \frac{y}{2}}$ n'est pas forcément borné près de 0. Notons $f(x^+)$ et $f(x^-)$ les limites à droite et à gauche de f en x . On va donc introduire la fonction g définie par

$$g(y) = \frac{f(x-y) - f(x^+) \chi_{[-\pi, 0[}(y) - f(x^-) \chi_{]0, \pi](y)}}{\sin \frac{y}{2}}$$

qui est cette fois bien continue par morceaux. En effet, par exemple si y tend vers 0^- , comme f est dérivable à droite en x et comme $\sin(y/2) \sim y/2$, on a que $g(y)$ tend vers $-2f'(x^+)$ (et idem si y tend vers 0^+). Du coup, on a que

$$\begin{aligned} P_N f(x) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} g(y) \sin((N + \frac{1}{2})y) dy + \frac{1}{2\pi} \int_{-\pi}^{\pi} (f(x^+) \chi_{y>0} + f(x^-) \chi_{y<0}) \frac{\sin((N + \frac{1}{2})\xi)}{\sin \frac{\xi}{2}} dy \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} g(y) \sin((N + \frac{1}{2})y) dy + \frac{f(x^-)}{2\pi} \int_{-\pi}^0 D_N(y) dy + \frac{f(x^+)}{2\pi} \int_0^{\pi} D_N(y) dy \end{aligned}$$

La première intégrale tend vers 0 d'après le lemme 4.4. Remarquons que D_N est pair et que $\frac{1}{2\pi} \int_{-\pi}^{\pi} D_N(y) dy = 1$ (ce qui se montre en considérant cette intégrale comme la somme partielle de Fourier de la fonction constante égale à 1). Donc les deux autres termes valent respectivement $f(x^-)/2$ et $f(x^+)/2$, ce qui montre le résultat. \square

Deux illustrations simples du théorème de Dirichlet sont données par les fonctions f et g du paragraphe 3.3. Dans ces exemples, la valeur de la série de Fourier en $x = 0$ est respectivement π et $1/2$, ce qui correspond bien à la moyenne des limites à gauche et à droite de 0 pour f .

4.2 Convergence uniforme

Nous avons vu la convergence au sens de la norme L^2 , c'est-à-dire au sens de l'énergie physique, ainsi que la convergence point par point. Ces convergences ne sont toutefois pas toujours satisfaisantes, en particulier à cause du phénomène de Gibbs, illustré dans les figures 4.5 et 4.7, et qui apparaît près des discontinuités de la fonction. Pour obtenir une meilleure convergence, il faut se placer dans le cas d'une fonction continue et de classe \mathcal{C}^1 sauf sur un nombre fini de points.

Théorème 4.5. Si f est une fonction de classe $\mathcal{C}^0(\mathbb{R}) \cap \mathcal{C}_{T\text{-per},m}^1(\mathbb{R})$, alors sa série de Fourier converge uniformément vers f dans le sens où

$$\sup_{x \in \mathbb{R}} \left| f(x) - \left(a_0(f) + \sum_{n=1}^N a_n(f) \cos\left(n \frac{2\pi}{T} x\right) + b_n(f) \sin\left(n \frac{2\pi}{T} x\right) \right) \right| \xrightarrow{N \rightarrow +\infty} 0 .$$

Démonstration : Commençons par constater qu'il suffit de montrer la convergence uniforme sur $[0, T]$ par T -périodicité. Il suffit en outre de montrer que la série de Fourier converge uniformément vers une fonction continue g . En effet, g sera forcément égale à f d'après le théorème 4.3 de Dirichlet. On va appliquer directement le théorème th-serie-fonctions et comme les cosinus et les sinus sont bornés sur $[0, 2\pi]$, il suffit de montrer que les séries $\sum |a_n(f)|$ et $\sum |b_n(f)|$ sont convergentes. Prenons le cas de la série $\sum |a_n|$ dans le cas $T = 2\pi$, les autres cas étant similaires. Soient $x_0 = 0, x_1, \dots, x_p = 2\pi$ les points de discontinuités de $f'(x)$ sur $[0, 2\pi]$. On a

$$\begin{aligned} a_n(f) &= \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nx) dx \\ &= \frac{1}{\pi} \int_{x_0}^{x_1} f(x) \cos(nx) dx + \frac{1}{\pi} \int_{x_1}^{x_2} f(x) \cos(nx) dx + \dots + \frac{1}{\pi} \int_{x_{p-1}}^{x_p} f(x) \cos(nx) dx \\ &= \left[f(x) \frac{\sin(nx)}{n\pi} \right]_{x_0}^{x_1} - \int_{x_0}^{x_1} f'(x) \frac{\sin(nx)}{n\pi} dx + \left[f(x) \frac{\sin(nx)}{n\pi} \right]_{x_1}^{x_2} - \int_{x_1}^{x_2} f'(x) \frac{\sin(nx)}{n\pi} dx \\ &\quad + \dots + \left[f(x) \frac{\sin(nx)}{n\pi} \right]_{x_{p-1}}^{x_p} - \int_{x_{p-1}}^{x_p} f'(x) \frac{\sin(nx)}{n\pi} dx \\ &= - \int_{x_0}^{x_1} f'(x) \frac{\sin(nx)}{n\pi} dx - \int_{x_1}^{x_2} f'(x) \frac{\sin(nx)}{n\pi} dx - \dots - \int_{x_{p-1}}^{x_p} f'(x) \frac{\sin(nx)}{n\pi} dx \\ &= - \int_0^{2\pi} f'(x) \frac{\sin(nx)}{n\pi} dx \end{aligned}$$

les termes du type $f(x_i) \frac{\sin(nx_i)}{n\pi}$ se simplifiant en cascade. On a donc que le coefficient de Fourier de la dérivée f' vaut $b_n(f') = -na_n(f)$. Comme f' est une fonction continue par morceaux, l'inégalité de Bessel montre que $\sum |b_n(f')|^2$ converge. Par ailleurs, $\sum 1/n^2$ est aussi une série convergente. On obtient alors par Cauchy-Schwarz que

$$\sum_{n=1}^N |a_n(f)| = \sum_{n=1}^N \frac{1}{n} |b_n(f')| \leq \sqrt{\sum_{n=1}^N \frac{1}{n^2}} \sqrt{\sum_{n=1}^N |b_n(f')|^2} \leq \sqrt{\sum_{n=1}^{\infty} \frac{1}{n^2}} \sqrt{\sum_{n=1}^{\infty} |b_n(f')|^2}$$

ce qui montre que la série $\sum_{n \geq 1} |a_n(f)|$ est convergente. □

Une bonne illustration de ce théorème se trouve dans la figure 4.6 où l'on voit bien que la série de Fourier approche très bien la fonction h_1 qui est continue et \mathcal{C}^1 par morceaux. Le phénomène de Gibbs n'apparaît pas pour de telle fonction d'après le théorème 4.5.

4.3 Retour sur la preuve de la convergence L^2

Nous allons donner quelques rapides arguments pour la preuve du théorème 4.2. Constata-tions d'abord que nous avons pris soin de ne pas utiliser son énoncé dans les démonstrations ci-dessus : nous avons simplement eu besoin de l'inégalité de Bessel. Pour démontrer la convergence L^2 , on commence par prendre une fonction f continue par morceaux et on la régularise en une fonction \tilde{f}_ε qui est continue et \mathcal{C}^1 par morceaux et telle que $\|f - \tilde{f}_\varepsilon\|_{L^2} \leq \varepsilon$. Cela se fait en considérant des moyennes locales de f du type

$$\tilde{f}_\varepsilon(x) = \int_{x-\varepsilon}^{x+\varepsilon} f(x)\rho_\varepsilon(x)dx$$

avec ρ_ε une bosse lisse et d'intégrale 1 qui est très concentrée autour de 0 (c'est le même procédé que pour la valeur principale, mais avec un gain de régularité en prenant ρ_ε lisse). On applique alors le théorème 4.5 et on obtient que la série de Fourier de \tilde{f}_ε tend uniformément vers elle. Il reste alors à voir que \tilde{f}_ε et f sont proches en norme L^2 et qu'il en est de même pour leurs coefficients de Fourier. Du coup, la série de Fourier de f va être proche d'elle au sens L^2 .

5 Série de Fourier en exponentielles complexes

Il existe une autre façon d'écrire les séries de Fourier, c'est de décomposer une fonction f comme une série d'exponentielles complexes

$$f(x) = \sum_{n \in \mathbb{Z}} c_n(f) e^{n \frac{2i\pi}{T} x}, \quad (4.3)$$

avec les coefficients de Fourier $c_n(f)$ donnés par la formule

$$c_n(f) = \frac{1}{T} \int_0^T f(x) e^{-n \frac{2i\pi}{T} x} .$$

Pour établir cette formule, on peut simplement partir du développement sur les sinus et cosinus (en prenant $T = 2\pi$ pour simplifier les notations) et écrire

$$\begin{aligned}
 P_N f(x) &= a_0 + \sum_{n=1}^N a_n \cos(nx) + b_n \sin(nx) \\
 &= a_0 + \sum_{n=1}^N \frac{a_n}{2} (e^{inx} + e^{-inx}) + \frac{b_n}{2i} (e^{inx} - e^{-inx}) \\
 &= a_0 + \sum_{n=1}^N \left(\frac{a_n}{2} + \frac{b_n}{2i} \right) e^{inx} + \left(\frac{a_n}{2} - \frac{b_n}{2i} \right) e^{-inx} \\
 &= a_0 + \sum_{n=1}^N \left(\frac{a_n}{2} - \frac{ib_n}{2} \right) e^{inx} + \left(\frac{a_n}{2} + \frac{ib_n}{2} \right) e^{-inx} \\
 &= \sum_{n=-N}^N c_n e^{inx}
 \end{aligned}$$

avec

$$c_0 = a_0 = \frac{1}{2\pi} \int_0^{2\pi} f(x) dx ,$$

si $n \geq 1$

$$c_n = a_n/2 - ib_n/2 = \frac{1}{2\pi} \int_0^{2\pi} f(x) (\cos(nx) - i \sin(nx)) dx = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} ,$$

et si $n \leq -1$

$$\begin{aligned}
 c_n &= a_{-n}/2 + ib_{-n}/2 = \frac{1}{2\pi} \int_0^{2\pi} f(x) (\cos(-nx) + i \sin(-nx)) dx \\
 &= \frac{1}{2\pi} \int_0^{2\pi} f(x) (\cos(nx) - i \sin(nx)) dx = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} .
 \end{aligned}$$

Comme il s'agit simplement d'une réécriture de la formule précédente de $P_N f$, le sens à donné à (4.3) est le même que celui des théorèmes 4.2, 4.3 et 4.5. On peut se demander s'il est aussi possible de voir cet écriture sous la forme d'une projection orthogonale sur une base orthonormale. La réponse est positive, mais pas avec les notions de produit scalaire et d'espace préhilbertien que nous avons vues. En effet, nous avons besoin d'une version du produit scalaire adaptée aux complexes : pour un nombre complexe z , z^2 n'est pas forcément positif et la quantité $\sqrt{\int_0^T f(x)^2 dx}$ n'est pas définie. Il faut donc remplacer le

produit scalaire L^2 que nous avons vu par

$$\forall f, g \in \mathcal{C}^0([0, T], \mathbb{C}) , \quad \langle f|g \rangle = \int_0^T f(x)\bar{g}(x) dx .$$

Il ne s'agit plus d'un produit scalaire car l'expression n'est pas linéaire par rapport à g : si on multiplie g par $\lambda \in \mathbb{C}$, $\langle f|g \rangle$ sera multiplié par $\bar{\lambda}$. On parle plutôt de forme *sesquilinéaire* et $\langle \cdot | \cdot \rangle$ est appelé un *produit hermitien*. On peut reprendre mutatis mutandis le cours précédent en remplaçant le produit scalaire par ce produit hermitien et cela conduira à écrire les séries de Fourier sous la forme (4.3).

Chapitre 5 : Applications aux EDP

1 Décompositions de fonctions sur $[0, L]$

Contrairement au chapitre précédent, nous considérons maintenant une fonction continue par morceaux sur un intervalle $[0, L]$ ($L > 0$) sans supposer de périodicité. Nous allons voir trois types de décompositions inspirées des séries de Fourier.

Théorème 5.1. *Soit $L > 0$ et $f \in \mathcal{C}_{pm}^0([0, L], \mathbb{R})$ une fonction continue par morceaux sur $[0, L]$. Alors, on a la décomposition*

$$f = \sum_{n \in \mathbb{Z}} c_n e^{in \frac{2\pi}{L} \cdot} = a_0 + \sum_{n \geq 1} a_n \cos\left(n \frac{2\pi}{L} \cdot\right) + b_n \sin\left(n \frac{2\pi}{L} \cdot\right) \quad (5.1)$$

avec

$$\begin{aligned} c_n &= \frac{1}{L} \int_0^L f(x) e^{-in \frac{2\pi}{L} x} dx & a_0 &= \frac{1}{L} \int_0^L f(x) dx \\ a_n &= \frac{2}{L} \int_0^L f(x) \cos(nx) dx & b_n &= \frac{2}{L} \int_0^L f(x) \sin(nx) dx \end{aligned}$$

la convergence dans (5.1) étant à prendre dans le sens du théorème de Parseval 4.2, c'est-à-dire de la convergence de l'énergie L^2 .

Si, en outre, on a que $f \in \mathcal{C}^0([0, L], \mathbb{R}) \cap \mathcal{C}_{pm}^1([0, L], \mathbb{R})$ est continue et de classe \mathcal{C}^1 par morceaux et que $f(0) = f(L)$, alors la convergence dans (5.1) est uniforme, c'est-à-dire dans le sens du théorème 4.5 et on a

$$f(x) = \sum_{n \in \mathbb{Z}} c_n e^{in \frac{2\pi}{L} x} = a_0 + \sum_{n \geq 1} a_n \cos\left(n \frac{2\pi}{L} x\right) + b_n \sin\left(n \frac{2\pi}{L} x\right) .$$

Démonstration : Il s'agit juste d'une application du chapitre précédent à la fonction $\tilde{f} \in \mathcal{C}_{L-per, pm}^0(\mathbb{R})$ définie par $\tilde{f}(x) = f(x)$ sur $[0, L]$ et \tilde{f} est L -périodique. Le point important est de voir que pour appliquer le théorème 4.5, il faut que \tilde{f} soit continue et donc que non seulement f soit continue, mais aussi que $f(0) = f(L)$. \square

En utilisant d'autres façons de prolonger f en \tilde{f} sur \mathbb{R} que celle de la démonstration ci-dessus, on obtient d'autres façons de décomposer f .

Théorème 5.2. Soit $L > 0$ et $f \in \mathcal{C}_{pm}^0([0, L], \mathbb{R})$ une fonction continue par morceaux sur $[0, L]$. Alors, on a la décomposition

$$f = d_0 + \sum_{n \geq 1} d_n \cos\left(n \frac{\pi}{L} \cdot\right) \quad (5.2)$$

avec

$$d_0 = \frac{1}{L} \int_0^L f(x) dx \quad \text{et} \quad d_n = \frac{2}{L} \int_0^L f(x) \cos\left(n \frac{\pi}{L} x\right) dx$$

la convergence dans (5.2) étant à prendre dans le sens du théorème de Parseval 4.2, c'est-à-dire de la convergence de l'énergie L^2 .

Si, en outre, on a que $f \in \mathcal{C}^0([0, L], \mathbb{R}) \cap \mathcal{C}_{pm}^1([0, L], \mathbb{R})$ est continue et de classe \mathcal{C}^1 par morceaux, alors la convergence dans (5.2) est uniforme, c'est-à-dire dans le sens du théorème 4.5 et on a

$$f(x) = d_0 + \sum_{n \geq 1} d_n \cos\left(n \frac{\pi}{L} x\right)$$

Démonstration : Comme pour la démonstration du théorème 5.1, l'idée est de prolonger f : on définit \tilde{f} comme la seule fonction $2L$ -périodique et paire telle que $\tilde{f}(x) = f(x)$ sur $[0, L]$ (autrement dit, on prolonge f par parité sur $[-L, 0]$ puis on finit le prolongement par périodicité). On peut appliquer le théorème 4.2 à \tilde{f} . En outre, comme \tilde{f} est paire, seuls les cosinus apparaissent dans sa série de Fourier. Si f est continue et \mathcal{C}^1 par morceaux, alors \tilde{f} est aussi continue et \mathcal{C}^1 par morceaux et on peut appliquer le théorème 4.5. \square

Théorème 5.3. Soit $L > 0$ et $f \in \mathcal{C}_{pm}^0([0, L], \mathbb{R})$ une fonction continue par morceaux sur $[0, L]$. Alors, on a la décomposition

$$f = \sum_{n \geq 1} e_n \sin\left(n \frac{\pi}{L} \cdot\right) \quad (5.3)$$

avec

$$e_n = \frac{2}{L} \int_0^L f(x) \sin\left(n \frac{\pi}{L} x\right) dx$$

la convergence dans (5.3) étant à prendre dans le sens du théorème de Parseval 4.2, c'est-à-dire de la convergence de l'énergie L^2 .

Si, en outre, on a que $f \in \mathcal{C}^0([0, L], \mathbb{R}) \cap \mathcal{C}_{pm}^1([0, L], \mathbb{R})$ est continue et de classe \mathcal{C}^1 par morceaux et que $f(0) = f(L) = 0$, alors la convergence dans (5.2) est uniforme, c'est-à-dire dans le sens du théorème 4.5 et on a

$$f(x) = \sum_{n \geq 1} e_n \sin\left(n \frac{\pi}{L} x\right)$$

Démonstration : L'idée est la même que celle de la démonstration du théorème 5.2 : on définit \tilde{f} comme la seule fonction $2L$ -périodique et impaire telle que $\tilde{f}(x) = f(x)$ sur $[0, L]$. On peut appliquer le théorème 4.2 à \tilde{f} et comme \tilde{f} est impaire, seuls les sinus apparaissent dans sa série de Fourier. Si f est continue et \mathcal{C}^1 par morceaux, pour que \tilde{f} soit aussi continue et \mathcal{C}^1 par morceaux et pour appliquer le théorème 4.5, il faut supposer en outre que $f(0) = f(L) = 0$. \square

2 Equation de la chaleur

2.1 Modélisation

Imaginons une tige de longueur L d'un matériau conduisant la chaleur. A l'instant initial $t = 0$, la température est donnée par une fonction u_0 définie sur $[0, L]$. Pour savoir comment évolue cette température, on utilise que le flux de chaleur entre deux éléments est proportionnel au gradient de température entre ces éléments (loi de Fourier). Ainsi, entre les segments de tige $[x, x + \delta x]$ et $[x + \delta x, x + 2\delta]$ (où l'on supposera la température u constante en première approximation), il y a un flux $a \frac{u(x+\delta x, t) - u(x, t)}{\delta x}$, où a est un coefficient de diffusion. A l'instant t , le bloc $[x, x + \delta x]$ contenant donc une quantité de chaleur $u(x, t)\delta$ et interagissant avec les blocs $[x + \delta x, x + 2\delta]$ et $[x - \delta x, x]$, on trouve que

$$\partial_t u(x, t)\delta x = a \left(\frac{u(x + \delta x, t) - u(x, t)}{\delta x} + \frac{u(x - \delta x, t) - u(x, t)}{\delta x} \right)$$

c'est-à-dire que

$$\partial_t u(x, t) = a \frac{u(x + \delta x, t) - 2u(x, t) + u(x - \delta, t)}{\delta x^2} .$$

En faisant tendre δx vers 0, on trouve l'équation de la chaleur établie par Joseph Fourier en 1811 :

$$\partial_t u(x, t) = a \partial_{xx}^2 u(x, t) . \tag{5.4}$$

Le même raisonnement en dimension plus grande (sur un bloc de matériau et non une tige) donne $\partial_t u(x, t) = a \Delta u(x, t)$ avec $\Delta = \sum_{i=1}^3 \partial_{x_i x_i}^2$.

Il faut noter que l'équation (5.4) ne suffit pas à elle seule pour décrire l'évolution de la température. En effet, on ne connaît pas le flux de chaleur au bord de la tige. Si la tige est supposée isolée thermiquement, alors on rajoute à l'équation de la chaleur des conditions aux bords de type « flux nul », appelées conditions de Neumann et qui s'écrivent $\partial_x u(0, t) = \partial_x u(L, t) = 0$. Si on suppose que la tige est reliée à un objet de température constante (disons égale à 0 pour fixer une référence), alors on ajoute des conditions aux bords, appelées conditions de Dirichlet et qui s'écrivent $u(0, t) = u(L, t) = 0$. Enfin, on peut supposer que la tige se referme sur elle-même pour former un cercle et on obtient des conditions périodiques $u(0, t) = u(L, t)$ et $\partial_x u(0, t) = \partial_x u(L, t)$.

2.2 Conditions aux bords de Neumann

D'après ce qu'on a vu au paragraphe précédent, on prend un $a > 0$ et une donnée initiale u_0 et on considère l'équation

$$\begin{cases} \partial_t u(x, t) = a \partial_{xx}^2 u(x, t) & (x, t) \in]0, L[\times]0, +\infty[\\ \partial_x u(0, t) = \partial_x u(L, t) = 0 & t > 0 \\ u(x, 0) = u_0(x) & x \in [0, L] \end{cases} \quad (5.5)$$

On va démontrer le théorème suivant.

Théorème 5.4. *Pour tout $u_0 \in \mathcal{C}_{pm}^0([0, L], \mathbb{R})$ continue par morceau, il existe une unique solution u de (5.5) qui est de classe \mathcal{C}^1 en temps et \mathcal{C}^2 en espace pour $(x, t) \in]0, L[\times]0, +\infty[$ et qui a pour donnée initiale u_0 dans le sens que $\|u(\cdot, t) - u_0\|_{L^2}$ tend vers 0 quand t tend vers 0.*

En outre, cette solution u est de classe \mathcal{C}^∞ pour $(x, t) \in]0, L[\times]0, +\infty[$ et est donnée par

$$u(x, t) = d_0(u_0) + \sum_{k \geq 1} d_k(u_0) e^{-a(k\pi/L)^2 t} \cos\left(k \frac{\pi}{L} x\right), \quad (5.6)$$

où $d_k(u_0)$ est le coefficient de la décomposition de u_0 suivant le théorème 5.2. En particulier, $u(x, t)$ tend uniformément et exponentiellement vite vers une température constante égale à la température moyenne de u_0 .

Le théorème 5.4 énonce de nombreuses propriétés. En premier lieu, il existe une solution au problème dans un certain sens (régularité suffisante pour l'EDP et condition initiale vérifiée au sens L^2) et il n'y a pas d'autres solutions qui vérifie (5.5) dans ce sens-là. Par ailleurs, cette solution est en fait plus régulière (de classe \mathcal{C}^∞) et on peut donner son expression exacte et son comportement asymptotique.

Démonstration du théorème 5.4 : concernant l'existence de la solution, il suffit de vérifier que (5.6) définit bien une solution avec les propriétés voulues. Notons que le terme $e^{-a(k\pi/L)^2 t}$ fait que, dès que $t > 0$, la série ainsi que les séries de toutes les dérivées partielles de tout ordre de $u(x, t)$ converge absolument et uniformément. C'est pour cela que la fonction u est bien régulière. En utilisant le théorème 3.14, on obtient bien que $\partial_t u(x, t) = a \partial_{xx}^2 u(x, t)$. Par ailleurs,

$$\begin{aligned} \|u_0 - u(t)\|_{L^2}^2 &= \sum_{k \geq 1} \frac{L}{2} |d_k(u_0)|^2 |1 - e^{-a(k\pi/L)^2 t}|^2 \\ &\leq \sum_{k=1}^K |d_k(u_0)|^2 |1 - e^{-a(k\pi/L)^2 t}|^2 + \sum_{k \geq K+1} |d_k(u_0)|^2 \end{aligned}$$

Le deuxième terme est aussi petit que voulu en prenant K grand, puis le premier est petit si t est petit. On a donc bien que $u(t)$ tend vers u_0 dans L^2 quand t tend vers 0.

Le comportement asymptotique de u est clair car, quand t tend vers $+\infty$,

$$|u(x, t) - d_0| \leq \sum_{k \geq 1} d_k(u_0) e^{-a(k\pi/L)^2 t} \leq e^{-a(\pi/L)^2(t-1)} \left(\sum_{k \geq 1} d_k(u_0) e^{-a(k\pi/L)^2} \right).$$

Pour obtenir l'unicité, imaginons que v soit une autre solution telle que décrite au début du théorème 5.4. On forme $w = u - v$ qui est aussi solution de (5.5) mais avec $u_0 = 0$. On considère l'énergie $E(t) = \int_0^L |w(x, t)|^2 dx$ et on vérifie que $E'(t) = -a \int_0^L |\partial_x w(x, t)|^2 dx \leq 0$. Donc $E(t)$ décroît en temps et comme $E(0) = 0$, c'est que $w \equiv 0$ pour tout temps et donc que $v = u$. \square

2.3 Autres conditions aux bords

Si on met d'autres conditions aux bords que Neumann, il suffit d'utiliser les décompositions particulières des théorèmes 5.3 et 5.1 à la place du théorème 5.2. Notons que ces décompositions sont chacune bien adaptées aux conditions aux bords considérées.

Théorème 5.5. *Pour tout $u_0 \in \mathcal{C}_{pm}^0([0, L], \mathbb{R})$ continue par morceau, il existe une unique solution u de*

$$\begin{cases} \partial_t u(x, t) = a \partial_{xx}^2 u(x, t) & (x, t) \in]0, L[\times]0, +\infty[\\ u(0, t) = u(L, t) = 0 & t > 0 \\ u(x, 0) = u_0(x) & x \in [0, L] \end{cases} \quad (5.7)$$

qui est de classe \mathcal{C}^1 en temps et \mathcal{C}^2 en espace pour $(x, t) \in]0, L[\times]0, +\infty[$ et qui a pour donnée initiale u_0 dans le sens que $\|u(\cdot, t) - u_0\|_{L^2}$ tend vers 0 quand t tend vers 0.

En outre, cette solution u est de classe \mathcal{C}^∞ pour $(x, t) \in]0, L[\times]0, +\infty[$ et est donnée par

$$u(x, t) = \sum_{k \geq 1} e_k(u_0) e^{-a(k\pi/L)^2 t} \sin\left(k \frac{\pi}{L} x\right), \quad (5.8)$$

où $e_k(u_0)$ est le coefficient de la décomposition de u_0 suivant le théorème 5.3. En particulier, $u(x, t)$ tend uniformément et exponentiellement vite vers 0.

Dans le cas des conditions périodiques, il faut plutôt s'imaginer sur un cercle. On note \mathbb{S}^1 le cercle de longueur 2π et une fonction u de classe \mathcal{C}^k sur \mathbb{S}^1 est une fonction de classe \mathcal{C}^k sur $[0, 2\pi]$ telle que $u^{(i)}(2\pi) = u^{(i)}(0)$ pour $i = 1, \dots, k$.

Théorème 5.6. *Pour tout $u_0 \in \mathcal{C}_{pm}^0(\mathbb{S}^1, \mathbb{R})$ continue par morceau, il existe une unique solution u de*

$$\begin{cases} \partial_t u(x, t) = a \partial_{xx}^2 u(x, t) & (x, t) \in \mathbb{S}^1 \times]0, +\infty[\\ u(x, 0) = u_0(x) & x \in [0, L] \end{cases} \quad (5.9)$$

qui est de classe C^1 en temps et C^2 en espace pour $(x, t) \in \mathbb{S}^1 \times]0, +\infty[$ et qui a pour donnée initiale u_0 dans le sens que $\|u(\cdot, t) - u_0\|_{L^2}$ tend vers 0 quand t tend vers 0.

En outre, cette solution u est de classe C^∞ pour $(x, t) \in \mathbb{S}^1 \times]0, +\infty[$ et est donnée par

$$u(x, t) = \sum_{k \in \mathbb{Z}} c_k(u_0) e^{-ak^2 t} e^{ikx}, \quad (5.10)$$

où $c_k(u_0)$ est le coefficient de la décomposition de u_0 suivant le théorème 5.1. En particulier, $u(x, t)$ tend uniformément et exponentiellement vite vers la température constante $c_0(u_0)$ qui est la température moyenne de la distribution de température initiale.

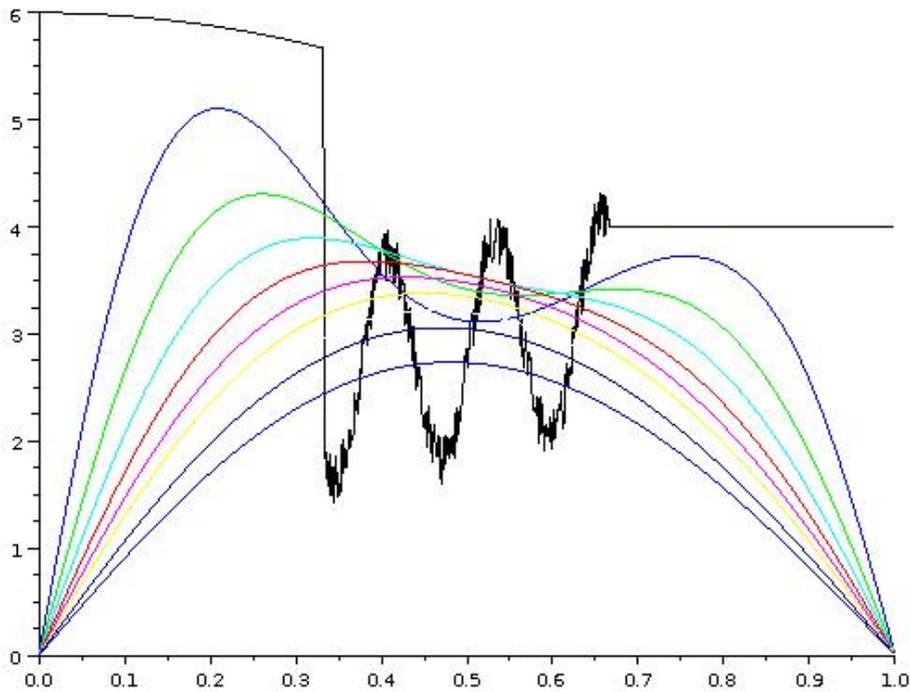


FIGURE 5.1 – Une simulation de l'équation de la chaleur sur le segment $[0, 1]$ avec des conditions aux bords de Dirichlet. On voit nettement le profil de température devenir très régulier et tendre vers 0 en prenant la forme du premier mode $\sin(\pi x)$.

3 Equation de Laplace

Imaginons un fil tendu entre deux points de même hauteur, qu'on supposera positionnés en $x = 0$ et $x = L$. L'écart entre la position du fil en $x \in]0, L[$ et $t \geq 0$ et la position horizontale sera noté $u(x)$ et sera supposé petit. La partie verticale de la tension du fil entre le segment $[x, x + \delta x]$ et le segment $[x + \delta x, x + 2\delta x]$ est (en première approximation) proportionnelle à la pente $(u(x + \delta x) - u(x))/\delta x$. De même, la partie verticale de la tension avec le segment $[x - \delta x, x]$ est proportionnelle à la pente $(u(x - \delta x) - u(x))/\delta x$. Supposons en outre que le fil est pesant, ou que l'on a accroché des masses dessus. Il s'exerce aussi sur le segment de fil un poids $-m(x)\delta x g$ où m est la masse linéique. Le bilan des forces donne qu'à l'équilibre

$$k \frac{u(x + \delta x) - u(x)}{\delta x} + k \frac{u(x - \delta x) - u(x)}{\delta x} - m(x)\delta x g = 0 ,$$

où k est un coefficient de tension. En faisant tendre δx vers 0 et en tenant compte des conditions aux bords, on obtient que la forme u du fil au repos est donnée par

$$\begin{cases} \partial_{xx}^2 u(x) = \frac{g}{k} m(x) & x \in]0, L[\\ u(0) = u(L) = 0 \end{cases} \quad (5.11)$$

Théorème 5.7. *Pour tout $m \in \mathcal{C}_{pm}^0([0, L], \mathbb{R})$ continue par morceau, on peut donner une solution formelle de (5.11) par*

$$u(x) = -\frac{g}{k} \sum_{n \geq 1} e_n(m) \left(\frac{L}{n\pi} \right)^2 \sin \left(n \frac{\pi}{L} x \right) , \quad (5.12)$$

où $e_n(m)$ est le coefficient de la décomposition de m suivant le théorème 5.3.

De plus, si $\sum |e_n(u_0)|$ converge, par exemple si m est continue, de classe \mathcal{C}^1 par morceaux et $m(0) = m(L) = 0$, alors la solution u ci-dessus est de classe \mathcal{C}^2 et vérifie (5.11) de façon rigoureuse. En outre, c'est la seule fonction à le faire.

Démonstration : On commence par noter que le coefficient de la série (5.12) est au moins en $1/n^2$ et engendre donc une série convergente de fonctions continues. D'où u est continue et vaut 0 en 0 et L . Le fait que u vérifie (5.11) au sens formelle vient simplement que $u''(x) = gm(x)/k$ uniquement si on s'autorise à dériver sous le signe somme, ce qui n'est pas possible en général (par exemple si m n'est pas continue en x_0 , alors u n'est pas deux fois dérivable en x_0). Pour pouvoir appliquer le théorème de dérivation 3.14, il faut que les coefficients $e_k(m)$ décroissent suffisamment vite, ce qui est le cas si m est continue, de classe \mathcal{C}^1 par morceaux et si $m(0) = m(L) = 0$. Dans ce cas, l'unicité de la solution vient que si v est une autre solution, on doit avoir $(u - v)''(x) = 0$ et donc $u - v$ est une droite et doit valoir 0 en 0 et L , d'où $u - v = 0$. \square

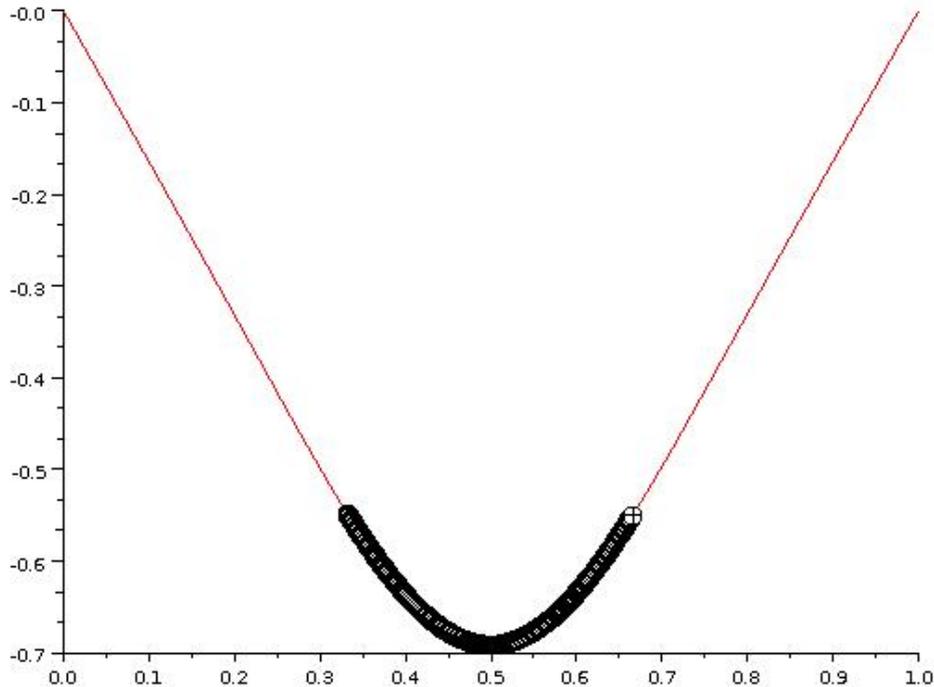


FIGURE 5.2 – Simulation pour une corde dont seul le tiers central est pesant, de masse linéique constante. Il a suffi des 5 premiers modes de Fourier pour obtenir cette approximation.

On notera que, contrairement à l'équation de la chaleur, les séries de Fourier ne donne ici qu'une solution dans un sens formel. De fait, si m n'est pas continue (comme par exemple dans la figure 5.2), on ne pourra pas trouver de fonction u deux fois dérivables telle que $u''(x) = m(x)$. Si m est seulement continue par morceaux, il faut trouver une façon de définir rigoureusement ce qu'est une solution de (5.11). C'est possible et c'est ce qu'on appelle une solution *au sens des distributions*, mais ce n'est pas le sujet de ce cours. Un des intérêts du théorème 5.7 est donc de donner une solution raisonnable à (5.11), même dans le cas où il n'y a pas de solutions au sens classique.

Notons aussi qu'il existe une formule plus simple pour la solution de (5.11) qui est

$$u(x) = \int_0^x \int_0^y \frac{g}{k} m(s) ds dy - x \int_0^1 \int_0^y \frac{g}{k} m(s) ds dy .$$

puisque $\int_0^x \int_0^y gm(s)/k ds dy$ est une primitive seconde de $gm(x)/k$ s'annulant en 0. Les

raisons pour lesquelles la représentation en série de Fourier reste pertinente sont :

- pour des calculs numériques, la décomposition en modes de Fourier peut-être plus rapide et efficace. Ainsi, la figure 5.2 est faite à partir de juste 5 modes de Fourier, donc 5 coefficients et donc 5 intégrales à calculer pour obtenir tout le tracé. A l’opposé, l’expression ci-dessus demande de calculer une intégrale différente pour chaque point x .
- pour un problème similaire en dimension 2 (une nappe tendue), seule la méthode de type Fourier reste transposable.

4 Equation des ondes

On regarde une corde vibrante de longueur L fixée à un support par ses extrémités. Les petits déplacements verticaux de la corde sont régis par l’équation des ondes

$$\begin{cases} \partial_{tt}^2 u(x, t) = \frac{k}{m} \partial_{xx}^2 u(x, t) & (x, t) \in]0, L[\times \mathbb{R}_+^* \\ u(0, t) = u(L, t) = 0 & t > 0 \\ u(x, 0) = u_0(x), \partial_t u(x, 0) = u_1(x) & x \in]0, L[\end{cases} \quad (5.13)$$

où k est un coefficient de tension et m la masse linéique de la corde. Cette équation se retrouve par la même modélisation que (5.11) en négligeant la pesanteur et en rajoutant le terme cinétique $m \delta x \partial_{tt}^2 u(x, t)$ correspondant à l’accélération. On notera que (5.13) contient deux données initiales : u_0 décrit la position de la corde à l’instant $t = 0$ et u_1 décrit l’impulsion initiale.

Théorème 5.8. *Soit u_0 une fonction continue et de classe \mathcal{C}^1 par morceaux, vérifiant $u_0(0) = u_0(L) = 0$ et soit u_1 une fonction continue par morceaux sur $[0, L]$. Alors, il existe une solution formelle de (5.13) données par*

$$u(x, t) = \sum_{n \geq 1} \left(e_n(u_0) \cos(n\omega t) + e_n(u_1) \frac{1}{n\omega} \sin(n\omega t) \right) \sin\left(n \frac{\pi}{L} x\right), \quad (5.14)$$

où $e_n(u_0)$ et $e_n(u_1)$ correspondent aux coefficients de Fourier de u_0 et u_1 dans la décomposition en sinus du théorème 5.3 et où $\omega = \frac{\pi}{L} \sqrt{\frac{k}{m}}$.

Cette solution est $\frac{2\pi}{\omega}$ -périodique et conserve l’énergie physique

$$E(t) = \frac{1}{2} \int_0^L k |\partial_x u(x, t)|^2 + m |\partial_t u(x, t)|^2 dx$$

Démonstration : Le fait que (5.14) définisse une solution formelle se vérifie facilement si on s’autorise à faire toutes les dérivations sous le signe somme. Voyons à quel point ces opérations sont autorisées. Les hypothèses sur u_0 et u_1 montrent que $\sum n^2 e_n^2(u_0)$ et

$\sum e_n^2(u_1)$ sont des séries convergentes d'après le théorème de Parseval et la démonstration du théorème de dérivation (théorèmes 4.2 et 3.14). On peut donc dériver une fois (5.14) sous le signe somme et considérer la convergence de la série obtenue au sens L^2 . Ici, on admet qu'il existe un théorème de dérivation qui est au théorème de Parseval, ce que le théorème de dérivation 3.14 est au théorème 3.11. Dans ce sens là, on peut par exemple vérifier que u satisfait aux données initiales. Elle satisfait aussi les conditions aux bords car la série de (5.14) converge uniformément dans le sens du théorème 3.11.

On note par contre, que la série dans (5.14) n'est pas dérivable deux fois car rien n'assure que les coefficients $n^2 e_n(u_0)$ et $n e_n(u_1)$ forment des séries convergentes.

Pour vérifier que l'énergie est constante, il suffit de la calculer (toujours en admettant qu'on peut utiliser une dérivation sous le signe somme avec des convergences de type Parseval). Pour simplifier les notations, prenons $L = \pi$ et $\omega = \sqrt{k/m}$ et rappelons que $\int_0^\pi \cos^2(nx) dx = \int_0^\pi \sin^2(nx) dx = \pi/2$. On a

$$\begin{aligned} E(t) &= \frac{1}{2} \sum_{n \geq 1} kn^2 \frac{\pi}{2} (e_n(u_0) \cos(n\omega t) + \frac{1}{n\omega} e_n(u_1) \sin(n\omega t))^2 \\ &\quad + m \frac{\pi}{2} (-n\omega e_n(u_0) \sin(n\omega t) + e_n(u_1) \sin(n\omega t))^2 \\ &= \frac{\pi}{4} \sum_{n \geq 1} kn^2 e_n^2(u_0) + m e_n^2(u_1) . \end{aligned}$$

□

Modélisation des instruments de musique :

L'équation des ondes en dimension un (5.13) intervient pour modéliser de nombreux instruments de musique. Considérons le cas d'une corde de guitare. On pince la corde et on la relâche, c'est-à-dire qu'on prend u_0 du type de h_a (voir figure 4.6), où a va correspondre à la position du doigt, et $u_1 = 0$. La corde vibre alors avec une fréquence $\omega = \frac{\pi}{L} \sqrt{\frac{k}{m}}$ qui est la fréquence du premier mode de Fourier, qu'on appelle *fondamentale*, et qui donne la hauteur de la note entendue. Les autres modes de Fourier oscillent à des fréquences $n\omega$, multiples de la fréquence fondamentale, et sont appelées *harmoniques*. Ces harmoniques vont surtout influencer le timbre de la note. Cette décomposition du son en une fréquence fondamentale ω et des harmoniques de fréquences multiples de la fondamentale est caractéristique des sons que le cerveau interprètera comme étant une note de musique.

Vu le théorème 5.8, la guitare émet donc bien des notes de musique et :

- si L diminue, la note devient plus aigüe. Si L est divisée par deux, la fréquence est doublée et la note entendue augmente d'une octave.
- si la tension k augmente, la note devient plus aigüe.
- plus on pince la corde près du bord ($a \rightarrow 0$ dans la figure 4.6), plus la deuxième harmonique sera forte par rapport à la fondamentale et plus le timbre semblera dur. La figure 5.3 illustre bien cette propriété.

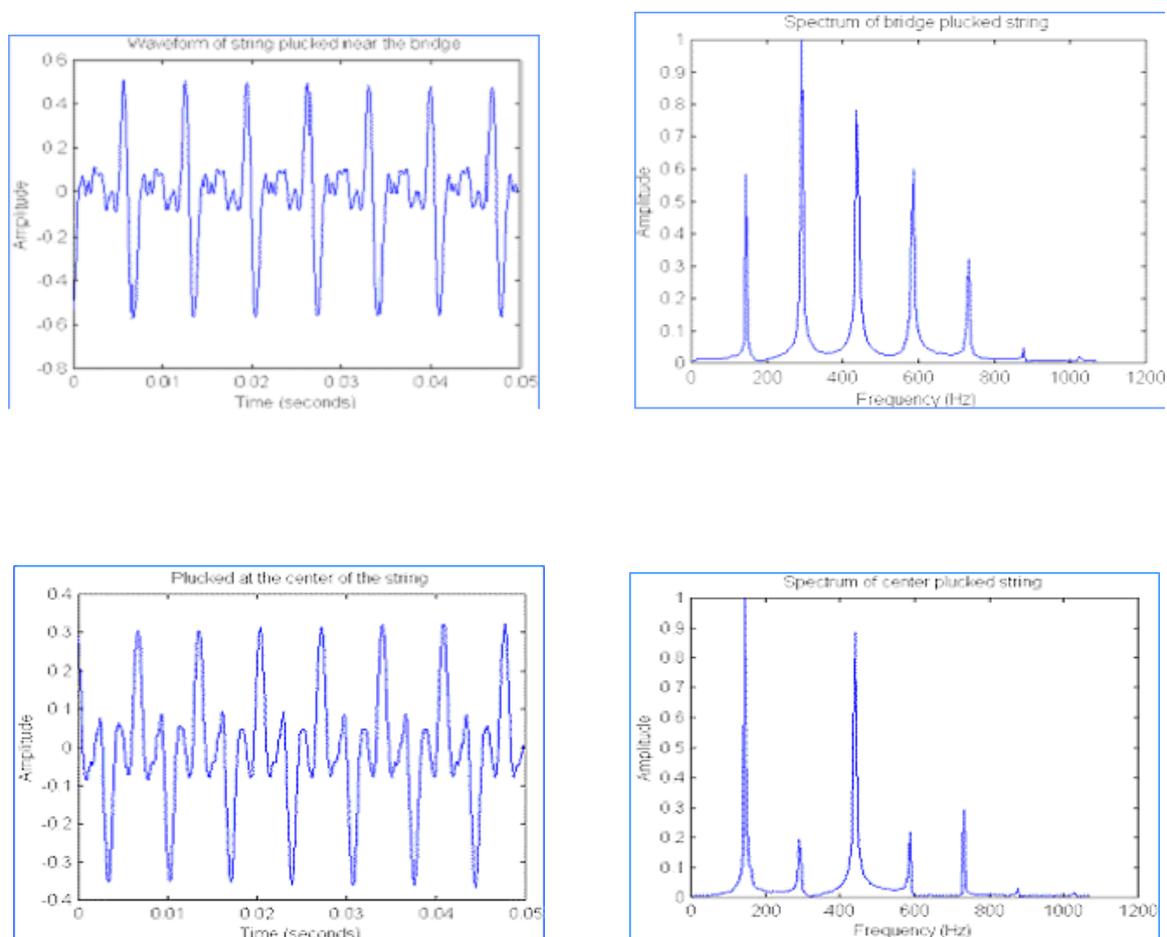


FIGURE 5.3 – A gauche, le son d’une corde de guitare pincée en tant que fonction du temps et à droite le même son vu dans l’espace des fréquences après une transformation de Fourier. On note bien la superposition d’une fréquence fondamentale et d’harmoniques de fréquences multiples. En haut, la corde est pincée près du bord et en bas, elle est pincée au milieu. La différence du spectre des harmoniques correspond bien aux calculs des coefficients de Fourier l’exemple h_a de la figure 4.6 quand $a \rightarrow 0$ ou quand $a = \pi/2$. Images prises sur <http://tehuilami.free.fr/tpe.htm>.