

MATHÉMATIQUE POUR LA PHYSIQUE II

B. Torrèsani
Université de Provence

Licence de Physique, L3

Année Universitaire 2011-12, second semestre



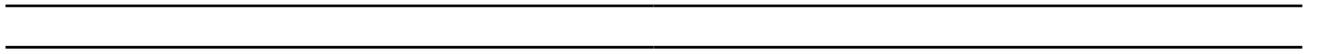


Table des matières

I	Préliminaires : calcul différentiel, calcul des variations	9
1	Calcul des variations	11
1.1	Calcul différentiel, optimisation	11
1.1.1	Dérivées partielles	11
1.1.2	Optimisation	15
1.1.3	Optimisation sous contrainte	18
1.2	Calcul des variations, équations d'Euler-Lagrange, équation de Beltrami	19
1.2.1	Equations d'Euler-Lagrange	19
1.2.2	Equation de Beltrami	21
1.3	Quelques exemples... et contre-exemples simples	22
1.3.1	Le brachistochrone	22
1.3.2	Le Principe de Fermat	25
1.3.3	La colonne de Cox et Overton...	25
1.4	Extensions à des cadres plus généraux	26
1.4.1	Le cas multidimensionnel ; exemples de l'équation de Laplace et de l'équation des ondes	26
1.4.2	Fonctionnelles dépendant de plusieurs fonctions	29
1.4.3	Fonctionnelles de dérivées d'ordre supérieur	30
1.4.4	Conditions aux bords naturelles	31
1.5	Formalisme Hamiltonien	32
1.6	Calcul des variations sous contrainte	33
II	Transformations intégrales	35
2	Transformation de Fourier finie	37
2.1	Préliminaires : équation de la chaleur finie	37
2.2	La transformation de Fourier finie (TFF)	38
2.2.1	Définition, Propriétés, Inversion	38
2.2.2	Une base particulière de \mathbb{C}^N	39
2.2.3	Convolution ; produit	43
2.2.4	Remarques sur la périodicité	44
2.2.5	Transformation de Fourier rapide ; le problème de la complexité	46

2.3	La transformation de Fourier discrète (TFD)	46
2.3.1	Définition, difficultés	46
2.3.2	Convolution et produit	47
2.3.3	Notions de signal et de filtrage numériques	48
2.3.4	Approximation d'une TFD par une TFF	48
2.3.5	Inversion : les séries de Fourier	50
3	Séries de Fourier	53
3.1	Fonctions trigonométriques	53
3.1.1	Polynômes trigonométriques	53
3.1.2	Orthogonalité	54
3.2	Séries de Fourier	54
3.2.1	Généralités	54
3.2.2	Une base orthonormée de $L_p^2([a, b])$	56
3.2.3	Un exemple	57
3.2.4	Quelques remarques	58
3.2.5	Problèmes de convergence des séries de Fourier	61
3.2.6	Le problème de l'extension	63
3.3	Retour sur la transformation de Fourier discrète	66
4	Intégrales de Fourier	69
4.1	La transformation de Fourier des fonctions intégrables	71
4.1.1	Propriétés élémentaires	73
4.1.2	Le théorème de Riemann-Lebesgue	75
4.1.3	Dérivabilité	76
4.1.4	Transformation de Fourier d'une dérivée	77
4.1.5	Le problème de l'inversion	78
4.1.6	La convolution	79
4.2	Les fonctions de carré intégrable	80
4.2.1	La transformation de Fourier dans l'espace de Schwartz $\mathcal{S}(\mathbb{R})$	81
4.2.2	Le passage à $L^2(\mathbb{R})$	82
4.2.3	Le théorème convolution-produit	83
4.2.4	L'inégalité de Heisenberg	83
4.3	Transformation de Fourier des fonctions de plusieurs variables	84
4.3.1	Transformation de Fourier des fonctions intégrables	85
4.3.2	Propriétés simples	85
4.3.3	Convolution-produit	86
4.3.4	Différentiation	87
4.3.5	Passage à $L^2(\mathbb{R}^n)$	87
5	La transformation de Laplace	91
5.1	Définition et premières propriétés	91
5.2	Propriétés élémentaires	93
5.3	La convolution finie	93
5.4	Propriétés de régularité	94
5.4.1	La transformée de Laplace est holomorphe	94
5.4.2	Transformation de Laplace et dérivation	95
5.4.3	Transformation de Laplace et intégration	97
5.5	Inversion de la transformation de Laplace	97
5.6	La transformation de Laplace bilatérale	101

III	Equations différentielles et aux dérivées partielles	105
6	Equations différentielles	107
6.1	Définitions, généralités	108
6.1.1	Généralités, équations linéaires	108
6.1.2	Equations homogènes à coefficients constants	109
6.1.3	Equations homogènes : variation de la constante	110
6.1.4	Facteur intégrant pour les équations linéaires du premier ordre	113
6.1.5	Equations du second ordre, Wronskien	114
6.2	La méthode de Fröbenius	116
6.2.1	Notion de point singulier, théorème de Fuchs	116
6.2.2	Un exemple : les fonctions de Bessel	118
6.2.3	Un autre exemple : les polynômes de Legendre	123
6.2.4	Autres exemples de polynômes orthogonaux	124
6.3	Problème de Sturm-Liouville	125
6.3.1	Généralités	125
6.3.2	Propriétés des fonctions propres et valeurs propres d'une équation de Sturm-Liouville	128
6.4	Résolution d'équations différentielles par transformation	129
6.4.1	Transformation de Fourier	129
6.4.2	Transformation de Laplace	130
7	Equations aux dérivées partielles	133
7.1	EDP du second ordre, classification	133
7.2	Problèmes elliptiques : le Laplacien	134
7.2.1	Le Laplacien	134
7.2.2	Conditions aux bords	136
7.2.3	Le problème de Dirichlet intérieur en dimension 2 pour l'équation de Laplace	137
7.2.4	Problème extérieur, problème dans une couronne	140
7.3	Problèmes paraboliques : phénomènes de diffusion	141
7.3.1	Exemples	141
7.3.2	Le cas de domaines infinis : transformations	142
7.3.3	Domaines bornés : conditions aux bords	142
7.3.4	Equation de la chaleur 1D sur un domaine borné	142
7.3.5	Problèmes inhomogènes : utilisation des bases "Sturm-Liouville"	148
7.4	Problèmes hyperboliques : phénomènes de propagation	149
7.4.1	Exemples : corde et poutre	149
7.4.2	Problèmes sans bord : la solution de d'Alembert	149
7.4.3	Corde vibrante : condition aux bords et résolution	151
7.4.4	Cas bi-dimensionnel : la membrane d'un tambour	153
IV	Distributions, fonctions de Green	157
8	Distributions et fonctions de Green	159
8.1	Distributions : définition, propriétés et exemples	160
8.1.1	Fonctions test et dualité	160
8.1.2	Distributions sur \mathbb{R}	160
8.1.3	Propriétés élémentaires	162
8.2	Distributions sur \mathbb{R}^n	163
8.3	Dérivation des distributions	165

8.3.1	Le cas unidimensionnel	165
8.3.2	Un exemple : la valeur principale de $1/t$	167
8.3.3	Le cas multidimensionnel	168
8.3.4	Surfaces fermées dans \mathbb{R}^2	170
8.3.5	Domaines fermés dans \mathbb{R}^3	171
8.4	Suites de distributions	174
8.5	Produit de convolution des fonctions et des distributions	175
8.5.1	Convolution des fonctions	175
8.5.2	Support d'une distribution	176
8.5.3	Convolution des distributions	177
8.5.4	Equations de convolution	178
8.5.5	Application aux problèmes de valeur initiale	180
8.6	Transformations de Laplace et Fourier des distributions	182
8.6.1	Espace de Schwartz et distributions tempérées	182
8.6.2	Transformation de Laplace des distributions	183
8.6.3	Transformée de Fourier des distributions tempérées	185
8.6.4	Propriétés	186
8.6.5	Transformation de Fourier et dérivation	186
8.6.6	Inversion de la transformation de Fourier	188
8.7	Application au calcul des fonctions de Green	188
8.7.1	Exemple simple : fonction de Green de l'opérateur de dérivation dans \mathcal{D}'_+	189
8.7.2	Oscillateur harmonique 1D	189
8.7.3	Un exemple multidimensionnel : l'opérateur de la chaleur	190

V Opérateurs linéaires 193

9 Opérateurs linéaires 195

9.1	Introduction, rappels	195
9.1.1	Formes sesquilinéaires, formes Hermitiennes	195
9.1.2	Bases, matrices	196
9.1.3	Rang et noyau	197
9.1.4	Réduction d'une forme Hermitienne	198
9.1.5	Espaces pré-Hilbertiens, espaces Hermitiens	200
9.2	Opérateurs linéaires dans les espaces Hermitiens	201
9.2.1	Définitions	201
9.2.2	Opérateur adjoint	203
9.2.3	Diagonalisation des opérateurs autoadjoints	204
9.2.4	Diagonalisation simultanée d'opérateurs auto-adjoints qui commutent	207
9.3	Opérateurs linéaires dans les espaces de Hilbert	208
9.3.1	Généralités	208
9.3.2	Adjoint d'un opérateur sur un espace de Hilbert	210
9.3.3	Spectre d'un opérateur sur un espace de Hilbert	212
9.4	Théorie spectrale des opérateurs compacts	216
9.5	Quelques mots sur les opérateurs non bornés	218
9.6	Eléments de théorie des perturbations	220
9.6.1	Perturbation d'un vecteur propre non-dégénéré d'un opérateur linéaire auto-adjoint sur un espace Hermitien	220
9.6.2	Le cas des valeurs propres dégénérées	222

VI	Probabilités	223
10	Introduction aux Probabilités	225
10.1	Modélisation probabiliste	225
10.1.1	Généralités	225
10.1.2	Calcul des probabilités	227
10.1.3	Exemples	233
10.2	Variables aléatoires	239
10.2.1	Généralités	239
10.2.2	Variables aléatoires discrètes	240
10.2.3	Variables aléatoires à densité	248
10.3	Variables aléatoires indépendantes, distribution conditionnelle	254
10.3.1	Variables aléatoires indépendantes	254
10.3.2	Sommes de variables aléatoires indépendantes	258
10.3.3	Variables aléatoires et probabilités conditionnelles	261
10.3.4	Espérance conditionnelle et prédiction	264
VII	Annexes	267
A	Mesure et convergence	A.1
A.1	Notion de mesure	A.1
A.2	Convergence	A.2
A.2.1	Convergence simple	A.2
A.2.2	Convergence uniforme	A.2
A.2.3	Convergence en moyenne	A.2
A.2.4	Convergence simple presque partout	A.3
B	Espaces fonctionnels	B.1
B.1	Espaces caractérisant la régularité	B.1
B.2	Les espaces $C^\infty(\mathbb{R})$, $\mathcal{D}(\mathbb{R})$ et $\mathcal{S}(\mathbb{R})$	B.2
B.3	Eléments de théorie de l'intégration	B.2
B.4	Les espaces de Lebesgue	B.4
B.5	Relations entre $\mathcal{S}(\mathbb{R})$, $\mathcal{D}(\mathbb{R})$ et $L^2(\mathbb{R})$	B.5
C	Elements de géométrie des courbes et surfaces	C.1
C.1	Courbes, courbes paramétrées	C.1
C.1.1	Courbes planes	C.1
C.1.2	Courbes gauches	C.2
C.2	Surfaces paramétrées dans l'espace	C.4
D	Théorie des perturbations	D.1
D.1	Comparaisons : notations de Landau	D.1
D.1.1	Négligeabilité	D.1
D.1.2	Equivalence	D.1
D.1.3	Domination	D.2
D.2	Perturbations, équations algébriques	D.2

Première partie

Préliminaires : calcul différentiel, calcul des variations

Calcul des variations

De nombreux problèmes de la physique peuvent être posés sous ce que l'on appelle une *formulation variationnelle*, c'est à dire comme la recherche d'un optimum parmi une famille de solutions possibles. C'est notamment le cas de la mécanique de Hamilton (principe de moindre action), mais aussi du principe de Fermat en optique et de multiples autres exemples.

Dans les problèmes d'optimisation simples, la solution (l'optimum) est la plupart du temps obtenu en cherchant les zéros de la dérivée de la fonction à optimiser, ou de son gradient dans les cas de dimension supérieure. Ceci fournit dans la plupart des cas des optima locaux, parmi lesquels on doit trouver le vrai optimum.

On va s'intéresser ici à la situation de dimension infinie, dans laquelle on sera amené à étendre la notion de gradient à ce cadre plus large.

1.1 Calcul différentiel, optimisation

Classiquement, le calcul des variations s'applique pour rechercher des minima ou maxima d'une fonction. Il consiste à partir d'une solution donnée, et examiner les solutions proches pour voir si elles sont meilleures ou moins bonnes. Ceci se fait à l'aide de la dérivation.

Il est utile de rappeler dans un premier temps les bases du calcul différentiel sur les fonctions de plusieurs variables.

1.1.1 Dérivées partielles

Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, et soit $U \subset \mathbb{R}^n$. Soit $\underline{x}_0 \in U$.

– f admet une dérivée en \underline{x}_0 , suivant $\underline{v} \in \mathbb{R}^n$ si la fonction d'une variable réelle

$$\varphi_{\underline{v}} : t \in \mathbb{R} \rightarrow \varphi_{\underline{v}}(t) = f(\underline{x}_0 + t\underline{v})$$

est dérivable en $t = 0$. On note alors

$$D_{\underline{v}}f(\underline{x}_0) = \lim_{t \rightarrow 0} \frac{1}{t} (f(\underline{x}_0 + t\underline{v}) - f(\underline{x}_0)) \quad (1.1)$$

la dérivée directionnelle, dans la direction \underline{v} de f .

– Etant donnée la base canonique $\mathcal{B} = \{\underline{e}_1, \dots, \underline{e}_n\}$ de \mathbb{R}^n , les dérivées partielles de f en \underline{x}_0 sont les nombres

$$\frac{\partial f}{\partial x_i}(\underline{x}_0) = D_{\underline{e}_i}f(\underline{x}_0). \quad (1.2)$$

– Les fonctions

$$\underline{x} \rightarrow \frac{\partial f}{\partial x_i}(\underline{x})$$

sont également appelées fonctions dérivées partielles de f .

DÉFINITION 1.1 Soit $U \subset \mathbb{R}^n$. $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est de classe C^1 sur U si les fonctions dérivées partielles $\partial f / \partial x_i$ existent et sont continues sur U .

La base du calcul différentiel est la formule de Taylor au premier ordre, donnée dans la proposition suivante

PROPOSITION 1.1 (TAYLOR) Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, de classe C^1 sur $U \subset \mathbb{R}^n$. Soit $\underline{x}_0 \in U$, notons $U_0 = \{\underline{h}, \underline{x}_0 + \underline{h} \in U\}$. Alors il existe une application $\epsilon : U_0 \rightarrow \mathbb{R}^m$, telle que

$$f(\underline{x}_0 + \underline{h}) = f(\underline{x}_0) + \sum_{j=1}^n h_j \frac{\partial f}{\partial x_j}(\underline{x}_0) + |\underline{h}| \epsilon(\underline{h}), \quad (1.3)$$

où $|\underline{h}| = \sqrt{h_1^2 + h_2^2 + \dots + h_n^2}$ est la norme Euclidienne de \underline{h} , et

$$\lim_{\underline{h} \rightarrow 0} \epsilon(\underline{h}) = 0. \quad (1.4)$$

Le second terme du développement forme ce que l'on appelle l'application différentielle de f en x_0 , définie plus précisément comme suit.

DÉFINITION 1.2 Avec les mêmes notations que ci-dessus, l'application

$$d_{\underline{x}_0} f : \underline{h} \in U_0 \rightarrow d_{\underline{x}_0} f(\underline{h}) = \sum_{j=1}^n h_j \frac{\partial f}{\partial x_j}(\underline{x}_0) \in \mathbb{R}^m \quad (1.5)$$

est appelée **différentielle** de f en \underline{x}_0 (ou **application tangente** à f en \underline{x}_0).

REMARQUE 1.1 On note généralement

$$x_j : \underline{h} = (h_1, \dots, h_n) \in \mathbb{R}^n \rightarrow h_j \in \mathbb{R}$$

l'application "projection" sur la j -ième composante. Pour tous $\underline{y} \in \mathbb{R}^n$ et $\underline{h} \in \mathbb{R}^n$, on a $x_j(\underline{y} + \underline{h}) = y_j + h_j$, et donc

$$d_{\underline{y}} x_j(\underline{h}) = h_j,$$

de sorte que l'on peut noter $dx_j = d_{\underline{y}} x_j$, et ainsi écrire la différentielle de f sous la forme classique

$$d_{\underline{x}_0} f = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\underline{x}_0) dx_j. \quad (1.6)$$

EXEMPLE 1.1 (LONGUEUR DU GRAPHE D'UNE FONCTION) On considère une fonction continûment différentiable $f : [a, b] \rightarrow \mathbb{R}$, et la fonction $\ell = u \rightarrow \ell(u)$ qui à $u \in [a, b]$ associe la longueur de la courbe définie par le graphe de f entre a et u . En approximant la courbe par une fonction affine par morceaux, affine sur des intervalles $[x_k, x_{k+1}]$ de longueur h , on obtient, par le théorème de Pythagore,

$$\begin{aligned}\ell(u) &= \lim_{h \rightarrow 0} \sum_k \sqrt{(x_{k+1} - x_k)^2 + (f(x_{k+1}) - f(x_k))^2} \\ &= \lim_{h \rightarrow 0} \sum_k (x_{k+1} - x_k) \sqrt{1 + \left(\frac{f(x_{k+1}) - f(x_k)}{x_{k+1} - x_k} \right)^2}.\end{aligned}$$

Le théorème des accroissements finis montre que pour tout k , il existe $\tau_k \in [x_k, x_{k+1}]$ tel que $f(x_{k+1}) - f(x_k) = f'(\tau_k)(x_{k+1} - x_k)$. En passant à la limite $h \rightarrow 0$, on obtient donc (par définition de l'intégrale de Riemann) l'expression de la fonction ℓ :

$$\ell(u) = \int_a^u \sqrt{1 + f'(x)^2} dx \quad (1.7)$$

et la longueur $L = \ell(b)$ de la courbe.

De plus,

$$\frac{d\ell}{dx}(x_0) = \sqrt{1 + f'(x_0)^2},$$

et la différentielle en x_0 de ℓ vaut donc

$$d_{x_0}\ell = \frac{d\ell}{dx}(x_0)dx.$$

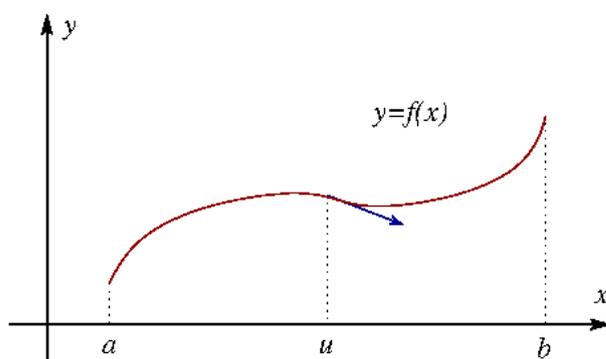


FIG. 1.1: Fonction d'une variable et longueur du graphe

DÉFINITION 1.3 Avec les mêmes définitions que ci-dessus, la **matrice Jacobienne** de f en \underline{x}_0 est la matrice

$$J_f(\underline{x}_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}. \quad (1.8)$$

Lorsque $m = n$, le déterminant de la matrice Jacobienne est appelé **Jacobien** de f en \underline{x}_0 .

La matrice Jacobienne permet une écriture synthétique de la formule de Taylor :

$$f(\underline{x}_0 + \underline{h}) = f(\underline{x}_0) + J_f(\underline{x}_0)\underline{h} + \underline{h}\epsilon(\underline{h}). \quad (1.9)$$

La différentielle et la matrice Jacobienne possèdent d'importantes propriétés vis à vis de la composition des fonctions.

THÉORÈME 1.1 Soient $U \subset \mathbb{R}^n$, $V \subset \mathbb{R}^m$, $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, de classe C^1 dans U , telle que $f(U) \subset V$, et $g : \mathbb{R}^m \rightarrow \mathbb{R}^p$, de classe C^1 dans V . Alors $g \circ f$ est de classe C^1 dans U , et on a

$$d_{\underline{x}_0}(g \circ f) = d_{f(\underline{x}_0)}g \circ d_{\underline{x}_0}f, \quad (1.10)$$

$$J_{g \circ f}(\underline{x}_0) = J_g(f(\underline{x}_0))J_f(\underline{x}_0). \quad (1.11)$$

Les dérivées d'ordre supérieur sont définies récursivement. Par exemple, étant donnée une fonction de plusieurs variables $f : \mathbb{R}^n \rightarrow \mathbb{C}$, on définit

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\underline{x}_0) = \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j}(\underline{x}_0). \quad (1.12)$$

Pour des dérivées d'ordres peu élevés, on utilisera également la notation suivante

$$f'_{x_i} = \frac{\partial f}{\partial x_i}, \quad f''_{x_i x_j} = \frac{\partial^2 f}{\partial x_i \partial x_j}, \dots$$

Le résultat essentiel est le théorème de Schwarz, qui stipule que lorsque toutes les dérivées partielles considérées sont continues, le résultat d'une dérivation multiple ne dépend pas de l'ordre dans lequel les dérivations sont effectuées. Plus précisément :

THÉORÈME 1.2 (SCHWARZ) Soit $U \subset \mathbb{R}^n$, soit $\underline{x}_0 \in U$, et soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Si f est de classe C^1 dans U , si $f''_{x_i x_j}$ et $f''_{x_j x_i}$ existent et sont continues en \underline{x}_0 , alors

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\underline{x}_0) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\underline{x}_0). \quad (1.13)$$

REMARQUE 1.2 Ce résultat semble évident. Cependant, il est important de noter qu'il n'est plus valable lorsque les hypothèses sont violées. Par exemple, prenons

$$f(x, y) = \begin{cases} \frac{xy^3}{x^2+y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon} \end{cases}$$

Les dérivées partielles premières sont :

$$\frac{\partial f}{\partial x}(x, y) = \begin{cases} y^3 \frac{y^2 - x^2}{(x^2 + y^2)^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon} \end{cases}, \quad \frac{\partial f}{\partial y}(x, y) = \begin{cases} xy^2 \frac{3x^2 + y^2}{(x^2 + y^2)^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon,} \end{cases}$$

de sorte que

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = 0 \quad \text{alors que} \quad \frac{\partial^2 f}{\partial y \partial x}(0, 0) = 1.$$

DÉFINITION 1.4 On dira que $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est de classe C^k dans $U \subset \mathbb{R}^n$ si toutes les dérivées partielles

$$\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}}$$

de degré $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_n \leq k$ existent et sont continues dans U .

1.1.2 Optimisation

Optimiser une fonction de plusieurs variables équivaut à en chercher les extrêma, c'est à dire les maxima et les minima.

DÉFINITION 1.5 Soit $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$.

- f admet un maximum local (resp. maximum local strict) en $\underline{x}_0 \in X$ si il existe un voisinage V de \underline{x}_0 , telle que pour tout $\underline{x} \in V$, $f(\underline{x}) \leq f(\underline{x}_0)$ (resp. $f(\underline{x}) < f(\underline{x}_0)$).
- f admet un minimum local (resp. minimum local strict) en $\underline{x}_0 \in X$ si il existe un voisinage V de \underline{x}_0 , telle que pour tout $\underline{x} \in V$, $f(\underline{x}) \geq f(\underline{x}_0)$ (resp. $f(\underline{x}) > f(\underline{x}_0)$).
- f admet un maximum global (resp. maximum global strict) en $\underline{x}_0 \in X$ si pour tout $\underline{x} \in X$, $f(\underline{x}) \leq f(\underline{x}_0)$ (resp. $f(\underline{x}) < f(\underline{x}_0)$).
- f admet un minimum global (resp. minimum global strict) en $\underline{x}_0 \in X$ si pour tout $\underline{x} \in X$, $f(\underline{x}) \geq f(\underline{x}_0)$ (resp. $f(\underline{x}) > f(\underline{x}_0)$).

Une notion centrale pour la recherche d'extrêma locaux est la notion de **point critique**. On dit que \underline{x}_0 est un point critique de la fonction f (supposée de classe C^1 dans un voisinage de \underline{x}_0) si pour tout $i = 1, \dots, n$, on a

$$\frac{\partial f}{\partial x_i}(\underline{x}_0) = 0, \quad i = 1, \dots, n,$$

ce que l'on note

$$\nabla f(\underline{x}_0) = 0.$$

Les points critiques caractérisent les extrêma locaux d'une fonction f à l'ordre 1. Cependant, le gradient ∇f de f ne permet pas de décider si un point critique est effectivement un extrêmu, ni si il s'agit d'un minimum ou d'un maximum. Il est nécessaire pour cela d'effectuer une étude à l'ordre deux. On utilise pour cela le résultat suivant, qui étend la proposition 1.1

THÉORÈME 1.3 (TAYLOR-YOUNG) Soient $U \subset \mathbb{R}^n$, $\underline{x}_0 \in U$ et $f : \mathbb{R}^n \rightarrow \mathbb{R}$, de classe C^2 dans U . On a alors pour tout $\underline{h} \in U$ tel que $\underline{x}_0 + \underline{h} \in U$,

$$f(\underline{x}_0 + \underline{h}) = f(\underline{x}_0) + \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(\underline{x}_0) + \frac{1}{2} \sum_{i,j=1}^n h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(\underline{x}_0) + |\underline{h}|^2 \epsilon(\underline{h}), \quad (1.14)$$

avec $\lim_{|\underline{h}| \rightarrow 0} \epsilon(\underline{h}) = 0$.

Le terme de second ordre dans le théorème ci-dessus fait intervenir la **Matrice Hessienne** (ou Hessienne) de f en \underline{x}_0 :

$$H_f(\underline{x}_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\underline{x}_0) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\underline{x}_0) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\underline{x}_0) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\underline{x}_0) & \frac{\partial^2 f}{\partial x_2^2}(\underline{x}_0) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\underline{x}_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\underline{x}_0) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\underline{x}_0) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\underline{x}_0) \end{pmatrix}. \quad (1.15)$$

La matrice Hessienne est réelle symétrique, elle est donc diagonalisable. Ses valeurs propres (qui sont réelles) possèdent une interprétation simple.

THÉORÈME 1.4 Soit $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction de classe C^2 dans U . Soit $\underline{x}_0 \in U$ un point critique de f .

1. Si la Hessienne $H_f(\underline{x}_0)$ est définie positive (i.e. si ses valeurs propres sont strictement positives), \underline{x}_0 est un minimum local de f .
2. Si la Hessienne $H_f(\underline{x}_0)$ est définie négative (i.e. si ses valeurs propres sont strictement négatives), \underline{x}_0 est un maximum local de f .
3. Si la Hessienne $H_f(\underline{x}_0)$ est indéfinie (avec des valeurs propres non nulles), \underline{x}_0 est un **point selle**.

Notons que si certaines valeurs propres de $H_f(\underline{x}_0)$ sont nulles, il n'est pas possible de conclure.

EXEMPLE 1.2 Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, définie par

$$f(x, y) = x^4 + y^4 - 4xy + 1,$$

et représentée en FIGURE 1.2.

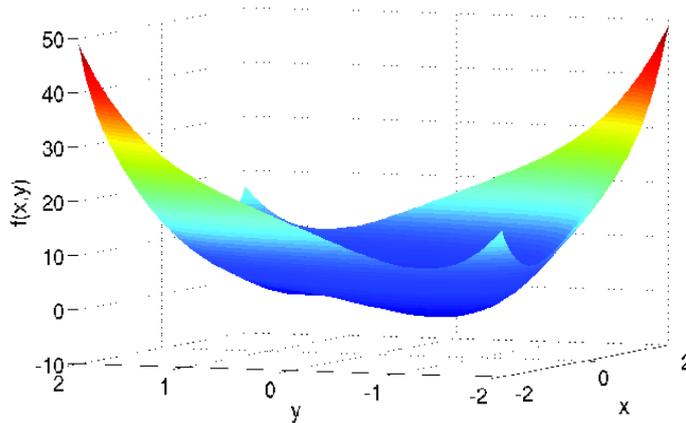


FIG. 1.2: Graph de la fonction $f(x, y) = x^4 + y^4 - 4xy + 1$

On a $f'_x(x, y) = 4x^3 - 4y$ et $f'_y(x, y) = 4y^3 - 4x$. Ainsi les points critiques (x, y) satisfont nécessairement $x^3 = y$ et $y^3 = x$, d'où les trois solutions $(0, 0)$, $(1, 1)$ et $(-1, -1)$. La matrice Hessienne est de la forme

$$H_f(x, y) = \begin{pmatrix} 12x^2 & -4 \\ -4 & 12y^2 \end{pmatrix}.$$

On en déduit aisément que $(0,0)$ est un point selle, alors que $(1,1)$ et $(-1,-1)$ sont des minima locaux.

EXEMPLE 1.3 (LOIS DE SNELL-DESCARTES EN DEUX DIMENSIONS) Un rayon lumineux se propage à la vitesse v_1 dans le milieu 1 (demi-plan supérieur dans la FIG. 1.3), et à la vitesse v_2 dans le milieu 2 (demi-plan inférieur). Le principe de Fermat précise que la lumière suit le trajet le plus économique en temps. En notant x l'abscisse du point où la trajectoire coupe l'interface, le temps nécessaire pour aller de (x_1, z_1) à (x_2, z_2) en passant par le point d'abscisse x vaut

$$T(x) = \frac{\sqrt{(x-x_1)^2 + z_1^2}}{v_1} + \frac{\sqrt{(x_2-x)^2 + z_2^2}}{v_2}.$$

Cette quantité est (localement) optimale quand sa dérivée par rapport à x s'annule, c'est à dire lorsque

$$\frac{x-x_1}{v_1\sqrt{(x-x_1)^2 + z_1^2}} - \frac{x_2-x}{v_2\sqrt{(x_2-x)^2 + z_2^2}} = 0,$$

ce qui conduit à la loi de Snell-Descartes

$$\frac{\sin \theta_1}{v_1} = \frac{\sin \theta_2}{v_2}.$$

On vérifie facilement qu'il s'agit effectivement d'un minimum. On obtient de la même façon la loi de

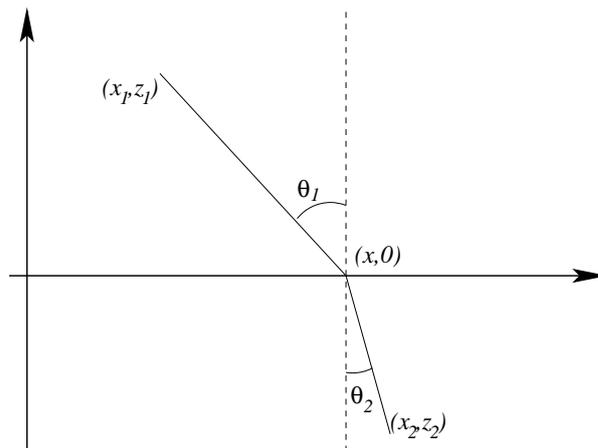


FIG. 1.3: Loi de Snell-Descartes

Snell-Descartes à la réflexion.

EXEMPLE 1.4 (LOIS DE SNELL-DESCARTES EN TROIS DIMENSIONS) Dans le cas tridimensionnel, la situation est similaire. Il s'agit cette fois de déterminer les coordonnées (x, y) dans le plan de l'interface où le rayon coupera celui-ci. Le temps de trajet vaut cette fois

$$T(x, y) = \frac{\sqrt{(x-x_1)^2 + (y-y_1)^2 + z_1^2}}{v_1} + \frac{\sqrt{(x_2-x)^2 + (y_2-y)^2 + z_2^2}}{v_2}.$$

Il s'agit cette fois d'optimiser par rapport à x et y simultanément, ce qui revient à annuler simultanément les dérivées de T par rapport à x et y , c'est à dire son gradient bidimensionnel. Ceci conduit

aux équations

$$\frac{x - x_1}{v_1 \sqrt{(x - x_1)^2 + (y - y_1)^2 + z_1^2}} - \frac{x_2 - x}{v_2 \sqrt{(x_2 - x)^2 + (y_2 - y)^2 + z_2^2}} = 0,$$

$$\frac{y - y_1}{v_1 \sqrt{(x - x_1)^2 + (y - y_1)^2 + z_1^2}} - \frac{y_2 - y}{v_2 \sqrt{(x_2 - x)^2 + (y_2 - y)^2 + z_2^2}} = 0,$$

qui impliquent immédiatement

$$\begin{pmatrix} x_2 - x \\ y_2 - y \end{pmatrix} = \frac{v_2 d_2}{v_1 d_1} \begin{pmatrix} x - x_1 \\ y - y_1 \end{pmatrix},$$

i.e. le point du plan de coordonnées (x, y) se trouve dans le segment compris entre (x_1, y_1) et (x_2, y_2) . On se ramène donc à un problème bidimensionnel, et le raisonnement ci-dessus s'applique directement.

1.1.3 Optimisation sous contrainte

Il arrive que l'on ait à rechercher des extrêma de certaines fonctions de plusieurs variables, auxquelles sont imposées un certain nombre de contraintes supplémentaires. On a dans ce cas recours à la méthode des multiplicateurs de Lagrange, qui est illustrée en FIG. 1.4 dans un exemple bidimensionnel. Dans cet exemple, on cherche à minimiser une certaine fonction F définie sur le plan, dont les lignes de niveau sont tracées en courbes (presque parallèles sur la figure), sous une contrainte prenant la forme $G(x) = C$ (une constante), représentée comme le bord d'un domaine sur la figure). Il apparaît clairement que l'optimum est obtenu lorsque les gradients de F et G sont confondus.

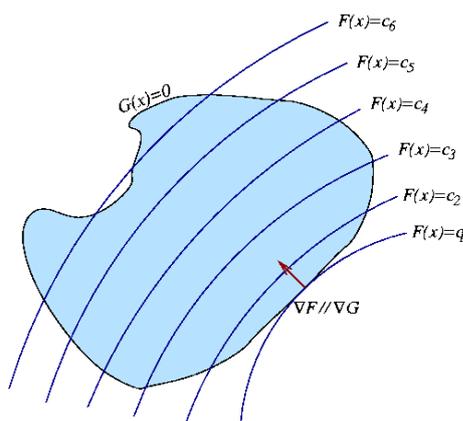


FIG. 1.4: Optimisation sous contrainte par multiplicateurs de Lagrange : les lignes sont les lignes de niveau de la fonction à optimiser, et la contrainte est représentée par le bord de la surface fermée.

Cet exemple peut se généraliser en dimension quelconque, où il prend une forme tout à fait similaire. Plus précisément, le résultat général est le suivant :

THÉORÈME 1.5 Soient f et g deux fonctions différentiables dans $\Omega \subset \mathbb{R}^d$. Soit $\underline{x}_0 \in \Omega$, tel que $\nabla g(\underline{x}_0) \neq 0$. Si f admet un extrémum en \underline{x}_0 , sous la contrainte $g(\underline{x}) = 0$, alors il existe λ_0 tel qu'en posant

$$F(\underline{x}, \lambda) = f(\underline{x}) + \lambda g(\underline{x}),$$

on ait

$$\nabla F(\underline{x}_0, \lambda_0) = 0, \quad \text{et} \quad \frac{\partial F}{\partial \lambda}(\underline{x}_0, \lambda_0) = 0.$$

EXEMPLE 1.5 Cherchons les extrêma de la fonction

$$f(x, y, z) = x^2 + y^2 + z^2$$

sous la contrainte

$$g(x, y, z) = x + y + z - 1 = 0.$$

Pour cela, on forme la fonction étendue

$$F(x, y, z, \lambda) = x^2 + y^2 + z^2 - \lambda(x + y + z - 1),$$

que l'on optimise. Ceci conduit aux équations

$$x = y = z = \lambda/2.$$

En imposant la contrainte $x + y + z = 1$, on obtient la valeur de λ , qui conduit à $x = y = z = 1/3$.

1.2 Calcul des variations, équations d'Euler-Lagrange, équation de Beltrami

Le calcul des variations constitue une généralisation du problème de l'optimisation à des fonctions dépendant non plus d'une ou plusieurs variables réelles, mais aux "fonctions de fonctions" (on parle alors de fonctionnelles). Dans cette section, on décrit assez rapidement les bases du calcul des variations, non seulement dans sa formulation la plus usuelle, qui conduit à la mécanique de Hamilton, mais aussi dans des cadres plus généraux.

1.2.1 Equations d'Euler-Lagrange

Le formalisme d'Euler-Lagrange étend l'approche précédente au cas de **fonctionnelles**, c'est à dire d'applications définies sur un espace de fonctions à valeurs dans \mathbb{R} (ou \mathbb{C})

$$\Phi : f_1, \dots, f_N \rightarrow \Phi(f_1, \dots, f_N) \in \mathbb{R} \quad (\text{ou } \mathbb{C})$$

c'est à dire d'applications associant une valeur numérique à une ou plusieurs fonctions f_1, \dots, f_N , et l'on recherche les fonctions qui minimisent cette fonctionnelle.

On part du cas simple d'une fonctionnelle Φ , dépendant des valeurs d'une unique fonction recherchée f et sa dérivée f' , sous la forme

$$\Phi[f] = \int_{x_1}^{x_2} F(x, f(x), f'(x)) dx. \quad (1.16)$$

La variable x est appelée **variable indépendante**, et les variables $f(x), f'(x)$ sont les **variables dépendantes**.

Le problème posé est de trouver la fonction f telle que le nombre $\Phi[f]$ soit minimal (ou maximal), sous contrainte que les valeurs de f aux bords x_1 et x_2 soient fixées. Dans le cas de des fonctions de plusieurs variables, rechercher les extrêma d'une fonction revient à chercher les points où la dérivée s'annule. C'est cette notion qu'il nous faut généraliser au cas des fonctionnelles. Nous allons rechercher les fonctions f telles que Φ ne varie pas lorsque l'on varie f de façon infinitésimale, dans une "direction" quelconque.

Etudier les variations de f dans une "direction" g (ici, g est donc une fonction) donnée revient à étudier la quantité (la variation)

$$D_g \Phi[f] = \left(\frac{d}{d\epsilon} \Phi[f + \epsilon g] \right)_{\epsilon=0} = \lim_{\epsilon \rightarrow 0} \frac{\Phi[f + \epsilon g] - \Phi[f]}{\epsilon}.$$

La quantité $D_g \Phi$ porte le nom de **dérivée fonctionnelle** de Φ dans la direction g ; l'ensemble des dérivées fonctionnelles de Φ dans "toutes" les directions g (à condition de préciser le sens qu'on donne à l'expression "toutes") peut être vu comme une généralisation de la notion de gradient au cas des fonctionnelles.

On se limite aux fonctions g telles que

$$g(x_1) = g(x_2) = 0.$$

Un calcul formel montre que ceci revient à écrire

$$\int_{x_1}^{x_2} \left(g(x) \frac{\partial F}{\partial f}(x, f(x), f'(x)) + g'(x) \frac{\partial F}{\partial f'}(x, f(x), f'(x)) \right) dx = 0,$$

et une intégration par parties conduit à

$$\int_{x_1}^{x_2} g(x) \left(\frac{\partial F}{\partial f}(x, f(x), f'(x)) + \frac{d}{dx} \frac{\partial F}{\partial f'}(x, f(x), f'(x)) \right) dx = 0.$$

On utilise ici la version unidimensionnelle du **Lemme fondamental du calcul variationnel**

LEMME 1.1 Soit $\Omega \subset \mathbb{R}^n$ un domaine ouvert de l'espace n -dimensionnel, de bord noté $\partial\Omega$, et soit $f \in C(\Omega)$, telle que

$$\int_{\Omega} f(\underline{x}) h(\underline{x}) d\underline{x} = 0$$

pour toute fonction $h \in C^2(\Omega)$ telle que $h(\underline{x}) = 0$ pour tout $\underline{x} \in \partial\Omega$. Alors $f(\underline{x}) = 0$ pour tout $\underline{x} \in \Omega$.

En prenant pour Ω l'intervalle $\Omega =]a, b[$, on en déduit alors le résultat principal :

THÉORÈME 1.6 Considérons la fonctionnelle Φ définie en (1.16), et soit $f \in C^2([a, b])$ un extrênum de cette fonctionnelle, avec des conditions aux bords fixées

$$f(x_1) = f_1, \quad f(x_2) = f_2.$$

Alors f satisfait nécessairement l'équation différentielle

$$\frac{\partial F}{\partial f} - \frac{d}{dx} \frac{\partial F}{\partial f'} = 0, \quad (1.17)$$

appelée **équation d'Euler-Lagrange** associée.

EXEMPLE 1.6 (LA LIGNE DROITE) Etant donné un chemin continu

$$y : x \in [x_1, x_2] \rightarrow y(x)$$

reliant $y_1 = y(x_1)$ à $y_2 = y(x_2)$, sa longueur s'écrit

$$\mathcal{L}[y] = \int_{x_1}^{x_2} \sqrt{1 + y'(x)^2} dx ,$$

c'est à dire $F(x, y, y') = \sqrt{1 + y'(x)^2}$: pas de dépendance explicite en x et y . L'équation d'Euler-Lagrange s'écrit donc

$$\frac{d}{dx} \frac{\partial F}{\partial y'} = \frac{d}{dx} \frac{y'(x)}{\sqrt{1 + y'(x)^2}} = \frac{y''(x)}{(1 + y'(x)^2)^{3/2}} = 0 ,$$

ce qui implique $y''(x) = 0$ pour tout x , i.e. $y(x) = ax + b$, où a et b sont déterminés par les conditions aux bords. On retrouve donc bien une trajectoire rectiligne entre les deux points.

Cet exemple était aussi intéressant car il s'agissait d'un cas où la fonction F ne dépend pas explicitement de f , mais seulement de f' . Alors, l'équation d'Euler-Lagrange prend la forme plus simple

$$\frac{\partial F}{\partial f'} = C^{te} .$$

1.2.2 Equation de Beltrami

Une autre situation dans laquelle les choses se simplifient est celle où F ne dépend pas explicitement de x . Dans ce cas, on peut écrire

$$\frac{d}{dx} F(f(x), f'(x)) = f'(x) \frac{\partial F}{\partial f}(f(x), f'(x)) + f''(x) \frac{\partial F}{\partial f'}(f(x), f'(x)) .$$

L'équation d'Euler-Lagrange implique alors que

$$f'(x) \frac{\partial F}{\partial f}(f(x), f'(x)) = f'(x) \frac{d}{dx} \frac{\partial F}{\partial f'}(f(x), f'(x)) = \frac{d}{dx} [f'(x) \frac{\partial F}{\partial f'}(f(x), f'(x))] - f''(x) \frac{\partial F}{\partial f'}(f(x), f'(x))$$

d'où

$$\begin{aligned} \frac{d}{dx} F(f(x), f'(x)) &= f'(x) \frac{d}{dx} \frac{\partial F}{\partial f'}(f(x), f'(x)) + f''(x) \frac{\partial F}{\partial f'}(f(x), f'(x)) \\ &= \frac{d}{dx} \left(f'(x) \frac{\partial F}{\partial f'}(f(x), f'(x)) \right) . \end{aligned}$$

On en déduit l'équation de Beltrami

PROPOSITION 1.2 Si la fonction F ne dépend pas explicitement de la variable indépendante x , alors l'équation d'Euler-Lagrange se ramène à l'équation de Beltrami

$$F(f(x), f'(x)) - f'(x) \frac{\partial F}{\partial f'}(f(x), f'(x)) = C^{te} . \quad (1.18)$$

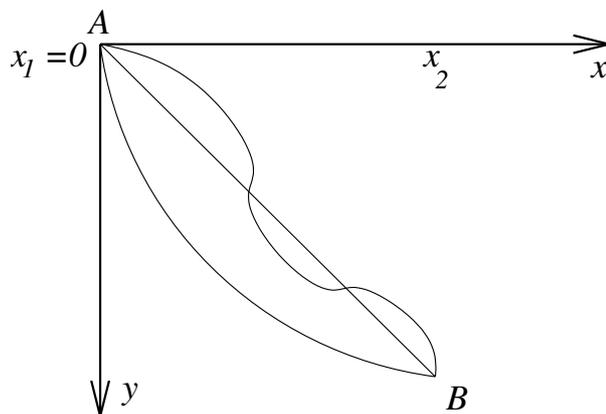


FIG. 1.5: Trajet le plus court (en temps) d'un point à un autre : trois trajectoires possibles, plus ou moins vraisemblables.

Ce résultat a souvent d'importantes implications lorsqu'il s'applique à un contexte physique. Il traduit le fait que lorsque la fonctionnelle à minimiser ne comporte pas de dépendance explicite dans la variable indépendante, alors il existe une **quantité conservée**, ici

$$F - f' \frac{\partial F}{\partial f'} .$$

Par exemple, dans le cas de l'application à la mécanique Hamiltonienne, nous verrons que lorsque le Lagrangien ne dépend pas explicitement du temps (qui est alors la variable indépendante), il existe une quantité conservée, le **Hamiltonien**, qui représente en fait l'énergie du système considéré.

EXEMPLE 1.6 (SUITE). Dans ce cas, on a $F(x, y, y') = F(y') = \sqrt{1 + y'^2}$, et la proposition ci-dessus s'applique. On écrit

$$F(y') - y' \frac{\partial F(y')}{\partial y'} = \sqrt{1 + y'^2} - \frac{y'^2}{\sqrt{1 + y'^2}} = \frac{1}{\sqrt{1 + y'^2}} .$$

Cette quantité est constante si et seulement si y' est constante ce qui nous conduit au même résultat.

1.3 Quelques exemples... et contre-exemples simples

1.3.1 Le brachistochrone

On considère un mobile de masse m , assujéti à glisser sans frottement sur une surface inclinée reliant un point initial A de coordonnées (x_1, z_1) à un point final B de coordonnées (x_2, z_2) , utilisant la seule force de la pesanteur. On suppose que la vitesse initiale du mobile est nulle, et on cherche la forme de la surface inclinée (donc la trajectoire du mobile) qui conduira au trajet le plus court en temps.

On note x la variable horizontale, qui sera notre variable indépendante, et $z = z(x)$ la variable verticale (dépendante), orientée vers le bas. On choisit le système de coordonnées de sorte que A soit placé à l'origine des abscisses et des côtes. L'énergie cinétique et l'énergie potentielle valent respectivement

$$T = \frac{1}{2} m v^2 , \quad V = -mgz + C .$$

La conservation de l'énergie totale implique que

$$E = T + V = C$$

de sorte que la vitesse varie avec l'altitude comme

$$v(z) = \sqrt{2gz}.$$

Le temps total nécessaire pour le trajet $x \in [x_1, x_2] \rightarrow z(x)$ vaut

$$\Phi[z] = \int_{x_1}^{x_2} \frac{ds(x)}{v(z(x))} = \frac{1}{\sqrt{2g}} \int_{x_1}^{x_2} \frac{\sqrt{1+z'(x)^2}}{\sqrt{z(x)}} dx$$

On est bien dans le cas de l'équation de Beltrami (pas de dépendance explicite de F dans la variable indépendante x), qui s'écrit donc

$$\frac{z'(x)^2}{\sqrt{z(x)(1+z'(x)^2)}} - \frac{\sqrt{1+z'(x)^2}}{\sqrt{z(x)}} = C.$$

Quelques manipulations donnent

$$C^2 z(x) (1+z'(x)^2) = 1,$$

d'où, en posant $k = 1/C^2$,

$$\frac{dz}{dx}(x) = \sqrt{\frac{k-z(x)}{z(x)}}.$$

Remarquons que $C^2 z \leq 1$, donc $0 \leq z \leq k$. On peut donc poser $z = k \sin^2 \phi$, d'où on déduit $\sqrt{(k-z)/z} = \cotan \phi$.

Pour exprimer x en fonction de ϕ , calculons

$$\frac{dx}{d\phi} = \frac{dx}{dz} \frac{dz}{d\phi} = 2k \sin^2 \phi = k[1 - \cos(2\phi)].$$

Ainsi, on en déduit (puisque pour $x = 0$ on a aussi $z = 0$)

$$x = k(\phi - \sin(2\phi)/2),$$

d'où on tire une forme paramétrique de la solution

$$\begin{cases} x &= \frac{k}{2}(2\phi - \sin(2\phi)) \\ z &= \frac{k}{2}(1 - \cos(2\phi)), \end{cases} \quad (1.19)$$

ce qui est l'équation paramétrique d'une cycloïde. On peut éliminer le paramètre ϕ , en posant $a = k/2$, et en remarquant que

$$2\phi = \cos^{-1} \left(\frac{a-z}{a} \right),$$

puis en insérant dans l'équation de x

$$x(z) = a \cos^{-1} \left(\frac{a-z}{a} \right) - a \sqrt{1 - \left(\frac{a-z}{a} \right)^2} = a \cos^{-1} \left(\frac{a-z}{a} \right) - \sqrt{2az - z^2}.$$

La valeur numérique de a est alors déterminée par la condition limite $x(z_2) = x_2$.

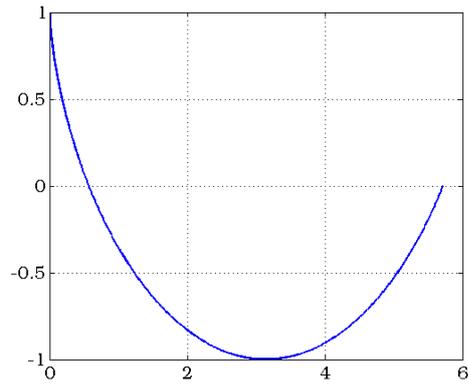


FIG. 1.6: La cycloïde.

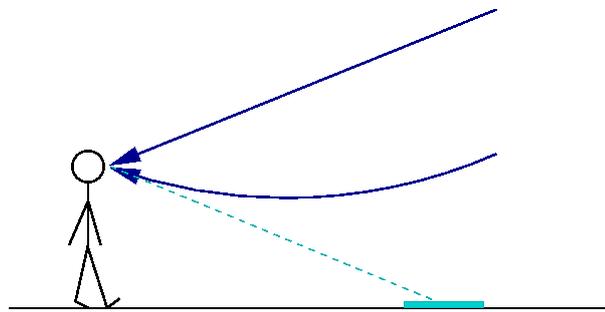


FIG. 1.7: Le phénomène du mirage : deux trajectoires de lumière du ciel (traits pleins) ; la trajectoire du bas est courbée par un gradient de densité (dû, par exemple, à un gradient de température) ; l'oeil perçoit une image virtuelle au sol, qui est en fait une image du ciel.

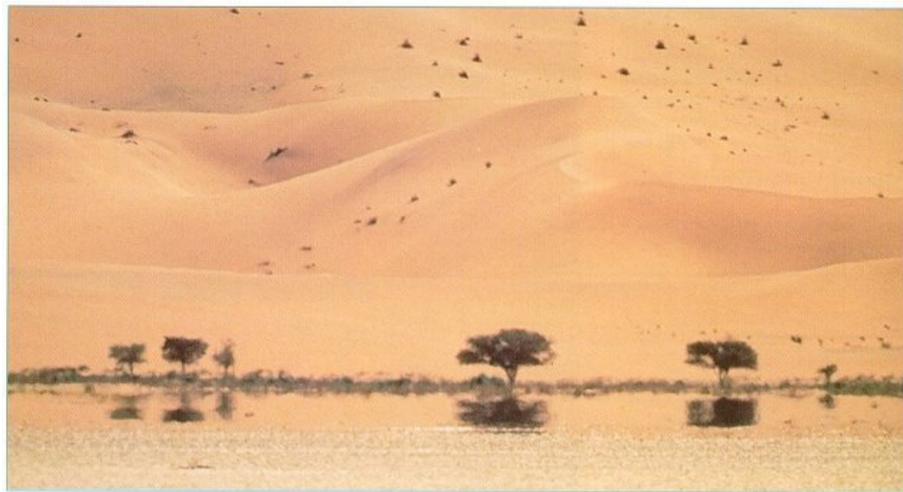


Fig. 2. Mirage inférieur (encore appelé mirage chaud) dans le désert du Namib, en Afrique australe. L'appareil photographique a enregistré deux images : celle correspondant à un trajet rectiligne (MO) et celle correspondant à un trajet avec réflexion totale (MRO).

D'après "Physique Chimie 2e" de A. Tomasino, etc. chez Nathan (1997)

FIG. 1.8: Un vrai mirage

1.3.2 Le Principe de Fermat

Retour sur les lois de Snell-Descartes, dans un cas où l'indice (et donc la vitesse de propagation de la lumière) dépend de la position : le problème des mirages, illustré en FIG. 1.7 et FIG. 1.8

Le temps de trajet dépend du chemin choisi ; en prenant $y = y(x)$, ou alternativement $x = x(y)$, on obtient pour le temps de trajet de la lumière entre deux points l'expression suivante

$$\int n ds = \int f(y) ds = \int f(y) \sqrt{1 + x'(y)^2} dy = \int f(y) \sqrt{1 + y'(x)^2} dx .$$

On privilégie y comme variable indépendante, car on n'a ainsi pas de dépendance explicite dans la variable dépendante x , ce qui simplifie l'équation d'Euler-Lagrange¹. Celle-ci donne

$$f(y) \frac{x'}{\sqrt{1 + x'^2}} = C ,$$

c'est à dire, en supposant $x' \geq 0$ (le cas $x' \leq 0$ se traite similairement, et conduit à la même solution)

$$x' = \frac{C}{\sqrt{f(y)^2 - C^2}} .$$

Pour aller plus loin, il est temps de faire un choix pour la variation de l'indice en fonction de l'altitude. Supposons une dépendance linéaire. Le choix particulier $f(y) = n_0(1 + \alpha y)$ donne

$$dx = \frac{C}{\alpha n_0} \frac{dy}{\sqrt{(y + 1/\alpha)^2 - (C/\alpha n_0)^2}}$$

en posant $u = y + 1/\alpha = (C/\alpha n_0) \cosh \theta$ on obtient $dy = \frac{C}{\alpha n_0} \sinh \theta d\theta$, d'où

$$dx = \left(\frac{C}{\alpha n_0} \right)^2 \frac{\sinh \theta d\theta}{\sqrt{(C/\alpha n_0)^2 \cosh^2 \theta - (C/\alpha n_0)^2}} = \frac{C}{\alpha n_0} d\theta ,$$

et finalement, une simple intégration donne $\theta = (\alpha n_0 / C)(x - x_0)$, d'où

$$y = -\frac{1}{\alpha} + \frac{C}{\alpha n_0} \cosh((\alpha n_0 / C)(x - x_0)) .$$

Il s'agit d'une **chaînette**, ou **caténoïde**, que l'on peut approximer par une parabole au second ordre, autour de x_0 . Les constantes d'intégration x_0 et C peuvent être déterminées en considérant les conditions aux bords.

1.3.3 La colonne de Cox et Overton...

ou même les plus grands peuvent se tromper...

Les équations d'Euler-Lagrange que nous avons dérivées plus haut donnent des conditions nécessaires que doit satisfaire l'optimum de la fonctionnelle considérée, *sous l'hypothèse que celui-ci soit deux fois différentiable*. Il existe des problèmes pour lesquels cette hypothèse n'est pas adaptée, au sens où il existe des solutions f ne la satisfaisant pas, mais donnant à la fonctionnelle une valeur meilleure que la meilleure solution deux fois différentiable.

Un exemple célèbre est donné par un problème étudié par Lagrange en 1783 (initialement formulé par Euler en 1744 puis Bernoulli) : trouver la forme optimale d'une colonne de hauteur et volume fixés, pouvant supporter une pression maximale au sommet. Lagrange eut le grand mérite de donner une formulation mathématique précise au problème, mais la solution qu'il proposa, et qui le conduit à une colonne cylindrique, était entachée d'un certain nombre d'erreur, en particulier une hypothèse de différentiabilité de la solution. Il a en fait fallu attendre 1992 (après plusieurs tentatives, chacune marquée par des erreurs dues à des hypothèses inadaptées) pour que Cox et Overton aboutissent à une solution rigoureuse à ce problème, qui s'est ainsi avéré bien plus difficile qu'il n'y paraissait.

¹On peut vérifier qu'en prenant x comme variable indépendante, l'équation de Beltrami conduit au même résultat.

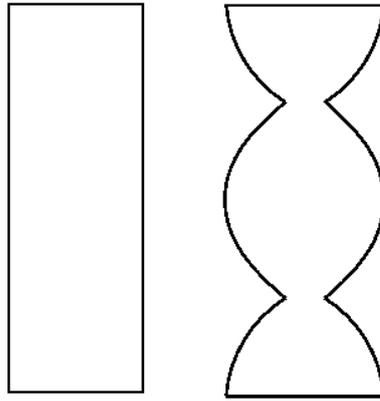


FIG. 1.9: Le problème d'Euler-Bernoulli, la solution de Lagrange (à gauche) et la solution de Cox et Overton (à droite)

1.4 Extensions à des cadres plus généraux

Le cas que nous avons considéré est en fait le cas le plus simple, qui peut être généralisé de diverses manières. Les cas d'intérêt les plus classiques concernent les situations multidimensionnelles, ainsi que les cas de fonctionnelles dépendant de plusieurs fonctions.

1.4.1 Le cas multidimensionnel ; exemples de l'équation de Laplace et de l'équation des ondes

La généralisation la plus immédiate concerne l'extension au cas d'une variable indépendante $\underline{x} \in \Omega \subset \mathbb{R}^d$ vectorielle : la fonctionnelle Φ considérée dépend alors des dérivées partielles de f par rapport aux différentes composantes de x . Pour simplifier, on notera

$$f'_{x_i} = \frac{\partial f}{\partial x_i}, \quad i = 1, \dots, d$$

la dérivée partielle de la fonction f par rapport à la variable x_i ,

$$f''_{x_i x_j} = \frac{\partial^2 f}{\partial x_i \partial x_j}, \quad i, j = 1, \dots, d$$

et ainsi de suite.

Prenons l'exemple bidimensionnel : on se donne un domaine du plan $\Omega \subset \mathbb{R}^2$, dont on note $\partial\Omega$ le bord, et on recherche les extrêma $f : \Omega \rightarrow \mathbb{R}$ de la fonctionnelle Φ définie par

$$\Phi[f] = \int_{\Omega} F(\underline{x}, f(\underline{x}), f'_{x_1}(\underline{x}), f'_{x_2}(\underline{x})) d\underline{x}.$$

Comme dans le cas plus simple, on doit étudier les variations de Φ dans les "directions" g

$$D_g \Phi[f] = \left(\frac{d}{d\epsilon} \Phi[f + \epsilon g] \right)_{\epsilon=0} = \lim_{\epsilon \rightarrow 0} \frac{\Phi(f + \epsilon g) - \Phi(f)}{\epsilon},$$

en se limitant aux fonctions g telles que $g(\underline{x}) = 0$ pour tout $\underline{x} \in \partial\Omega$. En suivant pas à pas le calcul précédent, on aboutit à

$$\int_{\Omega} g(\underline{x}) \frac{\partial F}{\partial f}(\underline{x}, f(\underline{x}), f'_{x_1}(\underline{x}), f'_{x_2}(\underline{x})) d\underline{x} + \int_{\Omega} g'_{x_1}(\underline{x}) \frac{\partial F}{\partial f'_{x_1}}(\underline{x}, f(\underline{x}), f'_{x_1}(\underline{x}), f'_{x_2}(\underline{x})) d\underline{x}$$

$$+ \int_{\Omega} g'_{x_2}(\underline{x}) \frac{\partial F}{\partial f'_{x_2}}(\underline{x}, f(\underline{x}), f'_{x_1}(\underline{x}), f'_{x_2}(\underline{x})) d\underline{x} = 0,$$

et une intégration par parties dans chacun des deux derniers termes, prenant en compte le fait que $g = 0$ sur $\partial\Omega$ conduit à

$$\int_{\Omega} g(\underline{x}) \left(\frac{\partial F}{\partial f}(\underline{x}, f(\underline{x}), f'_{x_1}(\underline{x}), f'_{x_2}(\underline{x})) + \frac{\partial}{\partial x_1} \frac{\partial F}{\partial f'_{x_1}}(\underline{x}, f(\underline{x}), f'_{x_1}(\underline{x}), f'_{x_2}(\underline{x})) \right. \\ \left. + \frac{\partial}{\partial x_2} \frac{\partial F}{\partial f'_{x_2}}(\underline{x}, f(\underline{x}), f'_{x_1}(\underline{x}), f'_{x_2}(\underline{x})) \right) d\underline{x} = 0.$$

Il suffit alors d'utiliser le lemme fondamental du calcul des variations (voir le lemme 1.1), pour en déduire

$$\frac{\partial F}{\partial f} - \frac{\partial}{\partial x_1} \frac{\partial F}{\partial f'_{x_1}} - \frac{\partial}{\partial x_2} \frac{\partial F}{\partial f'_{x_2}} = 0.$$

Plus généralement, on montre le résultat suivant

THÉORÈME 1.7 Soit $\Omega \subset \mathbb{R}^d$, et soit $f \in C^2(\Omega)$ un extrémum de la fonctionnelle Φ définie par

$$\Phi[f] = \int_{\Omega} F(\underline{x}, f(\underline{x}), f'_{x_1}(\underline{x}), \dots, f'_{x_d}(\underline{x})) d\underline{x},$$

avec conditions aux bords $f(\underline{x}) = f_0(\underline{x})$ fixé, pour tout $\underline{x} \in \partial\Omega$. Alors f satisfait nécessairement l'équation d'Euler-Lagrange correspondante

$$\frac{\partial F}{\partial f} - \sum_{i=1}^d \frac{\partial}{\partial x_i} \frac{\partial F}{\partial f'_{x_i}} = 0.$$

On obtient donc un résultat qui est une généralisation directe du Théorème 1.6.

On considère ci-dessous un certain nombre d'exemples d'applications de ce résultat.

Surfaces minimales

Le **problème de Plateau** (ou du film de savon) : que se passe-t-il lorsque l'on plonge un anneau de forme quelconque dans de l'eau savonneuse? Un film de savon se forme à l'intérieur de l'anneau, représentant une surface. La forme de cette surface peut être calculée, partant du principe que la nature cherche à minimiser l'aire de la surface ainsi créée.

On modélise cela de la façon suivante. on considère un domaine $\Omega \subset \mathbb{R}^2$ du plan, de bord $\partial\Omega$ et la hauteur du film de savon au point \underline{x} est notée $z = u(\underline{x})$, $\underline{x} \in \Omega$. L'anneau est modélisé comme une fonction $u_0 : \underline{x} \in \partial\Omega \rightarrow u_0(\underline{x}) \in \mathbb{R}$, et la surface du film de savon représenté par u s'écrit comme

$$\Phi[u] = \int_{\Omega} \sqrt{1 + u'_{x_1}(\underline{x})^2 + u'_{x_2}(\underline{x})^2} d\underline{x}. \quad (1.20)$$

Le problème posé est donc de minimiser Φ , avec la condition au bord

$$u = u_0 \quad \text{sur } \partial\Omega.$$

On a donc

$$F(\underline{x}, u(\underline{x}), u'_{x_1}(\underline{x}), u'_{x_2}(\underline{x})) = \sqrt{1 + u'_{x_1}(\underline{x})^2 + u'_{x_2}(\underline{x})^2}.$$

fonction qui ne dépend explicitement ni de la variable indépendante ni de u . Donc Calculons

$$\frac{\partial F}{\partial u'_{x_1}} = \frac{u'_{x_1}}{\sqrt{1 + u'_{x_1}(\underline{x})^2 + u'_{x_2}(\underline{x})^2}},$$

et une expression similaire pour la dérivée par rapport à u'_{x_2} . Calculons maintenant

$$\frac{\partial}{\partial x_1} \frac{\partial F}{\partial u'_{x_1}}(\underline{x}) = \frac{u''_{x_1 x_1} [1 + u'_{x_1}(\underline{x})^2 + u'_{x_2}(\underline{x})^2] - u'_{x_1}(\underline{x}) (u'_{x_1}(\underline{x}) u''_{x_1 x_1}(\underline{x}) + u'_{x_2}(\underline{x}) u''_{x_1 x_2}(\underline{x}))}{[1 + u'_{x_1}(\underline{x})^2 + u'_{x_2}(\underline{x})^2]^{3/2}},$$

et une expression similaire pour $\frac{\partial}{\partial x_2} \frac{\partial F}{\partial u'_{x_2}}(\underline{x})$.

L'équation d'Euler-Lagrange associée

$$\frac{\partial}{\partial x_1} \frac{\partial F}{\partial u'_{x_1}}(\underline{x}) + \frac{\partial}{\partial x_2} \frac{\partial F}{\partial u'_{x_2}}(\underline{x}) = 0$$

prend donc la forme

$$(1 + u'_{x_1}{}^2) u''_{x_2 x_2} - 2u'_{x_1} u'_{x_2} u''_{x_1 x_2} + (1 + u'_{x_2}{}^2) u''_{x_1 x_1} = 0 \quad (1.21)$$

Equation de Laplace

On cherche à déterminer le champ électromagnétique à l'intérieur d'un domaine $\Omega \subset \mathbb{R}^3$, dont les bords sont fixés à un potentiel ϕ_0 . On note ϕ le potentiel dont dérive le champ électrique E .

L'énergie électromagnétique engendrée par un potentiel ϕ quelconque dans un volume $\Omega \subset \mathbb{R}^3$ est de la forme

$$W[\phi] = \frac{\epsilon_0}{2} \int_{\Omega} |\nabla \phi(\underline{x})|^2 d\underline{x} = \int_{\Omega} F[\phi'_{x_1}(\underline{x}), \phi'_{x_2}(\underline{x}), \phi'_{x_3}(\underline{x})] d\underline{x},$$

avec

$$F[\phi'_{x_1}(\underline{x}), \phi'_{x_2}(\underline{x}), \phi'_{x_3}(\underline{x})] = \frac{\epsilon_0}{2} [\phi'_{x_1}(\underline{x})^2 + \phi'_{x_2}(\underline{x})^2 + \phi'_{x_3}(\underline{x})^2]$$

Pour un potentiel aux bords fixé, le potentiel intérieur est celui qui minimise cette énergie.

En variant $W[\phi]$ par rapport à ϕ , avec conditions aux bords fixées $\phi = \phi_0$ sur $\partial\Omega$, on obtient l'équation d'Euler-Lagrange

$$-\epsilon_0 \sum_{i=1}^3 \frac{\partial^2 \phi}{\partial x_i^2} = -\epsilon_0 \Delta \phi = 0.$$

C'est l'équation de Laplace. En l'absence de conditions aux bords, un potentiel nul (et donc un champ nul) serait une solution, ce qui n'est évidemment plus le cas en présence de conditions aux bords.

Equation des ondes

Soit $\phi : (\underline{x}, t) \in \Omega \subset \mathbb{R}^4 \rightarrow \phi(\underline{x}, t) \in \mathbb{R}$ une fonction de l'espace et du temps, représentant la valeur d'un champ (scalaire) au point \underline{x} et à l'instant t . On suppose que l'énergie associée à ce champ ϕ coïncide avec l'intégrale de la (pseudo)norme de Minkowsky du quadrivecteur des dérivées spatiales et temporelle du champ :

$$I[\phi] = \int_{\Omega} \left[|\nabla \phi(\underline{x}, t)|^2 - \frac{1}{c^2} \left(\frac{\partial \phi(\underline{x}, t)}{\partial t} \right)^2 \right] d\underline{x} dt.$$

En minimisant cette quantité par rapport à ϕ avec une condition aux bords donnée

$$\phi(\underline{x}, t) = \phi_0(\underline{x}, t), \quad \forall \underline{x} \in \partial\Omega$$

Il est facile de voir, par un calcul similaire au précédent, que l'équation d'Euler-Lagrange correspondante est alors l'équation des ondes

$$\frac{\partial^2}{\partial t^2} \phi(x, t) = \frac{1}{c^2} \Delta \phi(x, t),$$

Δ représentant le Laplacien spatial.

1.4.2 Fonctionnelles dépendant de plusieurs fonctions

Une autre extension possible est celle au cas de fonctionnelles non plus d'une mais de plusieurs variables dépendantes. Contrairement à la section précédente, c'est maintenant la variable dépendante qui est ici vectorielle, et non plus la variable indépendante.

Par exemple, supposons que Φ s'écrive sous la forme

$$\Phi[f_1, f_2] = \int F(x, f_1(x), f_2(x), f_1'(x), f_2'(x)) dx.$$

Pour trouver les extrêma d'une telle fonctionnelle, c'est à dire les couples (f_1, f_2) qui la minimisent ou la maximisent, il faut faire varier simultanément f_1 et f_2 , et annuler les dérivées fonctionnelles

$$\lim_{\epsilon \rightarrow 0} \frac{\Phi[f_1 + \epsilon g_1, f_2] - \Phi[f_1, f_2]}{\epsilon}, \quad \lim_{\epsilon \rightarrow 0} \frac{\Phi[f_1, f_2 + \epsilon g_2] - \Phi[f_1, f_2]}{\epsilon}$$

pour toutes fonctions g_1, g_2 de classe C^1 s'annulant sur les bords du domaine considéré.

Il est facile de vérifier que dans un tel cas de figure, on obtient non plus une mais deux équations d'Euler-Lagrange, donc un système d'équations de la forme

$$\begin{cases} \frac{\partial F}{\partial f_1} - \frac{d}{dx} \frac{\partial F}{\partial f_1'} = 0, \\ \frac{\partial F}{\partial f_2} - \frac{d}{dx} \frac{\partial F}{\partial f_2'} = 0. \end{cases}$$

Dans le cas de dimension quelconque, on montre similairement le résultat suivant

THÉORÈME 1.8 Soit Φ une fonctionnelle de plusieurs fonctions f_1, f_2, \dots, f_n , de la forme

$$\Phi[f_1, f_2, \dots, f_n] = \int F(x, f_1(x), f_2(x), f_1'(x), f_2'(x), \dots, f_n(x), f_n'(x)) dx,$$

où F est une fonction fixée, continûment différentiable.

Soit $(f_1, \dots, f_n) \in C^2(\mathbb{R}^n)$ un optimum de cette fonctionnelle, avec des conditions aux bords $f_k(x_1)$ et $f_k(x_2)$ fixées. Alors celui-ci satisfait nécessairement le système d'équations d'Euler-Lagrange

$$\frac{\partial F}{\partial f_k} - \frac{d}{dx} \frac{\partial F}{\partial f_k'} = 0, \quad k = 1, \dots, n.$$

EXEMPLE 1.7 La mécanique Hamiltonienne fournit de nombreux exemples de telles situations. Notons par exemple $q_1(t), q_2(t), q_3(t)$ les coordonnées d'un point matériel à l'instant t , et $\dot{q}_i(t), i = 1, 2, 3$ les dérivées temporelles de ces coordonnées. Si on note $\mathcal{L}(t, q_1(t), \dots, \dot{q}_3(t))$ la densité Lagrangienne associée, alors les trajectoires entre t_1 et t_2 sont données par les minima de l'intégrale d'action

$$\mathcal{A}[q] = \int_{t_1}^{t_2} \mathcal{L}(t, q_1(t), q_2(t), q_3(t), \dot{q}_1(t), \dot{q}_2(t), \dot{q}_3(t)) dt.$$

L'équation de Beltrami (Proposition 1.2) se généralise elle aussi de façon simple. Supposons que la fonction F ne dépende pas explicitement de la variable dépendante x , on peut alors écrire (en omettant les variables, pour simplifier les écritures) pour tout i

$$\frac{dF}{dx} = \sum_{i=1}^n \left(f_i' \frac{\partial F}{\partial f_i} + f_i'' \frac{\partial F}{\partial f_i'} \right).$$

Partant des équations d'Euler-Lagrange du THÉORÈME 1.7, multipliant la i -ième par f_i' et sommant sur i , on obtient alors

$$\begin{aligned} \sum_{i=1}^n f_i' \frac{\partial F}{\partial f_i} &= \sum_{i=1}^n f_i' \frac{d}{dx} \frac{\partial F}{\partial f_i'} \\ &= \frac{dF}{dx} - \sum_{i=1}^n f_i'' \frac{\partial F}{\partial f_i'}, \end{aligned}$$

d'où

$$\begin{aligned} \frac{dF}{dx} &= \sum_{i=1}^n \left(f_i'' \frac{\partial F}{\partial f_i'} + \frac{d}{dx} \frac{\partial F}{\partial f_i'} \right) \\ &= \frac{d}{dx} \left(\sum_{i=1}^n f_i' \frac{\partial F}{\partial f_i'} \right). \end{aligned}$$

On a donc montré

PROPOSITION 1.3 Avec les notations du théorème 1.8, si la fonction F ne dépend pas explicitement de la variable indépendante x , alors l'équation d'Euler-Lagrange se ramène à l'équation de Beltrami

$$F(f_1, \dots, f_n, f_1', \dots, f_n') - \sum_{i=1}^n f_i' \frac{\partial F}{\partial f_i'}(f_1, \dots, f_n, f_1', \dots, f_n') = C^{te}.$$

Dans ce cadre aussi, l'absence de dépendance dans la variable indépendante se traduit par l'existence d'une **quantité conservée**. On verra plus loin une interprétation de cette formule dans le cadre de la mécanique Hamiltonienne.

1.4.3 Fonctionnelles de dérivées d'ordre supérieur

Nous avons jusqu'à présent considéré uniquement le cas de fonctionnelles Φ dépendant d'une variable dépendante y et de sa dérivée y' (essentiellement motivés par la formulation Hamiltonienne de la mécanique classique). On peut en fait étendre le formalisme du calcul des variations à des cadres plus larges, comme par exemple le cadre où la fonctionnelle Φ dépend de dérivées d'ordre supérieur de la variable dépendante y . On est dans ce cas conduit à des équations différentielles d'ordre supérieur.

Par exemple, considérons la fonctionnelle Φ , définie par

$$\Phi[f] = \int_{x_1}^{x_2} F(x, f(x), f'(x), f''(x)) dx,$$

et soit $f \in C^4$, extrémum de cette fonctionnelle. Alors, pour tout $g \in C^2$, telle que $g(x_1) = g(x_2) =$

$g'(x_1) = g'(x_2) = 0$, on a

$$\begin{aligned} \frac{d}{d\epsilon} \Phi[f + \epsilon g] &= \int_{x_1}^{x_2} \left(g(x) \frac{\partial F}{\partial f}(x, f(x), f'(x), f''(x)) \right. \\ &\quad + g'(x) \frac{\partial F}{\partial f'}(x, f(x), f'(x), f''(x)) \\ &\quad \left. + g''(x) \frac{\partial F}{\partial f''}(x, f(x), f'(x), f''(x)) \right) dx \end{aligned}$$

Par intégrations par parties successives, et application du lemme fondamental, on obtient l'équation d'Euler-Lagrange correspondante

$$\frac{\partial F}{\partial f} - \frac{d}{dx} \frac{\partial F}{\partial f'} + \frac{d^2}{dx^2} \frac{\partial F}{\partial f''} = 0$$

On voit bien que cette équation sera généralement une équation différentielle d'ordre supérieur.

REMARQUE 1.3 Dans ce cadre, on voit intervenir des dérivées d'ordres supérieurs des fonctions considérées. Il est donc nécessaire en toute rigueur, de faire les hypothèses de différentiabilité correspondantes pour pouvoir énoncer un théorème correct.

1.4.4 Conditions aux bords naturelles

Les calculs que nous avons effectués jusqu'à présent concernaient les situations dans lesquelles les conditions aux bords sont totalement fixées. Il existe des situations dans lesquelles les conditions aux bords ne sont pas complètement fixées, et il est possible d'étendre l'analyse faite plus haut à ces cas de figure.

Prenons par exemple le cas où $f \in C^2([x_1, x_2])$ extrênum de

$$\Phi[f] = \int_{x_1}^{x_2} F(x, f(x), f'(x)) dx$$

sous la seule condition aux limites

$$f(x_1) = a.$$

Dans ce cas, la condition nécessaire d'existence d'un extrênum en f s'écrit

$$\frac{d}{d\epsilon} \Phi[f + \epsilon g] = 0$$

pour tout $g \in C^2(]x_1, x_2[)$ tel que $g(x_1) = 0$, sans autre condition en $x = x_2$ pour g . Ceci implique

$$\begin{aligned} \int_{x_1}^{x_2} g(x) \left[\frac{\partial F}{\partial f}(x, f(x), f'(x)) - \frac{d}{dx} \frac{\partial F}{\partial f'}(x, f(x), f'(x)) \right] dx \\ + g(x_2) \frac{\partial F}{\partial f'}(x_2, f(x_2), f'(x_2)) = 0. \end{aligned}$$

Ceci étant vérifié pour tout g , en particulier pour les g telles que $g(x_2) = 0$, on en déduit que l'équation d'Euler-Lagrange (1.17) reste valide. De plus, le cas $g(x_2) \neq 0$, quelconque, implique une seconde équation, appelée **condition aux bords naturelle**

$$\frac{\partial F}{\partial f'}(x_2, f(x_2), f'(x_2)) = 0. \quad (1.22)$$

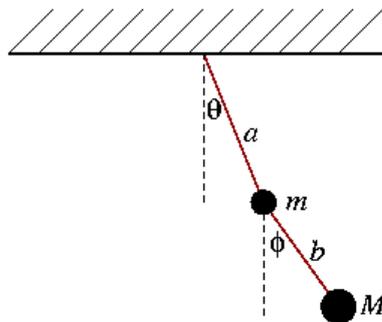


FIG. 1.10: Pendule double

1.5 Formalisme Hamiltonien

La mécanique Hamiltonienne est l'exemple le plus classique d'application du calcul variationnel, et des théorèmes 1.6, 1.7 et 1.8. Elle postule que la trajectoire suivie entre les instants t_1 et t_2 par un système en mouvement dans l'espace est toujours celle qui optimisera la quantité suivante, appelée **action**

$$J = \int_{t_1}^{t_2} L dt ,$$

où $L = T - V$ est le **Lagrangien**, égal à la différence entre l'énergie cinétique et l'énergie potentielle. C'est ce que l'on appelle le **principe de moindre action**. Le Lagrangien est en fait une fonctionnelle, qui dépend de la trajectoire suivie.

Supposons par exemple que le système dynamique considéré soit décrit par des coordonnées q_1, q_2, \dots, q_n et leurs dérivées $\dot{q}_1, \dot{q}_2, \dots, \dot{q}_n$, alors le Lagrangien prend la forme

$$L = L(q_1, q_2, \dots, q_n, \dot{q}_1, \dot{q}_2, \dots, \dot{q}_n) ,$$

et conduit aux équations du mouvement, appelées ici **équations de Lagrange**

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0, \quad i = 1, \dots, n .$$

REMARQUE 1.4 Ces équations du mouvement peuvent aussi être obtenues à partir des lois de Newton (et sont en fait équivalentes). Toutefois, pour des systèmes complexes, les équations de Lagrange sont généralement bien plus faciles à obtenir que ce que l'on pourrait faire à partir des équations de Newton.

EXEMPLE 1.8 (PENDULE DOUBLE) On considère le pendule double représenté en FIG. 1.10, et on note θ et ϕ les coordonnées angulaires décrivant les deux brins.

On choisit pour origine des hauteurs des deux masses leur hauteur d'équilibre. L'énergie potentielle est alors donnée par

$$V(\theta, \phi) = mga(1 - \cos(\theta)) + Mg(a(1 - \cos \theta) + b(1 - \cos \phi)) .$$

On montre facilement que l'énergie cinétique est quant à elle donnée par

$$T(\theta, \phi, \dot{\theta}, \dot{\phi}) = \frac{1}{2}ma^2\dot{\theta}^2 + \frac{1}{2}M(a^2\dot{\theta}^2 + b^2\dot{\phi}^2 + 2ab\dot{\theta}\dot{\phi}\sin(\theta + \phi)) .$$

Tous calculs faits, les équations d'Euler-Lagrange s'écrivent

$$\begin{aligned} \frac{d}{dt} ((m + M)a^2\dot{\theta} + Mab\dot{\phi}\sin(\theta + \phi)) + (M + m)ga \sin \theta &= 0 \\ \frac{d}{dt} (Mb^2\dot{\phi} + Mab\dot{\theta}\sin(\theta + \phi)) + Mgb \sin \phi &= 0 . \end{aligned}$$

Le Hamiltonien : Supposons maintenant que le Lagrangien ne dépende pas explicitement du temps. On peut alors utiliser la variante multidimensionnelle de l'identité de Beltrami (Proposition 1.3) et écrire les équations d'Euler-Lagrange sous la forme

$$\frac{d}{dt} \left(L - \sum_{i=1}^n \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} \right) = 0,$$

ce qui implique que

$$L - \sum_{i=1}^n \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} = C,$$

C étant une constante. Cette propriété est appelée **loi de conservation**. On définit le **Hamiltonien** du système par

$$H = \sum_{i=1}^n \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L, \quad (1.23)$$

et la loi de conservation stipule donc que le Hamiltonien est une quantité conservée.

1.6 Calcul des variations sous contrainte

Il est parfois possible d'utiliser le même type d'approche pour résoudre des problèmes variationnels lorsque des contraintes supplémentaires sont imposées à la solution. On a alors recours à la méthode des multiplicateurs de Lagrange, que nous avons déjà vu dans la section 1.1.3, et à une généralisation du théorème 1.5.

Plus précisément :

THÉORÈME 1.9 Soit $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, extrémum de la fonctionnelle

$$\Phi[f] = \int_{\Omega} F(\underline{x}, f(\underline{x}), \nabla f(\underline{x})) \, d\underline{x},$$

avec conditions aux bords $f = f_0$ sur $\partial\Omega$, et sous la contrainte

$$\Psi[f] := \int_{\Omega} G(\underline{x}, f(\underline{x}), \nabla f(\underline{x})) \, d\underline{x} = C,$$

C étant une constante fixée. Alors f est nécessairement solution des équations d'Euler-Lagrange associées à la fonctionnelle

$$\Gamma[f] = \Phi[f] - \lambda(\Psi[f] - C) = \int_{\Omega} [F(\underline{x}, f(\underline{x}), \nabla f(\underline{x})) - \lambda G(\underline{x}, f(\underline{x}), \nabla f(\underline{x}))] \, d\underline{x} - C.$$

Notons que C ne joue aucun rôle dans les équations d'Euler-Lagrange ci-dessus, et peut donc être supprimé. Il doit cependant être pris en compte (naturellement) lorsque la contrainte est imposée, ce qui permet de déterminer la valeur du multiplicateur λ .

On peut mettre ce résultat en application sur l'exemple suivant.

EXEMPLE 1.9 (CORDE PESANTE À L'ÉQUILIBRE) on se donne une corde pesante de masse linéique μ et de longueur L , dans le plan xOz , attachée à ses extrémités $A = (0,0)$ et $B = (a, z_1)$. La corde est supposée à l'équilibre, donc le problème se ramène à minimiser l'énergie potentielle

$$V = \int_0^a \mu g z(x) \sqrt{1 + z'(x)^2} \, dx,$$

sous la contrainte

$$L = \int_0^a \sqrt{1 + z'(x)^2} dx .$$

On introduit donc la fonctionnelle

$$\begin{aligned} J[z] &= \int_0^a \mu g z(x) \sqrt{1 + z'(x)^2} dx + \lambda \left(\int_0^a \sqrt{1 + z'(x)^2} dx - L \right) \\ &= \int_0^a (\mu g z(x) + \lambda) \sqrt{1 + z'(x)^2} dx - \lambda L , \end{aligned}$$

à minimiser par rapport à z et à λ . Il n'y a pas de dépendance explicite dans la variable indépendante x , on peut donc utiliser l'identité de Beltrami, qui s'écrit

$$(\mu g z + \lambda) \sqrt{1 + z'^2} - (\mu g z + \lambda) \frac{z'^2}{\sqrt{1 + z'^2}} = \frac{\mu g z + \lambda}{\sqrt{1 + z'^2}} = C .$$

Posons $u = z - \lambda/\mu g$. On a $u' = z'$ et l'équation devient $\mu g u / \sqrt{1 + u'^2} = C$. Cette forme suggère de rechercher u sous la forme $u = c \cosh((x - x_0)/c)$, ce qui impose $c = C/\mu g$. Finalement, la solution est une chaînette

$$z(x) = -\frac{\lambda}{\mu g} + c \cosh\left(\frac{x - x_0}{c}\right) .$$

Les constantes c et x_0 sont déterminées par les conditions aux limites, et le multiplicateur de Lagrange λ est quant à lui déterminé en imposant la contrainte.

La physique offre de nombreuses autres applications de ce principe variationnel sous contrainte. Pour n'en citer qu'un autre, on peut mentionner l'équation de Schrödinger.

EXEMPLE 1.10 (FORMULATION VARIATIONNELLE POUR L'ÉQUATION DE SCHRÖDINGER) L'équation de Schrödinger donnant les niveaux d'énergie d'un opérateur Hamiltonien

$$H\psi = E\psi , \quad H = -\frac{\hbar^2}{2m}\Delta + V ,$$

où V est une fonction à valeurs réelles, peut être obtenue comme équation d'Euler-Lagrange d'un problème variationnel (c'est d'ailleurs ainsi que Schrödinger la formule initialement). On considère pour cela la fonctionnelle

$$\Phi[\psi] = \int_{\mathbb{R}^3} \left[\frac{\hbar^2}{2m} |\nabla \psi(x)|^2 + V(x) |\psi(x)|^2 \right] dx .$$

On montre facilement que la fonction ψ qui minimise cette fonctionnelle, avec conditions aux limites nulles, et sous la contrainte de normalisation

$$\int_{\mathbb{R}^3} |\psi(x)|^2 dx = 1 ,$$

doit satisfaire l'équation de Schrödinger ci-dessus.

Deuxième partie

Transformations intégrales

2 Transformation de Fourier finie

2.1 Préliminaires : équation de la chaleur finie

Le problème de l'équation de la chaleur est un problème très classique de la physique. Il peut être résolu assez facilement dans un cadre très simple. Supposons que la variable temporelle soit une variable continue ($t \in \mathbb{R}^+$), mais considérons une variable spatiale discrète : $n \in \mathbb{Z}$. On considère maintenant un champ de température

$$T : (n, t) \rightarrow T_n(t) \in \mathbb{R},$$

et on suppose pour simplifier encore que le champ T est périodique en n : il existe un nombre entier positif N tel que pour tout n , on ait $T_{n+N}(t) = T_n(t)$ pour tout t . Ainsi, nous avons ramené notre problème à un problème plus simple ne faisant intervenir que N variables indépendantes.

La version correspondante de l'équation de la chaleur précédente est la famille d'équation différentielles couplées

$$\frac{dT_n(t)}{dt} = \alpha (T_{n+1}(t) - 2T_n(t) + T_{n-1}(t)), \quad n = 0, \dots, N-1, \quad (2.1)$$

et c'est ce système que nous voulons résoudre. On introduit pour cela les fonctions $(t, k) \rightarrow \hat{T}_k(t)$, définies par

$$\hat{T}_k(t) = \sum_{n=0}^{N-1} T_n(t) e^{-2i\pi kn/N}. \quad (2.2)$$

Calculons maintenant

$$\sum_{n=0}^{N-1} \frac{dT_n(t)}{dt} e^{-2i\pi kn/N} = \frac{d\hat{T}_k(t)}{dt}.$$

La même transformation appliquée au membre de droite de l'équation de la chaleur discrète conduit à

$$\begin{aligned} \sum_{n=0}^{N-1} (T_{n+1}(t) - 2T_n(t) + T_{n-1}(t)) e^{-2i\pi kn/N} &= \sum_{n=0}^{N-1} (e^{2i\pi k/N} - 2 + e^{-2i\pi k/N}) T_n(t) e^{-2i\pi kn/N} \\ &= -4 \sin^2\left(\frac{\pi k}{N}\right) \hat{T}_k(t). \end{aligned}$$

Ainsi, notre équation devient

$$\frac{\partial \hat{T}_k(t)}{\partial t} = -4\alpha \sin^2\left(\frac{\pi k}{N}\right) \hat{T}_k(t). \quad (2.3)$$

Il s'agit d'une famille de N équations différentielles découplées, dont les solutions sont connues :

$$\hat{T}_k(t) = \hat{T}_k(0) \exp \left\{ -4\alpha t \sin^2 \left(\frac{\pi k}{N} \right) \right\}. \quad (2.4)$$

Ainsi, si nous savons calculer les $T_n(t)$ à partir des $\hat{T}_k(t)$, nous aurons résolu notre problème.

La transformation $\{T_n, n = 0, \dots, N-1\} \rightarrow \{\hat{T}_k, k = 0, \dots, N-1\}$ est un exemple de ce que l'on appelle la transformation de Fourier finie. Cette dernière est inversible, comme nous allons maintenant le voir. Ainsi, nous pourrions, à partir des fonctions $t \rightarrow \hat{T}_k(t)$ données en (2.4), obtenir le champ de température $T_n(t)$ en tout point n et tout instant t .

2.2 La transformation de Fourier finie (TFF)

Le champ d'application de la transformation de Fourier finie est le domaine des suites de longueur finie N , que l'on notera sous la forme $\underline{f} = \{f_0, f_1, \dots, f_{N-1}\}$, où les nombres f_n sont des nombres *a priori* complexes (mais peuvent être réels, ou imaginaires purs). Une suite finie $\underline{f} \in \mathbb{C}^N$ peut également être considérée comme un vecteur de \mathbb{C}^N , dont les nombres f_n sont les composantes par rapport à une base orthonormée de référence. C'est ce langage là que nous emploierons dans un premier temps.

2.2.1 Définition, Propriétés, Inversion

Définition et propriétés

Commençons par introduire la transformation de Fourier finie :

DÉFINITION 2.1 Soit $\underline{f} = \{f_0, \dots, f_{N-1}\}$ un vecteur de \mathbb{C}^N . Sa transformée de Fourier finie est le vecteur $\hat{\underline{f}} = \{\hat{f}_0, \dots, \hat{f}_{N-1}\}$ défini par

$$\hat{f}_k = \sum_{n=0}^{N-1} f_n \exp \left\{ -2i\pi \frac{kn}{N} \right\}. \quad (2.5)$$

Il est facile de vérifier que la transformation

$$\mathcal{F} : \underline{f} \in \mathbb{C}^N \rightarrow \hat{\underline{f}} \in \mathbb{C}^N$$

est une transformation linéaire : pour tous $\underline{f}, \underline{g} \in \mathbb{C}^N$ et $\lambda, \mu \in \mathbb{C}$, on a

$$\mathcal{F}(\lambda \underline{f} + \mu \underline{g}) = \lambda \hat{\underline{f}} + \mu \hat{\underline{g}}.$$

Inversion

THÉORÈME 2.1 La TFF est inversible de la façon suivante : si \hat{f} est la TFF de $f \in \mathbb{C}^N$, on a alors

$$f_n = \frac{1}{N} \sum_{k=0}^{N-1} \hat{f}_k e^{2i\pi \frac{kn}{N}}. \quad (2.6)$$

On a de plus la formule de Parseval :

$$\frac{1}{N} \sum_{k=0}^{N-1} |\hat{f}_k|^2 = \sum_{n=0}^{N-1} |f_n|^2. \quad (2.7)$$

Preuve : Calculons tout d'abord la série géométrique : soit n un nombre entier, et supposons tout d'abord n ne soit pas un multiple entier de N . On a

$$\sum_{k=0}^{N-1} \exp \left\{ -2i\pi \frac{kn}{N} \right\} = \frac{1 - \exp \left\{ -2i\pi \frac{nN}{N} \right\}}{1 - \exp \left\{ -2i\pi \frac{n}{N} \right\}} = 0 \quad (2.8)$$

(le dénominateur est bien défini). Par contre, si n est un multiple entier de N , le terme générique de la série est constant et égal à 1. Par conséquent, on a

$$\sum_{k=0}^{N-1} \exp \left\{ -2i\pi \frac{kn}{N} \right\} = N \delta_{n[\text{mod } N], 0}. \quad (2.9)$$

Calculons maintenant le membre de droite de (2.6) :

$$\frac{1}{N} \sum_{k=0}^{N-1} \hat{f}_k e^{2i\pi \frac{kn}{N}} = \frac{1}{N} \sum_{k=0}^{N-1} \sum_{m=0}^{N-1} f_m e^{2i\pi \frac{k(n-m)}{N}} = \frac{1}{N} \sum_{m=0}^{N-1} f_m N \delta_{(m-n)[\text{mod } N], 0}$$

ce qui complète la preuve de (2.6). La preuve de (2.7) est très similaire. Il suffit de calculer

$$\frac{1}{N} \sum_{k=0}^{N-1} |\hat{f}_k|^2 = \frac{1}{N} \sum_{k=0}^{N-1} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f_m \bar{f}_n e^{2i\pi \frac{k(n-m)}{N}} = \frac{1}{N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} f_m \bar{f}_n N \delta_{(m-n)[\text{mod } N]},$$

ce qui prouve (2.7), et achève la preuve du théorème. ♠

2.2.2 Une base particulière de \mathbb{C}^N

\mathbb{C}^N est un espace vectoriel complexe de dimension N , muni du produit scalaire suivant : si $\underline{f}, \underline{g} \in \mathbb{C}^N$, on définit

$$\underline{f} \cdot \underline{g} = \sum_{n=0}^{N-1} \bar{f}_n g_n,$$

et on note $\|\cdot\|$ la norme correspondante : $\|\underline{f}\| = \sqrt{\underline{f} \cdot \underline{f}}$.

Noter la présence de la conjugaison complexe dans la définition du produit scalaire, nécessaire car on est en présence d'un espace vectoriel complexe. Ceci permet que la norme soit bien à valeurs réelles. Par contre, on a aussi, pour $\underline{f}, \underline{g} \in \mathbb{C}^N$, $\underline{f} \cdot \underline{g} = \overline{\underline{g} \cdot \underline{f}}$.

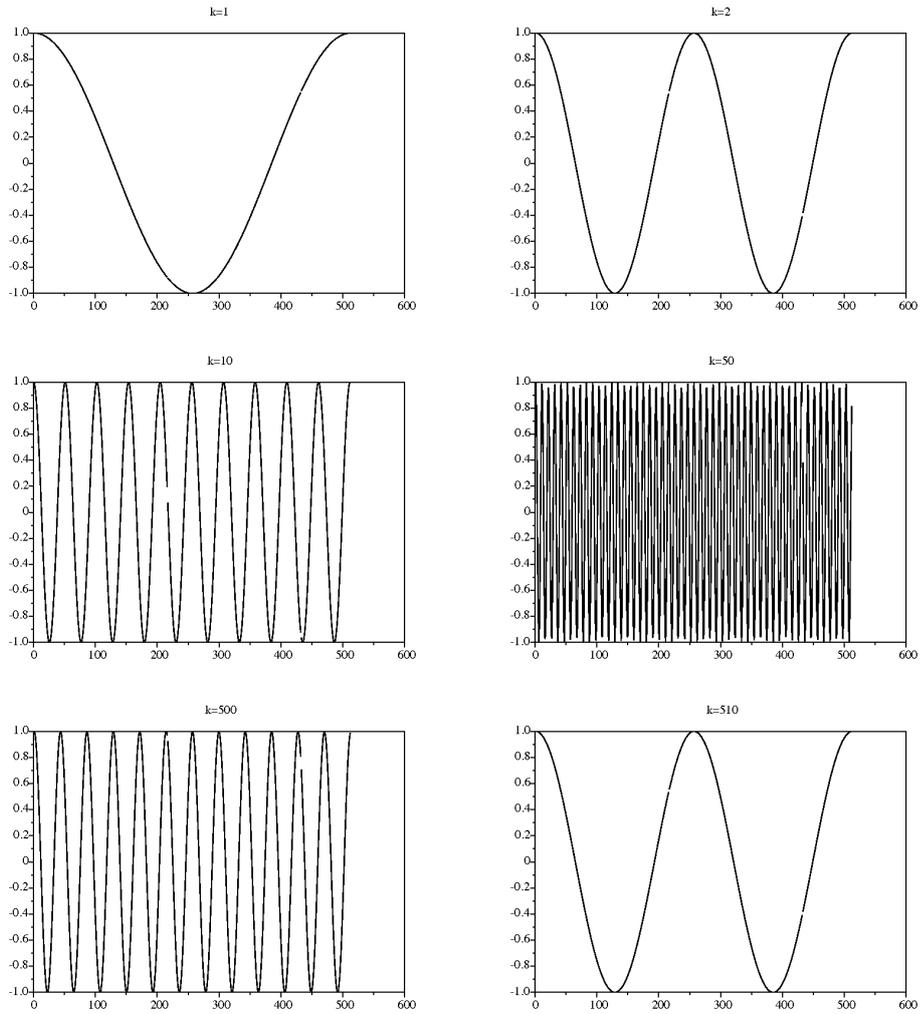


FIG. 2.1: Les vecteurs $\underline{\gamma}^k$ (partie réelle) pour différentes valeurs de k .

Il est facile de vérifier qu'il s'agit bien d'un produit scalaire. Introduisons maintenant les N vecteurs $\underline{\gamma}^k, k = 0, \dots, N-1$, définis par

$$\underline{\gamma}_n^k = \frac{1}{\sqrt{N}} \exp \left\{ 2i\pi \frac{kn}{N} \right\}, n = 0, \dots, N-1. \quad (2.10)$$

Les résultats précédents impliquent alors

COROLLAIRE 2.1 La famille $\{\underline{\gamma}^0, \dots, \underline{\gamma}^{N-1}\}$ est une base orthonormée de \mathbb{C}^N .

Preuve : Il suffit d'utiliser les résultats déjà obtenus : on a

$$\underline{\gamma}^k \cdot \underline{\gamma}^\ell = \frac{1}{N} \sum_{n=0}^{N-1} \exp \left\{ 2i\pi \frac{n(\ell-k)}{N} \right\} = \delta_{k,\ell}.$$

On a donc une famille de N vecteurs unitaires, orthogonaux deux à deux, dans un espace de dimension N , et c'est donc une base orthonormée. ♠

Il est alors facile de donner une interprétation de la transformation de Fourier finie. Etant donné un vecteur $\underline{f} \in \mathbb{C}^N$, les composantes de sa TFF $\hat{\underline{f}}$ sont de la forme

$$\hat{f}_k = \sqrt{N} \sum_{n=0}^{N-1} f_n \bar{\gamma}_n^k = \sqrt{N} \underline{\gamma}^k \cdot \underline{f}. \quad (2.11)$$

Ainsi, à un facteur multiplicatif (\sqrt{N}) près, les coefficients \hat{f}_k ne sont autres que les coordonnées de \underline{f} par rapport à la base des $\underline{\gamma}^k$. La formule d'inversion de la TFF peut aussi s'écrire

$$\underline{f} = \sum_{k=0}^{N-1} (\underline{\gamma}^k \cdot \underline{f}) \underline{\gamma}^k, \quad (2.12)$$

alors que la formule de Parseval n'est autre que la version correspondante du théorème de Pythagore

$$\|\underline{f}\|^2 = \sum_{k=0}^{N-1} |\underline{\gamma}^k \cdot \underline{f}|^2. \quad (2.13)$$

Il est important de bien saisir le rôle des vecteurs $\underline{\gamma}^k$. Comme on peut le voir sur la FIG. 2.1, le graphe des suites finies $\gamma_n^k, n = 0, \dots, N-1$ montre des oscillations dont la vitesse augmente avec k pour $k \in \{0, \dots, N/2\}$, puis diminue. Ainsi, la formule de décomposition (2.12) permet de décomposer la suite finie \underline{f} en une somme de contributions qui varient à des vitesses différentes. On donne à la variable k/N le nom de *fréquence*. Lorsque, pour une valeur de k proche de $N/2$, le nombre \hat{f}_k est grand, ceci signifie que la suite $\underline{\gamma}^k$ correspondante, qui varie rapidement, est affectée d'un gros coefficient dans la décomposition (2.12), et que la suite \underline{f} est elle aussi rapidement variable. Par contre, lorsque les coefficients \hat{f}_k pour k proche de $N/2$ sont petits, ceci signifie que la suite \underline{f} varie lentement. On dit qu'elle est plus "lisse", ou "régulière".

Ceci nous indique la marche à suivre lorsque l'on veut rendre une suite finie \underline{f} plus "lisse" : il suffit de diminuer les valeurs des coefficients \hat{f}_k pour les valeurs de k proches de $N/2$, en préservant les valeurs correspondant à des k proches de 0. On verra plus loin les effets de telles transformations.

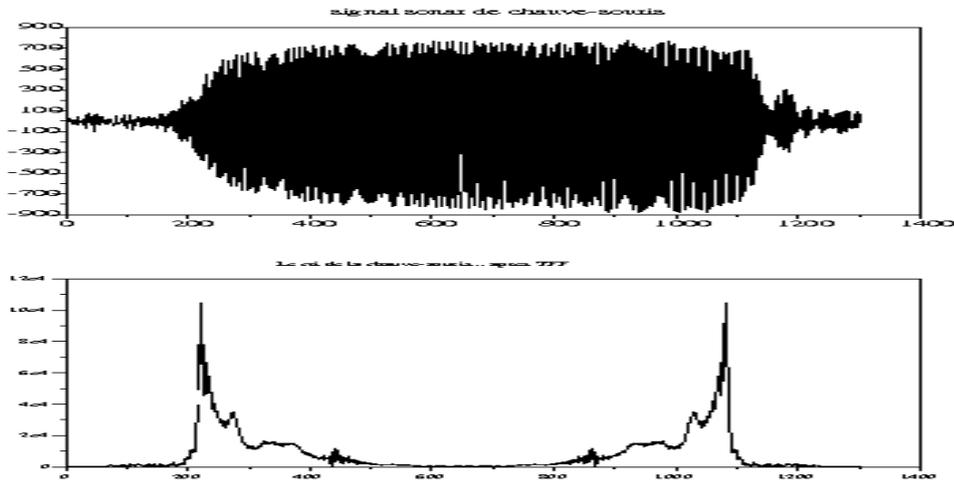


FIG. 2.2: Signal sonar de veille et de chasse de chauve souris, et sa TFF

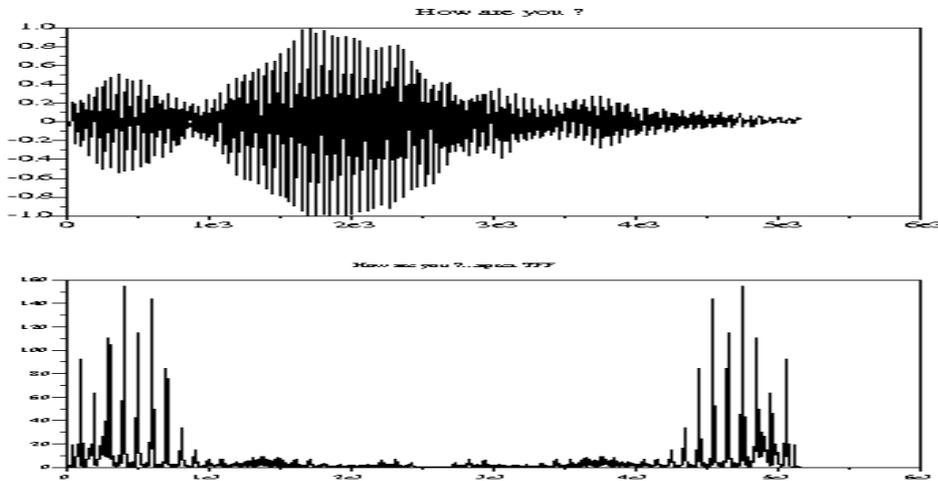


FIG. 2.3: Signal de parole et sa TFF.

Pour illustrer la transformation de Fourier finie, examinons quelques exemples. Le premier exemple est un exemple de signal acoustique (une onde de pression donc), plus précisément le signal SONAR émis par les chauve-souris (les chauve-souris étant pratiquement aveugles, elles utilisent les ultrasons pour s'orienter et détecter et localiser leurs proies). La FIG. 2.2 représente le signal lui-même, et sa transformée de Fourier. On voit que le signal est assez oscillant, avec des oscillations assez régulières. Ceci est confirmé par la transformée de Fourier, qui présente un pic bien marqué autour d'une fréquence donnée, qui est précisément la fréquence des oscillations régulières que l'on observe.

Le second exemple que nous considérons en FIG. 2.3 est lui aussi un signal acoustique, puisqu'il s'agit de signal de parole. Dans ce cas, la suite f (le signal) est aussi très rapidement oscillante, mais son graphe est plus complexe que le précédent. L'examen de la transformée de Fourier \hat{f} nous permet d'en comprendre la raison : le signal \hat{f} contient non plus une fréquence prépondérante, mais un grand nombre (une dizaine environ), qui sont toutes multiples d'une fréquence fondamentale. Ceci est spécifique du signal de parole, que l'on qualifie parfois de *signal harmonique*. Cette structure du signal de parole peut être expliquée grâce à des modèles physiques décrivant les vibrations des cordes vocales et la propagation du son dans le conduit vocal...

2.2.3 Convolution ; produit

La transformation de Fourier, sous toutes ses formes, est intimement liée au concept de produit de convolution. Nous allons le vérifier tout d'abord dans le cas simple de la TFF.

Etat donnés $\underline{f}, \underline{g} \in \mathbb{C}^N$, on définit leur produit de convolution $\underline{h} = \underline{f} * \underline{g} \in \mathbb{C}^N$ comme le vecteur de composantes

$$h_n = (\underline{f} * \underline{g})_n = \sum_{m=0}^{N-1} f_m g_{n-m \pmod{N}}. \quad (2.14)$$

Une propriété essentielle du produit de convolution est qu'il est symétrique (on dit aussi commutatif) :

$$\underline{f} * \underline{g} = \underline{g} * \underline{f}. \quad (2.15)$$

Calculons maintenant la TFF de h :

$$\begin{aligned} \hat{h}_k &= \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} f_m g_{n-m \pmod{N}} \exp \left\{ -2i\pi \frac{kn}{N} \right\} \\ &= \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} f_m g_{n-m \pmod{N}} \exp \left\{ -2i\pi \frac{km}{N} \right\} \exp \left\{ -2i\pi \frac{k(n-m)}{N} \right\} \\ &= \sum_{n=0}^{N-1} \sum_{j=0}^{N-1} f_m g_j \exp \left\{ -2i\pi \frac{km}{N} \right\} \exp \left\{ -2i\pi \frac{kj}{N} \right\} \\ &= \hat{f}_k \hat{g}_k. \end{aligned}$$

Ainsi, la transformée de Fourier finie du produit de convolution de deux vecteurs f et g est égale au vecteur dont les composantes sont égales au produit des composantes correspondantes de \hat{f} et \hat{g} . On note $\hat{h} = \hat{f}\hat{g}$ ce nouveau vecteur, appelé *produit simple* de \hat{f} et \hat{g} .

Attention : il ne faut pas confondre le vecteur $\hat{f}\hat{g}$ avec le produit scalaire $\hat{f} \cdot \hat{g}$, qui est un nombre (et pas un vecteur).

Considérons maintenant la TFF du produit simple de deux vecteurs. Pour cela, nous pouvons écrire en utilisant la formule d'inversion de la TFF

$$\begin{aligned} \widehat{\underline{f}\underline{g}}_k &= \sum_{n=0}^{N-1} f_n g_n \exp \left\{ -2i\pi \frac{kn}{N} \right\} \\ &= \frac{1}{N} \sum_{n=0}^{N-1} f_n \sum_{\ell=0}^{N-1} \hat{g}_\ell \exp \left\{ 2i\pi \frac{(\ell-k)n}{N} \right\} \\ &= \frac{1}{N} \sum_{\ell=0}^{N-1} \hat{g}_\ell \hat{f}_{\ell-k \pmod{N}} \\ &= \frac{1}{N} (\hat{f} * \hat{g})_k. \end{aligned}$$

Nous obtenons donc, à peu de choses près, un résultat "symétrique" du précédent : la TFF d'un produit simple est égale (à un facteur $1/N$ près) au produit de convolution des TFF. On résume les résultats obtenus dans la proposition suivante :

PROPOSITION 2.1 Soient $f, g \in \mathbb{C}^N$. Alors, on a, pour tout $k = 0, \dots, N-1$:

$$(\widehat{\underline{f} * \underline{g}})_k = \hat{f}_k \hat{g}_k \quad (2.16)$$

$$(\widehat{\underline{f}\underline{g}})_k = \frac{1}{N} (\hat{f} * \hat{g})_k. \quad (2.17)$$

Ce résultat, parfois appelé *théorème convolution-produit*, est un résultat que nous allons retrouver dans toutes les formes de transformation de Fourier (à de petites modifications près, à savoir des constantes multiplicatives différentes).

Nous sommes maintenant en position de terminer la résolution du problème posé au début de ce chapitre. On déduit de (2.4) la solution

$$T_n(t) = \frac{1}{N} \sum_{k=0}^{N-1} \hat{T}_k(0) \exp \left\{ -4\alpha t \sin^2 \left(\frac{\pi k}{N} \right) \right\} \exp \left\{ 2i\pi \frac{k}{N} \right\},$$

qui prend la forme de la TFF inverse d'un produit simple. Ainsi, la solution peut aussi s'écrire sous la forme d'un produit de convolution de la condition initiale $T(0) = \{T_n(0), n = 0, \dots, N-1\}$ par la TFF inverse de la suite

$$\hat{h}_k(t) = \left\{ \exp \left\{ -4\alpha t \sin^2 \left(\frac{\pi k}{N} \right) \right\}, k = 0, \dots, N-1 \right\}.$$

C'est là aussi une forme que nous allons retrouver souvent. On remarque que $\hat{h}_k(t)$ est proche de 1 pour k proche de 0 ou N , et vaut $\exp\{-4\alpha t\}$ pour $k = N/2$ (en supposant N pair).

Les solutions sont représentées en FIG. 2.4, pour différentes valeurs de t . On remarque que l'évolution temporelle consiste en un "lissage" progressif de la condition initiale, qui devient de plus en plus "étalée". C'est un comportement que l'on peut comprendre assez facilement, dans la mesure où l'équation (2.4) montre que dans l'espace des transformées de Fourier, l'évolution temporelle se traduit par une atténuation des composantes correspondant à k proche de $N/2$, c'est à dire les composantes qui varient le plus rapidement.

2.2.4 Remarques sur la périodicité

Dans la définition de la transformation de Fourier finie,

$$\hat{f}_k = \sum_{n=0}^{N-1} f_n \exp \left\{ -2i\pi \frac{kn}{N} \right\},$$

il est clair que \hat{f}_k peut également être défini pour des valeurs entières de k n'appartenant pas nécessairement à $\{0, \dots, N-1\}$. Par contre, on a pour tout $k \in \mathbb{Z}$

$$\hat{f}_{k+N} = \hat{f}_k \tag{2.18}$$

de sorte que \hat{f} peut être considéré soit comme un vecteur de \mathbb{C}^N , soit comme une suite infinie, périodique de période N . C'est pourquoi on rencontre souvent dans la littérature des définitions différentes de la TFF, par exemple, une définition dans laquelle \hat{f}_k est défini pour $k = 1 - N/2, \dots, N/2$ (quand N est pair).

Inversement, la formule

$$f_n = \frac{1}{N} \sum_{k=0}^{N-1} \hat{f}_k \exp \left\{ 2i\pi \frac{kn}{N} \right\}$$

permet de définir une suite infinie, N -périodique. On peut donc considérer f soit comme un vecteur de \mathbb{C}^N , soit comme une suite infinie, périodique de période N .

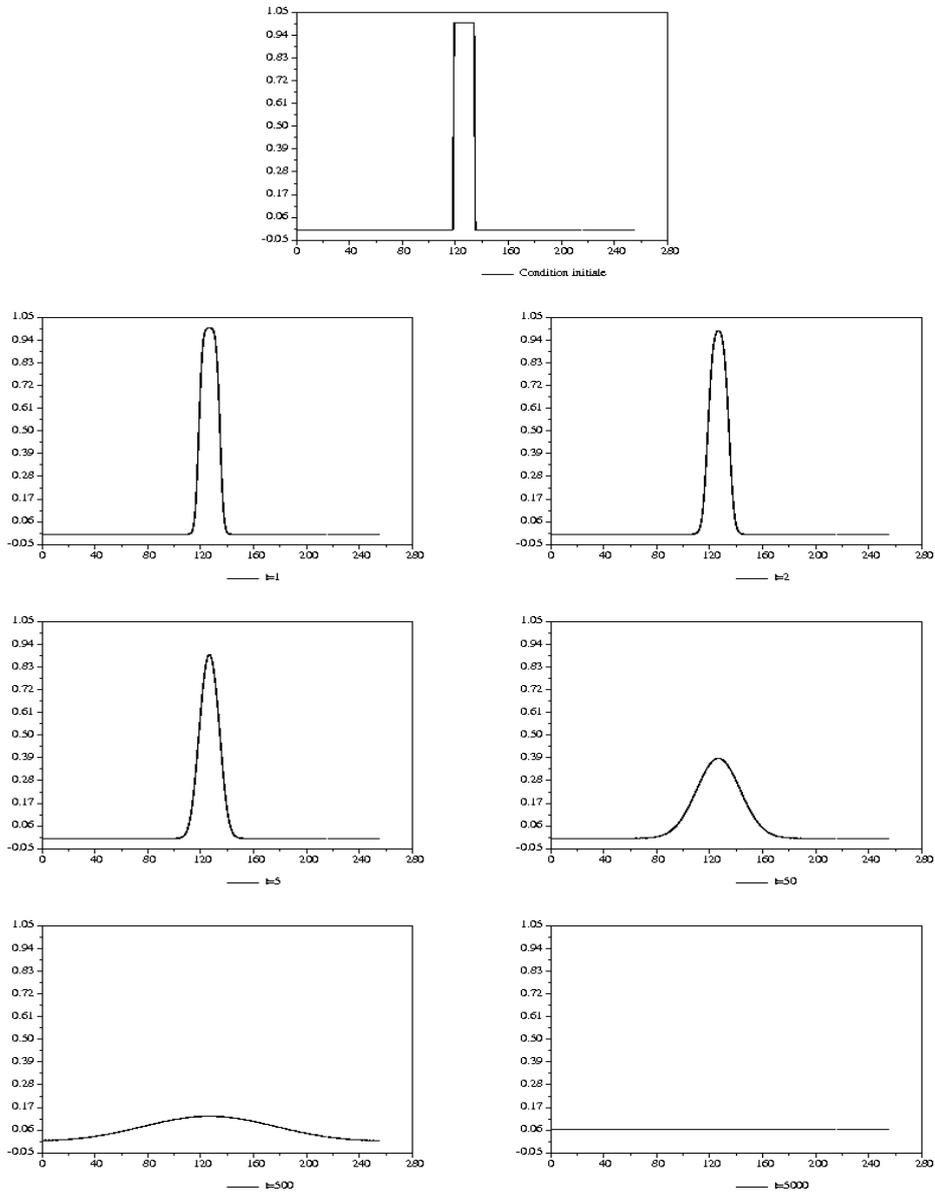


FIG. 2.4: Solution de l'équation de la chaleur, à des temps différents

2.2.5 Transformation de Fourier rapide ; le problème de la complexité

Lorsque l'on veut calculer une transformée de Fourier finie sur un ordinateur, il est nécessaire de se préoccuper du coût de cette transformation, c'est à dire d'essayer d'avoir une estimation du temps de calcul. Pour cela, un moyen simple consiste à compter le nombre d'opérations élémentaires (multiplications, additions, ou les deux) nécessaires au calcul. On appelle cela l'analyse de la *complexité* du calcul.

Dans le cas de la transformation de Fourier finie, le calcul de la complexité est assez simple. Pour faciliter les choses, on se limitera au calcul des multiplications (souvent plus coûteuses que les additions). Si on suppose que les nombres complexes $\exp\{-2i\pi kn/N\}$ ont été calculés une bonne fois pour toutes, le calcul de

$$\sum_{n=0}^{N-1} f_n \exp\left\{-2i\pi \frac{kn}{N}\right\}$$

fait intervenir N multiplications. Ce calcul est effectué N fois, ce qui conduit à N^2 multiplications. On dit que l'algorithme a une complexité en $O(N^2)$ (littéralement, de l'ordre de N^2).

Lorsque N est grand, le temps de calcul peut devenir exagérément long. Fort heureusement, il existe une façon d'organiser les calculs (un *algorithme*) plus efficace, permettant d'effectuer une TFF avec une complexité de $O(N \log_2 N)$. Il s'agit de la transformation de Fourier rapide (TFR, ou FFT en anglais).

Décrire l'algorithme de TFR dépasse le cadre de ce cours, mais il est important d'en connaître l'existence, et d'en mesurer l'importance. Prenons pour cela le cas $N = 2^{10} = 1024$. Alors $N^2 = 2^{20}$, alors que $N \log_2 N \approx 10 \times 2^{10}$, de sorte que l'algorithme de TFR est dans ce cas approximativement $2^{10}/10 \approx 102$ fois plus rapide. Ceci se passe de commentaire...

2.3 La transformation de Fourier discrète (TFD)

2.3.1 Définition, difficultés

Passons maintenant au cas de la dimension infinie : nous allons maintenant étudier la version de la transformation de Fourier adaptée au cas de "vecteurs infinis", c'est à dire adaptée au cas des suites. La dimension infinie pose toutefois des problèmes que l'on ne rencontre pas en dimension finie. Le premier a trait à la définition même de la transformation de Fourier : étant donnée une suite $f = \{f_n, n \in \mathbb{Z}\}$, sa transformée de Fourier est la fonction 2π périodique \hat{f} définie par

$$\hat{f}(\omega) = \sum_{n=-\infty}^{\infty} f_n e^{-in\omega}. \quad (2.19)$$

Cette série n'est pas nécessairement convergente, et il faut donc faire des hypothèses sur f pour s'assurer de l'existence de \hat{f} . L'hypothèse la plus simple consiste à supposer que la suite f considérée est sommable, c'est à dire telle que

$$\sum_{n=-\infty}^{\infty} |f_n| < \infty.$$

On dit alors que f appartient à l'espace $\ell^1(\mathbb{Z})$.

On vérifie facilement que $\ell^1(\mathbb{Z})$ est bien un espace linéaire, c'est à dire tel que pour tous $f, g \in \ell^1(\mathbb{Z})$, et $\lambda \in \mathbb{C}$, $f + g \in \ell^1(\mathbb{Z})$, et $\lambda f \in \ell^1(\mathbb{Z})$: $\sum_n |f_n + g_n| \leq \sum_n |f_n| + \sum_n |g_n| < \infty$, et $\sum_n |\lambda f_n| = |\lambda| \sum_n |f_n| < \infty$. On vérifie aussi que l'application $\|\cdot\|_1 : f \in \ell^1(\mathbb{Z}) \rightarrow \|f\|_1 \in \mathbb{R}$ définie par

$$\|f\|_1 = \sum_{n=-\infty}^{\infty} |f_n|$$

est une norme sur $\ell^1(\mathbb{Z})$: on a bien $\|f + g\|_1 \leq \|f\|_1 + \|g\|_1$ et $\|\lambda f\|_1 = |\lambda| \|f\|_1$ (voir ci dessus). De plus, $\|f\|_1 = 0$ implique que $f_n = 0$ pour tout n , donc $f = 0$.

DÉFINITION 2.2 Soit $f \in \ell^1(\mathbb{Z})$. Sa transformée de Fourier discrète est la fonction 2π périodique \hat{f} définie par (2.19). \hat{f} est une fonction bornée.

Le fait que \hat{f} soit bornée est facilement vérifié : en effet on a

$$|\hat{f}(\omega)| = \left| \sum_{n=-\infty}^{\infty} f_n e^{-in\omega} \right| \leq \sum_{n=-\infty}^{\infty} |f_n| = \|f\|_1 < \infty.$$

Comme son homologue finie, la transformation de Fourier discrète est une transformation linéaire. Nous verrons par la suite quelques propriétés plus “fines”, mais nous nous en tiendrons pour le moment à des propriétés élémentaires.

2.3.2 Convolution et produit

Le produit de convolution de deux suites est défini de façon similaire au produit de convolution des vecteurs. Etant données deux suites $f, g \in \ell^1(\mathbb{Z})$, on définit $h = f * g = g * f$ par

$$h_n = \sum_{m=-\infty}^{\infty} f_m g_{n-m}. \quad (2.20)$$

Il est facile de vérifier que $h \in \ell^1(\mathbb{Z})$:

$$\sum_n \left| \sum_m f_m g_{n-m} \right| \leq \sum_n \sum_m |f_m| |g_{n-m}| = \|f\|_1 \|g\|_1.$$

Par conséquent, sa transformée de Fourier discrète existe et est une fonction bornée. On a

$$\begin{aligned} \hat{h}(\omega) &= \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} f_m g_{m-n} e^{-i\omega n} \\ &= \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} f_m e^{-i\omega m} g_{n-m} e^{-i\omega(n-m)} \\ &= \hat{f}(\omega) \hat{g}(\omega). \end{aligned}$$

On a donc le résultat suivant :

PROPOSITION 2.2 Soient $f, g \in \ell^1(\mathbb{Z})$. Alors $f * g \in \ell^1(\mathbb{Z})$, et pour tout $\omega \in \mathbb{R}$,

$$\widehat{f * g}(\omega) = \hat{f}(\omega) \hat{g}(\omega). \quad (2.21)$$

On peut également montrer un résultat analogue dans le cas d’un produit simple : si f et g sont deux suites quelconques, leur produit simple est la suite notée fg , définie par $(fg)_n = f_n g_n$. La TFD de fg est alors égale à un produit de convolution “continu” des fonctions \hat{f} et \hat{g} . Cette propriété sera discutée en détails au chapitre suivant.

2.3.3 Notions de signal et de filtrage numériques

En théorie des communications, le support de l'information est appelé *signal*. Un signal peut prendre différentes formes, souvent la forme d'une fonction ou une suite. Dans le premier cas, il s'agit en général d'un signal "physique" : onde acoustique, électromagnétique, ... on parle dans ce cas de signal analogique. Dans le second cas, on parle de signal numérique.

L'opération de base en traitement du signal est le *filtrage* des signaux. Nous verrons par la suite des exemples plus précis de filtrage, mais nous pouvons d'ores et déjà voir les bases du filtrage numérique. Le filtrage numérique est généralement réalisé par convolution : le filtre est caractérisé par une suite $h = \{h_n, n \in \mathbb{Z}\} \in \ell^1(\mathbb{Z})$, appelée *réponse impulsionnelle* du filtre. Le filtre K_h est une application linéaire qui associe à toute suite f une suite $g = K_h f$, définie par

$$g_n = (K_h f)_n = (h * f)_n = \sum_{m=-\infty}^{\infty} h_m f_{n-m}.$$

Supposons $f \in \ell^1(\mathbb{Z})$. Comme nous l'avons vu, après transformation de Fourier, nous obtenons un produit simple :

$$\hat{g}(\omega) = \hat{h}(\omega) \hat{f}(\omega)$$

de sorte que le filtrage permet de modifier le contenu fréquentiel du signal f .

Un exemple de filtrage est présenté dans la FIG. 2.5. Il s'agit en fait d'un lissage, c'est à dire d'un filtrage qui a pour but de diminuer les valeurs de $\hat{f}(\omega)$ pour les grandes valeurs de ω . En l'occurrence, on travaille sur une suite finie \underline{f} assez "irrégulière" appelée *mouvement Brownien*, et le filtrage est effectué en remplaçant par des zéros toutes les valeurs de $\hat{f}(\omega)$ supérieures (en valeur absolue) à une valeur limite ω_c . ω_c est pris de plus en plus petit, de sorte que le lissage est de plus en plus important. Un tel filtrage peut être réalisé via un produit de convolution, mais est dans ce cas plus aisé à réaliser en travaillant directement sur les coefficients $\hat{f}(\omega)$.

2.3.4 Approximation d'une TFD par une TFF

En pratique, on ne peut calculer numériquement la transformée de Fourier discrète, et ce pour la simple raison que l'on ne peut utiliser sur ordinateur que des suites (ou vecteurs) finies. Le mieux que l'on puisse faire est alors de se rabattre sur la TFF.

Soit donc $f \in \ell^1(\mathbb{Z})$ une suite, et soit \hat{f} sa transformée de Fourier discrète. \hat{f} est donc une fonction 2π -périodique d'une variable réelle ω . Supposons maintenant que nous ne disposions pas de la suite f entière, mais d'une sous-suite finie, par exemple $\underline{f} = \{f_0, \dots, f_{N-1}\} \in \mathbb{C}^N$. L'idée la plus simple est alors de calculer une approximation \hat{f}_{app} de \hat{f} :

$$\hat{f}_{app}(\omega) = \sum_{n=0}^{N-1} f_n e^{-in\omega}. \quad (2.22)$$

On verra dans le chapitre suivant comment mesurer l'erreur commise en utilisant cette approximation. Par contre, la TFD étant une transformation linéaire, l'image de \mathbb{C}^N par TFD ne peut être de dimension supérieure à N , de sorte que considérer des valeurs $\hat{f}_{app}(\omega)$ pour une infinité de valeurs de ω n'a pas grand sens. On peut se limiter à N valeurs de ω , et le choix le plus simple consiste à considérer N valeurs régulièrement espacées entre 0 et 2π , soit

$$\omega_k = 2\pi \frac{k}{N}, k = 0, \dots, N-1.$$

On obtient alors

$$\hat{f}_{app}(\omega_k) = \sum_{n=0}^{N-1} f_n e^{-in\omega_k} = \sum_{n=0}^{N-1} f_n e^{-2i\pi kn/N} = \hat{f}_k, \quad (2.23)$$

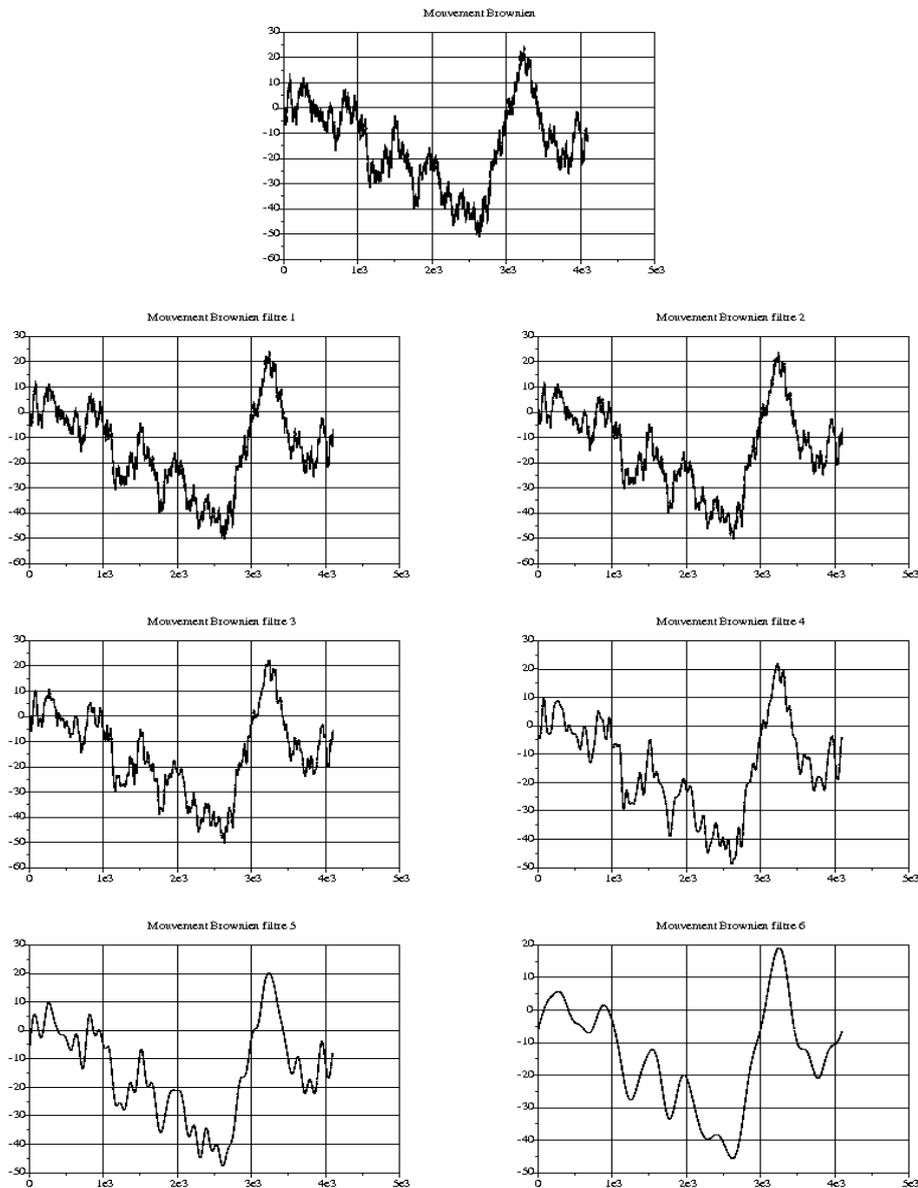


FIG. 2.5: Un exemple de signal "irrégulier" (mouvement Brownien, figure du haut), et des versions filtrées, de plus en plus "lissées".

de sorte que nous retombons naturellement sur la transformation de Fourier finie de la sous-suite finie $\underline{f} \in \mathbb{C}^N$.

Reste à vérifier que ceci caractérise effectivement \hat{f}_{app} . Pour cela, il suffit de calculer, pour $\omega \neq \pi k/N, k \in \mathbb{Z}$

$$\begin{aligned} \hat{f}_{app}(\omega) &= \frac{1}{N} \sum_{n=0}^{N-1} f_n e^{-in\omega} \\ &= \frac{1}{N} \sum_{n=0}^{N-1} \sum_{k=0}^{N-1} \hat{f}_k e^{2i\pi kn/N} e^{-in\omega} \\ &= \frac{1}{N} \sum_{k=0}^{N-1} \hat{f}_k \sum_{n=0}^{N-1} \exp\{-i(\omega - 2\pi k/N)\}^n \\ &= \frac{1}{N} \sum_{k=0}^{N-1} \hat{f}_k \frac{1 - \exp\{-iN\omega\}}{1 - \exp\{-i(\omega - \pi k/N)\}}, \end{aligned}$$

ce qui montre que l'on peut bien calculer $\hat{f}_{app}(\omega)$ pour tout ω à partir des $\hat{f}(\omega_k)$.

2.3.5 Inversion : les séries de Fourier

Nous avons vu que l'inversion de la transformation de Fourier finie est une opération assez simple. Le cas de la transformation de Fourier discrète est plus complexe. Cependant, le résultat suivant reste simple.

THÉORÈME 2.2 Soit $f \in \ell^1(\mathbb{Z})$, et soit \hat{f} sa transformée de Fourier discrète. Alors on a pour tout $n \in \mathbb{Z}$

$$f_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{f}(\omega) e^{in\omega} d\omega. \quad (2.24)$$

Preuve : Nous savons que \hat{f} est bornée. Donc l'intégrale du membre de droite de (2.24) est convergente, et nous pouvons calculer

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{f}(\omega) e^{in\omega} d\omega &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{m=-\infty}^{\infty} f_m e^{i(n-m)\omega} d\omega \\ &= \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} f_m \int_{-\pi}^{\pi} e^{i(n-m)\omega} d\omega \\ &= f_n, \end{aligned}$$

ce qui prouve le théorème. ♠

Notons que la formule (2.24) coïncide avec la définition des coefficients de Fourier :

$$f_n = a_n + ib_n,$$

avec

$$a_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{f}(\omega) \cos(n\omega) d\omega, \quad b_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{f}(\omega) \sin(n\omega) d\omega.$$

Nous verrons au chapitre suivant un lien plus étroit entre ces deux théories.

Nous avons donc maintenant à notre disposition une transformation de Fourier discrète, définie sur $\ell^1(\mathbb{Z})$, que nous savons inverser. Cependant, elle n'est pas tout à fait satisfaisante, pour plusieurs raisons. D'une part, l'hypothèse $f \in \ell^1(\mathbb{Z})$ est parfois trop restrictive. Et d'autre part, nous n'avons pas dans ce cadre d'interprétation simple en termes de base orthonormée. Nous verrons que la théorie des séries de Fourier permet de compléter ces manques.

Séries de Fourier

Nous avons étudié au chapitre précédent la transformation de Fourier des suites, finies ou infinies. Nous passons maintenant au cas des fonctions, qui est plus complexe. Pour simplifier quelque peu, on s'intéressera tout d'abord au cas des fonctions à support borné, ou aux fonctions périodiques, qui peuvent être traitées de façon similaires. La théorie de Fourier correspondante est la théorie des séries de Fourier.

La théorie des séries de Fourier est tout d'abord développée dans le contexte des *fonctions périodiques*. Rappelons qu'une fonction $t \rightarrow f(t)$ est périodique de période T si pour tout $t \in \mathbb{R}$, on a $f(t + T) = f(t)$. La théorie générale de Fourier montre que si f est une telle fonction, elle peut s'exprimer sous forme de combinaison linéaire infinie de sinusoides, de fréquences multiples d'une fréquence fondamentale égale à l'inverse de T . Toutefois, il est encore nécessaire ici de faire attention au sens que l'on donne à l'égalité, qui peut varier suivant le contexte.

3.1 Fonctions trigonométriques

3.1.1 Polynômes trigonométriques

Un exemple simple de fonction périodique de période T est l'oscillation ϵ_n , définie par

$$\epsilon_n(t) = e^{2in\pi t/T} \quad (3.1)$$

pour tout entier n . Il en va de même pour les fonctions

$$P(t) = \sum_{n \in I} c_n e^{2in\pi t/T},$$

où I est un index fini, et les $c_n, n \in I$ des nombres complexes quelconques. Quitte à compléter l'expression précédente par des coefficients c_n nuls, on peut mettre $P(t)$ sous la forme

$$P(t) = \sum_{n=-N}^N c_n e^{2in\pi t/T}, \quad (3.2)$$

que l'on appelle *polynôme trigonométrique de degré N* .

En décomposant l'exponentielle complexe en sinus et cosinus, on peut aussi mettre [3.2] sous la forme

$$P(t) = c_0 + \sum_{n=1}^N \left(a_n \cos\left(\frac{2n\pi t}{T}\right) + b_n \sin\left(\frac{2n\pi t}{T}\right) \right), \quad (3.3)$$

en posant

$$\begin{cases} a_n &= c_n + c_{-n} \\ b_n &= i(c_n - c_{-n}) \end{cases} \quad (3.4)$$

3.1.2 Orthogonalité

Soit \mathcal{P}_N l'espace vectoriel des polynômes trigonométriques de degré inférieur ou égal à N (qui est de dimension $2N + 1$). Cet espace peut être muni du produit scalaire suivant : $\forall P, Q \in \mathcal{P}_N$, on pose

$$\langle P, Q \rangle = \int_0^T \overline{P}(t) Q(t) dt . \quad (3.5)$$

On peut vérifier que ce produit possède bien toutes les propriétés d'un produit scalaire. Posons pour simplifier

$$e_n(t) = \frac{1}{\sqrt{T}} e^{2in\pi t/T} . \quad (3.6)$$

Il est facile de voir que

$$\langle e_m, e_n \rangle = \int_0^T \overline{e_m}(t) e_n(t) dt = \frac{1}{T} \int_0^T e^{2i(n-m)\pi t/T} dt = \delta_{n,m} , \quad (3.7)$$

c'est à dire que les fonctions $e_{2n\pi T}$ sont orthogonales deux à deux.

On dispose donc d'une famille de $2N + 1$ fonctions orthogonales, de norme égale à 1, dans un espace de dimension $2N + 1$. Donc,

PROPOSITION 3.1 La famille $\{e_n, n = -N, \dots, N\}$ est une base orthonormée de \mathcal{P}_N .

Par conséquent, étant donné un polynôme trigonométrique $P \in \mathcal{P}_N$, nous pouvons calculer les coefficients c_n de son développement par rapport à la base $\{e_n, n = -N, \dots, N\}$ grâce à la relation

$$\langle e_n, P \rangle = \sqrt{T} \sum_m c_m \langle e_n, e_m \rangle = \sqrt{T} c_n \|e_n\|^2 = \sqrt{T} c_n ,$$

et donc on a

$$c_n = \frac{1}{\sqrt{T}} \langle e_n, P \rangle = \frac{1}{T} \int_0^T P(t) e^{-2in\pi t/T} dt . \quad (3.8)$$

Le calcul de la norme quadratique fournit quant à lui la formule de Parseval, qui est essentiellement une version adaptée du théorème de Pythagore :

$$\frac{1}{T} \int_0^T |P(t)|^2 dt = \sum_n |c_n|^2 . \quad (3.9)$$

3.2 Séries de Fourier

3.2.1 Généralités

Jusqu'à présent, nous avons travaillé dans un contexte de dimension finie, que nous voulons maintenant étendre au cas d'espaces de fonctions autres que des polynômes trigonométriques, qui sont généralement des espaces de dimension infinie. Parmi les espaces fonctionnels classiques, les espaces de fonctions de module carré intégrable jouent un rôle primordial. Dans le cas qui nous intéresse, à savoir le cas des fonctions périodiques, il s'agit de l'espace $L_p^2([a, b])$, défini de la façon suivante :

$$L_p^2([a, b]) = \{f : \mathbb{R} \rightarrow \mathbb{C}, f \text{ périodique de période } b - a, \|f\| < \infty\} , \quad (3.10)$$

où $a < b$ sont deux réels, et où on a noté

$$\|f\|^2 = \int_a^b |f(t)|^2 dt . \quad (3.11)$$

Il est facile de vérifier qu'il s'agit bien d'un espace linéaire (si $f, g \in L_p^2([a, b])$, alors $f + g \in L_p^2([a, b])$, et $\lambda f \in L_p^2([a, b])$ pour tout $\lambda \in \mathbb{C}$. C'est de plus un espace normé (l'application $f \in L_p^2([a, b]) \rightarrow \|f\|$ est bien une norme). Cependant, ce qui différencie $L_p^2([a, b])$ de la majeure partie des espaces fonctionnels normés, c'est le fait que cet espace possède un produit scalaire : pour $f, g \in L_p^2([a, b])$, on définit

$$\langle f, g \rangle = \int_a^b \bar{f}(t)g(t) dt, \quad (3.12)$$

On peut en fait montrer que ceci munit $L_p^2([a, b])$ d'une structure d'espace de Hilbert.

La notion d'intégrale utilisée est l'intégrale de Lebesgue. Cette dernière coïncide dans la plupart des cas avec l'intégrale de Riemann, mais permet d'étendre la notion d'intégration à des fonctions pour lesquelles l'intégrale "à la Riemann" n'est pas définie. Des détails complémentaires sur l'intégrale de Lebesgue, et plus généralement sur les espaces fonctionnels "classiques" se trouvent en Annexe B.

REMARQUE 3.1 Pour être précis, l'espace $L_p^2([a, b])$ n'est pas un espace de fonctions à proprement parler, mais plutôt un espace de *classes d'équivalence* de fonctions : en effet, si $f \in L_p^2([0, 1])$, et si on définit g par $g(t) = f(t)$ pour tout $t \neq 1/2$, et $g(1/2) = f(1/2) + 1$, alors conformément à la définition de l'intégrale de Lebesgue, on a $\|f - g\| = 0$, de sorte que l'on ne peut dans $L_p^2([0, 1])$ discerner g de f^1 . Plus généralement, toute fonction g obtenue à partir de f en la modifiant sur un sous-ensemble de $[0, 1]$ de mesure nulle sera indiscernable de f , et donc identifiée à f . On dit parfois que les fonctions des espaces de Lebesgue L^p sont définies *modulo un ensemble de mesure nulle*.

Par la suite, on posera

$$T = b - a.$$

Il est immédiat de vérifier que les oscillations ϵ_n et les fonctions e_n appartiennent à $L_p^2([a, b])$, de sorte que pour tout $N \in \mathbb{Z}^+$, on a l'inclusion

$$\mathcal{P}_N \subset L_p^2([a, b]). \quad (3.13)$$

On peut maintenant se poser la question suivante :

Etant donné un $f \in L_p^2([a, b])$, quel est l'élément de \mathcal{P}_N qui en est le plus proche ? en d'autres termes, quelle est la projection orthogonale f_N de f sur \mathcal{P}_N ?

Il est facile de répondre à cette question en utilisant les outils à notre disposition dans $L_p^2([a, b])$, et en particulier la norme. La distance entre $f \in L_p^2([a, b])$ et un polynôme $P \in \mathcal{P}_N$, de la forme

$$P(t) = \sum_{-N}^N c_n \epsilon_n(t),$$

est naturellement définie par

$$d(f, P) = \|f - P\|. \quad (3.14)$$

¹Précisément, la norme $\|\cdot\|$ munit $L_p^2([0, 1])$ d'une topologie qui ne sépare pas f de g .

Calculons alors

$$\begin{aligned}
\|f - P\|^2 &= \left\langle \left(f - \sum_{-N}^N c_n \epsilon_n \right), \left(f - \sum_{-N}^N c_n \epsilon_n \right) \right\rangle \\
&= \|f\|^2 - 2\Re \left(\sum_{-N}^N c_n \langle f, \epsilon_n \rangle \right) + \sum_{n,m=-N}^N c_n \bar{c}_m \langle \epsilon_m, \epsilon_n \rangle \\
&= \|f\|^2 - 2\Re \left(\sum_{n=-N}^N \bar{d}_n c_n \right) + T \sum_{n=-N}^N |c_n|^2 \\
&= \left(\|f\|^2 - \frac{1}{T} \sum_{-N}^N |d_n|^2 \right) + \sum_{-N}^N \left| \sqrt{T} c_n - d_n / \sqrt{T} \right|^2,
\end{aligned}$$

où on a posé

$$d_n = \langle \epsilon_n, f \rangle = \int_a^b f(t) e^{-2i\pi n t / T} dt. \quad (3.15)$$

Les deux premiers termes (entre parenthèses) sont indépendants des coefficients c_n recherchés, de sorte que minimiser $\|f - P\|$ est équivalent à minimiser la seconde somme par rapport aux coefficients c_n . Il est clair que cette dernière est minimale (et nulle) si et seulement si

$$c_n = c_n(f) = \frac{1}{T} d_n = \frac{1}{T} \int_a^b f(t) e^{-2i\pi n t / T} dt. \quad (3.16)$$

On a alors

$$P(t) = f_N(t) = \sum_{n=-N}^N c_n(f) e^{2i\pi n t / T}. \quad (3.17)$$

De plus, on a également

$$\|f - f_N\|^2 = \|f\|^2 - T \sum_{-N}^N |c_n(f)|^2. \quad (3.18)$$

Nous avons donc montré :

PROPOSITION 3.2 1. La projection orthogonale F_N de f sur \mathcal{P}_N est donnée par

$$f_N(t) = \sum_{n=-N}^N c_n(f) e^{2i\pi n t / T}, \quad (3.19)$$

où les $c_n(f)$ sont définis en (3.16).

2. On a de plus l'inégalité de Bessel

$$\sum_{n=-N}^N |c_n(f)|^2 \leq \frac{1}{T} \|f\|^2. \quad (3.20)$$

3.2.2 Une base orthonormée de $L_p^2([a, b])$

Le cas de la transformation de Fourier finie était très simple, mais aussi très utile pour comprendre le statut de la transformation de Fourier. Non seulement, nous avons montré que la TFF est une transformation inversible de \mathbb{C}^N dans \mathbb{C}^N , mais que de plus la famille des suites finies $\gamma^k, k = 0, \dots, N-1$ est une base orthonormée de \mathbb{C}^N , et que la TFF n'est autre que l'application qui associe à tout élément

$\underline{f} \in \mathbb{C}^N$ la suite finie des coefficients $(\hat{f}_0, \dots, \hat{f}_{N-1})$ de son développement par rapport à cette base (à un facteur multiplicatif constant près). Il se trouve que ceci reste vrai lorsque l'on passe au cadre de l'espace $L_p^2([a, b])$ et des séries de Fourier associées.

En effet, nous avons déjà vu que les fonctions $t \rightarrow e_n(t)$ définies en (3.6) sont orthogonales deux à deux, et normées. Plus précisément, la Proposition 3.1 montre qu'elles forment des bases orthonormées de sous-espaces de dimension finie de $L_p^2([a, b])$. Qu'en est-il en dimension infinie? La réponse est apportée par le théorème suivant, donné sans démonstration :

THÉORÈME 3.1 *La famille des fonctions exponentielles e_n définies en (3.6) est une base orthonormée de $L_p^2([a, b])$. Pour tout $f \in L_p^2([a, b])$, on a*

$$\|f - f_N\|^2 = T \sum_{|n| > N} |c_n(f)|^2 \rightarrow 0 \text{ quand } N \rightarrow \infty, \quad (3.21)$$

où on a posé

$$f_N = \sum_{n=-N}^N c_n(f) e_n \quad (3.22)$$

De plus, la formule de Parseval s'écrit, pour toutes $f, g \in L_p^2([a, b])$

$$\sum_{-\infty}^{\infty} \overline{c_n(f)} c_n(g) = \frac{1}{b-a} \int_a^b \overline{f}(t) g(t) dt = \frac{1}{b-a} \langle f, g \rangle. \quad (3.23)$$

Ainsi, on a également la propriété suivante : pour tout $f \in L_p^2([a, b])$, la suite de ses coefficients de Fourier est de module carré sommable.

DÉFINITION 3.1 *Etant donnée $f \in L_p^2([a, b])$, la limite f_∞ des fonctions f_N définies en (3.22) est appelée série de Fourier de f .*

Ainsi, le théorème montre que pour tout f , $f_N \rightarrow f$ (au sens de $L_p^2([a, b])$). On parle alors de *convergence forte*.

REMARQUE 3.2 Ceci ne signifie nullement que pour tout $t \in [a, b]$, $f_N(t) \rightarrow f(t)$. On verra des contre-exemples plus loin. Par contre, une conséquence du théorème est que $f_N(t) \rightarrow f(t)$ pour presque tout t , c'est à dire pour tout t sauf éventuellement un ensemble de mesure nulle. En effet, si tel n'était pas le cas, alors on aurait $\|f_N - f\| \not\rightarrow 0$ quand $N \rightarrow \infty$.

3.2.3 Un exemple

Considérons le cas de la fonction $f \in L_p^2([0, 2\pi])$, définie par

$$f(t) = \begin{cases} 1 & \text{si } 0 \leq t < \pi \\ -1 & \text{si } \pi \leq t < 2\pi, \end{cases}$$

(appelée *fonction de Haar*) et complétée par périodicité. Un calcul explicite donne

$$c_0(f) = \frac{1}{2\pi} \int_0^{2\pi} f(t) dt = 0,$$

et pour $n \neq 0$,

$$c_n(f) = \frac{1}{2\pi} \int_0^{2\pi} f(t)e^{-int} dt = \frac{1}{2\pi} \left[\int_0^{\pi} e^{-int} dt - \int_{\pi}^{2\pi} e^{-int} dt \right] = \frac{1}{\pi} \frac{1 - (-1)^n}{in}.$$

Ainsi

$$c_n(f) = \begin{cases} 0 & \text{si } n \text{ est pair} \\ \frac{2}{in\pi} & \text{si } n \text{ est impair.} \end{cases}$$

On a par exemple

$$\begin{aligned} f_1(t) &= \frac{4}{\pi} \sin(\pi t) \\ f_3(t) &= \frac{4}{\pi} \left(\sin(\pi t) + \frac{1}{3} \sin(3\pi t) \right) \\ f_5(t) &= \frac{4}{\pi} \left(\sin(\pi t) + \frac{1}{3} \sin(3\pi t) + \frac{1}{5} \sin(5\pi t) \right). \end{aligned}$$

Il est généralement instructif de s'intéresser à l'erreur d'approximation $\|f - f_N\|$. Dans notre cas, on a

$$\|f - f_{2K-1}\|^2 = 4\pi \sum_{m=K}^{\infty} \left| \frac{2}{(2m+1)\pi} \right|^2 = \frac{4}{\pi} \sum_{m=K}^{\infty} \frac{1}{(m+1/2)^2} \leq \frac{4}{\pi} \sum_{m=K}^{\infty} \frac{1}{m^2}.$$

Or, on a

$$\int_{m-1}^m \frac{dt}{t^2} = \frac{1}{m-1} - \frac{1}{m} = \frac{1}{m(m-1)} \geq \frac{1}{m^2},$$

de sorte que

$$\|f - f_{2K-1}\|^2 \leq \frac{4}{\pi} \int_{K-1}^{\infty} \frac{dt}{t^2} = \frac{4}{\pi} \frac{1}{K-1}.$$

L'erreur d'approximation $\|f - f_N\|$ décroît donc comme $1/\sqrt{N}$. Nous verrons par la suite des conditions permettant d'estimer *a priori* l'erreur d'approximation en fonction de propriétés génériques (essentiellement des propriétés de régularité) de f .

Il est également intéressant d'étudier les propriétés de convergence ponctuelle de la série de Fourier de f . On sait que $f_N(t)$ converge vers $f(t)$ pour presque tout t . Par contre, on a également $f_N(0) = 0$ pour tout N , de sorte que $f_N(0) \rightarrow 0 \neq f(0)$. Voici donc un contre-exemple à la convergence ponctuelle. Par contre, on peut aussi remarquer que $\lim_{N \rightarrow \infty} f_N(0) = 0 = (\lim_{t \rightarrow 0^+} f(t) + \lim_{t \rightarrow 0^-} f(t))/2$, ce qui est en fait un résultat que nous allons montrer un peu plus loin dans un cadre général.

L'exemple est représenté en FIG. 3.1. On voit clairement que les approximations successives convergent (lentement) vers la fonction étudiée (en haut à gauche). La lenteur de la convergence s'explique par le fait que $\|f - f_N\|$ se comporte comme $1/\sqrt{N}$. On observe également des oscillations qui apparaissent au voisinage des discontinuités de la fonction. Ces discontinuités sont en fait les caractéristiques de f les plus difficiles à approximer par des sinusoides, et se traduisent donc par l'apparition d'oscillations. Ce phénomène est appelé *phénomène de Gibbs*.

Ce phénomène est toujours présent dans l'exemple de la fonction "dents de scie" de la FIG. 3.2, mais est considérablement atténué. En effet, dans ce cas, la fonction considérée étant une fonction continue, on peut montrer que ses coefficients de Fourier $c_n(f)$ décroissent comme $1/n^2$, de sorte que $\|f - f_N\|$ se comporte comme $1/N\sqrt{N}$. La convergence est ainsi plus rapide.

3.2.4 Quelques remarques

1. Soit $f \in L_p^2([a, b])$. Ses coefficients de Fourier sont définis par (3.16); cependant, compte tenu de la périodicité supposée de f , on peut également prendre (en posant $T = b - a$)

$$c_n(f) = \frac{1}{T} \int_{-T/2}^{T/2} f(t)e^{-2i\pi nt/T} dt = \frac{1}{T} \int_I f(t)e^{-2i\pi nt/T} dt,$$

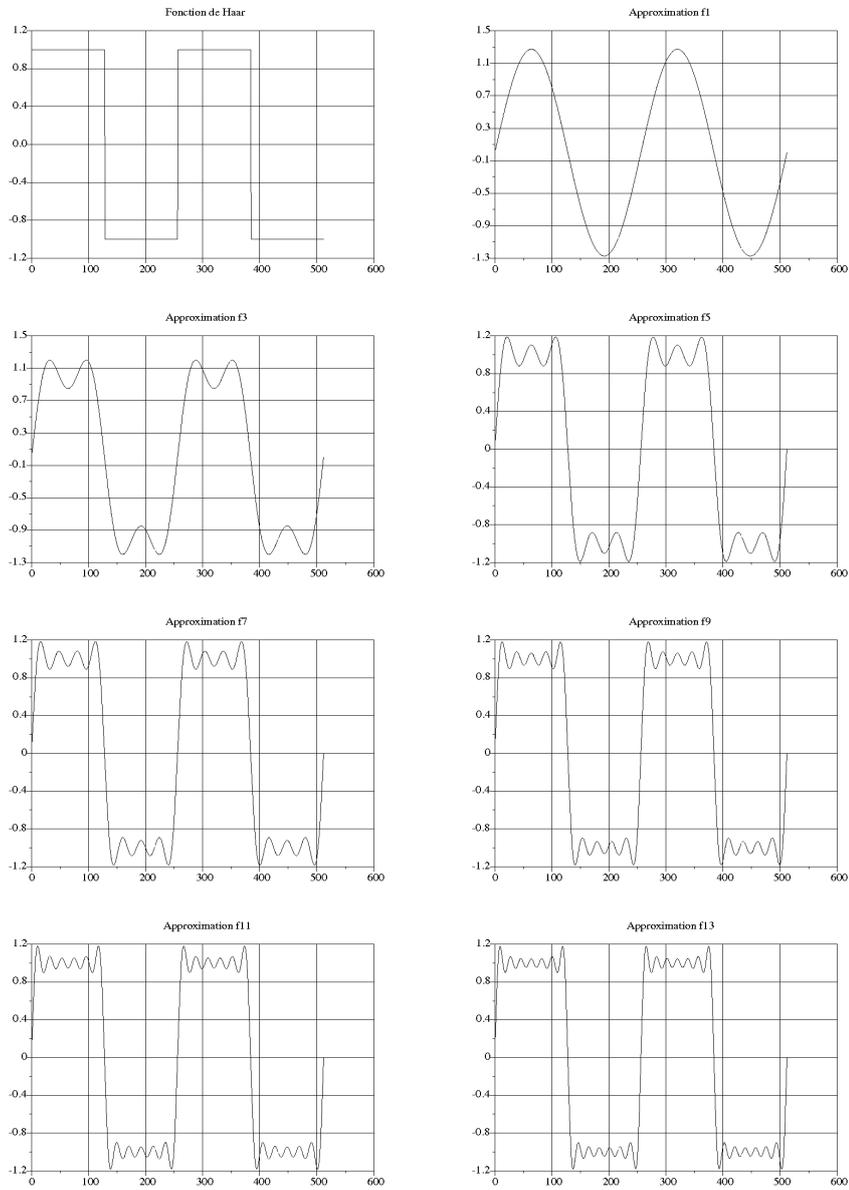


FIG. 3.1: Séries de Fourier tronquées de la fonction de Haar : Le phénomène de Gibbs.

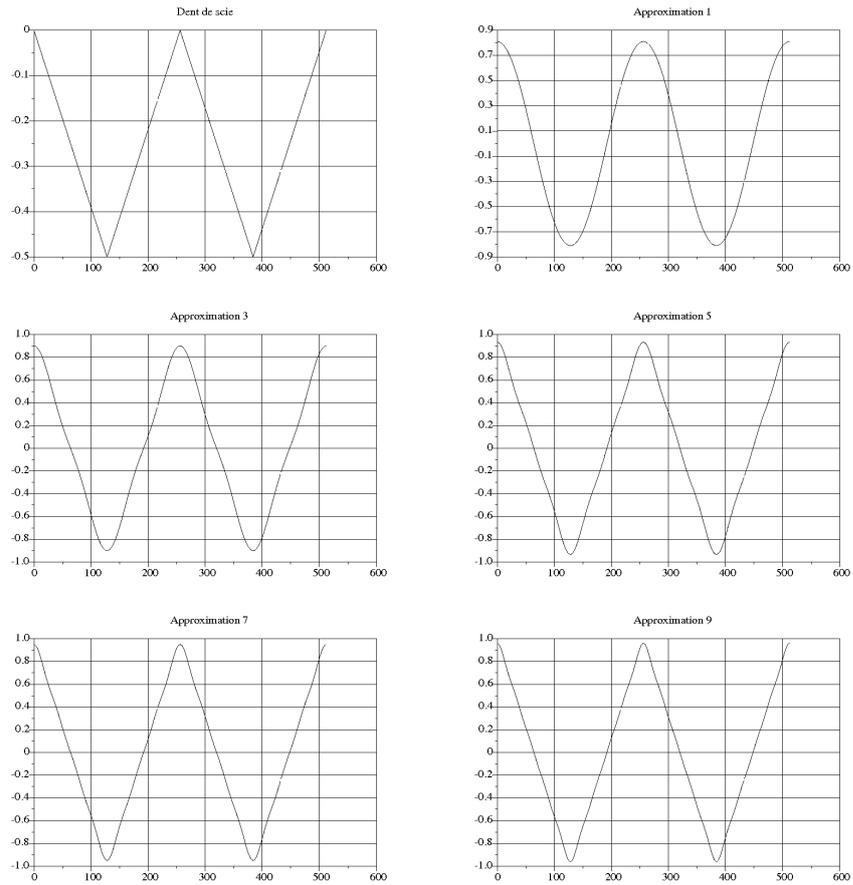


FIG. 3.2: Séries de Fourier tronquées d'une fonction "dents de scie" : Le phénomène de Gibbs.

où I est n'importe quel intervalle de longueur $b - a$.

2. Considérons une fonction $f \in L_p^2([a, b])$, et supposons que f soit paire : $f(-t) = f(t)$ pour tout t . Alors, on a

$$c_{-n}(f) = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{2i\pi nt/T} dt = \frac{1}{T} \int_{-T/2}^{T/2} f(-t) e^{-2i\pi nt/T} dt = c_n(f).$$

De plus, on peut aussi écrire

$$c_n(f) = \frac{1}{2}(c_n(f) + c_{-n}(f)) = \frac{1}{T} \int_{-T/2}^{T/2} f(t) \cos(2\pi nt/T) dt,$$

et

$$f_N(t) = \sum_{n=-N}^N c_n(f) e^{2i\pi nt/T} = \frac{a_0(f)}{2} + \sum_{n=1}^N a_n(f) \cos(2\pi nt/T),$$

où $a_n(f) = 2c_n(f)$.

3. Similairement, supposons maintenant que f soit impaire : $f(-t) = -f(t)$ pour tout t . On a alors

$$f_N(t) = \sum_{n=1}^N b_n(f) \sin(2\pi nt/T),$$

où

$$b_n(f) = 2c_n(f) = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \sin(2\pi nt/T) dt.$$

3.2.5 Problèmes de convergence des séries de Fourier

La théorie L^2 ne donne que des indications relativement limitées sur le comportement des coefficients de Fourier, et sur la convergence des séries de Fourier. Par exemple, la formule de Parseval implique que si $f \in L_p^2([a, b])$, alors $|c_n(f)| \rightarrow 0$ quand $n \rightarrow \infty$. On sait également que la série de Fourier de f converge (fortement) vers f . Ceci entraîne la convergence presque partout, mais pas nécessairement la convergence ponctuelle.

Il est donc utile de se poser de tels problèmes dans des cadres fonctionnels différents. Un cadre intéressant est le cadre L^1 , plus large (dans le cas des fonctions périodiques) que le cadre L^2 . En effet, on a

LEMME 3.1 Soit $f \in L_p^2([a, b])$. Alors $f \in L_p^1([a, b])$.

Preuve : Supposons $f \in L_p^2([a, b])$, et calculons

$$\int_a^b |f(t)| dt \leq \sqrt{\int_a^b dt} \sqrt{\int_a^b |f(t)|^2 dt} < \infty,$$

la première inégalité étant simplement l'inégalité de Cauchy-Schwarz. Ceci montre le lemme. ♠
Commençons par le résultat important suivant, appelé *Théorème de Riemann-Lebesgue*, qui joue un rôle central dans la preuve de nombreux résultats, et que nous retrouverons aussi dans d'autres cadres.

THÉORÈME 3.2 (RIEMANN-LEBESGUE) Soit $f \in L^1([a, b])$. Alors

$$\int_a^b f(t) e^{-int} dt \longrightarrow 0 \text{ quand } n \rightarrow \infty. \quad (3.24)$$

La démonstration utilise le résultat classique suivant

PROPOSITION 3.3 $C^1([a, b])$ est dense dans $L^1([a, b])$.

Ceci signifie que pour tout $f \in L^1([a, b])$, et pour tout $\epsilon > 0$, il existe $g_\epsilon \in C^1([a, b])$ tel que $\|f - g_\epsilon\|_1 \leq \epsilon$.

Preuve du Théorème 3.2 : Supposons dans un premier temps que $f \in C^1([a, b])$. Alors, par intégration par parties, on a

$$\int_a^b f(t)e^{-int} dt = -\frac{1}{in}[f(b)e^{-inb} - f(a)e^{-ina}] + \frac{1}{in} \int_a^b f'(t)e^{-int} dt,$$

qui tend bien vers 0 quand $n \rightarrow \infty$. On utilise maintenant la densité de $C^1([a, b])$ dans $L^1([a, b])$: si $f \in L^1([a, b])$, pour tout $\epsilon > 0$, il existe $g_\epsilon \in C^1([a, b])$ tel que $\int_a^b |f(t) - g_\epsilon(t)| dt < \epsilon/2$. Alors

$$\int_a^b f(t)e^{-int} dt \leq \int_a^b |f(t) - g_\epsilon(t)| dt + \left| \int_a^b g_\epsilon(t)e^{-int} dt \right|.$$

Pour tout ϵ , il existe N tel que la seconde intégrale soit plus petite que $\epsilon/2$ pour $|n| \geq N$. Ceci complète la démonstration. ♠

Une conséquence importante de ce résultat est le résultat suivant de convergence ponctuelle. Pour simplifier, on se limitera dans cette section au cas particulier $b = -a = \pi$. Etant donné un point t_0 , on notera

$$f(t_{0+}) = \lim_{t \rightarrow t_0, t \geq t_0} f(t), \quad f(t_{0-}) = \lim_{t \rightarrow t_0, t \leq t_0} f(t)$$

quand ces limites existent.

THÉORÈME 3.3 (DIRICHLET) Soit $f \in L^1_p([-\pi, \pi])$. Soit t_0 un point tel que les limites $f(t_{0+})$ et $f(t_{0-})$ existent, de même que les dérivées à gauche et à droite de f en t_0 . Alors quand $N \rightarrow \infty$,

$$f_N(t_0) \rightarrow \frac{1}{2}(f(t_{0+}) + f(t_{0-})). \quad (3.25)$$

Preuve : Commençons par évaluer la somme partielle $f_N(t_0)$:

$$\begin{aligned} f_N(t_0) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \left(\sum_{n=-N}^N e^{in(t_0-t)} \right) dt \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \frac{\sin(N+1/2)(t_0-t)}{\sin(t_0-t)/2} dt \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t_0+s) \frac{\sin(N+1/2)s}{\sin s/2} ds \\ &= \frac{1}{\pi} \int_0^{\pi} (f(t_0+s) + f(t_0-s)) \frac{\sin(N+1/2)s}{\sin s/2} ds \end{aligned}$$

En posant $\mu = (f(t_{0+}) + f(t_{0-}))/2$, on a donc

$$f_N(t_0) - \mu = \frac{1}{\pi} \int_0^{\pi} \phi(s) \sin(N+1/2)s ds, \quad (3.26)$$

où on a posé

$$\phi(s) = \frac{f(t_0 + s) - f(t_{0+})}{\sin s/2} + \frac{f(t_0 - s) - f(t_{0-})}{\sin s/2}.$$

Puisque f est supposée différentiable à droite et à gauche en t_0 , $\phi(s)$ admet une limite finie en $s \rightarrow 0$. Donc, il existe $\alpha > 0$ tel que ϕ soit bornée sur $]0, \alpha[$:

$$|\phi(s)| \leq C \quad \text{pour tout } s \in]0, \alpha[.$$

ϕ est donc intégrable sur $]0, \alpha[$. Comme $f \in L^1([-\pi, \pi])$, ϕ est également intégrable sur $]\alpha, \pi[$, et donc sur $[0, \pi]$. Il suffit alors d'appliquer le théorème de Riemann-Lebesgue à (3.26) pour conclure. ♠

EXEMPLE 3.1 Prenons l'exemple de $f \in L^2_p([-\pi, \pi])$ définie par $f(t) = \chi_{[0, \pi]}(t) - \chi_{[-\pi, 0]}(t)$. Un calcul immédiat donne $c_0(f) = 0$, et $c_n(f) = (1 - (-1)^n)/(in\pi)$ pour $n \neq 0$. On obtient également $\sum_{-\infty}^{\infty} c_n(f) \exp\{int\} = 0$ pour $t = k\pi, k \in \mathbb{Z}$, ce qui coïncide bien avec le résultat du théorème précédent.

REMARQUE 3.3 En particulier, si de plus f est continue en t_0 , $f_N(t_0) \rightarrow f(t_0)$ quand $N \rightarrow \infty$.

REMARQUE 3.4 Il est possible de prouver des versions plus "fines" de ce résultat, ainsi que des résultats de convergence uniforme des séries de Fourier. On pourra se référer au livre de Zygmund (*Trigonometric series*) pour plus de détails.

3.2.6 Le problème de l'extension

Le théorème de Riemann-Lebesgue montre en particulier que si $f \in L^1_p([a, b])$, alors $|c_n(f)| \rightarrow 0$ quand $n \rightarrow \pm\infty$. En pratique, on s'intéresse non seulement à la convergence ponctuelle, mais aussi à la vitesse de convergence. Or cette dernière est directement liée à la régularité de f , comme le montre le lemme suivant.

LEMME 3.2 Soit $f \in C^r(\mathbb{R})$, périodique de période $b - a$; alors il existe une constante K telle que

$$|c_n(f)| \leq Kn^{-r}.$$

Preuve : une intégration par parties donne

$$c_n(f) = \frac{1}{b-a} \int_a^b f(t) e^{-2i\pi \frac{nt}{b-a}} dt = \frac{1}{2i\pi n} \int_a^b f'(t) e^{-2i\pi \frac{nt}{b-a}} dt,$$

car f étant continue, on a en particulier $f(b) = f(a)$. Similairement, on a

$$c_n(f) = \frac{1}{b-a} \left(\frac{b-a}{2i\pi n} \right)^r \int_a^b f^{(r)}(t) e^{-2i\pi \frac{nt}{b-a}} dt.$$

Cette dernière intégrale étant bornée par hypothèse, on en déduit le lemme. ♠

REMARQUE 3.5 Ceci permet d'estimer la vitesse de convergence des sommes partielles :

$$\|f - f_N\|^2 = (b-a) \sum_{|n|>N} |c_n(f)|^2 \leq 2(b-a)K^2 \sum_{|n|>N} n^{-2r} \leq K'N^{1-2r}.$$

Cependant, on utilise également les séries de Fourier pour développer des fonctions à support borné. Et il y a là une différence importante avec le cas des fonctions périodiques. En effet, le développement en série de Fourier d'une fonction à support borné n'est *pas unique*, comme on va le voir. Soit $f \in L^2(\mathbb{R})$, à support borné dans l'intervalle $[a, b]$. Soit f_p la fonction périodique de période $b - a$, qui coïncide avec f sur $[a, b]$, définie par

$$f_p(t) = \sum_{k=-\infty}^{\infty} f(t - k(b - a)) .$$

Il est clair que $f_p \in L^2_p([a, b])$. On peut donc la développer en série de Fourier, et écrire, puisque $f = f_p \chi_{[a, b]}$

$$f = \lim_{N \rightarrow \infty} f_N , \tag{3.27}$$

où la limite est toujours à prendre au sens de L^2 (c'est à dire $\lim_{N \rightarrow \infty} \|f - f_N\| = 0$), et où

$$f_N(t) = \left(\sum_{n=-N}^N c_n(f) e^{2i\pi n t / (b-a)} \right) \chi_{[a, b]}(t) , \tag{3.28}$$

et où les coefficients de Fourier $c_n(f) = c_n(f_p)$ sont toujours définis par

$$c_n(f) = \frac{1}{b-a} \int_a^b f(t) e^{-2i\pi n t / (b-a)} dt . \tag{3.29}$$

La série de Fourier de la fonction f_p est unique. Cependant, le passage de f à f_p n'est pas la seule possibilité. Il existe de multiples alternatives, dont on va donner deux exemples ci dessous. Considérons tout d'abord la fonction g , définie sur l'intervalle $[2a - b, b]$ par

$$g(t) = \begin{cases} f(t) & \text{si } t \in [a, b] \\ f(2a - t) & \text{si } t \in [2a - b, a] . \end{cases}$$

g est une fonction symétrique par rapport à a , et on peut considérer sa périodisée, de période $2(b - a)$

$$g_p(t) = \sum_{k=-\infty}^{\infty} g(t - 2k(b - a)) .$$

g_p admet un développement en série de Fourier

$$g_p(t) = \sum_n c_n(g) e^{i\pi n t / (b-a)} ,$$

où

$$c_n(g) = \frac{1}{2(b-a)} \int_{2a-b}^b g_p(t) e^{i\pi n t / (b-a)} dt .$$

Cependant, on peut aussi écrire, pour $n \neq 0$,

$$c_n(g) = e^{-in\pi a / (b-a)} \frac{1}{b-a} \int_a^b f(t) \cos \left(\pi n \frac{t-a}{b-a} \right) dt ,$$

et

$$c_0(g) = c_0(f) .$$

Posons maintenant

$$A_n(f) = \frac{2}{b-a} \int_a^b f(t) \cos \left(\pi n \frac{t-a}{b-a} \right) dt . \tag{3.30}$$

Il est clair que $A_{-n}(f) = A_n(f)$; la série de Fourier de g_p s'écrit maintenant

$$g_p(t) = \frac{1}{2}A_0 + \frac{1}{2} \sum_{n \neq 0} A_n(f) e^{i\pi n(t-a)/(b-a)}.$$

Ceci nous donne directement une autre série de Fourier pour la fonction f :

$$f = \lim_{N \rightarrow \infty} f_N^{(C)}, \quad (3.31)$$

où les sommes partielles $f_N^{(C)}$ sont définies par

$$f_N^{(C)}(t) = \left(\frac{1}{2}A_0 + \sum_{n=1}^N A_n(f) \cos \left(\pi n \frac{t-a}{b-a} \right) \right) \chi_{[a,b]}(t). \quad (3.32)$$

Une autre possibilité consiste à considérer une extension non pas symétrique (comme l'est la fonction g) de f , mais une extension antisymétrique h , définie par

$$h(t) = \begin{cases} f(t) & \text{si } t \in [a, b] \\ -f(2a-t) & \text{si } t \in [2a-b, a]. \end{cases}$$

Il est alors facile de vérifier que la même procédure conduit à un développement en série de sinus : on a

$$f = \lim_{N \rightarrow \infty} f_N^{(S)}, \quad (3.33)$$

où les sommes partielles $f_N^{(S)}$ sont définies par

$$f_N^{(S)}(t) = \left(\sum_{n=1}^N B_n(f) \sin \left(\pi n \frac{t-a}{b-a} \right) \right) \chi_{[a,b]}(t), \quad (3.34)$$

et où les coefficients $B_n(f)$ sont donnés par

$$B_n(f) = \frac{2}{b-a} \int_a^b f(t) \sin \left(\pi n \frac{t-a}{b-a} \right) dt. \quad (3.35)$$

La question qui se pose alors est celle du choix de la série à utiliser. Dans certaines applications, par exemple pour le codage des signaux ou des images, on a intérêt à privilégier la vitesse de décroissance des coefficients du développement de f . Nous avons vu plus haut que celui-ci est directement lié à la régularité, non pas de f elle-même, mais de la fonction périodique utilisée dans le développement, c'est à dire ici f_p , ou g_p , ou la fonction équivalente dans le cas du développement en série de sinus.

Or, même si f est une fonction continue, il est rare qu'elle soit telle que $f(b) = f(a)$. Donc f_p est discontinue, et les coefficients $c_n(f)$ n'ont aucune raison de décroître assez vite quand n est grand. Par contre, si f est continue, alors il est facile de voir que g_p est continue également, de sorte que les coefficients $A_n(f)$ ont toutes les chances de décroître plus rapidement que les coefficients $c_n(f)$.

C'est pourquoi on utilise souvent les séries de cosinus dans les codeurs de signaux comme ceux employés dans les standards de communication (comme JPEG ou MPEG par exemple).

EXEMPLE 3.2 Comme illustration de cet fait, prenons la fonction $f(t) = \chi_{[0,\pi]}$. Un calcul immédiat montre que $c_n(f) = \delta_{n,0}$. Par contre, les coefficients $B_n(f)$ se comportent comme $1/n$; Le développement en série de sinus est donc très inapproprié dans ce cas, comme on peut le voir en FIG. 3.3, avec l'approximation par cosinus (qui est exacte, et identique à la série de Fourier usuelle, et ne comporte qu'un terme) et l'approximation par une série de sinus comportant 10 termes. La série de sinus convergera toujours vers 0 en $t = 0$ et en $t = \pi$.

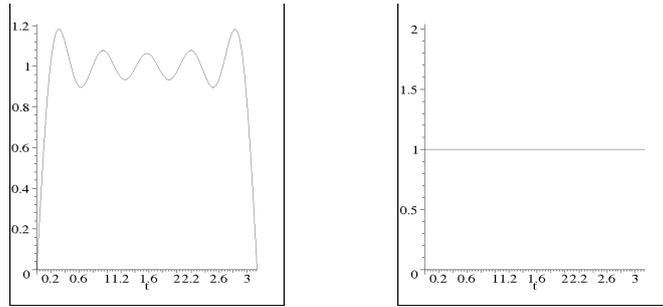


FIG. 3.3: Diverses séries de Fourier décrivant $\chi_{[0,1]}$: série de sinus, série de cosinus.

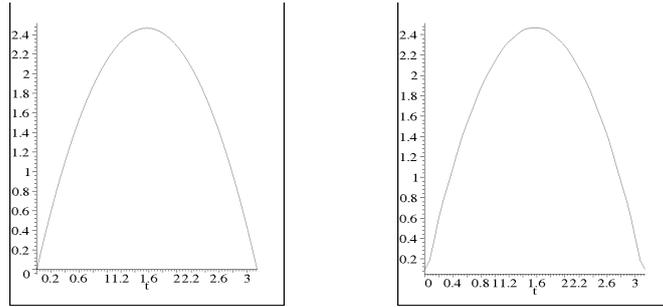


FIG. 3.4: Diverses séries de Fourier décrivant un arc de parabole : arc de parabole, et sa série de Fourier usuelle.

EXEMPLE 3.3 On considère l'exemple de la fonction

$$f(t) = t(\pi - t)$$

définie sur $[0, \pi]$. Un calcul explicite montre que ses coefficients de Fourier sont donnés par

$$c_n(f) = \frac{1}{\pi} \int_0^\pi f(t)e^{-2int} dt = -\frac{1}{2n^2}.$$

Par contre, on a aussi

$$A_n(f) = -\frac{2}{n^2} (1 - (-1)^n),$$

et

$$B_n(f) = -\frac{4}{\pi n^3} (1 - (-1)^n).$$

Donc, dans ce cas particulier, le développement en série de sinus est le plus économique. Les figures 3.4 et 3.5 représentent les approximations obtenues avec ces 3 développements, respectivement $f, f_{10}, f_{10}^{(S)}$ et $f_{10}^{(C)}$. La figure 3.6 représente l'erreur d'approximation dans les 3 cas : $f - f_5, f - f_5^{(S)}$ et $f - f_5^{(C)}$.

3.3 Retour sur la transformation de Fourier discrète

Nous avons vu au chapitre précédent comment associer à une suite de $\ell^1(\mathbb{Z})$ une fonction périodique bornée, dont les coefficients de Fourier ne sont autres que les éléments de la suite. Nous sommes maintenant en position de développer une théorie similaire, valable dans le cadre $\ell^2(\mathbb{Z})$.

La famille de fonctions $\omega \rightarrow e^{in\omega} / \sqrt{2\pi}$ étant une base orthonormée de $L^2([-\pi, \pi])$, on déduit immédiatement de la discussion précédente le résultat suivant.

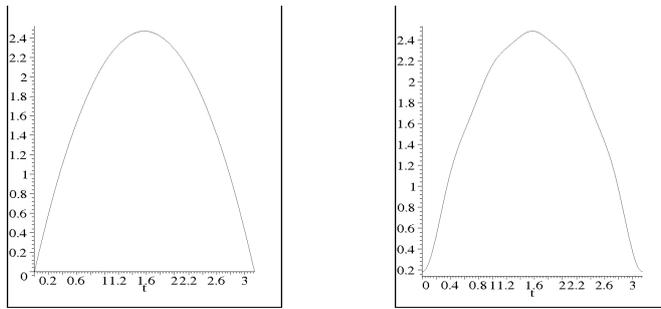


FIG. 3.5: Diverses séries de Fourier décrivant un arc de parabole (suite) : série de sinus, série de cosinus.

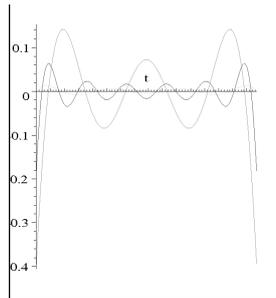


FIG. 3.6: Diverses séries de Fourier décrivant un arc de parabole (suite) : erreurs de reconstruction.

THÉORÈME 3.4 La transformation $s \in \ell^2(\mathbb{Z}) \rightarrow \frac{1}{\sqrt{2\pi}}\hat{s}$, où \hat{s} est la transformée de Fourier discrète de s , est une isométrie bijective de $\ell^2(\mathbb{Z})$ sur $L^2([-\pi, \pi])$. On a la formule de Parseval : pour toutes $u, v \in \ell^2(\mathbb{Z})$,

$$\sum_n \bar{u}_n v_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \bar{\hat{u}}(\omega) \hat{v}(\omega) d\omega . \quad (3.36)$$

De plus, la transformation inverse s'écrit

$$s_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} \hat{s}(\omega) e^{in\omega} d\omega . \quad (3.37)$$

Intégrales de Fourier

On a vu jusqu'à présent plusieurs versions différentes des transformations de Fourier : transformations de Fourier finie et discrète, et séries de Fourier. Dans tous les cas, les propriétés essentielles suivantes étaient vérifiées :

- La transformation de Fourier est une transformation linéaire, inversible pour peu que l'on se place dans un espace approprié.
- *Théorème convolution-produit* : Le produit simple et le produit de convolution possèdent des propriétés remarquables vis à vis de la transformation de Fourier : à une constante multiplicative près, la transformée de Fourier d'un produit simple est égale au produit de convolution des transformées de Fourier. Et la transformée de Fourier d'un produit de convolution est égale (toujours à une constante multiplicative près) au produit simple des transformées de Fourier.
- La transformation de Fourier finie, et la décomposition en séries de Fourier peuvent s'interpréter comme des décompositions par rapport à une base orthonormée de l'espace considéré, le corollaire en étant l'existence de la formule de Parseval.

On s'intéresse maintenant à une autre version de la transformation de Fourier, adaptée cette fois aux fonctions définies sur \mathbb{R} (ou \mathbb{R}^n). On va maintenant voir que ces propriétés restent vraies pour l'essentiel dans le cas de la transformation de Fourier intégrale, à ceci près que de nouvelles difficultés interviennent. Ces difficultés ont trait à la définition même de la transformée de Fourier, qui est définie par une intégrale, qui n'est pas obligatoirement convergente même dans des cas simples. Ceci implique qu'il va falloir faire des hypothèses supplémentaires pour nous assurer l'existence de la transformée de Fourier des fonctions considérées. De même, la transformation de Fourier est inversible, à condition de se placer dans des espaces de fonctions bien choisis. Le théorème convolution-produit reste valable, à condition que toutes les fonctions considérées soient bien définies. On verra aussi que l'interprétation de la décomposition en termes de décomposition par rapport à une base orthonormée n'est plus correcte stricto sensu, mais qu'il reste néanmoins un analogue de la formule de Parseval, qui porte dans ce cas le nom de *formule de Plancherel*.

Avant d'entrer dans le vif du sujet, il est utile de se pencher sur un exemple qui montre l'utilité de la transformation de Fourier.

Préliminaires : circuit RC

On considère le circuit décrit dans la figure 4.1, composé d'une résistance R et d'une capacité C , entretenu par une tension dépendant du temps $t \rightarrow u(t)$. On note $Q(t)$ la charge du condensateur, $i(t)$ l'intensité du courant.

En notant $v(t) = Q(t)/C$ la tension aux bornes du condensateur, la loi d'Ohm s'écrit

$$Ri(t) + v(t) = u(t),$$

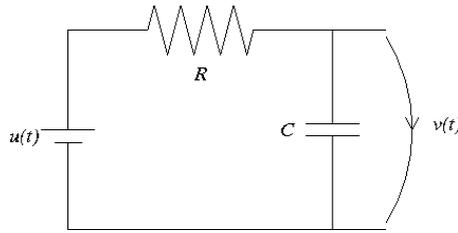


FIG. 4.1: Le filtre RC

ce qui entraîne, puisque $i(t) = Q'(t) = Cv'(t)$, que la tension $v(t)$ satisfait à l'équation différentielle ordinaire

$$RC v'(t) + v(t) = u(t).$$

Pour résoudre cette dernière, il est utile d'introduire la fonction auxiliaire $w(t) = v(t)e^{t/RC}$. On a donc

$$w'(t) = \frac{1}{RC} e^{t/RC} u(t)$$

et la solution est

$$w(t) = \frac{1}{RC} \int_{-\infty}^t e^{s/RC} u(s) ds.$$

Par conséquent,

$$v(t) = \frac{1}{RC} \int_{-\infty}^t e^{-(t-s)/RC} u(s) ds = \int_{-\infty}^{\infty} h(t-s) u(s) ds,$$

où nous avons posé

$$h(t) = \Theta(t) e^{-t/RC}$$

$\Theta(t)$ étant la fonction d'Heaviside :

$$\Theta(t) = \begin{cases} 1 & \text{si } t \geq 0 \\ 0 & \text{sinon.} \end{cases}$$

Nous voyons donc que la solution possède une forme bien particulière, celle d'un produit de convolution, que nous allons maintenant étudier plus avant.

Le produit de convolution de deux fonctions est défini de la façon suivante.

DÉFINITION 4.1 *Etant données deux fonctions f et g , la fonction h , définie par*

$$t \rightarrow h(t) = \int_{-\infty}^{\infty} f(s)g(t-s)ds \quad (4.1)$$

(quand l'intégrale est bien définie) est appelée produit de convolution de f et g , et notée $h = f \star g$.

Etant donnée une fonction g , l'application linéaire K_g qui à toute fonction f associe la fonction $K_g f$, définie par

$$K_g f = f \star g,$$

est appelé *opérateur de convolution* par g .

Énonçons maintenant quelques propriétés simples des produits de convolution. Les propriétés suivantes découlent directement de la définition.

1. Commutativité : étant données deux fonctions f et g , alors $f \star g = g \star f$.
2. Associativité : $(f \star g) \star h = f \star (g \star h)$.
3. Distributivité : $f \star (g_1 + g_2) = (f \star g_1) + (f \star g_2)$.

Une autre propriété caractéristique du produit de convolution nous ramène aux fonctions oscillantes $\epsilon_\omega(t)$. Supposons que f soit une fonction intégrable, et calculons

$$\begin{aligned}(f \star \epsilon_\omega)(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(s) e^{i\omega(t-s)} ds \\ &= \left(\int_{-\infty}^{\infty} f(s) e^{-i\omega s} ds \right) \epsilon_\omega(t),\end{aligned}$$

où il est facile de vérifier que l'intégrale de la dernière ligne ci-dessus est absolument convergente (voir ci-dessous).

Nous retrouvons ici une propriété simple : pour toute fonction f , l'action de K_f sur la fonction oscillante

$$\epsilon_\omega : t \in \mathbb{R} \rightarrow e^{i\omega t}$$

se ramène à une multiplication par une constante (finie) $\int f(s) e^{-i\omega s} ds$. Cette constante n'est autre que la transformée de Fourier de la fonction f (à un facteur près), dont nous donnons maintenant la définition.

La transformation de Fourier

DÉFINITION 4.2 Etant donnée une fonction f , sa transformée de Fourier est une fonction, notée \hat{f} ou encore $\mathcal{F}f$, d'une variable réelle notée ω , définie par

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \quad (4.2)$$

pour toute valeur de ω pour laquelle cette intégrale est convergente.

REMARQUE 4.1 La convergence de l'intégrale ci-dessus est loin d'être un fait acquis dans nombre de situations. Pour s'en assurer, il est nécessaire de faire des hypothèses supplémentaires. Dans certains cas où $\hat{f}(\omega)$ n'existe pas pour tout ω , il est encore possible de construire une théorie pour la transformation de Fourier. Ce sont ces aspects que nous abordons ci-dessous.

On a donc introduit la *transformation de Fourier* \mathcal{F} , qui est une application linéaire agissant sur des espaces de fonctions. On utilisera également la *transformation de Fourier conjuguée* $\overline{\mathcal{F}}$, définie par

$$[\overline{\mathcal{F}}\varphi](t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi(\omega) e^{i\omega t} d\omega. \quad (4.3)$$

On verra que lorsque la transformation de Fourier est inversible, la transformation inverse est donnée par $\overline{\mathcal{F}}$. Mais avant cela, il faut tout d'abord analyser les propriétés élémentaires de \mathcal{F} .

4.1 La transformation de Fourier des fonctions intégrables

Pour donner un sens précis à la transformation de Fourier que nous venons de voir, il est nécessaire de faire des hypothèses supplémentaires sur la fonction étudiée. Il est utile à ce point d'introduire

l'espace des fonctions intégrables (ou absolument intégrables) $L^1(\mathbb{R})$, défini de la façon suivante. Etant donnée une fonction intégrable f , on lui associe le nombre

$$\|f\|_1 = \int_{-\infty}^{\infty} |f(t)| dt. \quad (4.4)$$

L'espace $L^1(\mathbb{R})$ est alors défini par

$$L^1(\mathbb{R}) = \{f : \mathbb{R} \rightarrow \mathbb{C}, \|f\|_1 < \infty\}. \quad (4.5)$$

$L^1(\mathbb{R})$ est le premier des espaces de fonctions que nous rencontrons. On peut vérifier que c'est bien un espace vectoriel, c'est à dire que toute combinaison linéaire d'éléments de $L^1(\mathbb{R})$ est élément de $L^1(\mathbb{R})$. On verra plus loin des propriétés plus fines de $L^1(\mathbb{R})$, ainsi que d'autres espaces du même type.

Supposons donc que f soit une fonction intégrable. Alors, il vient immédiatement que

$$|\hat{f}(\omega)| = \frac{1}{\sqrt{2\pi}} \left| \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \right| \leq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |f(t)| dt = < \infty$$

Par conséquent, $\hat{f}(\omega)$ existe pour tout $\omega \in \mathbb{R}$, et la transformée de Fourier de f est bien définie, et bornée. L'espace $L^1(\mathbb{R})$ est donc un premier cadre naturel pour définir la transformation de Fourier, et étudier ses propriétés essentielles (puisque l'on n'a pas à se soucier du problème de convergence des intégrales).

REMARQUE 4.2 De même, on montre avec les mêmes arguments que si $\varphi \in L^1(\mathbb{R})$, alors $\overline{\mathcal{F}\varphi}$ est une fonction bornée. En fait, les applications \mathcal{F} et $\overline{\mathcal{F}}$ ont des propriétés tout à fait similaires.

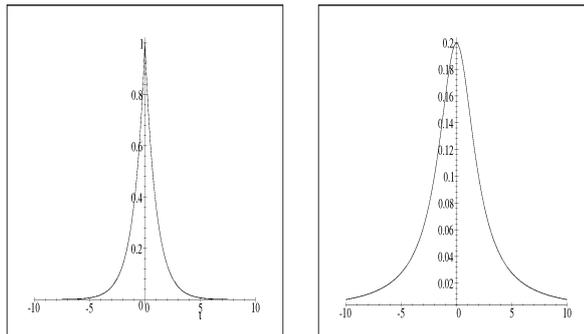


FIG. 4.2: Exponentielle décroissante (à gauche), et sa transformée de Fourier (à droite)

EXEMPLE 4.1 Considérons l'exemple de la fonction $t \rightarrow f(t) = \exp\{-\lambda|t|\}$. Cette fonction est intégrable, et sa transformée de Fourier est donc bien définie en tant que fonction bornée. On peut donc calculer

$$\begin{aligned} \hat{f}(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\lambda|t|} e^{-i\omega t} dt \\ &= \frac{1}{\sqrt{2\pi}} \left(\int_0^{\infty} e^{-(\lambda+i\omega)t} dt + \int_{-\infty}^0 e^{(\lambda-i\omega)t} dt \right) \\ &= \frac{1}{\sqrt{2\pi}} \left(\frac{1}{\lambda+i\omega} + \frac{1}{\lambda-i\omega} \right) = \frac{1}{\sqrt{2\pi}} \frac{2\lambda}{\lambda^2 + \omega^2} \end{aligned}$$

Il s'agit bien d'une fonction bornée, qui est de plus continue et tend vers zéro à l'infini. Nous verrons plus loin qu'il s'agit d'une propriété générale des transformées de Fourier des fonctions intégrables. Les graphes de f et \hat{f} sont donnés en FIG. 4.2.

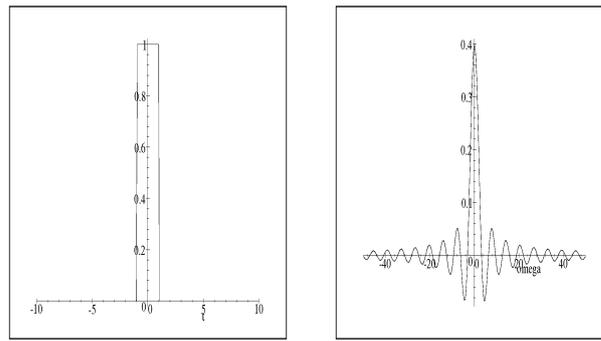


FIG. 4.3: Fonction caractéristique (à gauche), et sa transformée de Fourier (à droite)

EXEMPLE 4.2 Prenons maintenant le cas de $t \rightarrow f(t) = \chi_{[-1,1]}(t)$, la fonction caractéristique de l'intervalle $[-1, 1]$. Un calcul direct montre que

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-1}^1 e^{-i\omega t} dt = \frac{1}{\sqrt{2\pi}} \left(\frac{e^{-i\omega}}{-i\omega} - \frac{e^{i\omega}}{i\omega} \right) = \frac{1}{\sqrt{2\pi}} \frac{\sin \omega}{\omega}.$$

Cette dernière fonction est appelée *sinus cardinal*, et joue un rôle important dans bon nombre d'applications. Notons qu'elle est également bornée, continue (la continuité en $\omega = 0$ vient du fait que $\sin x \sim x$ quand $x \rightarrow 0$), et tend vers 0 quand $\omega \rightarrow \pm\infty$.

Les graphes de f et \hat{f} sont donnés en FIG. 4.3.

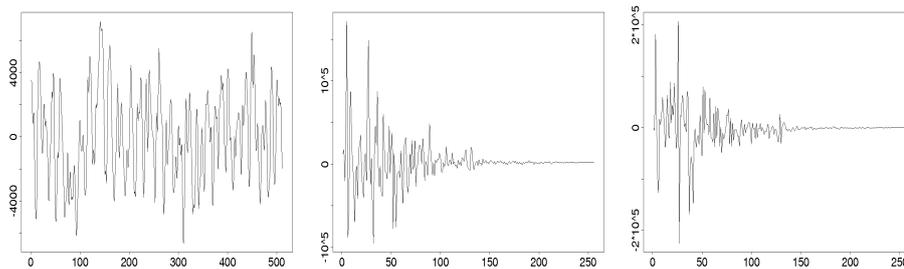


FIG. 4.4: Signal sonar émis par un dauphin (à gauche), et sa transformée de Fourier : partie réelle (au centre) et partie imaginaire (à droite)

EXEMPLE 4.3 Naturellement, la transformation de Fourier n'est pas réservée aux fonctions dont on connaît une expression analytique. En pratique, on peut calculer (sur ordinateur si nécessaire) la transformée de Fourier de n'importe quelle fonction issue d'une mesure physique (on appelle généralement de telles fonctions des *signaux*). Un exemple est montré en FIG. 4.4, qui représente un son émis par un dauphin (sonar), et sa transformée de Fourier (parties réelle et imaginaire).

4.1.1 Propriétés élémentaires

Commençons par quelques propriétés de nature "algébrique".

Linéarité

Soient f_1 et f_2 deux fonctions intégrables. Alors, il est clair que pour tous $\alpha, \beta \in \mathbb{C}$, la fonction $\alpha f_1 + \beta f_2$ est elle aussi intégrable. Un calcul élémentaire montre alors que pour tout $\omega \in \mathbb{R}$,

$$\mathcal{F}(\alpha f_1 + \beta f_2)(\omega) = \alpha \hat{f}_1(\omega) + \beta \hat{f}_2(\omega) \tag{4.6}$$

On dit que la transformation de Fourier est une transformation *linéaire*.

Symétrie Hermitienne

Soit $f \in L^1(\mathbb{R})$. Si f est à valeurs réelles, alors

$$\hat{f}(-\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{i\omega t} dt = \overline{\hat{f}(\omega)} \quad (4.7)$$

On dit alors que la transformée de Fourier \hat{f} possède la *symétrie Hermitienne*. Similairement, supposons que f possède la symétrie Hermitienne, c'est à dire que pour tout $t \in \mathbb{R}$, on ait $f(-t) = \overline{f(t)}$. Alors

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(-t)e^{i\omega t} dt = \overline{\hat{f}(\omega)}, \quad (4.8)$$

de sorte que $\hat{f}(\omega)$ est cette fois une fonction à valeurs réelles.

Comportement vis à vis des translations et des modulations

Considérons maintenant une fonction intégrable f . On définit la *translatée* de f par la quantité $b \in \mathbb{R}$ comme la fonction $g \in L^1(\mathbb{R})$ définie par $g(t) = f(t - b)$. On a alors, par un simple changement de variables

$$\hat{g}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t - b)e^{-i\omega t} dt = \frac{1}{\sqrt{2\pi}} e^{i\omega b} \int_{-\infty}^{\infty} f(t - b)e^{-i\omega(t-b)} dt = e^{i\omega b} \hat{f}(\omega) \quad (4.9)$$

On dit que \hat{g} est une version *modulée* de \hat{f} . Similairement, si $h \in L^1(\mathbb{R})$ définie par $h(t) = e^{i\eta t} f(t)$ est une version modulée de la fonction intégrable f , alors on a

$$\hat{h}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{-i(\omega-\eta)t} dt = \hat{f}(\omega - \eta), \quad (4.10)$$

de sorte que la transformée de Fourier de h est une version translatée de $\hat{f}(\omega)$.

Comportement vis à vis des dilatations

Soit f une fonction intégrable, et soit \hat{f} sa transformée de Fourier. Si a est une constante réelle, on considère une fonction $t \rightarrow f_a(t)$, dilatée de f du facteur a , définie par

$$f_a(t) = f\left(\frac{t}{a}\right).$$

Alors, on a, par un changement de variable $u = t/a$,

$$\hat{f}_a(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f\left(\frac{t}{a}\right) e^{-i\omega t} dt = \frac{a}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(u) e^{-ia\omega u} du = a\hat{f}(a\omega).$$

Ainsi, la transformée de Fourier de la copie dilatée d'une fonction n'est autre qu'une copie dilatée (d'un rapport inverse) de la transformée de Fourier de la fonction originale.

4.1.2 Le théorème de Riemann-Lebesgue

Le théorème de Riemann-Lebesgue est l'un des premiers résultats "fins" concernant la transformation de Fourier. Il précise d'une part la régularité (ici la continuité) et la décroissance de la transformée de Fourier d'une fonction de $L^1(\mathbb{R})$.

Ces deux notions sont fondamentales, et souvent difficiles à appréhender, car on peut leur associer différentes définitions. Sans entrer dans les détails, il faut dire que la régularité d'une fonction va de pair avec la vitesse de la décroissance à l'infini de sa transformée de Fourier. En d'autres termes plus imagés, plus une fonction "varie lentement" (ce que l'on peut prendre comme une idée acceptable de fonction régulière), plus $\hat{f}(\omega)$ tend rapidement vers zéro quand $\omega \rightarrow \infty$.

THÉORÈME 4.1 (RIEMANN-LEBESGUE) Soit $f \in L^1(\mathbb{R})$. Alors sa transformée de Fourier \hat{f} est bornée, et uniformément continue : pour tout $\epsilon > 0$, il existe δ tel qu'en tout point $\omega \in \mathbb{R}$

$$|\hat{f}(\omega + \delta) - \hat{f}(\omega)| \leq \epsilon \quad (4.11)$$

De plus, $\hat{f}(\omega)$ tend vers zéro quand ω tend vers $\pm\infty$.

Preuve : Nous savons déjà que si $f \in L^1(\mathbb{R})$, alors \hat{f} est bornée. Passons à la continuité. La preuve est relativement simple. Commençons par calculer, pour un ω quelconque,

$$\begin{aligned} \hat{f}(\omega + \delta) - \hat{f}(\omega) &= \frac{1}{\sqrt{2\pi}} \int f(t) \left(e^{-i(\omega+\delta)t} - e^{-i\omega t} \right) dt \\ &= \frac{-2i}{\sqrt{2\pi}} \int f(t) \sin\left(\frac{t\delta}{2}\right) e^{-it\delta/2} e^{-i\omega t} dt \end{aligned}$$

Le fait que f soit intégrable implique qu'il existe T tel que $\int_{-T}^T |f(t)| dt + \int_T^\infty |f(t)| dt \leq \sqrt{2\pi}\epsilon/4$. Nous n'avons donc plus qu'à nous préoccuper de l'intégrale entre $-T$ et T . Mais dans cet intervalle, nous savons que $|\sin(t\delta/2)| \leq |t\delta/2| \leq T\delta/2$. Donc, nous avons

$$\begin{aligned} \left| \hat{f}(\omega + \delta) - \hat{f}(\omega) \right| &\leq \frac{\epsilon}{2} + \frac{1}{\sqrt{2\pi}} \frac{T\delta}{2} \int_{-T}^T |f(t)| dt \\ &\leq \frac{\epsilon}{2} + \frac{1}{\sqrt{2\pi}} \frac{T\delta}{2} \|f\|_1 \end{aligned}$$

Il suffit maintenant de choisir δ de sorte que $T\delta\|f\|_1 \leq \sqrt{2\pi}\epsilon$, et on obtient bien l'estimation souhaitée (4.11). Notons au passage que le δ obtenu ne dépend pas de ω . La première partie du théorème est donc montrée.

Le fait que $\hat{f}(\omega) \rightarrow 0$ quand $\omega \rightarrow \pm\infty$ résulte de la densité des fonctions constantes par morceaux dans l'espace $L^1(\mathbb{R})$: pour toute fonction $f \in L^1(\mathbb{R})$, il existe une suite de fonctions g_n , constantes par morceaux, telle que pour tout ϵ fixé, on ait $\|f - g_n\|_1 \leq \epsilon$ pour un n assez grand. Il suffit donc d'étudier le comportement de la fonction caractéristique d'un intervalle ; prenons $g = \chi_{[a,b]}$. Alors, on a $\hat{g}(\omega) = (e^{-i\omega b} - e^{-i\omega a})/i\omega\sqrt{2\pi}$, qui tend bien vers 0 quand $|\omega| \rightarrow \infty$. De même, la transformée de Fourier de toute fonction intégrable constante par morceaux tend vers 0 à l'infini. Pour conclure, il suffit de remarquer que pour tout ω , on a par hypothèse $|\hat{f}(\omega) - \hat{g}_n(\omega)| \leq \|f - g_n\|_1$. Comme pour tout n , $\hat{g}_n(\omega) \rightarrow 0$ quand $|\omega| \rightarrow \infty$, et comme $\|f - g_n\|_1$ peut être rendu aussi proche de 0 que ce que l'on veut, on en déduit que $\hat{f}(\omega) \rightarrow 0$ quand $|\omega| \rightarrow \infty$. ♠

EXEMPLE 4.4 Dans les deux exemples que nous avons vus plus haut (à savoir $f(t) = e^{-\lambda|t|}$ et $f(t) = \chi_{[-1,1]}(t)$), nous avons affaire à une fonction de $L^1(\mathbb{R})$, et on vérifie directement la continuité et la décroissance à l'infini de \hat{f} .

EXEMPLE 4.5 *Contre exemple* : Remarquons que l'hypothèse d'intégrabilité est importante. Prenons l'exemple suivant :

$$f(t) = \frac{1}{t+i}.$$

Un calcul simple par la méthode des résidus montre que pour $\omega < 0$, $\hat{f}(\omega) = 0$, alors que pour $\omega > 0$, $\hat{f}(\omega) = i\sqrt{2\pi}e^{-\omega}$. Dans ce cas, $f \notin L^1(\mathbb{R})$, et ceci se traduit par le fait que \hat{f} est discontinue en 0.

EXEMPLE 4.6 *Un autre contre exemple* : Prenons l'exemple suivant :

$$f(t) = \cos\left(\frac{t^2}{2}\right).$$

Bien que cette fonction n'appartienne pas à $L^1(\mathbb{R})$ (elle ne tend absolument pas vers 0 à l'infini), il est quand même possible de lui associer une transformée de Fourier. Le résultat est

$$\hat{f}(\omega) = \cos\left(\frac{\omega^2}{2} - \frac{\pi}{4}\right).$$

Dans ce cas aussi, $f \notin L^1(\mathbb{R})$, et ceci se traduit cette fois par le fait que $\hat{f}(\omega)$ ne tend pas vers 0 quand $\omega \rightarrow \pm\infty$.

4.1.3 Dérivabilité

Nous allons maintenant nous focaliser quelque peu sur les propriétés de régularité des transformées de Fourier des fonctions intégrables. Le théorème de Riemann-Lebesgue ci-dessus nous a déjà permis de vérifier la continuité de la transformée de Fourier des fonctions de $L^1(\mathbb{R})$. Nous allons voir que sous des hypothèses appropriées, on peut montrer des propriétés de régularité plus forte. Començons par la différentiabilité.

PROPOSITION 4.1 Soit f une fonction intégrable, telle que la fonction $t \rightarrow tf(t)$ soit elle aussi intégrable. Alors sa transformée de Fourier \hat{f} admet en tout point une dérivée continue, qui n'est autre que la transformée de Fourier de la fonction $t \rightarrow -itf(t)$.

Preuve : Calculons donc

$$\begin{aligned} \hat{f}(\omega + \delta) - \hat{f}(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{-i\omega t} (e^{-i\omega\delta} - 1) dt \\ &= \frac{\delta}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (-itf(t)) \frac{\sin(\delta t/2)}{\delta t/2} e^{-i\delta t/2} e^{-i\omega t} dt \end{aligned}$$

En divisant des deux cotés par δ , et en utilisant des arguments similaires aux précédents, il est facile de montrer que la limite du quotient ainsi obtenu, quand $\delta \rightarrow 0$, existe. Comme $\lim_{u \rightarrow 0} \sin(u)/u = 1$, ceci montre la proposition. ♠

En itérant ce résultat, on montre que si f est une fonction intégrable, telle que les fonctions $t \rightarrow tf(t), \dots, t^n f(t)$ sont intégrables alors \hat{f} admet n dérivées continues, qui ne sont autres que les transformées de Fourier des fonctions $t \rightarrow (-it)^k f(t), k = 1, \dots, n$.

PROPOSITION 4.2 Soit $f : t \rightarrow f(t)$ une fonction intégrable, telle que les fonctions $t \rightarrow tf(t)$, $t \rightarrow t^2f(t), \dots, t \rightarrow t^n f(t)$ soient elles aussi intégrables. Alors sa transformée de Fourier \hat{f} admet en tout point des dérivées d'ordre $1, 2, \dots, n$ continues, qui ne sont autres que les transformées de Fourier des fonctions $t \rightarrow (-it)^k f(t)$.

REMARQUE 4.3 Notons que le résultat ainsi obtenu est celui que l'on obtient en dérivant sous le signe somme. La démonstration que nous avons faite donne une justification à cette opération.

4.1.4 Transformation de Fourier d'une dérivée

On vient de voir que la transformée de Fourier de la fonction $t \rightarrow (-it)^k f(t)$ (si celle-ci est intégrable) est la dérivée k -ième de \hat{f} . On se pose maintenant la question "duale", à savoir, "quelle est la transformée de la dérivée d'une fonction ?" Nous allons montrer le résultat suivant :

PROPOSITION 4.3 Soit $f : t \rightarrow f(t)$ une fonction intégrable, et supposons que les dérivées $f^{(k)}, k = 1, \dots, n$ de f existent presque partout et sont intégrables. Alors pour tout $k \leq n$, la fonction $f^{(k)}$ a pour transformée de Fourier la fonction $\omega \rightarrow (i\omega)^k \hat{f}(\omega)$.

Preuve : Pour démontrer ce résultat, considérons une fonction $t \rightarrow f(t)$ satisfaisant aux hypothèses du théorème, et montrons tout d'abord que $f(t) \rightarrow 0$ quand $t \rightarrow \infty$. Supposons que $\lim_{t \rightarrow \infty} f(t)$ existe. Alors comme f est intégrable, cette limite est nécessairement nulle. Reste à vérifier $\lim_{t \rightarrow \infty} f(t)$ existe. Pour cela, écrivons $f(t) = f(0) + \int_0^t f'(s) ds$. Comme f' est supposée intégrable, $\lim_{t \rightarrow \infty} \int_0^t f'(s) ds$ existe, et donc $\lim_{t \rightarrow \infty} f(t)$ existe. De même, on montre que $\lim_{t \rightarrow -\infty} f(t)$ existe.

On peut donc intégrer par parties dans l'intégrale qui définit la transformée de Fourier de f' , et on obtient

$$\int_{-\infty}^{\infty} f'(t) e^{-i\omega t} dt = \left[e^{-i\omega t} f(t) \right]_{-\infty}^{\infty} + i\omega \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt$$

et dont

$$(\mathcal{F}f')(\omega) = i\omega \hat{f}(\omega) \quad (4.12)$$

De façon plus générale, en intégrant par parties autant de fois qu'il le faut, on obtient de même, pour $k \leq n$,

$$(\mathcal{F}f^{(k)})(\omega) = (i\omega)^k \hat{f}(\omega), \quad (4.13)$$

ce qui est le résultat souhaité. ♠

Ce résultat, combiné au théorème de Riemann-Lebesgue, a le corollaire important suivant :

COROLLAIRE 4.1 Soit f une fonction intégrable, dont les dérivées $f^{(k)}, k = 1, \dots, n$ existent presque partout et sont intégrables. Alors il existe une constante C telle que

$$|\hat{f}(\omega)| \leq \frac{C}{1 + |\omega|^n}. \quad (4.14)$$

Preuve : D'après le théorème de Riemann-Lebesgue, on sait que la transformée de Fourier de f est bornée (donc, il existe C_1 telle que $|\hat{f}(\omega)| \leq C_1$ pour tout ω), et que la transformée de Fourier $\omega \rightarrow (i\omega)^n \hat{f}(\omega)$ de $f^{(n)}$ tend vers zéro quand $\omega \rightarrow \pm\infty$. Donc, a fortiori, il existe une constante C_2 telle que $|\hat{f}(\omega)| \leq C_2/|\omega|^n$. La borne (4.14) est une conséquence de ces deux bornes. Il suffit en effet d'écrire,

pour $|\omega| \geq 1$, que $|\hat{f}(\omega)| \leq 2C_2/(|\omega|^n + |\omega|^n) \leq 2C_2/(1 + |\omega|^n)$; de même, pour $|\omega| \leq 1$, on utilise $|\hat{f}(\omega)| \leq 2C_1/(1 + 1) \leq 2C_1/(1 + |\omega|^n)$; on obtient alors (4.14) en prenant $C = \max(2C_1, 2C_2)$. ♠

Ce résultat permet de préciser quelque peu l'assertion faite plus tôt. Il montre clairement que plus une fonction est régulière (c'est à dire dans ce cas, plus elle possède de dérivées continues), plus sa transformée de Fourier tend vers 0 rapidement à l'infini.

4.1.5 Le problème de l'inversion

Il s'agit d'un problème délicat. Il suffit pour s'en convaincre de remarquer que la transformation de Fourier étant définie à l'aide de l'intégrale de Lebesgue, deux fonctions différant sur un ensemble de mesure nulle ont la même transformée de Fourier. Par exemple, si f est une fonction intégrable, et si la fonction \tilde{f} est telle que $\tilde{f}(t) = f(t)$ sauf pour un nombre fini de valeurs de l'argument t , f et \tilde{f} ont la même transformée de Fourier. Par conséquent, il faut s'attendre à devoir imposer des hypothèses supplémentaires si on veut inverser la transformation de Fourier.

Nous avons déjà introduit en (4.3) l'application linéaire $\overline{\mathcal{F}}$. Pour les mêmes raisons que précédemment, $(\overline{\mathcal{F}}g)$ définit une fonction bornée dès que $g \in L^1(\mathbb{R})$. Nous allons maintenant montrer que sous certaines conditions, $\overline{\mathcal{F}}$ est effectivement l'inverse de la transformation de Fourier, c'est à dire que l'on peut écrire

$$f(t) = (\overline{\mathcal{F}}\hat{f})(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\omega)e^{i\omega t} d\omega$$

Cependant, pour les raisons évoquées plus haut, ce "résultat" est à prendre avec précautions. Nous allons en fait montrer le résultat suivant :

THÉORÈME 4.2 Soit f une fonction telle que $f, \hat{f} \in L^1(\mathbb{R})$. Supposons que f soit continue en $t = t_0$. Alors on a

$$f(t_0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\omega)e^{i\omega t_0} d\omega \tag{4.15}$$

Ce résultat utilise le résultat suivant, appelé *lemme d'échange*

LEMME 4.1 Soient $f, g \in L^1(\mathbb{R})$ deux fonctions intégrables. Alors on a l'égalité

$$\int_{-\infty}^{\infty} f(t)\hat{g}(t)dt = \int_{-\infty}^{\infty} \hat{f}(s)g(s)ds \tag{4.16}$$

Preuve du lemme : Puisque f et g sont intégrables, \hat{f} et \hat{g} sont bornées, et les produits $f\hat{g}$ et $\hat{f}g$ sont intégrables. Les deux intégrales dans (4.16) sont donc convergentes, et on peut écrire

$$\int_{-\infty}^{\infty} f(t)\hat{g}(t)dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)g(s)e^{-ist} dsdt .$$

Le lemme de Fubini s'applique, et l'intégration par rapport à t donne précisément (4.16). ♠

Preuve du théorème : Considérons la famille de fonctions intégrables définies par

$$g_n(t) = e^{-|t|/n} . \tag{4.17}$$

Leur transformée de Fourier est donnée par

$$\hat{g}_n(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-|t|/n} e^{-i\omega t} dt = \frac{1}{\sqrt{2\pi}} \frac{2n}{1 + n^2\omega^2} \tag{4.18}$$

La formule d'échange donne alors

$$\int_{-\infty}^{\infty} \hat{f}(\omega) g_n(\omega) e^{i\omega t} d\omega = \int_{-\infty}^{\infty} f(s) \hat{g}_n(s-t) ds \quad (4.19)$$

Intéressons nous tout d'abord au terme de gauche de cette égalité. Nous savons que $\lim_{n \rightarrow \infty} g_n(t) = 1$ pour tout t , et que par ailleurs, $|\hat{f}(\omega) g_n(\omega) e^{i\omega t}| \leq |\hat{f}(\omega)|$, qui est fini. Le théorème de convergence dominée de Lebesgue (voir plus loin) nous assure donc que pour tout t ,

$$\int_{-\infty}^{\infty} \hat{f}(\omega) g_n(\omega) e^{i\omega t} d\omega \rightarrow \int_{-\infty}^{\infty} \hat{f}(\omega) e^{i\omega t} d\omega \text{ quand } n \rightarrow \infty.$$

Passons maintenant au membre de droite, et notons tout d'abord que $\int_{-\infty}^{\infty} \hat{g}_n(u) du = \sqrt{2\pi}$. Calculons

$$\int_{-\infty}^{\infty} f(s+t) \hat{g}_n(s) ds - \sqrt{2\pi} f(t) = \int_{-\infty}^{\infty} [f(s+t) - f(t)] \hat{g}_n(s) ds. \quad (4.20)$$

Supposons que f soit continue en $t = t_0$. Donc, pour tout $\epsilon > 0$, il existe un $\rho > 0$ tel que $s \leq \rho$ entraîne $|f(t_0 + s) - f(t_0)| \leq \epsilon$. En traitant de façon différente les valeurs $s \leq \rho$ et $s > \rho$ dans le membre de droite de (4.20), nous sommes amenés à considérer

$$\left| \int_{|s| \leq \rho} [f(s+t_0) - f(t_0)] \hat{g}_n(s) ds \right| \leq \epsilon \int_{|s| \leq \rho} g_n(s) ds \leq \epsilon,$$

que nous pouvons rendre aussi petit que nous voulons. Quant à l'autre terme, nous avons à considérer

$$\left| f(t_0) \int_{|s| > \rho} |\hat{g}_n(s)| ds \right| = |f(t_0)| \frac{4n}{\sqrt{2\pi}} \int_{\rho}^{\infty} \frac{ds}{1+n^2s^2} = \frac{4}{\sqrt{2\pi}} \left(\frac{\pi}{2} - \arctan(\rho n) \right) |f(t_0)|$$

qui tend vers zéro quand $n \rightarrow \infty$, et

$$\left| \int_{|s| > \rho} f(t_0 + s) \hat{g}_n(s) ds \right| \leq \hat{g}_n(\rho) \|f\|_1$$

qui tend lui aussi vers zéro quand $n \rightarrow \infty$. Ainsi, la limite du membre de droite de (4.19), au point de continuité $t = t_0$, n'est autre que $f(t_0)$. Ceci achève la preuve du théorème. ♠

4.1.6 La convolution

Le produit de convolution, brièvement évoqué plus haut, joue un rôle central. Il est donc important d'étudier ses propriétés. Supposons que f soit une fonction intégrable, et que \hat{g} soit elle aussi une fonction intégrable. Par conséquent, les fonctions définies par

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \quad (4.21)$$

$$g(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{g}(\omega) e^{i\omega t} d\omega \quad (4.22)$$

sont bornées. De plus, la fonction h définie par

$$h(t) = (f \star g)(t) = \int_{-\infty}^{\infty} f(s) g(t-s) ds \quad (4.23)$$

existe pour tout $t \in \mathbb{R}$, et est elle aussi bornée. Si on remplace $g(t)$ par sa définition dans l'expression de h , on obtient

$$h(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(s) e^{i\omega(t-s)} \hat{g}(\omega) d\omega ds, \quad (4.24)$$

que l'on peut encore écrire en utilisant le lemme de Fubini (l'intégrale étant absolument convergente, c'est licite) et la définition de la transformation de Fourier

$$h(t) = \int_{-\infty}^{\infty} \hat{g}(\omega) \hat{f}(\omega) e^{i\omega t} d\omega \quad (4.25)$$

h s'écrit ainsi (à un facteur $\sqrt{2\pi}$ près) comme la transformée de Fourier inverse du produit $\hat{f}\hat{g}$. Remarquons que cette transformée est bien définie, dans la mesure où \hat{f} est bornée, et \hat{g} est intégrable par hypothèse.

De même, la fonction produit $fg : t \rightarrow (fg)(t) = f(t)g(t)$ est elle aussi intégrable ; on peut donc calculer

$$\begin{aligned} [\mathcal{F}(fg)](\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)g(t)e^{-i\omega t} dt \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t)\hat{g}(\eta)e^{i(\eta-\omega)t} d\eta dt \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{g}(\eta)\hat{f}(\omega-\eta)d\eta \end{aligned}$$

et vaut donc (à une constante près) le produit de convolution $f \star g(\omega)$. Nous avons donc montré :

PROPOSITION 4.4 (CONVOLUTION-PRODUIT) Si $f, \hat{g} \in L^1(\mathbb{R})$, et si on définit \hat{f} et g par $\hat{f} = \mathcal{F}f$ et $g = \overline{\mathcal{F}\hat{g}}$, alors

$$(f \star g) = \sqrt{2\pi} \overline{\mathcal{F}}(\hat{f}\hat{g}) \quad (4.26)$$

et

$$\mathcal{F}(fg) = \frac{1}{\sqrt{2\pi}} (\hat{f} \star \hat{g}) \quad (4.27)$$

Ce résultat est d'une importance pratique considérable, en particulier pour tout ce qui concerne la résolution des équations différentielles et équations aux dérivées partielles. Il est aussi utile d'un point de vue calculatoire, comme nous le montre l'exemple suivant.

EXEMPLE 4.7 Considérons les deux fonctions suivantes : $f(t) = e^{-t^2/2a^2}$, et $g(t) = e^{-t^2/2b^2}$, et supposons que nous ayons à calculer leur produit de convolution $f \star g$. Nous savons que la transformée de Fourier de la fonction h définie par $h(t) = e^{-t^2/2}$ est $\hat{h}(\omega) = e^{-\omega^2/2}$. Donc, $\hat{f}(\omega) = ae^{-a^2\omega^2/2}$ et $\hat{g}(\omega) = be^{-b^2\omega^2/2}$, et $\hat{f}(\omega)\hat{g}(\omega) = abe^{-(a^2+b^2)\omega^2/2}$. Le résultat précédent donne donc $(f \star g)(t) = \sqrt{2\pi} \frac{ab}{\sqrt{a^2+b^2}} e^{-t^2/2(a^2+b^2)}$. Nous avons pu ainsi éviter un calcul (simple mais sans intérêt) d'intégrales Gaussiennes.

4.2 Les fonctions de carré intégrable

Le cadre des fonctions de carré intégrables est souvent un cadre naturel pour formuler les problèmes de physique. Il est donc intéressant d'étudier les propriétés des transformées de Fourier des fonctions de carré intégrable. Cependant, il faut remarquer immédiatement que contrairement au cas précédent, la transformée de Fourier d'une fonction de carré intégrable existe en tout point. En effet, supposer $f \in L^2(\mathbb{R})$ ne permet pas *a priori* d'assurer immédiatement que contrairement au cas précédent, la transformée de Fourier d'une fonction de carré intégrable existe en tout point. En effet, supposer $f \in L^2(\mathbb{R})$ ne permet pas *a priori* d'assurer que l'intégrale

$$\int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt$$

converge pour toute valeur de t . Les transformées de Fourier des fonctions de carré intégrable qui ne sont pas intégrables ne sont pas bornées en général, et seront à manipuler avec précaution.

Il existe plusieurs manières (équivalentes) de définir la transformation de Fourier sur $L^2(\mathbb{R})$. La plupart sont basées sur une construction sur un sous-espace dense de $L^2(\mathbb{R})$, et étendent la définition par passage à la limite. C'est cette approche que nous emploierons, en nous basant sur l'espace $\mathcal{S}(\mathbb{R})$ de Schwartz.

4.2.1 La transformation de Fourier dans l'espace de Schwartz $\mathcal{S}(\mathbb{R})$

L'espace de Schwartz $\mathcal{S}(\mathbb{R})$ possède une propriété remarquable : la transformée de Fourier d'une fonction appartenant à $\mathcal{S}(\mathbb{R})$ est elle aussi une fonction de $\mathcal{S}(\mathbb{R})$. Pour nous en convaincre, commençons par faire les remarques suivantes :

1. Si $f \in \mathcal{S}(\mathbb{R})$, alors pour tout polynôme $t \rightarrow P(t)$, la fonction $Pf : t \rightarrow P(t)f(t)$ est elle aussi une fonction de $\mathcal{S}(\mathbb{R})$.
2. Si $f \in \mathcal{S}(\mathbb{R})$, alors $f' \in \mathcal{S}(\mathbb{R})$, et plus généralement, toutes les dérivées $f^{(k)}$ de f sont des fonctions de $\mathcal{S}(\mathbb{R})$.
3. $\mathcal{S}(\mathbb{R}) \subset L^1(\mathbb{R})$.

On déduit de ces remarques que la transformation de Fourier est bien définie sur $\mathcal{S}(\mathbb{R})$: toute fonction de $\mathcal{S}(\mathbb{R})$ possède une transformée de Fourier bornée. De plus, si $f \in \mathcal{S}(\mathbb{R})$, alors pour tout k , $t^k f(t)$ est aussi dans $\mathcal{S}(\mathbb{R})$ et est ainsi intégrable ; donc $\hat{f} \in C^k(\mathbb{R})$, et ce pour tout k . Donc $f \in C^\infty(\mathbb{R})$. Le même raisonnement s'applique à toutes les dérivées de f : $f^{(k)} \in \mathcal{S}(\mathbb{R})$ implique que la fonction $\omega \rightarrow \omega^k \hat{f}(\omega)$ est bornée, et ce quel que soit k . Donc nous avons montré que la transformée de Fourier $\omega \rightarrow \hat{f}(\omega)$ de toute fonction $f \in \mathcal{S}(\mathbb{R})$ est elle même une fonction de $\mathcal{S}(\mathbb{R})$.

Inversement, $\hat{f} \in \mathcal{S}(\mathbb{R}) \subset L^1(\mathbb{R})$, et la discussion précédente s'applique tout aussi bien à \hat{f} . Par conséquent, nous avons montré

THÉORÈME 4.3 *La transformation de Fourier est une bijection entre $\mathcal{S}(\mathbb{R})$ et $\mathcal{S}(\mathbb{R})$.*

Soient $f, g \in \mathcal{S}(\mathbb{R})$, et posons $h(\omega) = \overline{\hat{g}(\omega)}$. Alors un calcul simple montre que $\hat{h}(t) = \overline{g(t)}$. Appliquons la formule d'échange à f et h . Ceci nous donne la *formule de Plancherel-Parseval* :

$$\int_{-\infty}^{\infty} f(t) \overline{g(t)} dt = \int_{-\infty}^{\infty} \hat{f}(\omega) \overline{\hat{g}(\omega)} d\omega, \quad (4.28)$$

et en particulier dans le cas $g = f$:

$$\int_{-\infty}^{\infty} |f(t)|^2 dt = \int_{-\infty}^{\infty} |\hat{f}(\omega)|^2 d\omega, \quad (4.29)$$

(il découle de la discussion ci-dessus que toutes ces intégrales sont convergentes. Par conséquent, dans ce cas, nous avons aussi $f \in L^2(\mathbb{R})$ et $\hat{f} \in L^2(\mathbb{R})$). Nous avons donc montré :

PROPOSITION 4.5 *La transformation de Fourier est une isométrie de $\mathcal{S}(\mathbb{R})$ sur $\mathcal{S}(\mathbb{R})$.*

4.2.2 Le passage à $L^2(\mathbb{R})$

Le passage au cadre $L^2(\mathbb{R})$ se fait en utilisant la densité de $\mathcal{S}(\mathbb{R})$ dans $L^2(\mathbb{R})$, et le théorème général d'analyse fonctionnelle suivant :

THÉORÈME 4.4 Soient E et F deux espaces vectoriels normés, F étant un espace complet, et soit $E' \subset E$ un sous-espace dense de E . Soit $\Phi' : E' \rightarrow \Phi'(E') \subset F$ une application isométrique bijective. Alors, il existe une unique application $\Phi : E \rightarrow F$, isométrique et bijective de E sur F , qui prolonge Φ' (c'est à dire telle que sa restriction à E' coïncide avec Φ').

En considérant le cas $E = F = L^2(\mathbb{R})$ et $E' = \mathcal{S}(\mathbb{R})$, $\Phi' = \mathcal{F}$ (la transformation de Fourier sur $\mathcal{S}(\mathbb{R})$), et en utilisant les résultats obtenus dans la sous-section précédente, on obtient directement le résultat important suivant :

THÉORÈME 4.5 (TRANSFORMATION DE FOURIER SUR $L^2(\mathbb{R})$) 1. La transformation de Fourier sur $\mathcal{S}(\mathbb{R})$ se prolonge en une isométrie bijective de $L^2(\mathbb{R})$ sur $L^2(\mathbb{R})$. De plus, on a la formule de Parseval-Plancherel : $\forall f, g \in L^2(\mathbb{R})$,

$$\int_{-\infty}^{\infty} f(t)\overline{g(t)}dt = \int_{-\infty}^{\infty} \hat{f}(\omega)\overline{\hat{g}(\omega)}d\omega . \quad (4.30)$$

2. Si $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, les deux définitions de la transformée de Fourier \hat{f} coïncident.

Ce dernier résultat, pour important qu'il soit, n'est pas constructif, au sens où la transformation de Fourier dans $L^2(\mathbb{R})$ n'y est construite que par un argument de passage à la limite abstrait. Il est complété par la proposition suivante, qui prouve que la transformée de Fourier d'une fonction de $L^2(\mathbb{R})$ s'obtient (aux points où elle est bien définie) via le calcul usuel.

PROPOSITION 4.6 Si $f \in L^2(\mathbb{R})$, $\hat{f}(\omega)$ peut s'obtenir comme

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \lim_{T \rightarrow \infty} \int_{-T}^T f(t)e^{-i\omega t} dt \quad (4.31)$$

Preuve : Il suffit de remarquer que pour toute fonction $f \in L^2(\mathbb{R})$, on peut écrire

$$\lim_{T \rightarrow \infty} \|f - f\chi_{[-T, T]}\| = 0 ,$$

c'est à dire que $f \in L^2(\mathbb{R})$ est arbitrairement bien approximée par des fonctions $f\chi_{[-T, T]}$, pour T assez grand. Or ces dernières appartiennent à $L^1(\mathbb{R})$, donc leur transformée de Fourier est bien définie. D'après la formule de Plancherel, on a ainsi $\lim_{T \rightarrow \infty} \|\hat{f} - \widehat{f\chi_{[-T, T]}}\| = \lim_{T \rightarrow \infty} \|f - f\chi_{[-T, T]}\| = 0$, ce qui conclut la preuve de la proposition. ♠

EXEMPLE 4.8 Considérons la fonction f définie par

$$f(t) = \frac{1}{t+i} .$$

Cette fonction n'est pas intégrable, mais elle appartient à $L^2(\mathbb{R})$. Sa transformée de Fourier est facilement obtenue grâce à la méthode des résidus, et on obtient

$$\hat{f}(\omega) = i\sqrt{2\pi}e^{-\omega}\Theta(\omega),$$

où Θ est la fonction de Heaviside. \hat{f} est bien de carré intégrable, mais est discontinue.

4.2.3 Le théorème convolution-produit

Le théorème convolution-produit que nous avons déjà rencontré auparavant s'étend sans difficulté au cadre $L^2(\mathbb{R})$. En effet, soient $f, g \in L^2(\mathbb{R})$. Alors $\hat{f}, \hat{g} \in L^2(\mathbb{R})$, et les produits $fg, \hat{f}\hat{g} \in L^1(\mathbb{R})$. Les calculs que nous avons faits précédemment sont toujours corrects, et on a

PROPOSITION 4.7 (CONVOLUTION-PRODUIT DANS $L^2(\mathbb{R})$) Si $f, g \in L^2(\mathbb{R})$, alors

$$(f \star g)(\omega) = \sqrt{2\pi} \overline{\mathcal{F}(\hat{f}\hat{g})}(\omega) \quad (4.32)$$

et

$$\mathcal{F}(fg)(\omega) = \frac{1}{\sqrt{2\pi}} (\hat{f} \star \hat{g})(\omega) \quad (4.33)$$

4.2.4 L'inégalité de Heisenberg

Nous avons vu jusqu'à présent des relations liant une fonction à sa transformée de Fourier, et réciproquement. Nous allons maintenant voir une relation d'un type quelque peu différent, qui a de fortes implications pratiques, en mécanique quantique et en traitement du signal en particulier. Il s'agit d'inégalités, dues à W. Heisenberg, qui montrent qu'une fonction ne peut être aussi localisée simultanément en temps et en fréquence que ce que l'on pourrait le vouloir. Il faut pour cela introduire des mesures de localisation. Étant donnée une fonction $f \in L^2(\mathbb{R})$, on introduit sa moyenne μ_f par

$$\mu_f = \frac{1}{\|f\|^2} \int_{-\infty}^{\infty} t|f(t)|^2 dt, \quad (4.34)$$

et son écart quadratique moyen σ_f par

$$\sigma_f^2 = \frac{1}{\|f\|^2} \int_{-\infty}^{\infty} (t - \mu_f)^2 |f(t)|^2 dt, \quad (4.35)$$

pour peu que ces deux intégrales soient convergentes. On définit de même les quantités analogues dans le domaine fréquentiel : $\mu_{\hat{f}}$ et $\sigma_{\hat{f}}$. Les nombres μ sont des mesures de localisation, alors que les σ sont des mesures de "concentration". On a donc alors :

THÉORÈME 4.6 Soit $f \in C^1(\mathbb{R})$, telle que les fonctions $f(t), f'(t)$ et $tf(t)$ soient de carré intégrable. Alors on a

$$\sigma_f \sigma_{\hat{f}} \geq \frac{1}{2}. \quad (4.36)$$

Preuve : Supposons tout d'abord que $\mu_f = \mu_{\hat{f}} = 0$. La preuve est une conséquence de l'inégalité de Cauchy-Schwarz : écrivons

$$\int_{-\infty}^{\infty} t \frac{d}{dt} |f(t)|^2 dt = [t|f(t)|^2]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} |f(t)|^2 dt$$

Il est possible de montrer que sous les hypothèses du théorème, $\lim_{t \rightarrow \pm\infty} t|f(t)|^2 = 0$. Donc, en prenant la valeur absolue et en développant la dérivée, on a

$$\|f\|^2 \leq \left| \int_{-\infty}^{\infty} t f(t) \overline{f'(t)} dt \right| + \left| \int_{-\infty}^{\infty} t f'(t) \overline{f(t)} dt \right|$$

En appliquant l'inégalité de Cauchy-Schwarz à ces deux termes, on aboutit à

$$\|f\|^2 \leq 2 \sqrt{\int_{-\infty}^{\infty} t^2 |f(t)|^2 dt \int_{-\infty}^{\infty} |f'(t)|^2 dt}.$$

Or, nous savons que la transformée de Fourier de f' n'est autre que la fonction $\omega \rightarrow i\omega \hat{f}(\omega)$. En utilisant la formule de Plancherel, nous obtenons

$$\int_{-\infty}^{\infty} |f'(t)|^2 dt = \int_{-\infty}^{\infty} \omega^2 |\hat{f}(\omega)|^2 d\omega = \|f\|^2 \sigma_f^2.$$

On a bien le résultat désiré, dans le cas particulier $\mu_f = \mu_{\hat{f}} = 0$

Pour le cas général, considérons la fonction g définie par

$$g(t) = e^{-i\mu_f t} f(t + \mu_f).$$

Un calcul immédiat montre que $\mu_g = \mu_{\hat{g}} = 0$, de sorte que l'on peut appliquer à g le résultat que nous venons de montrer. Or, on a $\|g\| = \|f\|$,

$$\sigma_g^2 = \frac{1}{\|g\|^2} \int_{-\infty}^{\infty} t^2 |f(t + \mu_f)|^2 dt = \sigma_f^2,$$

et

$$\sigma_{\hat{g}}^2 = \frac{1}{\|\hat{g}\|^2} \int_{-\infty}^{\infty} \omega^2 |\hat{f}(t + \mu_{\hat{f}})|^2 d\omega = \sigma_{\hat{f}}^2.$$

Ceci conclut la démonstration. ♠

4.3 Transformation de Fourier des fonctions de plusieurs variables

La majorité des résultats obtenus dans le cas des fonctions d'une variable réelle se généralisent sans difficulté majeure au cas multidimensionnel. On notera génériquement

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

les vecteurs de \mathbb{R}^n (ou \mathbb{C}^n). La transformation de Fourier \mathcal{F} associe à toute fonction de n variables $\underline{x} \rightarrow f(\underline{x})$ une fonction de n variables $\underline{k} \rightarrow \hat{f}(\underline{k})$, définie formellement par

$$\hat{f}(\underline{k}) = [\mathcal{F}f](\underline{k}) = (2\pi)^{-(n/2)} \int_{\mathbb{R}^n} f(\underline{x}) e^{-i\mathbf{k} \cdot \underline{x}} d\underline{x}, \tag{4.37}$$

alors que la transformation de Fourier inverse (quand elle est définie) s'obtient via

$$[\overline{\mathcal{F}}\hat{f}](\underline{x}) = (2\pi)^{-(n/2)} \int_{\mathbb{R}^n} \hat{f}(\underline{k}) e^{i\mathbf{k} \cdot \underline{x}} d\underline{k}, \tag{4.38}$$

Il est clair que les mêmes problèmes de définition qu'en dimension un se posent dans ce cas également. On se contentera ici de passer en revue les résultats importants dans le cas des fonctions intégrables et de carré intégrable

4.3.1 Transformation de Fourier des fonctions intégrables

Les espaces de Lebesgue en dimension n sont définis de façon similaire à leurs homologues unidimensionnels : étant donnée une fonction de n variables f , on lui associe le nombre

$$\|f\|_p = \left(\int_{\mathbb{R}^n} |f(\underline{x})|^p d\underline{x} \right)^{1/p}, \quad (4.39)$$

quand cette intégrale est convergente. Les espaces $L^p(\mathbb{R}^n)$ sont alors définis par

$$L^p(\mathbb{R}^n) = \{f : \mathbb{R}^n \rightarrow \mathbb{C}, \|f\|_p < \infty\}. \quad (4.40)$$

Il est clair que si $f \in L^1(\mathbb{R}^n)$, la fonction \hat{f} définie par (4.37) est bien définie pour tout $k \in \mathbb{R}^n$. Plus généralement, on a l'analogie n -dimensionnel du théorème de Riemann-Lebesgue :

THÉORÈME 4.7 (RIEMANN-LEBESGUE) Soit $f \in L^1(\mathbb{R}^n)$. Alors $\hat{f}(\underline{k})$ tend vers zéro quand $|\underline{k}|$ tend vers ∞ . De plus, la fonction \hat{f} est bornée et continue.

La démonstration est une généralisation directe de la démonstration dans le cas unidimensionnel. L'inversion de la transformation de Fourier se heurte aux mêmes obstacles que dans le cas unidimensionnel : si $f \in L^1(\mathbb{R}^n)$, on sait que \hat{f} est bornée, mais ceci n'est pas suffisant pour assurer que $\overline{\mathcal{F}}\hat{f}$ existe. Si par contre on suppose en outre que $\hat{f} \in L^1(\mathbb{R}^n)$, alors $\overline{\mathcal{F}}\hat{f}$ existe bien, et on montre l'analogie du théorème 4.2 :

THÉORÈME 4.8 Soit $f \in L^1(\mathbb{R}^n)$ une fonction telle que $\hat{f} \in L^1(\mathbb{R}^n)$ aussi. En tout point $\underline{x} = \underline{x}_0$ où f est continue, on a

$$f(\underline{x}_0) = [\overline{\mathcal{F}}\hat{f}](\underline{x}_0) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} \hat{f}(\underline{k}) e^{i\underline{k} \cdot \underline{x}_0} d\underline{k}. \quad (4.41)$$

4.3.2 Propriétés simples

Translations, modulations

Soit $f \in L^1(\mathbb{R}^n)$. On définit la *translatée* de f par la quantité $\underline{b} \in \mathbb{R}^n$ comme la fonction $g \in L^1(\mathbb{R}^n)$ définie par $g(\underline{x}) = f(\underline{x} - \underline{b})$. On vérifie immédiatement que

$$\hat{g}(\underline{k}) = e^{i\underline{k} \cdot \underline{b}} \hat{f}(\underline{k}) \quad (4.42)$$

\hat{g} est une version *modulée* de \hat{f} . Similairement, si $\underline{x} \rightarrow h(\underline{x}) = e^{i\underline{k}_0 \cdot \underline{x}} f(\underline{x})$ est une version modulée de $f \in L^1(\mathbb{R}^n)$, alors on a

$$\hat{h}(\underline{k}) = \hat{f}(\underline{k} - \underline{k}_0), \quad (4.43)$$

de sorte que la transformée de Fourier de h est une version translatée de \hat{f} .

Dilatations

Soit $f \in L^1(\mathbb{R}^n)$, et soit $a \neq 0$ est une constante réelle. On considère une fonction f_a , dilatée de f du facteur a , définie par

$$f_a(\underline{x}) = f\left(\frac{\underline{x}}{a}\right).$$

Alors, on a, par un changement de variable $\underline{y} = \underline{x}/a$,

$$\hat{f}_a(\underline{k}) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} f\left(\frac{\underline{x}}{a}\right) e^{-i\underline{k}\cdot\underline{x}} d\underline{x} = \frac{a^n}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} f(\underline{y}) e^{-i\underline{k}\cdot\underline{y}} d\underline{y} = a^n \hat{f}(a\underline{k}).$$

De nouveau, la transformée de Fourier de la copie dilatée d'une fonction n'est autre qu'une copie dilatée (d'un rapport inverse) de la transformée de Fourier de la fonction originale.

Transformations linéaires

Plus généralement, Considérons une matrice $n \times n$ A , inversible. Etant donnée une fonction $f \in L^1(\mathbb{R}^n)$, on lui associe la fonction $g \in L^1(\mathbb{R}^n)$ définie par

$$g(\underline{x}) = f(A\underline{x}).$$

On voit alors facilement qu'en notant tA la transposée de A , et $|A|$ son déterminant, et par un changement de variables $\underline{y} = A\underline{x}$, on a

$$\begin{aligned} \hat{g}(\underline{k}) &= (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(A\underline{x}) e^{-i\underline{k}\cdot\underline{x}} d\underline{x} \\ &= |A|^{-1} (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(\underline{y}) e^{-i\underline{k}\cdot A^{-1}\underline{y}} d\underline{y} \\ &= |A|^{-1} \hat{f}\left({}^t(A^{-1})\underline{k}\right). \end{aligned}$$

Ce résultat est une généralisation du précédent (dilatation). En effet, si nous prenons pour A une matrice diagonale égale à a fois la matrice identité, alors $f(A\underline{x}) = f(a\underline{x})$. De plus, $|A| = a^n$, et ${}^t(A^{-1})$ n'est autre que a^{-1} fois la matrice identité. Ceci nous redonne exactement ce que nous avons obtenu plus haut.

EXEMPLE 4.9 Prenons un autre exemple : $n = 2$, et $a = r_\theta$, la matrice de rotation d'angle θ :

$$r_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

On vérifie immédiatement que $r_\theta^{-1} = r_{-\theta}$ (l'inverse d'une rotation d'angle θ est une rotation d'angle $-\theta$), et que donc

$${}^t((r_\theta)^{-1}) = r_\theta.$$

Comme $|r_\theta| = 1$, on en déduit que si g est définie par $g(\underline{x}) = f(r_\theta\underline{x})$, alors $\hat{g}(\underline{k}) = \hat{f}(r_\theta\underline{k})$.

4.3.3 Convolution-produit

Le théorème convolution-produit que nous avons vu dans le cas unidimensionnel s'étend sans difficulté au cas n -dimensionnel. La seule différence est une modification de la constante multiplicative.

PROPOSITION 4.8 (CONVOLUTION-PRODUIT) Si $f, \hat{g} \in L^1(\mathbb{R}^n)$, et si on définit \hat{f} et g par $\hat{f} = \mathcal{F}f$ et $g = \overline{\mathcal{F}\hat{g}}$, alors

$$(f \star g) = (2\pi)^{n/2} \overline{\mathcal{F}\left(\hat{f}\hat{g}\right)} \tag{4.44}$$

et

$$\mathcal{F}(fg) = (2\pi)^{-n/2} \left(\hat{f} \star \hat{g}\right). \tag{4.45}$$

4.3.4 Différentiation

De nouveau, la situation est similaire au cas unidimensionnel, et on peut d'ailleurs utiliser les résultats obtenus en dimension 1 pour les étendre au cas multidimensionnel. Par exemple, soit $f \in L^1(\mathbb{R}^n)$, et supposons que la dérivée f par rapport à la première variable x_1 soit intégrable. Alors, en posant

$$g(\underline{x}) = \frac{\partial f}{\partial x_1}(\underline{x})$$

on peut calculer

$$\hat{g}(\underline{k}) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} \frac{\partial f}{\partial x_1}(\underline{x}) e^{-ik \cdot \underline{x}} d\underline{x} = ik_1 \hat{f}(\underline{k}).$$

Plus généralement, étant donnée une fonction f de n variables, et un n -uplet de nombres entiers positifs $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$, on notera

$$f^{(\alpha_1 \alpha_2 \dots \alpha_n)} = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \frac{\partial^{\alpha_2}}{\partial x_2^{\alpha_2}} \dots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} f. \quad (4.46)$$

On montre :

PROPOSITION 4.9 1. Soit $f \in L^1(\mathbb{R}^n)$, telle que ses dérivées partielles $f^{(\beta_1 \beta_2 \dots \beta_n)}$, pour $\beta_1 \leq \alpha_1, \dots, \beta_n \leq \alpha_n$ sont intégrables. Alors pour tout n -uplet $\{\beta_1, \beta_2, \dots, \beta_n\}$, avec $\beta_1 \leq \alpha_1, \dots, \beta_n \leq \alpha_n$, on a

$$[\mathcal{F} f^{(\beta_1 \beta_2 \dots \beta_n)}](\underline{k}) = (ik_1)^{\beta_1} (ik_2)^{\beta_2} \dots (ik_n)^{\beta_n} \hat{f}(\underline{k}). \quad (4.47)$$

2. Soit $f \in L^1(\mathbb{R}^n)$, telle que pour tout n -uplet $\{\beta_1, \beta_2, \dots, \beta_n\}$, avec $\beta_k \leq \alpha_k, k = 1, \dots, n$, la fonction $g_{\beta_1 \dots \beta_n}$ définie par

$$g_{\beta_1 \dots \beta_n}(\underline{x}) = (-ix_1)^{\beta_1} (-ix_2)^{\beta_2} \dots (-ix_n)^{\beta_n} f(\underline{x}) \quad (4.48)$$

soit elle aussi intégrable. Alors sa transformée de Fourier \hat{f} admet en tout point des dérivées partielles $\frac{\partial^{\beta_1}}{\partial k_1^{\beta_1}} \dots \frac{\partial^{\beta_n}}{\partial k_n^{\beta_n}} \hat{f}$ continues, qui ne sont autres que les transformées de Fourier des fonctionse $g_{\beta_1 \dots \beta_n}$:

$$\frac{\partial^{\beta_1}}{\partial k_1^{\beta_1}} \dots \frac{\partial^{\beta_n}}{\partial k_n^{\beta_n}} \hat{f} = \mathcal{F} g_{\beta_1 \dots \beta_n}. \quad (4.49)$$

4.3.5 Passage à $L^2(\mathbb{R}^n)$

La construction de la transformation de Fourier multidimensionnelle se fait comme dans le cas unidimensionnel : on construit tout d'abord la transformation de Fourier sur un sous-espace dense de $L^2(\mathbb{R}^n)$ (par exemple, la version n -dimensionnelle $\mathcal{S}(\mathbb{R}^n)$ de l'espace de Schwartz, puis on utilise la densité de ce dernier pour étendre le résultat à $L^2(\mathbb{R}^n)$. Le résultat est similaire, et est résumé dans le théorème suivant :

THÉORÈME 4.9 Soit $f \in L^2(\mathbb{R}^n)$. Alors la fonction $\hat{f} : \underline{k} \in \mathbb{R}^n \rightarrow \hat{f}(\underline{k})$, définie par

$$\hat{f}(\underline{k}) = (2\pi)^{-n/2} \lim_{\rho \rightarrow \infty} \int_{|\underline{x}| \leq \rho} f(\underline{x}) e^{-i\underline{k} \cdot \underline{x}} d\underline{x} \quad (4.50)$$

est de carré intégrable également. Plus précisément, l'application $f \in L^2(\mathbb{R}^n) \rightarrow \hat{f} \in L^2(\mathbb{R}^n)$ est bijective, et on a la formule de Plancherel :

$$\int_{\mathbb{R}^n} |\hat{f}(\underline{k})|^2 d\underline{k} = \int_{\mathbb{R}^n} |f(\underline{x})|^2 d\underline{x}. \quad (4.51)$$

Le produit de convolution en dimension n est une généralisation directe du produit de convolution unidimensionnel :

$$(f \star g)(\underline{x}) = \int_{\mathbb{R}^n} f(\underline{y}) g(\underline{x} - \underline{y}) d\underline{y}, \quad (4.52)$$

quand cette intégrale est convergente. Le théorème convolution-produit prend cette fois la forme

PROPOSITION 4.10 (CONVOLUTION-PRODUIT DANS $L^2(\mathbb{R}^n)$) Si $f, g \in L^2(\mathbb{R}^n)$, alors

$$(f \star g) = (2\pi)^{n/2} \overline{\mathcal{F}(\hat{f}\hat{g})}, \quad (4.53)$$

et

$$\mathcal{F}(fg) = (2\pi)^{-n/2} (\hat{f} \star \hat{g}). \quad (4.54)$$

REMARQUE 4.4 Notons au passage que si $f, g \in L^2(\mathbb{R}^n)$, alors il en va de même pour leur transformée de Fourier \hat{f} et \hat{g} . On sait de plus, d'après Cauchy-Schwarz, que $fg \in L^1(\mathbb{R}^n)$ et $\hat{f}\hat{g} \in L^1(\mathbb{R}^n)$. Donc, d'après le théorème de Riemann-Lebesgue, $\overline{\mathcal{F}(\hat{f}\hat{g})}$ et $\mathcal{F}(fg)$ sont des fonctions continues. Comme le produit de convolution de deux fonctions de carré intégrable définit également une fonction continue, les deux égalités de la Proposition 4.10 sont des égalités entre fonctions continues, valables point par point.

FORMULAIRE

On rappelle les définitions de la transformation de Fourier \mathcal{F} et de la transformation de Fourier conjuguée $\overline{\mathcal{F}}$.

$$\hat{f}(\omega) = [\mathcal{F}f](\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt .$$

$$\check{\varphi}(t) = [\overline{\mathcal{F}}\varphi](t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \varphi(s) e^{ist} ds .$$

Quelques propriétés utiles sont listées dans la TABLE 4.1 ci-dessous. Les conditions d'application de ces résultats sont également indiquées. Les propriétés sont pour la plupart énoncées dans le cas de fonctions intégrables, mais restent valides dans d'autres cadres.

$h(t)$	$\hat{h}(\omega)$	hypothèse, commentaire
$f(t)e^{i\lambda t}$	$\hat{f}(\omega - \lambda)$	$\lambda \in \mathbb{R}$
$f(t - b)$	$e^{-i\omega b} \hat{f}(\omega)$	$b \in \mathbb{R}$
$f\left(\frac{t}{a}\right)$	$a \hat{f}(a\omega)$	$a \in \mathbb{R}^*$
$f'(t)$	$i\omega \hat{f}(\omega)$	si $f' \in L^1(\mathbb{R})$
$f^{(m)}(t)$	$(i\omega)^m \hat{f}(\omega)$	si $f^{(m)} \in L^1(\mathbb{R})$
$(-it)^k f(t)$	$\hat{f}^{(k)}(\omega)$	si $t \rightarrow t^k f(t)$ est intégrable
$(f * g)(t)$	$\sqrt{2\pi} \hat{f}(\omega) \hat{g}(\omega)$	Voir Propositions 4.4 et 4.7
$f(t)g(t)$	$\frac{1}{\sqrt{2\pi}} (\hat{f} * \hat{g})(\omega)$	Voir Propositions 4.4 et 4.7

TAB. 4.1: Propriétés utiles de la transformation de Fourier unidimensionnelle

En dimension n , les applications \mathcal{F} et $\overline{\mathcal{F}}$ sont définis par

$$\hat{f}(\underline{k}) = [\mathcal{F}f](\underline{k}) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(\underline{x}) e^{-i\underline{k} \cdot \underline{x}} d\underline{x} ,$$

et

$$\check{\varphi}(\underline{x}) = [\overline{\mathcal{F}}\varphi](\underline{x}) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} \varphi(\underline{k}) e^{i\underline{k} \cdot \underline{x}} d\underline{k} ,$$

L'analogie de la table précédente en dimension quelconque est donnée dans la TABLE 4.2.

$h(\underline{x})$	$\hat{h}(\underline{k})$	hypothèse, commentaire
$f(\underline{x})e^{i\underline{k}_0 \cdot \underline{x}}$	$\hat{f}(\underline{k} - \underline{k}_0)$	$\underline{k}_0 \in \mathbb{R}^n$
$f(\underline{x} - \underline{x}_0)$	$e^{-i\underline{k} \cdot \underline{x}_0} \hat{f}(\underline{k})$	$\underline{x}_0 \in \mathbb{R}^n$
$f\left(\frac{\underline{x}}{a}\right)$	$a^n \hat{f}(a\underline{k})$	$a \in \mathbb{R}^*$
$f(A\underline{x})$	$ A ^{-1} \hat{f}({}^t(A^{-1})\underline{k})$	A matrice $n \times n$, non singulière
$\frac{\partial f}{\partial x_\ell}(\underline{x})$	$ik_\ell \hat{f}(\underline{k})$	si $\frac{\partial f}{\partial x_\ell} \in L^1(\mathbb{R})$
$\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \dots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} f$	$(ik_1)^{\alpha_1} \dots (ik_n)^{\alpha_n} \hat{f}(\underline{k})$	si $\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \dots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} f \in L^1(\mathbb{R}^n)$
$(-ix_1)^{\alpha_1} \dots (-ix_n)^{\alpha_n} f(\underline{x})$	$\frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \dots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} \hat{f}(\underline{k})$	si $\underline{x} \rightarrow (x_1)^{\alpha_1} \dots (x_n)^{\alpha_n} f(\underline{x})$ est L^1
$(f * g)(\underline{x})$	$(2\pi)^{n/2} \hat{f}(\underline{k}) \hat{g}(\underline{k})$	Voir Propositions 4.8 et 4.10
$f(\underline{x})g(\underline{x})$	$(2\pi)^{-n/2} (\hat{f} * \hat{g})(\omega)$	Voir Propositions 4.8 et 4.10

TAB. 4.2: Propriétés utiles de la transformation de Fourier multidimensionnelle

5 La transformation de Laplace

La transformation de Laplace (TL) apparaît comme une extension de la transformation de Fourier des fonctions. Cette dernière est parfois difficile à manipuler dans le cas de certaines fonctions ne possédant pas de transformée de Fourier bornée ; on a maintenant tendance à résoudre ce problème en utilisant la théorie des distributions. Cependant, la transformation de Laplace fournit une alternative relativement simple.

La transformation de Laplace permet de s'abstraire simplement de ce problème dans de nombreuses situations. Elle est couramment utilisée dans certains domaines comme le traitement du signal, ou les équations différentielles. Dans les deux cas, on utilise sa propriété essentielle, qui est de transformer les dérivations en multiplications par une variable complexe.

La transformation de Laplace existe en deux versions : la TL unilatérale, adaptée à des fonctions définies sur le demi-axe positif, et la TL bilatérale, définie sur \mathbb{R} . On se limitera ici à la TL unilatérale.

5.1 Définition et premières propriétés

Il est utile d'introduire à ce point la notion de fonction localement intégrable.

DÉFINITION 5.1 On dit qu'une fonction f est localement intégrable sur \mathbb{R} (resp. \mathbb{R}^+) si elle est intégrable sur tout intervalle borné I de \mathbb{R} (resp. \mathbb{R}^+) : on écrit alors $f \in L^1_{loc}(\mathbb{R})$ (resp. $f \in L^1_{loc}(\mathbb{R}^+)$).

DÉFINITION 5.2 (TRANSFORMATION DE LAPLACE) La transformée de Laplace d'une fonction $f \in L^1_{loc}(\mathbb{R}^+)$ est la fonction $p \in \mathbb{C} \rightarrow F(p) = (\mathcal{L}f)(p)$ d'une variable complexe p , définie par

$$F(p) = (\mathcal{L}f)(p) = \int_0^{\infty} f(t)e^{-pt} dt, \quad (5.1)$$

quand cette intégrale converge.

La variable $p = \gamma + i\omega$ est une variable complexe. Le domaine des valeurs de p pour lesquelles cette intégrale est convergente pose déjà quelques problèmes. Deux remarques importantes sont à faire.

1. L'intégrabilité de $|e^{-pt}f(t)|$ ne dépend pas de p lui-même, mais seulement de sa partie réelle $\gamma = \Re(p)$.

2. Si pour un certain γ_0 , la fonction $t \rightarrow e^{-\gamma_0 t} f(t)$ est absolument intégrable, alors $t \rightarrow e^{-\gamma t} f(t)$ est elle aussi absolument intégrable pour tout $\gamma \geq \gamma_0$.

Ceci nous conduit au résultat suivant :

THÉORÈME 5.1 Soit $f \in L^1_{loc}(\mathbb{R}^+)$. Il existe un $s \in \overline{\mathbb{R}}$ tel que pour tout $p \in \mathbb{C}$, $\Re(p) > s$, $F(p)$ existe (c'est à dire que l'intégrale dans l'équation (5.1) est convergente). Si $\Re(p) = s$, on ne peut pas conclure dans le cas général.

DÉFINITION 5.3 Le nombre s défini ci-dessus est appelé abscisse d'intégrabilité (ou abscisse de sommabilité) de F .

Il est utile de considérer quelques exemples.

1. $f(t) = \Theta(t)$ (la fonction de Heaviside) ; alors $F(p)$ est bien définie pour $\Re(p) > 0$: l'intégrale est convergente (par contre, l'intégrale n'est pas absolument convergente pour $\Re(p) < 0$). Donc $s = 0$; et si $\Re(p) > 0$, on a

$$F(p) = \int_0^{\infty} e^{-pt} dt = \frac{1}{p}.$$

Notons que la fonction $p \in \mathbb{C} \rightarrow 1/p$ est bien définie pour tout $p \in \mathbb{C}^*$; cependant, elle n'est la transformée de Laplace de la fonction Θ que dans le demi-plan $\mathbb{R}^+ \times \mathbb{R}$, domaine de définition de cette dernière.

2. $f(t) = t^n \Theta(t)$: de nouveau, $s = 0$. Si $\Re(p) > 0$, on a

$$F(p) = \int_0^{\infty} t^n e^{-pt} dt = \frac{n!}{p^{n+1}}.$$

3. $f(t) = e^{\lambda t}$, où $\lambda \in \mathbb{R}$. L'intégrale définissant $F(p)$ est convergente dès que $\Re(p) > \lambda$. Donc $s = \lambda$, et si $\Re(p) > s$, on voit immédiatement que

$$F(p) = \frac{1}{p - \lambda}.$$

4. $f(t) = e^{-t^2}$; alors l'intégrale définissant $F(p)$ est convergente quel que soit $p \in \mathbb{C}$, et par conséquent $s = -\infty$.
5. $f(t) = e^{t^2}$; alors l'intégrale définissant $F(p)$ n'est jamais convergente, et donc $s = \infty$.

Plus généralement, on utilise souvent des critères simples qui assurent l'existence de la transformée de Laplace.

DÉFINITION 5.4 Soit $f : \mathbb{R}^+ \rightarrow \mathbb{C}$.

1. f est à croissance lente si il existe deux constantes $A \in \mathbb{R}^+$ et $m \in \mathbb{N}$ telles que pour tout $t \in \mathbb{R}^+$, on ait

$$|f(t)| \leq A|t|^m. \quad (5.2)$$

2. f est à croissance au plus exponentielle si il existe deux constantes réelles $A > 0$ et a telles que pour tout t on ait

$$|f(t)| \leq Ae^{-at}. \quad (5.3)$$

PROPOSITION 5.1 Soit $f \in L^1_{loc}(\mathbb{R}^+)$.

1. Si f est à croissance lente et vérifie (5.2), alors l'abscisse d'intégrabilité s de sa transformée de Laplace vérifie $s \leq 0$.
2. Si f est à croissance au plus exponentielle et vérifie (5.3), alors l'abscisse d'intégrabilité s de sa transformée de Laplace est tel que $s \leq a$.

Preuve : La preuve est immédiate dans les deux cas. Dans le premier cas il suffit d'insérer l'estimation (5.2) dans la définition de la transformée de Laplace, et de voir que

$$|F(p)| \leq A \int_0^\infty t^m |e^{pt}| dt,$$

qui converge dès que $\Re(p) > 0$, car l'exponentielle l'emporte alors sur t^m . La preuve dans le second cas est identique. ♠

Les hypothèses faites dans cette dernière proposition sont suffisamment générales pour couvrir un grand nombre de cas d'intérêt.

5.2 Propriétés élémentaires

La transformation de Laplace possède un certain nombre de propriétés simples, conséquences directes de la définition. Nous donnons ci-dessous les plus simples.

1. *Linéarité* : Si $f, g \in L^1_{loc}(\mathbb{R}^+)$ ont pour transformées de Laplace respectives $F(p)$ et $G(p)$, et abscisses d'intégrabilité s_f et s_g , soit h la fonction définie par $h(t) = f(t) + g(t)$. Alors $s_h = \max(s_f, s_g)$ si $s_f \neq s_g$, et $s_h \leq s_f$ si $s_f = s_g$. De plus, si $p \rightarrow H(p)$ est la TL de h , alors $H(p) = F(p) + G(p)$.
2. *Translation* : Soit $f \in L^1_{loc}(\mathbb{R}^+)$, d'abscisse d'intégrabilité s_f , et soit g définie par $g(t) = f(t - b)$, où $b \in \mathbb{R}^+$. Alors $g \in L^1_{loc}(\mathbb{R}^+)$ a pour abscisse d'intégrabilité $s_g = s_f$, et sa transformée de Laplace est donnée par

$$G(p) = e^{-bp}F(p), \quad \Re(p) > s_g.$$

3. *Modulation* : Soit $f \in L^1_{loc}(\mathbb{R}^+)$, d'abscisse d'intégrabilité s_f , et soit g définie par $g(t) = e^{at}f(t)$, où $a \in \mathbb{C}$. Alors $g \in L^1_{loc}(\mathbb{R}^+)$ a pour abscisse d'intégrabilité $s_g = s_f + \Re(a)$, et sa transformée de Laplace est donnée par

$$G(p) = F(p - a), \quad \Re(p) > s_g.$$

4. *Lien avec la transformation de Fourier* : Soit $f \in L^1_{loc}(\mathbb{R}^+)$, d'abscisse d'intégrabilité s et soit F sa transformée de Laplace. Pour tout $\gamma \in \mathbb{R}$, $\gamma > s$, soit $f_\gamma(t) = f(t)e^{-\gamma t}$. Alors, on a pour $p = \gamma + i\omega$, $\gamma > s$:

$$F(\gamma + i\omega) = \int_0^\infty f_\gamma(t)e^{-i\omega t} dt = \sqrt{2\pi} \widehat{f}_\gamma(\omega).$$

Cette dernière propriété va jouer un rôle important dans le problème d'inversion de la transformation de Laplace.

5.3 La convolution finie

Etant données deux fonctions f, g définies et localement intégrables sur \mathbb{R}^+ , on définit naturellement leur produit de convolution par

$$(f * g)(t) = \int_0^t f(\tau)g(t - \tau)d\tau. \quad (5.4)$$

Il est facile de vérifier que si $f, g \in L^1_{loc}(\mathbb{R}^+)$, alors $f * g \in L^1_{loc}(\mathbb{R}^+)$ aussi, de sorte que l'on peut s'intéresser à sa transformée de Laplace.

PROPOSITION 5.2 Soient $f, g \in L^1_{loc}(\mathbb{R}^+)$, d'abscisses d'intégrabilité respectifs s_f et s_g . Soit $h = f * g$. Alors la transformée de Laplace de h est donnée par

$$H(p) = F(p)G(p), \quad \Re(p) > s_h, \quad (5.5)$$

et est bien définie pour $\Re(p) > \max(s_f, s_g)$.

Preuve : Il suffit de considérer

$$\int_0^\infty \int_0^t f(s)g(t-s) ds e^{-pt} dt = \int_0^\infty \int_0^t f(s)e^{-ps}g(t-s)e^{-p(t-s)} ds dt.$$

Un changement de variables $u = t - s$, suivi du lemme de Fubini, permet d'obtenir (5.5), à condition que $F(p)$ et $G(p)$ soient bien définies, ce qui est assuré dès que $\Re(p) > \max(s_f, s_g)$. Ceci conclut la preuve. ♠

Ce résultat trouvera son intérêt quand on considèrera la transformation de Laplace inverse. Cette propriété est aussi utile pour évaluer certaines intégrales, comme le montre l'exemple suivant.

EXEMPLE 5.1 On peut montrer que la transformée de Laplace de la fonction de Bessel $t \rightarrow J_0(t)$ est la fonction $p \rightarrow F(p) = 1/\sqrt{p^2 + 1}$, définie pour $\Re(p) > 0$. Donc, $F(p)^2 = 1/(1 + p^2)$, qui n'est autre que la transformée de Laplace de la fonction $t \rightarrow \sin t$. Par conséquent (cela découle en fait de l'inversibilité de la transformation de Laplace que nous allons voir plus loin), on en déduit

$$\int_0^t J_0(s)J_0(t-s) ds = \sin t.$$

5.4 Propriétés de régularité

5.4.1 La transformée de Laplace est holomorphe

L'un des intérêts de la transformation de Laplace est que dans le domaine où elle est définie, elle est holomorphe. Ceci permet alors d'utiliser des techniques d'intégration dans le plan complexe, en particulier pour le calcul de la transformation de Laplace inverse que nous allons voir plus bas.

Plus précisément, on a le résultat suivant :

THÉORÈME 5.2 Soit $f \in L^1_{loc}(\mathbb{R}^+)$, et soit s l'abscisse d'intégrabilité de sa transformée de Laplace $F = \mathcal{L}f$. Alors, F est holomorphe dans le domaine

$$\mathcal{D} = \{p \in \mathbb{C}, \Re(p) > s\}. \quad (5.6)$$

Pour tout entier positif m , la dérivée m -ième de F est donnée par

$$F^{(m)}(p) = \int_0^\infty (-t)^m f(t)e^{-pt} dt. \quad (5.7)$$

Ainsi, la dérivée de la transformée de Laplace d'une fonction $t \rightarrow f(t)$ n'est autre que la transformée de Laplace de la fonction $t \rightarrow -tf(t)$. On verra plus loin l'utilité de cette remarque pour le calcul de transformées de Laplace inverses.

Preuve : Commençons par étudier l'abscisse d'intégrabilité de la fonction $t \rightarrow t^m f(t)$, noté s' . Pour $t \geq 1$, on a $|f(t)e^{-pt}| \leq t^m |f(t)e^{-pt}|$. Donc l'intégrabilité de $t \rightarrow t^m |f(t)e^{-pt}|$ dans $[1, \infty[$ implique l'intégrabilité de $t \rightarrow |f(t)e^{-pt}|$ dans $[1, \infty[$, et donc dans $[0, \infty[$. Par conséquent, on a $s' \geq s$.

Inversement, pour tout $\epsilon > 0$, on peut toujours trouver $t_0 \geq 0$ tel que pour tout $t \geq t_0$, on ait $t^m |f(t)e^{-pt}| \leq |f(t)e^{-(p-\epsilon)t}|$. L'argument précédent montre donc que pour tout $\epsilon > 0$, on a $s' \leq s + \epsilon$. Par conséquent, $s' = s$.

On peut donc dériver sous le signe somme, et on obtient bel et bien l'expression voulue. L'holomorphie est donnée par le case $m = 1$. ♠

EXEMPLE 5.2 Reprenons le cas particulier de la fonction

$$f(t) = \cos(\lambda t).$$

Cette fonction n'admet pas de transformée de Fourier dans le sens usuel (la théorie des distributions permet d'en définir une), mais possède une transformée de Laplace. Un calcul explicite donne

$$F(p) = \frac{1}{2} \int_0^\infty \left(e^{-(p-i\lambda)t} + e^{-(p+i\lambda)t} \right) dt = \frac{1}{2} \left(\frac{1}{p-i\lambda} + \frac{1}{p+i\lambda} \right) = \frac{p}{p^2 + \lambda^2},$$

à condition de se limiter à $p > 0$ (c'est à dire que l'on a $s = 0$). La fonction que l'on obtient est bien holomorphe dans le demi-plan (ouvert) $\mathbb{R}^+ + i\mathbb{R}$. Par contre, noter les deux pôles en $p = \pm i\lambda$.

REMARQUE 5.1 L'exemple précédent est aussi intéressant pour la raison suivante. La fonction $p \rightarrow p/(p^2 + \lambda^2)$ est en fait bien définie dès que $p \neq \pm i\lambda$. Cependant, cette fonction n'est transformée de Laplace de $t \rightarrow \cos \lambda t$ que dans le domaine ouvert $\mathbb{R}^+ + i\mathbb{R}$ (la TL n'étant pas définie ailleurs). Dans le reste du plan complexe (privé de $\pm i\lambda$ bien sûr), elle n'est que le prolongement analytique de la TL du cosinus.

COROLLAIRE 5.1 La transformée de Laplace d'une fonction $f \in \mathcal{L}_{loc}^1(\mathbb{R}^+)$ est analytique dans son domaine de définition : $\forall p_0 \in \mathbb{C}, \Re(p_0) > s_f$, le développement en série entière de F autour de p_0

$$F(p) = \sum_{n=0}^{\infty} a_n (p - p_0)^n$$

converge pour tout p tel que $\Re(p) > s_f$.

5.4.2 Transformation de Laplace et dérivation

Nous avons déjà vu que la transformation de Fourier se comporte de façon remarquable vis à vis des opérateurs différentiels (ce qui était d'ailleurs le point de départ du travail de J. Fourier). La transformation de Laplace possède des propriétés similaires, à une petite différence près, qui vient du fait que l'on ne travaille que sur le demi axe réel positif¹ : les valeurs de f et ses dérivées à l'origine interviennent, comme conséquence d'intégrations par parties successives. Pour cela, il faut que ces valeurs soient définies, ce qui suppose des hypothèses de régularité sur la fonction étudiée.

¹Nous verrons que la TL bilatérale est plus simple à cet égard.

PROPOSITION 5.3 Soit $f \in L^1_{loc}(\mathbb{R}^+)$, telle que toutes ses dérivées $f^{(k)}$ d'ordre $k < m$ soient continues et localement intégrables. Supposons de plus qu'il existe deux constantes $a \in \mathbb{R}$ et $t_0 > 0$ telles que pour tout $j = 0, 1, \dots, m-1$ et $t \geq t_0$, on ait

$$|f^{(j)}(t)| \leq A_j e^{at} \quad (5.8)$$

pour une certaine constante A_j . Alors pour tout $p \in \mathbb{C}$, $\Re(p) > a$, la transformée de Laplace de la dérivée d'ordre m de f est donnée par

$$[\mathcal{L}f^{(m)}](p) = p^m F(p) - p^{m-1}f(0) - p^{m-2}f'(0) - p^{m-3}f''(0) - \dots - f^{(m-1)}(0). \quad (5.9)$$

Preuve : Commençons par le cas $m = 1$. Il suffit de calculer, pour $\Re(p) > a$,

$$\begin{aligned} [\mathcal{L}f'](p) &= \int_0^\infty f'(t)e^{-pt} dt \\ &= [f(t)e^{-pt}]_0^\infty + p \int_0^\infty f(t)e^{-pt} dt \\ &= pF(p) - f(0). \end{aligned}$$

La disparition du terme tout intégré évalué en $t = \infty$ résulte de (5.8), et du fait que $\Re(p) > a$. Pour m quelconque, il suffit d'itérer comme suit

$$[\mathcal{L}f^{(m)}](p) = p[\mathcal{L}f^{(m-1)}](p) - f^{(m-1)}(0),$$

qui donne (5.9) par récurrence. Ceci conclut la preuve. ♠

Cette dernière propriété est souvent utilisée pour la résolution d'équations différentielles ou d'équations aux dérivées partielles.

EXEMPLE 5.3 Prenons l'exemple de l'équation de Poisson

$$\Delta u = f, \quad (5.10)$$

où f est un second membre fixé, et $\Delta = d^2/dx^2$ est le Laplacien unidimensionnel. Pour obtenir une solution unique, on sait qu'il est nécessaire de faire des hypothèses supplémentaires, et d'adjoindre à cette équation deux conditions additionnelles. On suppose donc que u, u' et f satisfont la condition (5.8) (il est possible de démontrer dans un cadre plus général l'existence de solutions satisfaisant à de telles conditions), et que f et f' sont continues. Supposons aussi par exemple que

$$u(0) = u_0, \quad u'(0) = v_0,$$

où u_0 et v_0 sont deux nombres fixés. On note U la transformée de Laplace de u , et F la transformée de Laplace de f . Les hypothèses faites assurent l'existence (et l'analyticité) de F dans un domaine $\{p \in \mathbb{C}, \Re(p) > s_u\}$, avec $s_u < \infty$. on se ramène alors à

$$p^2 U(p) - pu_0 - v_0 = F(p),$$

soit encore, pour $p \neq 0$,

$$U(p) = \frac{F(p) + v_0}{p^2} + \frac{u_0}{p},$$

dans un domaine de valeurs de p bien choisi : la fonction U ainsi obtenue est analytique dans le domaine $\{p \in \mathbb{C}, \Re(p) > \max(s_f, 0)\}$. Cette équation permet d'obtenir une solution u à partir de f , par transformation de Laplace inverse (que nous verrons plus loin).

On peut toutefois utiliser ce que l'on sait déjà, c'est à dire les propriétés de linéarité de la transformation de Laplace, sa relation avec le produit de convolution ainsi que les quelques transformées de Laplace déjà vues.

Ainsi, on sait que l'original de Laplace de la fonction $p \rightarrow 1/p$ est la fonction de Heaviside Θ , et que l'original de Laplace de $p \rightarrow 1/p^2$ est la fonction $t \rightarrow t\Theta(t)$. Par ailleurs, l'original de Laplace de la fonction $p \rightarrow F(p)/p^2$ est le produit de convolution de f par la fonction $t \rightarrow t\Theta(t)$. On en déduit donc

$$u(t) = \Theta(t) \left(u_0 + v_0 t + \int_0^t s f(t-s) ds \right).$$

5.4.3 Transformation de Laplace et intégration

PROPOSITION 5.4 Soit $f \in L^1_{loc}(\mathbb{R}^+)$, telle qu'il existe des constantes $t_0 \geq 0$, $a \in \mathbb{R}$ et $A > 0$ vérifiant, pour tout $t \geq t_0$

$$|f(t)| \leq A e^{at}. \quad (5.11)$$

Soit g la fonction définie par

$$g(t) = \int_0^t f(u) du. \quad (5.12)$$

Alors, pour $\Re(p) > a$, la transformée de Laplace de g est donnée par

$$G(p) = \frac{1}{p} F(p). \quad (5.13)$$

Preuve : Il résulte des hypothèses que l'abscisse d'intégrabilité s de f est tel que $s \leq a$. Comme $f \in L^1_{loc}(\mathbb{R}^+)$, g est continue et dérivable, et on a aussi, pour tout $t \geq t_0$,

$$|g(t)| \leq A' e^{at},$$

pour une certaine constante positive A' . Il suffit alors d'appliquer le théorème précédent, qui donne

$$F(p) = pG(p) - g(0) = pG(p),$$

et ceci prouve la proposition. ♠

5.5 Inversion de la transformation de Laplace

Contrairement à la transformation de Fourier, la transformation de Laplace ne permet pas de décomposer une fonction en une superposition de "fonctions élémentaires" (sinus et cosinus). Cependant, la transformation de Laplace inverse a une expression simple, et facilement calculable dans bon nombre de cas.

Une première approche possible consiste à utiliser certaines expressions connues pour les transformées de Laplace de certaines fonctions élémentaires, comme on va le voir dans l'exemple suivant.

EXEMPLE 5.4 Considérons la fonction f dont la transformée de Laplace est donnée par

$$F(p) = \frac{b}{(p-a)^2 + b^2}, \quad \Re(p) > a,$$

où a et b sont des nombres réels, $b \neq 0$. Une décomposition en éléments simples de F donne

$$F(p) = \frac{1}{2} \left(\frac{1}{p - (a + ib)} - \frac{1}{p - (a - ib)} \right),$$

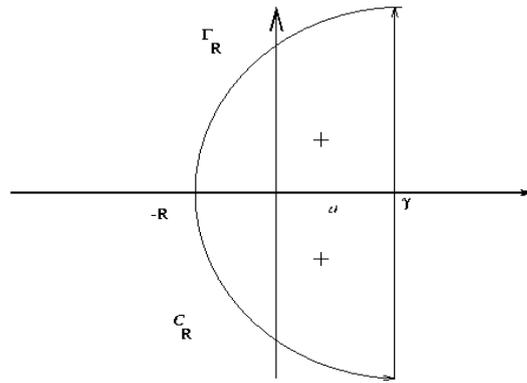


FIG. 5.1: Contour d'intégration. Le contour complet est noté C_R , et le demi-cercle de rayon R est Γ_R . Les deux pôles sont indiqués par des croix (+).

et ces deux termes sont transformées de Laplace de fonctions connues. On en déduit donc directement

$$f(t) = \frac{1}{2} \left(e^{(a+ib)t} - e^{(a-ib)t} \right) = e^{at} \sin(bt) .$$

THÉORÈME 5.3 Soit $f \in L^1_{loc}(\mathbb{R}^+)$, d'abscisse d'intégrabilité s , et soit $F(p)$ sa transformée de Laplace, définie pour $\Re(p) > s$. Pour tout $\gamma \in \mathbb{R}, \gamma > s$ on définit la fonction g_γ par $g_\gamma(t) = f(t)e^{-\gamma t}$. Alors, si $\widehat{g_\gamma} \in L^1(\mathbb{R})$, on a

$$f(t) = \frac{1}{2i\pi} \int_{\gamma-i\infty}^{\gamma+i\infty} F(p)e^{pt} dp . \quad (5.14)$$

Preuve : Le résultat est une conséquence de la relation entre transformation de Laplace et transformation de Fourier. Soit $\gamma \in \mathbb{R}, \gamma > s$, et soit $p = \gamma + i\omega$. Si $\widehat{g_\gamma} \in L^1(\mathbb{R})$, alors on a

$$g_\gamma(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{g_\gamma}(\omega) e^{i\omega t} d\omega = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\gamma + i\omega) d\omega .$$

En utilisant l'holomorphie de F dans le domaine $p \in \mathbb{C}, \Re(p) > s$ et la relation entre f et g_γ , on a bien

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\gamma + i\omega) e^{\gamma + i\omega t} d\omega = \frac{1}{2i\pi} \int_{\gamma-i\infty}^{\gamma+i\infty} F(p) e^{pt} dp ,$$

ce qui est le résultat désiré. ♠

Par conséquent, on fait généralement appel à la formule des résidus pour évaluer une transformée de Laplace inverse.

REMARQUE 5.2 Il est important de signaler que le choix de γ est crucial. γ doit obligatoirement être choisi dans le demi-plan où la transformée de Laplace est holomorphe. Un autre choix conduit inévitablement à un résultat erroné.

EXEMPLE 5.4 (SUITE) Reprenons l'exemple de la fonction f dont la transformée de Laplace est donnée par

$$F(p) = \frac{b}{(p-a)^2 + b^2} , \quad \Re(p) > a ,$$

où a et b sont des nombres réels, $b \neq 0$. F a 2 pôles simples en $a + ib$ et $a - ib$. On choisit un contour d'intégration comme donné sur la FIGURE 5.1, où $C_R = [\gamma - iR, \gamma + iR] \cup \Gamma_R$. Il est clair que pour

$\Re(p) < a$, on a $\lim_{|p| \rightarrow \infty} (p - a)F(p)e^{pt} = 0$ (car $t > 0$), de sorte que l'intégrale sur Γ_R tend vers 0 quand $R \rightarrow \infty$ (on peut aussi utiliser le critère de Carlsaw et Jaeger ci-dessous dans ce cas). On peut donc écrire

$$\begin{aligned} f(t) &= \lim_{R \rightarrow \infty} \frac{1}{2i\pi} \int_{C_R} F(p)e^{pt} dt \\ &= (\text{Res}(a + ib) + \text{Res}(a - ib)) \\ &= \frac{1}{2i} \left(e^{(a+ib)t} - e^{(a-ib)t} \right) \\ &= e^{at} \sin(bt). \end{aligned}$$

Dans ce type de problème, l'intégrale sur l'arc de cercle Γ_R tend vers 0 lorsque $R \rightarrow \infty$. On peut soit vérifier ceci directement, soit invoquer des conditions suffisantes sur la transformée de Laplace F assurant que cette intégrale tend réellement vers 0. On peut souvent utiliser pour cela le *Lemme de Jordan*. Une condition adaptée à la transformation de Laplace a aussi été donnée par Carlsaw et Jaeger :

LEMME 5.1 (CARSLAW ET JAEGER) Si il existe des constantes positives $C > 0$, $R_0 > 0$ et $k > 0$ telles que pour p appartenant au demi plan $\Re(p) < a$ et $|p| > R_0$, on ait

$$|F(p)| < C|p|^{-k}$$

Alors l'intégrale sur le contour Γ_R tend vers 0 quand $R \rightarrow \infty$.

EXEMPLE 5.5 Considérons un problème de conduction de chaleur dans un solide unidimensionnel semi-infini ($x > 0$), dont la température initiale est nulle, et dont l'extrémité $x = 0$ est maintenue à une température constante $T_0 > 0$. Ce système est schématisé en Fig. 5.2 ci dessous.

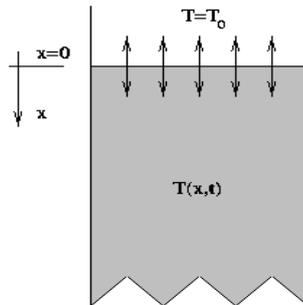


FIG. 5.2: Le milieu semi-infini

L'évolution spatio-temporelle de ce champ $u(x, t)$ est régie par l'équation de la chaleur

$$\frac{\partial u}{\partial t}(x, t) = \alpha \frac{\partial^2 u}{\partial x^2}(x, t), \quad (5.15)$$

que l'on complète par les conditions initiales et aux limites

$$u(0, t) = T_0 \quad \forall t \geq 0 \quad (5.16)$$

$$u(x, 0) = 0 \quad \forall x > 0. \quad (5.17)$$

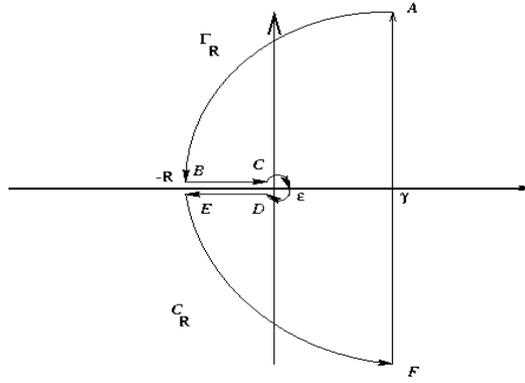


FIG. 5.3: Contour d'intégration pour le calcul de la solution de l'équation de la chaleur à partir de sa transformée de Laplace. Le contour complet est noté C_R , et le demi-cercle de rayon R est Γ_R .

On suppose que u est une fonction continue de la variable t , et que les fonctions $u, \partial u / \partial t$ sont bornées (donc satisfont la condition (5.8)), ce qui nous assure de pouvoir utiliser les résultats de la Proposition 5.3. Par transformation de Laplace (par rapport à la variable t), en notant

$$U(x, p) = \int_0^{\infty} u(x, t) e^{-pt} dt, \quad (5.18)$$

on se ramène à l'équation

$$pU(x, p) = \alpha \frac{\partial^2 U}{\partial x^2}(x, p), \quad (5.19)$$

dont la solution générale est de la forme

$$U(x, p) = C_1(p) e^{-x\sqrt{p/\alpha}} + C_2(p) e^{x\sqrt{p/\alpha}}.$$

Si l'on impose à la solution d'être bornée quand $x \rightarrow \infty$, on a nécessairement $C_2(p) = 0 \forall p$. Donc

$$C_1(p) = U(p, 0) = T_0 \int_0^{\infty} e^{-pt} dt = \frac{T_0}{p}.$$

Donc,

$$U(x, p) = \frac{T_0}{p} e^{-x\sqrt{p/\alpha}}, \quad (5.20)$$

et le problème sera résolu dès que l'on aura l'original de Laplace de cette fonction, qui est holomorphe dans le demi plan ouvert $\Re(p) > 0$. Soit donc $\gamma \in \mathbb{R}^+$, et considérons l'intégrale

$$I_a(t) = \frac{1}{2i\pi} \int_{\gamma-i\infty}^{\gamma+i\infty} e^{-a\sqrt{p}} e^{pt} \frac{dp}{p}. \quad (5.21)$$

La présence du terme \sqrt{p} nous oblige à faire un choix de détermination pour la racine carrée. On introduit donc une coupure, par exemple sur le demi axe réel négatif \mathbb{R}^- .

Considérons le contour Γ_R (voir FIG. 5.3). L'intégrand étant holomorphe à l'intérieur de Γ_R , l'intégrale sur Γ_R est nulle, de sorte que l'on peut écrire

$$I_a(t) = \lim_{R \rightarrow \infty; \epsilon \rightarrow 0} \frac{1}{2i\pi} \left(\int_B^A + \int_C^B + \int_D^C + \int_E^D + \int_F^E \right),$$

et il faut maintenant évaluer chacun des termes.

- A la limite $\epsilon \rightarrow 0$, l'intégrale sur le cercle (CD) s'évalue grâce à la méthode des résidus (pôle en 0), et on obtient

$$\int_D^C e^{-a\sqrt{p}} e^{pt} \frac{dp}{p} = 2i\pi \text{Res}_0(e^{-a\sqrt{p}} e^{pt}) = 2i\pi .$$

- Pour l'intégrale sur le segment (CB), on pose $p = ue^{i\pi}$, avec $u > 0$. On a alors

$$\lim_{R \rightarrow \infty; \epsilon \rightarrow 0} \int_C^B e^{-a\sqrt{p}} e^{pt} \frac{dp}{p} = \int_0^\infty e^{-ia\sqrt{u}} e^{-ut} \frac{du}{u} .$$

- Pour l'intégrale sur le segment (DE), on doit poser $p = ue^{-i\pi}$, avec $u > 0$. On a alors

$$\lim_{R \rightarrow \infty; \epsilon \rightarrow 0} \int_E^D e^{-a\sqrt{p}} e^{pt} \frac{dp}{p} = - \int_0^\infty e^{ia\sqrt{u}} e^{-ut} \frac{du}{u} .$$

- Reste à se charger de la contribution des arcs de cercle, à la limite $R \rightarrow \infty$. On applique ici le critère de Carslaw et Jaeger. Sur le quart de cercle supérieur, on pose $p = ue^{i\theta}$, avec $\theta \in [\pi/2, \pi]$. On a donc $\sqrt{p} = \sqrt{u}(\cos(\theta/2) + i \sin(\theta/2))$, et $\Re(\sqrt{p}) \geq 0$. Sur le quart de cercle inférieur, on pose $p = ue^{i\theta}$, avec $\theta \in [-\pi/2, -\pi]$. On a donc $\theta/2 \in [-\pi/4, -\pi/2]$, et de nouveau $\Re(\sqrt{p}) \geq 0$. Par conséquent, sur l'arc de cercle complet, $|\exp\{-a\sqrt{p}\}/p| \leq 1/|p|$, et le critère de Carslaw et Jaeger s'applique. Donc les intégrales sur les deux quarts de cercle (AB) et (CD) tendent vers 0 quand $R \rightarrow \infty$.

Mettant ces résultats ensembles, nous avons donc à calculer

$$I_a(t) = 1 - \frac{1}{\pi} \int_0^\infty e^{-tu} \sin(a\sqrt{u}) \frac{du}{u} = 1 - \frac{2}{\pi} \int_0^\infty e^{-tv^2} \sin(av) \frac{dv}{v} = 1 - \frac{2}{\pi} J_a(t) .$$

Pour finir, on remarque que

$$\frac{d}{da} J_a(t) = \int_0^\infty e^{-tv^2} \cos(av) dv = \frac{1}{4} \int_{-\infty}^\infty e^{-tv^2} (e^{iav} + e^{-iav}) dv = \frac{1}{2} \sqrt{\frac{\pi}{t}} e^{-a^2/4t} ,$$

de sorte que

$$J_a(t) = J_0 + \int_0^a J_u(t) du = \frac{1}{2} \sqrt{\frac{\pi}{t}} \int_0^a e^{-u^2/4t} du = \sqrt{\pi} \int_0^{a/2\sqrt{t}} e^{-u^2} du = \frac{\pi}{2} \text{erf} \left(\frac{a}{2\sqrt{t}} \right) .$$

On en déduit l'intégrale recherchée

$$I_a = \text{erfc} \left(\frac{a}{2\sqrt{t}} \right) .$$

Donc, la solution est finalement

$$T(x, t) = T_0 \text{erfc} \left(\frac{x}{2\sqrt{\alpha t}} \right) . \quad (5.22)$$

Quelques graphes de la solution, pour diverses valeurs de t , se trouvent en FIG. 5.4.

5.6 La transformation de Laplace bilatérale

La transformée de Laplace bilatérale s'adresse aux fonctions définies sur la droite réelle, et non plus sur la demi droite. Le cadre mathématique approprié est le cadre des fonctions localement intégrables sur l'axe réel $L_{loc}^1(\mathbb{R})$.

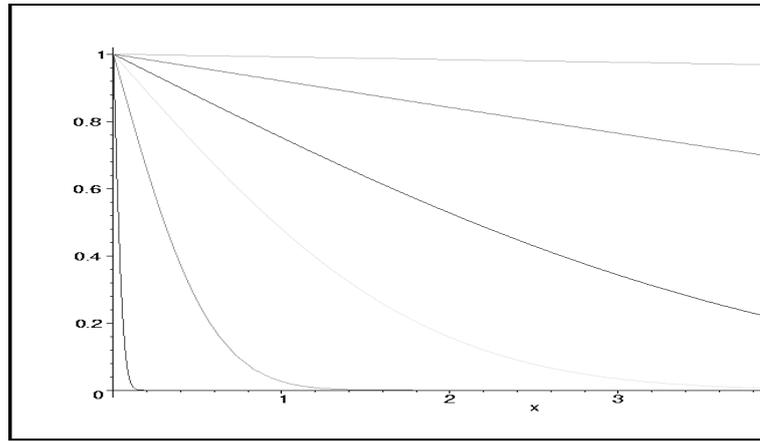


FIG. 5.4: Profils de température en fonction de la variable d'espace x , solutions de l'équation de la chaleur ($\alpha = 1$, $T_0 = 1$). Solutions pour des temps $t = 0,01$, $t = 0,1$, $t = 1$, $t = 5$, $t = 50$ et $t = 5000$.

DÉFINITION 5.5 La transformée de Laplace bilatérale d'une fonction $f \in L^1_{loc}(\mathbb{R})$ est la fonction $p \in \mathbb{C} \rightarrow F(p) = (\mathcal{L}f)(p)$ d'une variable complexe p , définie par

$$F(p) = (\mathcal{L}f)(p) = \int_{-\infty}^{\infty} f(t)e^{-pt} dt, \quad (5.23)$$

quand cette intégrale converge.

La transformation de Laplace bilatérale possède des propriétés très semblables à celles de la transformation de Laplace unilatérale. La différence essentielle vient de leurs domaines de définition respectifs.

Supposons que pour deux nombres réels x_0 et x_1 , avec $x_0 < x_1$ la fonction $t \rightarrow e^{-pt}f(t)$ soit intégrable pour $\Re(p) = x_0$ et $\Re(p) = x_1$. Alors, on voit immédiatement que cette fonction est également intégrable pour tout p tel que $x_0 \leq \Re(p) \leq x_1$. En notant s_0 et s_1 le minimum et le maximum respectivement des valeurs de x_0 et x_1 telles que la propriété précédente soit vraie, on aboutit au résultat suivant, qui généralise le cas unilatéral.

THÉORÈME 5.4 1. Soit $f \in L^1_{loc}(\mathbb{R})$. Il existe deux nombres réels s_1 et s_2 tels que l'intégrale définissant $F(p)$ soit convergente pour tout p tel que $\Re(p) \in]s_0, s_1[$. Pour $\Re(p) \notin]s_0, s_1[$, l'intégrale est divergente, et pour $\Re(p) = s_0$ ou $\Re(p) = s_1$, on ne peut pas conclure dans le cas général.

2. La fonction $p \rightarrow F(p)$ est holomorphe dans le domaine $p \in \mathbb{C}, \Re(p) \in]s_0, s_1[$.

Les propriétés essentielles de la transformation de Laplace bilatérale sont essentiellement des paraphrases des propriétés que nous avons vues dans le cas unilatéral. On ne s'étendra pas sur ces pro-

priétés. On insistera plutôt sur deux propriétés importantes, c'est à dire la forme de la transformée de Laplace d'une dérivée, et la formule d'inversion.

PROPOSITION 5.5 Soit $f \in C^{m-1}(\mathbb{R})$, telle que sa dérivée m -ième $f^{(m)}$ appartienne à $L^1_{loc}(\mathbb{R})$. Supposons en outre qu'il existe quatre constantes $-\infty < t_b \leq t_a < \infty$ et $a < b$, et $2m$ constantes positives A_j, B_j telles que pour tout $j = 0, \dots, m-1$

$$|f^{(j)}(t)| \leq B_j e^{bt} \quad \text{pour } t \leq t_b \quad (5.24)$$

$$|f^{(j)}(t)| \leq A_j e^{at} \quad \text{pour } t \geq t_a \quad (5.25)$$

Alors, pour tout p tel que $\Re(p) \in]a, b[$ on a

$$(\mathcal{L}f^{(m)})(p) = p^m F(p). \quad (5.26)$$

La formule d'inversion quant à elle est comme dans le cas unilatéral une conséquence de la formule d'inversion de la transformée de Fourier.

THÉORÈME 5.5 Soit $f \in C^{m-1}(\mathbb{R})$, et soit $D_f = \{p \in \mathbb{C}, s_0 < \Re(p) < s_1\}$ le domaine de définition et d'analyticité de sa transformée de Laplace bilatérale F . Alors, en posant $p = \gamma + i\omega$, si pour tout γ fixé dans $]s_0, s_1[$, la transformée de Fourier \hat{g}_γ de la fonction $g_\gamma : t \rightarrow f(t)e^{-\gamma t}$ est dans $L^1(\mathbb{R})$, alors pour presque tout $t \in \mathbb{R}$ on a

$$f(t) = \frac{1}{2i\pi} \int_{\gamma-i\infty}^{\gamma+i\infty} F(p)e^{pt} dp. \quad (5.27)$$

Les transformées de Laplace bilatérales inverses se calculent généralement en utilisant la méthode des résidus.

Formulaire

Quelques propriétés utiles de la transformation de Laplace unilatérale sont listées dans la Table 5.1. Dans cette table, la fonction f est une fonction localement intégrable sur \mathbb{R}^+ , et l'abscisse d'intégrabilité de sa transformée de Laplace est noté s_f . Dans certains cas, certaines hypothèses supplémentaires sont nécessaires pour donner un sens aux propriétés listées.

$g(t)$	$G(p)$	s_g	Hypothèse
$f(t)e^{at}$	$F(p-a)$	$s_f + \Re(a)$	$f \in L^1_{loc}(\mathbb{R})$
$f(t-b)$	$e^{bp}F(p)$	s_f	$f \in L^1_{loc}(\mathbb{R})$
$f'(t)$	$pF(p) - f'(0)$	s_f	$f, f' \in L^1_{loc}(\mathbb{R})$
$f^{(m)}(t)$	$p^m F(p) - \sum_0^{m-1} p^{m-k-1} f^{(k)}(0)$	s_f	$f, f', \dots, f^{(m)} \in L^1_{loc}(\mathbb{R})$
$\int_0^t f(\tau) d\tau$	$\frac{1}{p} F(p)$	s_f	$f \in L^1_{loc}(\mathbb{R})$
$\int_0^t f_1(\tau) f_2(t-\tau) d\tau$	$F_1(p) F_2(p)$	$\max(s_{f_1}, s_{f_2})$	$f_1, f_2 \in L^1_{loc}(\mathbb{R})$

TAB. 5.1: Propriétés utiles de la transformation de Laplace unilatérale

Troisième partie

Equations différentielles et aux dérivées partielles

Equations différentielles

Les équations différentielles et les équations aux dérivées partielles constituent une composante fondamentale de la physique, dans la mesure où elles sont utiles pour décrire de nombreux systèmes physiques. On se limitera ici à certaines classes d'équations différentielles qui admettent des solutions relativement simples à obtenir.

Dans le cas de fonctions d'une variable réelle, une équation différentielle d'ordre n d'indéterminée f est une relation de la forme

$$F(x, f(x), f'(x), \dots, f^{(n)}(x)) = 0, \quad (6.1)$$

où on a noté $f^{(n)}$ la dérivée n -ième de f :

$$f^{(n)}(x) = \frac{d^n f}{dx^n}(x). \quad (6.2)$$

Résoudre une équation différentielle revient à trouver **toutes** ses solutions.

Une solution d'une équation différentielle définit dans un repère (généralement choisi orthonormé) une courbe, appelée **courbe intégrale** de l'équation. Résoudre l'équation revient donc à déterminer la courbe intégrale correspondante.

Lorsque l'équation différentielle se met sous la forme

$$f^{(n)}(x) = G(x, f(x), \dots, f^{(n-1)}(x)),$$

on dit qu'elle est sous **forme résolue**. Lorsque G ne dépend pas explicitement de x , on parle d'**équation différentielle autonome**.

Une équation différentielle est généralement complétée de conditions supplémentaires, qui peuvent garantir l'unicité de la solution. Par exemple, quand une équation différentielle sous forme résolue est complétée par des conditions initiales (aussi appelées **conditions de Cauchy**)

$$f(x_0) = f_0, \quad f'(x_0) = f_1, \quad \dots \quad f^{(n-1)}(x_0) = f_{n-1},$$

pour un certain x_0 donné (souvent $x_0 = 0$). Sous des conditions assez peu restrictives sur G , les conditions de Cauchy garantissent l'existence de solutions locales (c'est à dire au voisinage de x_0) de l'équation. Plus précisément, dans le cas $n = 1$

THÉORÈME 6.1 (CAUCHY-LIPSCHITZ) Supposons que pour tout x fixé, G soit telle que

$$|G(x, y_2) - G(x, y_1)| \leq K|y_2 - y_1|,$$

dans un certain voisinage de x_0 , pour une certaine constante K (on dit que G est localement Lipschitzienne en x_0). Alors il existe une et une seule solution maximale (c'est à dire qui n'est la restriction d'aucune autre) satisfaisant une condition de Cauchy donnée.

Ce théorème se généralise à des équations d'ordre supérieur (toujours sous forme résolue, avec conditions de Cauchy).

On recourt parfois à des conditions aux bords (par exemple, $f(x_0) = y_0$ et $f(x_1) = y_1$ pour une équation du second ordre). De tels problèmes peuvent très bien n'avoir aucune solution ou au contraire une infinité de solutions.

6.1 Définitions, généralités

6.1.1 Généralités, équations linéaires

La grande majorité des équations différentielles n'admettent pas de solution explicite. Par exemple, l'équation différentielle

$$\cos(x) \left(\frac{d^3 f(x)}{dx^3} \right)^{2.5} + x^7 \sqrt{f(x)} = \pi e^{-x}$$

est très probablement difficile à résoudre explicitement. L'une des raisons est que cette équation différentielle est **non-linéaire** ; ainsi, étant donnée une solution f , $2f$ n'est pas solution.

DÉFINITION 6.1 1. Une équation différentielle d'ordre n est dite **linéaire** si elle peut se mettre sous la forme

$$a_0(x)f(x) + a_1(x)f'(x) + a_2(x)f''(x) + \dots + a_n(x)f^{(n)}(x) = b(x), \quad (6.3)$$

a_0, a_1, \dots, a_n et b étant des fonctions fixées.

2. Cette équation est **homogène** si $b(x) = 0$ pour tout x .

3. Elle est à **coefficients constants** si a_0, a_1, \dots, a_n et b sont indépendants de x .

Il est facile de démontrer que des combinaisons linéaires de solutions d'une équation différentielle linéaire homogène sont toujours solutions de cette équation. En fait, on montre le résultat fondamental suivant

THÉORÈME 6.2 L'ensemble des solutions d'une équation différentielle linéaire homogène d'ordre n est un espace vectoriel de dimension n .

Ainsi, une équation différentielle de ce type est loin d'admettre une unique solution. L'unicité peut être retrouvée en imposant des conditions supplémentaires. Dans le cas d'une équation d'ordre n , on a besoin de n équations linéaires supplémentaires pour spécifier une solution unique. Ceci est généralement effectué en imposant des conditions aux bords, ou conditions aux limites .

Les équations homogènes jouent en fait un rôle fondamental. En effet, considérons une équation différentielle linéaire inhomogène, comme en (6.3), et soient f_1 et f_2 deux solutions de cette équation. On voit alors facilement que $f_1 - f_2$ est solution de l'équation différentielle **homogène**

$$a_0(x)f(x) + a_1(x)f'(x) + a_2(x)f''(x) + \cdots + a_n(x)f^{(n)}(x) = 0. \quad (6.4)$$

Inversement, étant donnée une solution de (6.3), on peut lui ajouter n'importe quelle solution de l'équation homogène correspondante, le résultat étant toujours solution de (6.3). On a donc montré

PROPOSITION 6.1 *La solution générale d'une équation différentielle linéaire inhomogène peut toujours s'écrire comme la somme de la solution générale de l'équation homogène correspondante, et d'une solution particulière de l'équation inhomogène.*

On dit que l'ensemble des solutions d'une équation différentielle linéaire inhomogène est un **espace affine**.

EXEMPLE 6.1 Considérons l'exemple simple

$$(x^2 + 1)y'(x) + 3xy(x) = x.$$

L'équation homogène correspondante s'écrit sous la forme

$$\frac{y'(x)}{y(x)} = \frac{-3x}{x^2 + 1},$$

et s'intègre facilement, pour donner

$$\ln |y(x)| = -\frac{3}{2} \ln(x^2 + 1) + \ln |\lambda|,$$

d'où la solution générale de l'équation homogène

$$y_0(x) = \lambda(x^2 + 1)^{-3/2}.$$

Par ailleurs, une solution apparente (à défaut d'être évidente) de l'équation inhomogène est donnée par

$$y_p(x) = \frac{1}{3}.$$

On en déduit la solution générale de l'équation inhomogène

$$y(x) = \frac{1}{3} + \lambda(x^2 + 1)^{-3/2}.$$

On obtient en fait une famille à un paramètre de solutions.

6.1.2 Equations homogènes à coefficients constants

Les équations différentielles linéaires à coefficients constants se traitent souvent de façon simple. Prenons le cas d'une équation du second ordre homogène

$$af''(x) + bf'(x) + cf(x) = 0. \quad (6.5)$$

En recherchant des solutions sous la forme $f(x) = e^{\alpha x}$, cette équation se transforme en une équation algébrique

$$a\alpha^2 + b\alpha + c = 0,$$

qui se résout par la méthode des radicaux. Dans ce cas, les racines sont de la forme

$$\alpha_{\pm} = \frac{1}{2a} \left(-b \pm \Delta^{1/2} \right),$$

où $\Delta = b^2 - 4ac$ est le discriminant, et où $\Delta^{1/2}$ doit être compris comme $i\sqrt{-\Delta}$ si $\Delta < 0$. Supposons que $\Delta \neq 0$. Alors on a deux solutions linéairement indépendantes de cette équation :

$$f_{\pm}(x) = e^{\alpha_{\pm}x},$$

qui engendrent bien l'espace vectoriel de dimension 2 des solutions de l'équation. La solution générale est alors de la forme

$$f(x) = A_+ e^{\alpha_+ x} + A_- e^{\alpha_- x},$$

les constantes A_{\pm} étant déterminées par les conditions aux limites.

Le cas limite $\Delta = 0$ est un peu plus complexe. En effet, il y a dans ce cas une seule racine

$$\alpha_0 = -\frac{b}{2a}$$

à l'équation caractéristique, qui ne fournit qu'une seule solution de l'équation homogène

$$f_0(x) = A_0 e^{\alpha_0 x}.$$

Or on sait d'après le Théorème 6.2 qu'il existe deux solutions linéairement indépendantes de l'équation de départ (puisqu'elle est du second ordre). On verra un peu plus loin une méthode (basée sur le **Wronskien**, permettant d'obtenir une seconde solution. Dans notre cas, cette approche se simplifie, et revient à rechercher une seconde solution, sous la forme particulière

$$f_1(x) = u(x)f_0(x).$$

Comme $f_0(x)$ est déjà solution de l'équation homogène (6.5), en insérant la forme particulière de f_1 dans (6.5), on aboutit à une nouvelle équation que doit satisfaire u :

$$a(u''(x)f_0(x) + 2u'(x)f_0'(x)) + bu'(x)f_0(x) = 0,$$

soit en insérant la forme particulière de f_0 et en simplifiant par $e^{\alpha_0 x}$, on aboutit à la forme particulièrement simple

$$au''(x) = 0,$$

d'où

$$u(x) = \lambda x + \mu.$$

Ainsi, on obtient la solution générale de l'équation (6.5), dans le cas particulier $b^2 = 4ac$,

$$f(x) = (\lambda x + \mu)e^{\alpha_0 x},$$

6.1.3 Equations homogènes : variation de la constante

La méthode de **variation de la constante**, ou **méthode de Laplace** permet de trouver une solution particulière d'équations inhomogènes, dans les cas où il n'en existe pas d'évidente. L'idée est de se baser sur la solution générale de l'équation homogène, et de « faire varier les constantes » dans le sens suivant.

Considérons une équation linéaire d'ordre n comme en (6.3), et supposons que l'on ait déjà obtenu la solution générale de l'équation homogène (6.4), celle-ci engendrant un espace de dimension n . La solution générale s'écrit alors comme combinaison linéaire de n « solutions élémentaires »

$$y_0(x) = \alpha_1 y_{0,1}(x) + \alpha_2 y_{0,2}(x) + \cdots + \alpha_n y_{0,n}(x).$$

La méthode de Laplace revient à rechercher une solution particulière de (6.3) comme combinaison des solutions $y_{0,k}$, avec coefficients **variables** (d'où l'expression « variation de la constante ») :

$$y_p(x) = \alpha_1(x)y_{0,1}(x) + \alpha_2(x)y_{0,2}(x) + \cdots + \alpha_n(x)y_{0,n}(x).$$

En insérant cette forme dans l'équation inhomogène (6.3), on aboutit à une nouvelle équations différentielle, impliquant uniquement les fonctions $\alpha_1, \dots, \alpha_n$. Cette unique équation ne suffit (sauf dans le cas $n = 1$) pas à caractériser ces fonctions, et il faut la compléter par d'autres.

On ne traitera pas ici le cas général, et on se limitera aux cas $n = 1$ ou 2.

Dans le cas $n = 1$, partant d'une équation

$$f'(x) + a(x)f(x) = g(x), \quad (6.6)$$

partons d'une solution f_0 de l'équation homogène, et soit $f_1(x) = \alpha(x)f_0(x)$. Alors, $f_1'(x) = a(x)f_0'(x) + \alpha'(x)f_0(x)$. En insérant cette solution dans l'équation inhomogène, et en utilisant l'équation homogène pour simplifier l'expression ainsi obtenue, on se ramène à

$$\alpha'(x)f_0(x) = g(x),$$

d'où la solution

$$\alpha(x) = \alpha(x_0) + \int_{x_0}^x \frac{g(y)}{f_0(y)} dy, \quad (6.7)$$

d'où on déduit la solution de l'équation inhomogène. Reprenons l'exemple précédent.

EXEMPLE 6.1 (SUITE). Dans ce cas, on recherche une solution sous la forme

$$y(x) = \alpha(x)(1+x^2)^{-3/2}.$$

On a alors, compte tenu du fait que y_0 est solution de l'équation homogène,

$$(1+x^2)y'(x) + 3xy(x) = \alpha'(x)(1+x^2)^{-1/2}$$

Imposer que y soit solution de l'équation inhomogène implique que

$$\alpha'(x) = x\sqrt{1+x^2},$$

d'où

$$\alpha(x) = \lambda + \int_0^x t\sqrt{1+t^2} dt = \lambda + \frac{1}{2} \int_0^{x^2} \sqrt{1+u} du = \lambda + \frac{1}{3}(1+x^2)^{3/2},$$

λ étant une constante d'intégration. Finalement, on obtient bien la solution générale

$$y(x) = \alpha(x)(1+x^2)^{-3/2} = \lambda(1+x^2)^{-3/2} + \frac{1}{3}.$$

Le cas des équations du second ordre est un peu plus complexe, car l'espace vectoriel engendré par les solutions de l'équation homogène est de dimension 2. On peut alors imposer une condition supplémentaire sur λ_1 et λ_2 . On choisit en général

$$\lambda_1'(x)f_1(x) + \lambda_2'(x)f_2(x) = 0,$$

arbitraire... mais efficace.

Partons d'une équation de la forme

$$f''(x) + a(x)f'(x) + b(x)f(x) = g(x),$$

considérons deux solutions f_1 et f_2 de l'équation homogène, et recherchons une solution de l'équation inhomogène sous la forme

$$f(x) = \lambda_1(x)f_1(x) + \lambda_2(x)f_2(x).$$

On a alors

$$f'(x) = \lambda_1(x)f_1'(x) + \lambda_1'(x)f_1(x) + \lambda_2(x)f_2'(x) + \lambda_2'(x)f_2(x) = \lambda_1(x)f_1'(x) + \lambda_2(x)f_2'(x),$$

compte tenu de la condition supplémentaire imposée, et donc

$$f''(x) = \lambda_1(x)f_1''(x) + \lambda_1'(x)f_1'(x) + \lambda_2(x)f_2''(x) + \lambda_2'(x)f_2'(x).$$

En insérant cela dans l'équation de départ, on aboutit pour tout x au système linéaire d'équations en $(\lambda_1'(x), \lambda_2'(x))$

$$\begin{cases} \lambda_1'(x)f_1'(x) + \lambda_2'(x)f_2'(x) = g(x) \\ \lambda_1'(x)f_1(x) + \lambda_2'(x)f_2(x) = 0. \end{cases} \quad (6.8)$$

Si f_1 et f_2 ne sont pas proportionnelles, le déterminant de ce système, appelé **Wronskien**,

$$W(x) = f_1(x)f_2'(x) - f_2(x)f_1'(x) = f_1(x)^2 \frac{d}{dx} \left(\frac{f_2(x)}{f_1(x)} \right)$$

est non nul. Le système (6.8) admet donc une unique solution notée $(L_1(x), L_2(x))$, d'où on déduit, en prenant des primitives

$$(\lambda_1(x), \lambda_2(x)) = \left(\int L_1(x)dx, \int L_2(x)dx \right).$$

Notons que le choix de la constante d'intégration dans ces primitives n'a pas d'importance.

EXEMPLE 6.2 On considère l'équation du second ordre

$$f''(x) + f'(x) - 2f(x) = \frac{1}{\cosh(x)}.$$

L'équation homogène est à coefficients constants, et se prête donc bien à des solutions sous forme d'exponentielles. L'équation caractéristique

$$\alpha^2 + \alpha - 2 = 0$$

admet les deux racines $\alpha_1 = 1$ et $\alpha_2 = -2$, d'où la solution générale de l'équation homogène

$$f_0(x) = \lambda_1 e^x + \lambda_2 e^{-2x}.$$

On utilise maintenant la variation de la constante. La condition supplémentaire prend ici la forme

$$\lambda_1'(x)e^x + \lambda_2'(x)e^{-2x} = 0,$$

d'où on tire

$$\lambda_2'(x) = -\lambda_1'(x)e^{3x}.$$

En reportant cela dans l'autre équation, on obtient

$$\lambda_1'(x) (f_1'(x) - e^{3x} f_2'(x)) = \frac{1}{\cosh(x)},$$

c'est à dire

$$\lambda_1'(x) = \frac{2}{3} \frac{1}{e^{2x} + 1}.$$

On a donc

$$\begin{aligned}
 \lambda_1(x) &= \lambda_1(0) + \frac{2}{3} \int_0^x \frac{dy}{e^{2y} + 1} \\
 &= \lambda_1(0) + \frac{2}{3} \int_1^{e^x} \frac{du}{u(u^2 + 1)} \\
 &= \lambda_1(0) + \frac{2}{3} \int_1^{e^x} \left(\frac{1}{u} - \frac{u}{u^2 + 1} \right) du \\
 &= \lambda_1(0) + \left[\frac{2}{3} \ln |u| - \frac{1}{3} \ln(u^2 + 1) \right]_1^{e^x} \\
 &= C_1 + \frac{1}{3} \ln(1 + e^{-2x}),
 \end{aligned}$$

où C_1 est une constante. Un calcul similaire donne

$$\lambda_2(x) = \lambda_2(0) + \frac{2}{3} \int_1^{e^x} \frac{u^2 du}{u^2 + 1} = \lambda_2(0) + \frac{2}{3} \int_1^{e^x} \left(1 - \frac{1}{u^2 + 1} \right) du$$

d'où on déduit

$$\lambda_2(x) = C_2 + \frac{2}{3} (\text{Arc tan}(e^x) - e^x),$$

pour une certaine constante C_2 . De là on déduit la solution particulière, puis la solution générale de l'équation inhomogène.

6.1.4 Facteur intégrant pour les équations linéaires du premier ordre

De façon générale, un facteur intégrant pour une équation différentielle est une fonction telle que la multiplication terme à terme de l'équation par cette fonction permet d'en simplifier l'intégration. On se limitera ici aux équations du premier ordre (on rencontrera de nouveau un facteur intégrant en discutant la théorie de Sturm-Liouville).

Considérons une équation différentielle du premier ordre, sous forme résolue

$$\frac{dy}{dx}(x) = \psi(x, y).$$

Dans les cas où ψ ne dépend pas explicitement de y , la solution est de la forme

$$y(x) = C + \int_{x_0}^x \psi(z) dz,$$

où $C = y(x_0)$ est une constante. Dans le cas général, il existe une famille de solutions, dépendant d'un paramètre C , données de façon implicite

$$\Xi(y, x, C) = 0,$$

pour une certaine fonction Ξ .

On se limite ici aux équations du premier ordre linéaires. La forme la plus générale est

$$y'(x) + a(x)y(x) = b(x), \quad (6.9)$$

a et b étant des fonctions continues dans un intervalle $[x_0, x_1]$, et $y_0 = y(x_0)$ est une condition au bord donnée.

DÉFINITION 6.2 Un facteur intégrant pour l'équation (6.9) est une fonction $\alpha : x \in [x_0, x_1] \rightarrow \alpha(x) \in \mathbb{C}$ telle que (6.9) puisse s'écrire

$$\frac{d}{dx} (\alpha(x)y(x)) = \alpha(x)b(x) . \quad (6.10)$$

On peut remarquer que le facteur intégrant est défini à une constante multiplicative près.

Supposant connu un facteur intégrant α , une solution de (6.9) est alors explicite : $\alpha(x)y(x) = \alpha(x_0)y(x_0) + \int_{x_0}^x \alpha(u)b(u) du$, d'où

$$y(x) = \frac{\alpha(x_0)y_0}{\alpha(x)} + \frac{1}{\alpha(x)} \int_{x_0}^x \alpha(u)b(u) du . \quad (6.11)$$

Pour utiliser ce résultat, il faut déterminer un facteur intégrant, qui doit donc satisfaire

$$\frac{d\alpha}{dx}(x) = \alpha(x)a(x) ,$$

de sorte que la solution est

$$\alpha(x) = \alpha(x_0) \exp \left\{ \int_{x_0}^x a(u) du \right\} \quad (6.12)$$

(on peut fixer la constante multiplicative pour que $\alpha(x_0) = 1$).

EXEMPLE 6.3 On considère l'équation différentielle du premier ordre

$$y'(x) + xy(x) = x ,$$

avec la condition $y(x_0) = y_0$. On vérifie aisément que le facteur intégrant est de la forme

$$\alpha(x) = \exp \{ (x^2 - x_0^2)/2 \}$$

(donc $\alpha(x_0) = 1$), de sorte que la solution s'écrit finalement

$$y(x) = e^{-(x^2-x_0^2)/2} \left[y_0 + \int_{x_0}^x ze^{(z^2-x_0^2)/2} dz \right] = 1 + (y_0 - 1)e^{-(x^2-x_0^2)/2} .$$

6.1.5 Equations du second ordre, Wronskien

Le Wronskien, que nous avons déjà rencontré, permet de construire, à partir d'une solution d'une équation différentielle linéaire homogène, une seconde solution linéairement indépendante. Le principe général est, comme souvent, de rechercher une solution $y_2(x)$ sous la forme $f(x)y_1(x)$, où y_1 est la solution déjà connue.

DÉFINITION 6.3 Soient y_1 et y_2 deux solutions linéairement indépendantes de l'équation différentielle linéaire homogène du second ordre

$$\frac{d^2y}{dx^2}(x) + \alpha(x)\frac{dy}{dx}(x) + \beta(x)y(x) = 0 .$$

Le Wronskien (ou déterminant de Wronski) de y_1 et y_2 est la fonction $x \rightarrow W(x)$ définie par

$$W(x) = y_1(x)\frac{dy_2}{dx}(x) - y_2(x)\frac{dy_1}{dx}(x) . \quad (6.13)$$

Soit donc W le Wronskien, et calculons

$$\begin{aligned} W'(x) &= y_1'(x)y_2'(x) + y_1(x)y_2''(x) - y_1''(x)y_2'(x) - y_1'(x)y_2(x) \\ &= y_1(x)y_2''(x) - y_1''(x)y_2(x) \\ &= y_1(x) [-\alpha(x)y_2'(x) - \beta(x)y_2(x)] - y_2(x) [-\alpha(x)y_1'(x) - \beta(x)y_1(x)] \\ &= -\alpha(x)W(x) \end{aligned}$$

Le Wronskien vérifie l'équation du premier ordre

$$\frac{dW}{dx}(x) = -\alpha(x)W(x),$$

de sorte que sa forme est

$$W(x) = W(x_0) \exp \left\{ - \int_{x_0}^x \alpha(z) dz \right\}. \quad (6.14)$$

Ce résultat est appelé **Théorème de Liouville**.

REMARQUE 6.1 Le Wronskien admet une interprétation géométrique simple. Dans l'espace des phases, c'est à dire l'espace engendré par les vecteurs $(y(t), y'(t))$, le Wronskien associé à deux solutions (y_1, y_1') et (y_2, y_2') est l'aire du parallélogramme engendré par ces deux vecteurs du plan (soit le double de l'aire du triangle). Le théorème de Liouville décrit donc l'évolution temporelle de cette aire. Notons en particulier que si $\alpha = 0$, c'est à dire si l'équation ne contient pas de terme d'amortissement, cette aire est indépendante du temps.

Le Wronskien permet aussi de construire la solution générale de l'équation différentielle, connaissant une solution particulière. Supposant connue y_1 , on a en effet

$$\frac{y_1(x)y_2'(x) - y_1'(x)y_2(x)}{y_1(x)^2} = \frac{W(x)}{y_1(x)^2},$$

d'où

$$\frac{d}{dx} \left(\frac{y_2(x)}{y_1(x)} \right) = \frac{W(x)}{y_1(x)^2}.$$

On en déduit par intégration

$$y_2(x) = \frac{y_2(x_0)}{y_1(x_0)} y_1(x) + y_1(x) \int_{x_0}^x \frac{W(u)}{y_1(u)^2} du = y_1(x) \left(C_1 + C_2 \int_{x_0}^x \frac{W(u)}{y_1(u)^2} du \right). \quad (6.15)$$

Ainsi, connaissant une première solution d'une équation homogène du second ordre, on peut en déduire une seconde, linéairement indépendante ; il suffit pour cela de calculer au préalable de Wronskien via l'équation (6.14), puis d'utiliser (6.15).

EXEMPLE 6.4 Soit l'équation différentielle

$$y'' - 2y' + y = 0, \quad x \in \mathbb{R}.$$

il est immédiat que la méthode de l'équation caractéristique donne la solution

$$y_1(x) = e^x.$$

Un calcul explicite donne un Wronskien

$$W(x) = e^{2x},$$

d'où on déduit

$$y_2(x) = C_1 e^x + C_2 e^x \int_{x_0}^x du,$$

et donc une seconde solution linéairement indépendante

$$y_2(x) = x e^x.$$

6.2 La méthode de Fröbenius

Une approche classique pour résoudre des équations différentielles, consiste à rechercher des solutions sous forme de séries. Lorsque les équations sont à coefficients réguliers (c'est à dire suffisamment différentiables), les séries à considérer sont des séries entières. Par exemple, pour l'équation

$$y' - y = 0,$$

si l'on recherche des solutions sous forme de série entière¹ $y(x) = \sum_0^\infty a_n x^n$, on obtient en remplaçant terme à terme

$$\sum_1^\infty n a_n e^{n-1} - \sum_0^\infty a_n x^n = 0,$$

soit par un changement d'indice de sommation

$$\sum_0^\infty [a_n - (n+1)a_{n+1}]x^n, \quad \forall x$$

ce qui implique que tous les termes soient nuls, c'est à dire

$$a_{n+1} = \frac{a_n}{n+1} = \frac{a_{n-1}}{(n+1)n} = \dots = \frac{a_0}{n!}.$$

On en déduit la solution

$$y(x) = a_0 \sum_0^\infty \frac{x^n}{n!} = a_0 e^x.$$

Dans le cas où les coefficients de l'équation ne sont plus aussi réguliers, cette technique se généralise comme on va le voir, et porte le nom de **méthode de Fröbenius**. Il est nécessaire d'introduire tout d'abord un peu de terminologie.

On se focalise ici sur le cas des équations linéaires du second ordre, homogènes, que l'on met sous la forme générique

$$y''(x) + p(x)y'(x) + q(x)y(x) = 0. \quad (6.16)$$

Les équations inhomogènes se traitent de la façon usuelle, en résolvant tout d'abord l'équation homogène associée, et en ajoutant à la solution générale de celle-ci une solution particulière de l'équation inhomogène.

6.2.1 Notion de point singulier, théorème de Fuchs

La méthode de Fröbenius consiste à rechercher des solutions sous forme de série, soit une série entière, soit une série de Laurent, soit encore une série faisant intervenir des exposants non-entiers. La différence entre ces deux situations tient aux propriétés de régularité des coefficients (variables) de l'équation.

- DÉFINITION 6.4** 1. Un point x_0 est un point **ordinaire** (ou **régulier**) pour cette équation si $y''(x)$ reste fini en $x = x_0$ dès que $y'(x_0)$ et $y(x_0)$ sont finis. Sinon x_0 est dit **singulier**.
2. Un point singulier à distance finie x_0 est **régulier** pour l'équation si $(x - x_0)p(x)$ et $(x - x_0)^2 q(x)$ sont finis en x_0 . Dans le cas contraire, le point singulier est dit **essentiel**.

¹Pourquoi faire simple, quand on peut faire compliqué ?

Rechercher un possible point régulier ou singulier en x_0 revient donc à étudier le comportement de p et q au voisinage de x_0 . Un point x_0 est régulier si $p(x_0)$ et $q(x_0)$ sont finis. Il est singulier inessentiel si $(x - x_0)p(x)$ et $(x - x_0)^2q(x)$ admettent une limite finie lorsque $x \rightarrow x_0$.

REMARQUE 6.2 La définition ci-dessus n'a de sens que pour l'étude des points sur l'axe réel. Pour étudier le comportement des points à l'infini, il est nécessaire de faire le changement de variable $x' = 1/x$, et d'étudier le comportement en $x' \rightarrow 0$.

Les théorèmes de Fuchs donnent des conditions simples et explicites pour l'existence de solutions d'équations de type (6.16). Considérons tout d'abord les cas où les coefficients sont réguliers. Il est alors possible de les développer en série entière au voisinage d'un point régulier x_0 , et d'en déduire une solution elle-même sous forme de série entière. Plus précisément, on a le résultat suivant.

THÉORÈME 6.3 On considère l'équation différentielle (6.16), avec conditions $y(x_0) = f_0$ et $y'(x_0) = g_0$. Alors si les fonctions p et q , considérées comme fonctions d'une variable complexe z , sont régulières dans un cercle de rayon r_0 centré sur x_0 dans \mathbb{C} , il existe une unique solution de l'équation développable en série entière autour de x_0 et obéissant aux conditions données.

Dans ces conditions, étant donné un point régulier x_0 , on peut rechercher des solutions de l'équation différentielle sous la forme

$$y(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k .$$

Dans le cas plus complexe où les conditions de régularité sur p et q ne sont pas remplies, les solutions peuvent prendre une forme un peu plus compliquée.

Supposons que x_0 soit un point singulier régulier. L'équation différentielle (6.16) conduit à

$$(x - x_0)^2 y''(x) + (x - x_0)a(x)y'(x) + b(x)y(x) = 0 ,$$

où $a(x) = (x - x_0)p(x)$ et $b(x) = (x - x_0)^2q(x)$ sont deux fonctions régulières.

Au voisinage de x_0 , on approxime cette équation par l'équation approchée

$$(x - x_0)^2 y''(x) + (x - x_0)a(x_0)y'(x) + b(x_0)y(x) = 0 ,$$

dont les solutions sont de la forme

$$A_1(x - x_0)^{\nu_1} , \quad A_2(x - x_0)^{\nu_2} ,$$

où A_1 et A_2 sont deux constantes d'intégration, et ν_1 et ν_2 sont les racines de l'équation indiciale

$$\nu(\nu - 1) + \nu a(x_0) + b(x_0) = 0 .$$

Ceci conduit à rechercher des solutions de la forme

$$y_1(x) = (x - x_0)^{\nu_1} f_1(x) , \quad y_2(x) = (x - x_0)^{\nu_2} f_2(x) ,$$

où f_1 et f_2 sont deux fonctions régulières (c'est à dire développables en série entière en x_0)... à l'exception du cas où $\nu_1 - \nu_2$ est entier, voir ci-dessous.

Dans le cas où x_0 est un point singulier essentiel, les choses sont encore plus complexes, mais on montre qu'il est encore possible de trouver des solutions sous forme de série infinie.

Plus précisément, le résultat est le suivant :

THÉORÈME 6.4 (THÉORÈME DE FUCHS) Si x_0 est un pôle ou un point singulier essentiel de p ou q , c'est à dire si p et q peuvent être exprimés sous forme de série de Laurent

$$\begin{cases} p(z) = \sum_{k=-\infty}^{\infty} a_k (z - x_0)^k \\ q(z) = \sum_{k=-\infty}^{\infty} b_k (z - x_0)^k, \end{cases} \quad (6.17)$$

alors l'équation homogène admet deux solutions linéairement indépendantes, qui au voisinage de x_0 peuvent être représentées comme

$$\begin{cases} f_1(x) = (x - x_0)^{\nu_1} \sum_{k=-\infty}^{\infty} c_k (x - x_0)^k \\ f_2(x) = (x - x_0)^{\nu_2} \sum_{k=-\infty}^{\infty} d_k (x - x_0)^k, \end{cases} \quad (6.18)$$

et dans le cas dégénéré ($\nu_1 - \nu_2$ entier)

$$\begin{cases} f_1(x) = (x - x_0)^{\nu_1} \sum_{k=-\infty}^{\infty} c_k (x - x_0)^k \\ f_2(x) = (x - x_0)^{\nu_2} \sum_{k=-\infty}^{\infty} d_k (x - x_0)^k + \alpha f_1(x) \ln(x - x_0). \end{cases} \quad (6.19)$$

REMARQUE 6.3 Dans le résultat ci-dessus, si le point singulier au voisinage duquel on effectue le développement est inessentiel, on peut toujours se ramener à une somme sur $k \in \mathbb{N}$, par exemple dans le cas non dégénéré

$$\begin{cases} f_1(x) = (x - x_0)^{\nu_1} \sum_{k=0}^{\infty} c_k (x - x_0)^k \\ f_2(x) = (x - x_0)^{\nu_2} \sum_{k=0}^{\infty} d_k (x - x_0)^k, \end{cases}$$

Ainsi, étant donnée une équation différentielle linéaire d'ordre deux, homogène, la procédure à suivre est la suivante :

- Mettre l'équation sous la forme (6.16).
- Analyser les propriétés des fonctions $p(x)$ et $q(x)$.
 - Si elles sont régulières, on peut rechercher une solution sous forme de série entière.
 - Si elles sont singulières en x_0 , mais que $\lim_{x \rightarrow x_0} (x - x_0)p(x)$ et $\lim_{x \rightarrow x_0} (x - x_0)^2 q(x)$ existent, on peut rechercher une solution sous la forme d'une série $\sum_{k=0}^{\infty} a_k (x - x_0)^{k+\nu}$.
 - Dans le cas contraire, il faut rechercher une solution sous la forme d'une série doublement infinie $\sum_{k=-\infty}^{\infty} a_k (x - x_0)^{k+\nu}$.

Nous allons voir ci-dessous un certain nombre d'exemples.

6.2.2 Un exemple : les fonctions de Bessel

Les fonctions de Bessel apparaissent naturellement dans le cadre de l'étude de l'équation de Helmholtz bidimensionnelle exprimée en coordonnées polaires, c'est à dire l'équation aux dérivées partielles

$$(\Delta + k^2)\psi(r, \theta) = 0, \quad (6.20)$$

et donnent un premier exemple d'utilisation de la méthode de séparation des variables pour les équations aux dérivées partielles. En exprimant le Laplacien en coordonnées polaires

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2},$$

et en recherchant des solutions à variables séparées

$$\psi(r, \theta) = R(r)\Theta(\theta), \quad r \in \mathbb{R}^+, \theta \in [0, 2\pi]$$

on obtient la forme particulière

$$\frac{r^2}{R(r)} \frac{d^2 R(r)}{dr^2} + \frac{r}{R(r)} \frac{dR(r)}{dr} + \frac{k^2 r^2}{R(r)} = -\frac{1}{\Theta(\theta)} \frac{d^2 \Theta(\theta)}{d\theta^2} = p^2,$$

p étant une constante, a priori complexe (dont on verra qu'elle doit être réelle).

L'équation angulaire se résout facilement, et a pour solution

$$\Theta(\theta) = Ae^{ip\theta} + Be^{-ip\theta}.$$

Maintenant, notons que nous devons nécessairement avoir

$$\Theta(2\pi) = \Theta(0), \quad \Theta'(2\pi) = \Theta'(0),$$

ce qui implique $A + B = Ae^{2ip\pi} + Be^{2ip\pi}$ et $ip(A - B) = ip(Ae^{2ip\pi} - Be^{2ip\pi})$, et donc la restriction $p = n \in \mathbb{Z}$. (notons que si la constante p avait eu une partie imaginaire non nulle, les solutions auraient été non bornés).

Passons à l'équation radiale

$$r^2 R''(r) + rR'(r) + (k^2 r^2 - n^2)R(r) = 0. \quad (6.21)$$

Elle s'écrit sous la forme simple

$$\frac{d}{dr} \left(r \frac{dR}{dr}(r) \right) + \left(k^2 r - \frac{p^2}{r} \right) R(r) = 0,$$

ou en posant $x = kr$, et $y(x) = R(x/k)$

$$x \frac{d}{dx} \left(x \frac{dy}{dx}(x) \right) + (x^2 - p^2) y(x) = 0 \quad (6.22)$$

Cette équation est appelée **équation de Bessel** d'ordre p . On se propose de la résoudre maintenant en utilisant les séries de Fröbenius.

On voit facilement que $x = 0$ est un point singulier régulier. On peut également montrer l'existence d'un point singulier essentiel à l'infini. Supposons donc une solution de la forme

$$y(x) = \sum_{k=0}^{\infty} a_k x^{k+\nu}.$$

En insérant cette forme particulière dans l'équation, on obtient

$$\sum_{k=0}^{\infty} a_k (k + \nu)^2 x^{k+\nu} + (x^2 - p^2) \sum_{k=0}^{\infty} a_k x^{k+\nu} = 0.$$

Le terme de plus bas degré (c'est à dire x^ν) nous donne

$$a_0(\nu^2 - p^2) = 0,$$

soit $a_0 = 0$ ou $\nu = \pm p$. Le terme en $x^{\nu+1}$ donne

$$a_1((\nu + 1)^2 - p^2) = 0,$$

donc soit $a_0 = 0$ ou $\nu + 1 = \pm p$.

Choisissons $\nu = p$, et donc $a_1 = 0$. Le terme générique donne quant à lui

$$a_k = \frac{-a_{k-2}}{(k+p)^2 - p^2} = \frac{-a_{k-2}}{k(k+2p)},$$

et donc tous les coefficients a_k d'indice impair sont nuls. On a donc

$$a_{2k} = -\frac{a_{2k-2}}{4k(k+p)} = \frac{(-1)^k a_0 \Gamma(p+1)}{2^{2k} \Gamma(k+1) \Gamma(k+p+1)},$$

où Γ est la **fonction Gamma** d'Euler, définie par l'intégrale

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t}, \quad t \notin \mathbb{Z}^-, \quad (6.23)$$

et qui vérifie (comme le montre une simple intégration par parties)

$$\Gamma(x+1) = x\Gamma(x), \quad \forall x \notin \mathbb{Z}^-.$$

On a donc obtenu une solution particulière, qui dépend d'une constante a_0 . Si on impose une normalisation de la forme

$$a_0 = \frac{1}{2^p \Gamma(p+1)},$$

on aboutit à l'expression suivante pour la solution, appelée fonction de Bessel d'ordre p et notée J_p

$$J_p(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{\Gamma(k+1)\Gamma(k+p+1)} \left(\frac{x}{2}\right)^{2k+p}. \quad (6.24)$$

REMARQUE 6.4 On a choisi ici le cas $p = \nu$. On peut se convaincre facilement que $p = -\nu$ ne donne pas de nouvelle solution. En effet, on a pour tout x

$$J_{-p}(x) = (-1)^p J_p(x).$$

Nous avons aussi fait un autre choix, lorsque nous avons posé $a_1 = 0$, choix qui nous a conduit à une seule solution (pour p fixé). Or l'espace des solutions est bidimensionnel. Il est facile d'obtenir une seconde solution en utilisant le Wronskien. Les fonctions correspondantes sont appelées **fonctions de Neumann**, et notées N_p .

Le Wronskien, défini par

$$W(x) = J_p(x)N_p'(x) - N_p(x)J_p'(x),$$

est donné par

$$W(x) = W(x_1) e^{-\int_{x_1}^x P(z) dz},$$

où $P(x) = 1/x$, d'où une solution possible pour le Wronskien est

$$W(x) = \frac{1}{x}.$$

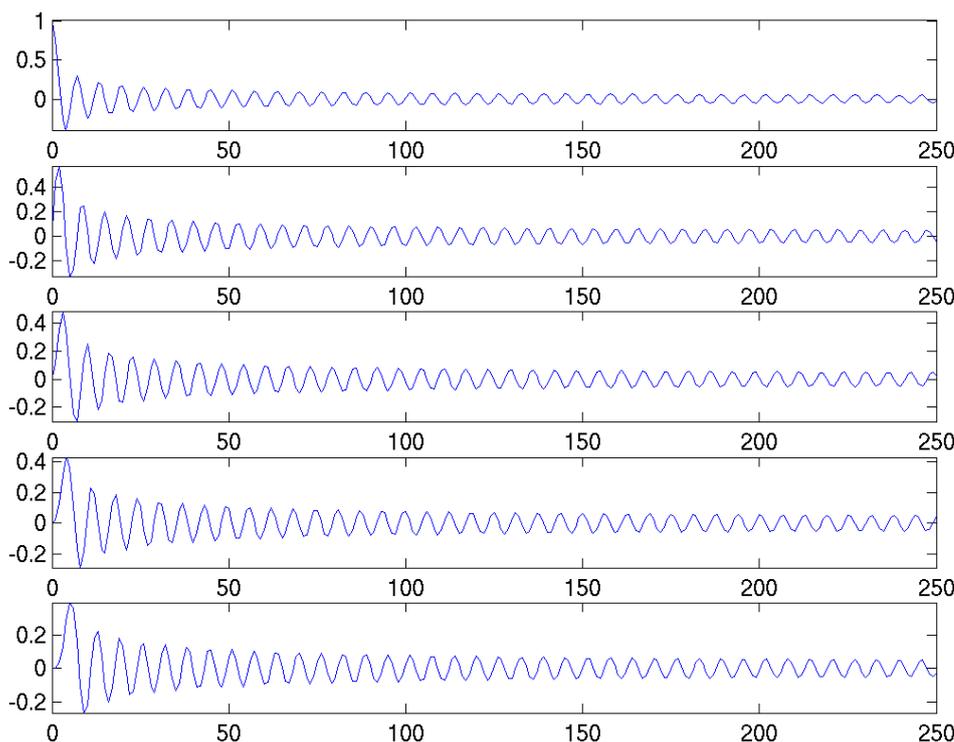
On a alors une autre solution, linéairement indépendante, de la forme

$$y_2(x) = J_p(x) \int_{x_0}^x \frac{\exp\left\{-\int_{x_1}^z P(s) ds\right\}}{J_p(z)^2} dz = J_p(x) \int_{x_0}^x \frac{dz}{zJ_p(z)^2}. \quad (6.25)$$

Pour un certain choix (conventionnel) des constantes, on obtient ainsi une seconde solution appelée **fonction de Neumann**, notée $N_p(x)$. On montre que N_p diverge logarithmiquement à l'origine.

On introduit parfois aussi les **fonctions de Hankel** de première et deuxième espèce, définies par

$$\begin{cases} H_p^{(1)}(x) = & J_p(x) + iN_p(x) \\ H_p^{(2)}(x) = & J_p(x) - iN_p(x) \end{cases} \quad (6.26)$$

FIG. 6.1: Les fonctions de Bessel de J_0 à J_4 .

Orthogonalité et complétude des fonctions de Bessel

Les fonctions de Bessel de J_0 à J_4 sont tracées en FIG. 6.1. Comme on peut le voir, ce sont des fonctions extrêmement oscillantes. Il s'avère que leurs zéros jouent un rôle prépondérant, comme on va le voir. Notons α_{pn} le $(n + 1)$ -ième zéro de la fonction de Bessel J_p :

$$J_p(\alpha_{pn}) = 0, \quad n = 0, 1, \dots$$

On a alors la propriété d'orthogonalité suivante :

PROPOSITION 6.2 Soit a un réel positif, et posons

$$k_{pn} = \alpha_{pn}/a .$$

Alors on a

$$\int_0^a J_p(k_{pn}r) J_p(k_{pm}r) r dr = \delta_{mn} \int_0^a J_p(k_{pn}r)^2 r dr \quad (6.27)$$

Preuve : Posons $R_n(r) = J_p(k_{pn}r)$. Alors on a $R_n(a) = 0$; de plus, il résulte de l'équation de Bessel que

$$\frac{d}{dr} \left(r \frac{dR_n}{dr}(r) \right) = - \left(k_{pn}^2 r - \frac{p^2}{r} \right) R_n(r) .$$

Par conséquent, on peut écrire

$$\begin{aligned} \int_0^a \left[\frac{d}{dr} \left(r \frac{dR_n(r)}{dr} \right) \right] R_m(r) dr &= - \int_0^a \left(k_{pn}^2 r - \frac{p^2}{r} \right) R_n(r) R_m(r) dr \\ \int_0^a \left[\frac{d}{dr} \left(r \frac{dR_m(r)}{dr} \right) \right] R_n(r) dr &= - \int_0^a \left(k_{pm}^2 r - \frac{p^2}{r} \right) R_m(r) R_n(r) dr . \end{aligned}$$

Notons aussi que grâce à deux intégrations par parties, on a

$$\begin{aligned} \int_0^a \left[\frac{d}{dr} \left(r \frac{dR_n(r)}{dr} \right) \right] R_m(r) dr &= \left[R_m(r) r \frac{dR_n(r)}{dr} \right]_0^a - \int_0^a r R'_m(r) R'_n(r) dr \\ &= - \left[R_n(r) r \frac{dR_m(r)}{dr} \right]_0^a + \int_0^a \left[\frac{d}{dr} \left(r \frac{dR_m(r)}{dr} \right) \right] R_n(r) dr \\ &= \int_0^a \left[\frac{d}{dr} \left(r \frac{dR_m(r)}{dr} \right) \right] R_n(r) dr , \end{aligned}$$

de sorte que l'on peut écrire, par soustraction

$$(k_{pn}^2 - k_{pm}^2) \int_0^a R_m(r) R_n(r) r dr = 0 ,$$

d'où on déduit le résultat. ♠

REMARQUE 6.5 Comme on peut le voir, le calcul ci-dessus a été possible grâce à une forme particulière donnée à l'équation de Bessel, appelée forme de Sturm-Liouville

$$\frac{d}{dr} r \frac{dR}{dr}(r) - \frac{p^2}{r} R(r) = -k^2 r R(r) ,$$

que l'on peut voir comme une équation aux valeurs propres généralisée (par la présence du facteur r dans le membre de droite) pour l'opérateur

$$L = \frac{d}{dr} r \frac{d}{dr} - \frac{p^2}{r} .$$

Pour ce qui est de la normalisation, elle peut être obtenue explicitement (calculs non reproduits ici). On pose

$$\mathcal{J}_{pn}(r) = \frac{J_p(k_{pn}r)}{\int_0^a J_p(k_{pn}r)^2 dr} ,$$

et on obtient ainsi une famille orthonormale dans l'espace de Hilbert

$$L^2([0, a], r dr) = \left\{ f : [0, a] \rightarrow \mathbb{C} \int_0^a |f(r)|^2 r dr < \infty \right\} .$$

Par ailleurs, il est possible de montrer que cette famille est complète dans $L^2([0, a], r dr)$:

$$\int_0^a f(r) \mathcal{J}_{pn}(r) r dr = 0 \quad \forall n \implies f = 0 .$$

Ainsi, on a le résultat suivant :

THÉORÈME 6.5 La famille des fonctions de Bessel $\{\mathcal{J}_{pn}, n = 0, 1, \dots\}$ est une base orthonormée de $L^2([0, a], r dr)$.

Ainsi, pour toute fonction $f \in L^2([0, a], r dr)$, on peut écrire, quel que soit p ,

$$f = \sum_{n=0}^{\infty} \langle f, \mathcal{J}_{pn} \rangle \mathcal{J}_{pn} . \tag{6.28}$$

Autres propriétés des fonctions de Bessel

1. **Fonction génératrice** : En posant

$$g(x, t) = e^{x(t-1/t)/2},$$

il est possible de montrer que les fonctions de Bessel d'ordre entier s'obtiennent via

$$g(x, t) = \sum_{n=-\infty}^{\infty} J_n(x) t^n. \quad (6.29)$$

En effet, en développant les exponentielles en série entière, on obtient

$$g(x, t) = \sum_{r=0}^{\infty} \left(\frac{x}{2}\right)^r \frac{t^r}{r!} \sum_{s=0}^{\infty} \left(\frac{x}{2}\right)^s \frac{(-t)^{-s}}{s!},$$

de sorte qu'en posant $n = r - s$, on obtient

$$g(x, t) = \sum_{n=-\infty}^{\infty} \left(\sum_{s=0}^{\infty} \frac{(-1)^s}{s!(n+s)!} \left(\frac{x}{2}\right)^{n+2s} \right) t^n,$$

d'où le résultat (on a utilisé ici le fait que $1/n! = 0$ pour tout n entier négatif).

2. **Relations de récurrence** : en différenciant la fonction génératrice par rapport à t , on obtient la relation

$$J_{n-1}(x) + J_{n+1}(x) = \frac{2n}{x} J_n(x). \quad (6.30)$$

De même, en dérivant par rapport à x et en identifiant les termes, on aboutit à

$$J_{n-1}(x) - J_{n+1}(x) = 2J'_n(x). \quad (6.31)$$

On en déduit en particulier

$$J_{n-1}(x) = \frac{n}{x} J_n(x) + J'_n(x) \quad (6.32)$$

$$J_{n+1}(x) = \frac{n}{x} J_n(x) - J'_n(x). \quad (6.33)$$

6.2.3 Un autre exemple : les polynômes de Legendre

On considère le Laplacien en coordonnées sphériques (r, θ, φ) , et l'équation de Laplace

$$\frac{1}{r^2} \frac{\partial}{\partial r} (r^2 F) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial F}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 F}{\partial \varphi^2} = 0. \quad (6.34)$$

En recherchant des solutions sous forme de fonctions à variables séparées

$$F(r, \theta, \varphi) = R(r)\Theta(\theta)\Phi(\varphi),$$

et en supposant que la fonction Φ soit constante, il est possible de montrer (voir le chapitre suivant) qu'en introduisant la variable $x \in [-1, 1]$ et la fonction y telles que

$$x = \cos \theta, \quad y(x) = \Theta(\theta),$$

l'équation en θ se ramène à une équation différentielle de la forme

$$(1 - x^2)y''(x) - 2xy'(x) + Cy(x) = 0, \quad x \in [-1, 1]. \quad (6.35)$$

Cette équation est appelée **équation de Legendre**, et se résout par la méthode de Fröbenius. On peut montrer que l'équation possède deux points singuliers réguliers en $x = \pm 1$. Ses solutions dépendent de la constante C . On peut montrer

THÉORÈME 6.6 Supposons qu'il existe $\ell \in \mathbb{N}$ tel que

$$C = \ell(\ell + 1).$$

Alors,

1. il existe une solution de l'équation de Legendre qui s'écrit sous forme d'un polynôme de degré ℓ , noté

$$P_\ell(x) = \sum_{k=0}^{\ell} a_k x^k.$$

2. Les polynômes de Legendre forment une base orthogonale dans l'espace de Hilbert $L^2([-1, 1])$:

$$\int_{-1}^1 P_k(x) P_\ell(x) dx = \frac{1}{\ell + 1/2} \delta_{k\ell},$$

et pour tout $f \in L^2([-1, 1])$,

$$f(u) = (\ell + 1/2) \int_{-1}^1 f(x) P_\ell(x) dx P_\ell(u).$$

Dans le cas général, c'est à dire si C n'est pas de la forme $C = \ell(\ell + 1)$, les solutions de l'équation de Legendre ne sont plus des polynômes, mais sont des séries infinies.

Tout comme les fonctions de Bessel, et bien d'autres systèmes de polynômes ou fonctions spéciales, les polynômes de Legendre peuvent être obtenus à partir d'une fonction génératrice, ici la fonction

$$g(x, t) = \frac{1}{\sqrt{1 - 2tx + t^2}}.$$

Il suffit pour s'en convaincre d'utiliser le développement en série entière de $(1 + \epsilon)^{-1/2}$, qui converge pour $|\epsilon| < 1$ (pour nous ici, $|t| < 1$).

On peut en déduire une relation de récurrence

$$(n + 1)P_{n+1}(x) - 2nxP_n(x) + (n - 1)P_{n-1}(x) = 0.$$

On montre également la **formule de Rodrigues** :

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n$$

6.2.4 Autres exemples de polynômes orthogonaux

La même approche peut se transposer à d'autres situations, c'est à dire à d'autres équations différentielles, que l'on résout de façon similaire. Par exemple, on peut obtenir les **polynômes de Laguerre** à partir de l'équation différentielle

$$xy''(x) + (1 - x)y'(x) = \lambda y(x), \quad x \in \mathbb{R}^+$$

(noter le point singulier régulier en $x = 0$). Ils peuvent également être définis par la **formule de Rodrigues**

$$L_n(x) = \frac{e^x}{n!} \frac{d^n}{dx^n} (e^{-x} x^n),$$

ou une fonction génératrice, via

$$\frac{e^{-xt/(1-t)}}{1-t} = \sum_{n=0}^{\infty} L_n(x)t^n,$$

et sont orthogonaux par rapport au produit scalaire

$$\langle f|g \rangle = \int_0^{\infty} f(x)g(x)e^{-x} dx.$$

Ils apparaissent notamment en mécanique quantique dans la partie radiale de la solution de l'équation de Schrödinger pour un atome à un électron.

Les **polynômes d'Hermite** sont quant à eux obtenus comme solutions de l'équation

$$y''(x) - 2xy'(x) = 2\lambda y(x), \quad x \in \mathbb{R}.$$

Ils peuvent également s'écrire via une **formule de Rodrigues**

$$H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} \left(e^{-x^2/2} \right),$$

où à l'aide de la fonction génératrice

$$\exp(xt - t^2/2) = \sum_{n=0}^{\infty} He_n(x) \frac{t^n}{n!},$$

et sont orthogonaux par rapport au produit scalaire

$$\langle f|g \rangle = \int_{-\infty}^{\infty} f(x)g(x)e^{-x^2/2} dx.$$

On peut trouver de multiples autres exemples d'intérêt physique. Ils sont toutefois souvent des cas particuliers d'une théorie plus générale, que nous allons brièvement évoquer plus loin.

6.3 Problème de Sturm-Liouville

Le cas des polynômes de Legendre que nous avons vus plus haut correspond à un cas particulier d'un problème plus général, appelé **problème de Sturm-Liouville**, qu'on décrit ici sous une forme quelque peu simplifiée.

6.3.1 Généralités

On considère, sur un intervalle $[a, b]$, les équations du second ordre, de la forme

$$Ly(x) = a_0(x) \frac{d^2 y}{dx^2}(x) + a_1(x) \frac{dy}{dx}(x) + a_2(x)y(x) = 0, \quad (6.36)$$

où $a_0 \in C^2([a, b])$, $a_1 \in C^1([a, b])$ et $a_2 \in C([a, b])$ sont des fonctions à valeurs réelles, telles que a_0 ne s'annule pas sur $[a, b]$, sauf éventuellement en un nombre fini de points.

Pour espérer existence et unicité de solutions de cette équation ou d'équations associées (voir ci-dessous), on doit la compléter par deux conditions supplémentaires, des **conditions aux bords**. On choisit généralement des conditions aux bords de la forme $y(a) = y(b) = 0$, ou plus généralement des conditions qui assurent que $a_0(x)y(x)y'(x) = 0$ en $x = a$ et $x = b$. Plaçons nous dans le cas le plus simple, et limitons l'analyse à l'espace $C_0^2([a, b])$ des fonctions deux fois continûment différentiables qui s'annulent en a et b .

L est un **opérateur différentiel**, dont l'adjoint L^* est obtenu de la façon habituelle : par définition, $\forall u, v \in C_0^2([a, b])$,

$$\begin{aligned}
 \langle L^*u, v \rangle &= \langle u, Lv \rangle \\
 &= \int_a^b u(x) [a_0(x)v''(x) + a_1(x)v'(x) + a_2(x)v(x)] dx \\
 &= [a_0(x)u(x)v'(x)]_a^b - \int_a^b (a_0u)'(x)v'(x) dx + [a_1(x)u(x)v(x)]_a^b \\
 &\quad - \int_a^b (a_1u)'(x)v(x) dx + \int_a^b a_2(x)u(x)v(x) dx \\
 &= - [(a_0u)'(x)v(x)]_a^b + \int_a^b (a_0u)''(x)v(x) dx \\
 &\quad - \int_a^b (a_1u)'(x)v(x) dx + \int_a^b a_2(x)u(x)v(x) dx \\
 &= \int_a^b [(a_0u)''(x) - (a_1u)'(x) + a_2(x)u(x)] v(x) dx,
 \end{aligned}$$

où on a utilisé des intégrations par parties, et le fait que les fonctions u, v considérées ainsi que leurs dérivées s'annulent en a et b . On a donc montré que l'adjoint de L est donné par

$$L^*y(x) = \frac{d^2}{dx^2} (a_0(x)y(x)) - \frac{d}{dx} (a_1(x)y(x)) + a_2(x)y(x), \quad (6.37)$$

d'où on déduit

LEMME 6.1 L est auto-adjoint si et seulement si les fonctions a_0 et a_1 sont telles que

$$a_0'(x) = a_1(x).$$

Preuve : il suffit de calculer

$$(a_0y)'' = a_0y'' + 2a_0'y' + a_0''y, \quad (a_1y)' = a_1y' + a_1'y$$

et d'insérer ce résultat dans l'expression de L^* . ♠

Dans ces conditions, L peut aussi se mettre sous la forme

$$L = \frac{d}{dx} p \frac{d}{dx} + q, \quad (6.38)$$

ou encore

$$[Ly](x) = \frac{d}{dx} \left[p(x) \frac{dy}{dx}(x) \right] + q(x)y(x) = 0, \quad (6.39)$$

avec $p(x) = a_0(x)$ et $q(x) = a_2(x)$.

REMARQUE 6.6 Les équations de type Sturm-Liouville peuvent généralement s'obtenir suite à des problèmes d'optimisation de fonctionnelles du type

$$J[y] = \frac{1}{2} \int_a^b [p(x)y'(x)^2 + q(x)y(x)^2] dx$$

avec conditions aux bords $y(a) = y(b) = 0$, et une contrainte de normalisation

$$\int_a^b w(x)y(x)^2 dx = 1.$$

REMARQUE 6.7 Dans le cas général, une équation telle que (6.36) peut toujours se mettre sous la forme (6.39), en la multipliant terme à terme par le facteur intégrant

$$I(x) = \frac{1}{a_0(x)} \exp \left\{ \int_{x_0}^x \frac{a_1(z)}{a_0(z)} dz \right\}$$

et en posant

$$p(x) = \exp \left\{ \int_{x_0}^x \frac{a_1(z)}{a_0(z)} dz \right\},$$

et

$$q(x) = \frac{a_2(x)}{a_0(x)} \exp \left\{ \int_{x_0}^x \frac{a_1(z)}{a_0(z)} dz \right\}.$$

DÉFINITION 6.5 Une équation

$$Ly(x) = \lambda w(x)y(x),$$

où L est un opérateur différentiel du second ordre auto-adjoint, et w est une fonction positive sur $[a, b]$ (sauf éventuellement en un nombre fini de points où elle peut s'annuler) est appelée **équation de Sturm-Liouville**

$$\left[\frac{d}{dx} p(x) \frac{d}{dx} \right] y(x) + q(x)y(x) = \lambda w(x)y(x). \quad (6.40)$$

Lorsque, pour des conditions aux bords données, cette équation admet des solutions y_n pour certaines valeurs λ_n spécifiques, y_n est appelée fonction propre, et λ_n valeur propre.

EXEMPLE 6.5 Prenons le cas particulier de l'équation des ondes sur $[0, 1]$, c'est à dire $p(x) = 1$, $q(x) = 0$ et $w(x) = 1$ pour tout $x \in [0, 1]$. L'équation de Sturm-Liouville prend la forme

$$y''(x) = \lambda y(x),$$

dont on a déjà vu qu'elle peut se résoudre via l'équation caractéristique

$$\alpha^2 = \lambda.$$

Les solutions sont de la forme

$$y(x) = A \exp\{\lambda^{1/2}x\} + B \exp\{-\lambda^{1/2}x\}.$$

En imposant les conditions aux bords

$$y(0) = y(1) = 0,$$

on aboutit à $B = -A$, et $\sin(\lambda^{1/2}) = 0$, dont les solutions sont $\lambda^{1/2} = ik\pi$. Les valeurs propres sont donc de la forme

$$\lambda = -k^2\pi^2, \quad k \in \mathbb{Z}^+,$$

et les fonctions propres correspondantes sont

$$y_k(x) = \sin(k\pi x).$$

Ainsi, dans ce cas particulier, la base orthonormale associée au problème de Sturm-Liouville considéré est une base trigonométrique. On considère généralement les bases associées à des problèmes de Sturm-Liouville comme des généralisations des bases de Fourier.

6.3.2 Propriétés des fonctions propres et valeurs propres d'une équation de Sturm-Liouville

Les fonctions propres et les valeurs propres de ces équations possèdent des propriétés importantes. Notamment, on va montrer que les valeurs propres sont toujours réelles, et que les fonctions propres associées sont orthogonales, au sens du produit scalaire

$$\langle f|g \rangle_w = \int_a^b \bar{f}(x)g(x)w(x)dx$$

définissant l'espace

$$L^2([a, b], w) = \left\{ f : [a, b] \rightarrow \mathbb{C}, \int_a^b |f(x)|^2 w(x) dx < \infty \right\},$$

dont on peut montrer qu'il s'agit d'un espace de Hilbert.

Pour cela, soient λ_m et λ_n deux valeurs propres, et y_m et y_n les fonctions propres associées. On a alors

$$\begin{aligned} \left[\frac{d}{dx} p(x) \frac{d}{dx} \right] y_m(x) &= -(q(x) - \lambda_m w(x)) y_m(x), \\ \left[\frac{d}{dx} p(x) \frac{d}{dx} \right] y_n(x) &= -(q(x) - \lambda_n w(x)) y_n(x). \end{aligned}$$

Multipliant complexe conjuguée de la première équation par $y_n(x)$ et la seconde équation par $\bar{y}_m(x)$, on obtient par intégration

$$\int_a^b \left(\bar{y}_m(x) [p(x)y_n'(x)]' - y_n(x) [p(x)\bar{y}_m'(x)]' \right) dx = (\lambda_n - \bar{\lambda}_m) \int_a^b \bar{y}_m(x)y_n(x)w(x)dx$$

Après intégration par parties, en utilisant les conditions aux bords pour éliminer les termes tout intégrés, on obtient

$$(\lambda_n - \bar{\lambda}_m) \int_a^b \bar{y}_m(x)y_n(x)w(x)dx = 0,$$

d'où on déduit :

– Deux fonctions propres y_n et y_m correspondant à des valeurs propres différentes λ_n et λ_m sont orthogonales par rapport au produit scalaire de $L^2([a, b], w)$:

$$\langle y_m|y_n \rangle_w = \int_a^b \bar{y}_m(x)y_n(x)w(x)dx = 0 \quad \text{si } m \neq n.$$

– Les valeurs propres λ_n sont réelles.

Il est possible de montrer de plus que

– Les valeurs propres λ_n ne sont pas bornées, plus précisément

$$\lim_{n \rightarrow \infty} \lambda_n = \infty,$$

– les fonctions propres sont complètes dans $L^2([a, b], w)$: une fonction $f \in L^2([a, b], w)$ telle que $\langle f|y_n \rangle_w = 0$ est nécessairement nulle : $\|f\|_w = 0$.

Par conséquent, on a

PROPOSITION 6.3 Les fonctions propres associées à une équation de Sturm-Liouville (6.39) forment une base orthogonale de l'espace $L^2([a, b], w)$. En les normalisant de sorte que

$$\|y_n\|_w^2 = \langle y_n|y_n \rangle_w = 1,$$

on a donc, $\forall f \in L^2([a, b], w)$,

$$f = \sum_n \langle y_n|f \rangle_w y_n. \quad (6.41)$$

Notons que ceci implique directement que $\forall f \in L^2([a, b], w)$,

$$Lf = \sum_n \langle y_n | Lf \rangle_w y_n = \sum_n \langle Ly_n | f \rangle_w y_n = \sum_n \lambda_n \langle y_n | f \rangle_w y_n, \quad (6.42)$$

expression qui sera souvent utile pour la résolution d'équations aux dérivées partielles.

6.4 Résolution d'équations différentielles par transformation

Dans certaines situations, notamment pour ce qui concerne les équations différentielles à coefficients constants, il est possible d'exploiter les propriétés remarquables de certaines transformations vis à vis de la différentiation. C'est particulièrement le cas de la transformation de Fourier et de la transformation de Laplace. La première s'utilise préférentiellement dans des situations où l'équation différentielle est considérée dans un domaine non-borné (on parle alors de problème aux limites). La seconde est plus adaptée aux problèmes dits *de Cauchy*, c'est à dire sur \mathbb{R}^+ .

6.4.1 Transformation de Fourier

Etant donnée une fonction f , on sait que la transformée de Fourier de sa dérivée p -ième s'exprime simplement via la transformée de Fourier de la fonction elle même :

$$\widehat{f^{(p)}}(\omega) = (i\omega)^p \hat{f}(\omega),$$

où on a défini la transformée de Fourier par

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt, \quad (6.43)$$

à condition que la fonction f soit de classe C^p sur \mathbb{R} , et que ses dérivées tendent vers zéro à l'infini.

REMARQUE 6.8 L'hypothèse $f \in C^p$ est importante pour que cette relation soit correcte. Si elle n'est pas satisfaite, des termes supplémentaires doivent être pris en compte.

Ainsi, la transformation de Fourier transforme dérivation en multiplication point par point, et donc équation différentielle en équation algébrique.

EXEMPLE 6.6 Considérons l'équation différentielle ordinaire

$$-f''(x) + a^2 f(x) = g(x),$$

où $g \in L^2(\mathbb{R})$ est une fonction fixée. On complète cette équation par les conditions aux limites

$$\lim_{x \rightarrow \pm\infty} f(x) = 0.$$

En faisant l'hypothèse que f est de classe C^2 , on obtient par transformation de Fourier

$$(\omega^2 + a^2) \hat{f}(\omega) = \hat{g}(\omega),$$

d'où on déduit la solution (dans l'espace de Fourier

$$\hat{f}(\omega) = \frac{\hat{g}(\omega)}{\omega^2 + a^2}.$$

Une transformation de Fourier inverse (utilisant le théorème des résidus, ou une table de transformées de Fourier) conduit alors à la solution

$$f(x) = \frac{1}{2a} \int_{-\infty}^{\infty} e^{-a|y|} g(x-y) dy.$$

Dans l'exemple ci-dessus, ainsi que dans de nombreux autres exemples, la transformation de Fourier transforme une équation différentielle

$$Lf = g$$

(où L est un opérateur différentiel, à coefficients constants) en équation algébrique

$$\Omega(\omega)\hat{f}(\omega) = \hat{g}(\omega).$$

Ensuite, si la fonction $\omega \rightarrow 1/\Omega(\omega)$ a de bonnes propriétés, cette équation permet d'exprimer \hat{f} en fonction de \hat{g} , et finalement f en fonction de g . Notons que \hat{f} s'exprimant comme produit simple dans le domaine de Fourier, f s'exprimera comme produit de convolution de g avec une certaine fonction (appelée *réponse impulsionnelle*, ou *fonction de Green* de L), transformée de Fourier inverse de $1/\Omega$. On obtient ainsi des solutions de la forme

$$f(x) = \int h(y)f(x-y) dy,$$

où h est une transformée de Fourier inverse de $1/\Omega$. Il est à noter que la détermination de la fonction de Green h doit exploiter les conditions aux bords ou aux limites associées à l'équation différentielle considérée. Nous verrons cela plus en détails dans le chapitre consacré aux distributions et aux fonctions de Green.

6.4.2 Transformation de Laplace

De même, dans le cas d'une fonction définie sur \mathbb{R}^+ (comme on en rencontre dans les problèmes de valeur initiale), elles aussi de classe C^p sur \mathbb{R}^+ , on a recours à la transformée de Laplace

$$F(p) = [\mathcal{L}f](p) = \int_0^\infty f(t)e^{-pt} dt, \Re(p) > s_f \quad (6.44)$$

qui vérifie

$$\mathcal{L}f^{(m)}(p) = p^m F(p) - p^{m-1}f(0) - p^{m-2}f'(0) - \dots - f^{(m-1)}(0).$$

A l'instar de la transformation de Fourier, la transformation de Laplace a la propriété marquante de transformer une équation différentielle à coefficients constants en équation algébrique.

EXEMPLE 6.7 Prenons l'exemple de l'équation de Poisson

$$\Delta u = f, \quad (6.45)$$

où f est un second membre fixé, et $\Delta = d^2/dx^2$ est le Laplacien unidimensionnel. Pour obtenir une solution unique, on sait qu'il est nécessaire de faire des hypothèses supplémentaires, et d'adjoindre à cette équation deux conditions additionnelles. On suppose donc que u, u' et f satisfont la condition (6.44) (il est possible de démontrer dans un cadre plus général l'existence de solutions satisfaisant à de telles conditions), et que f et f' sont continues. Supposons aussi par exemple que

$$u(0) = u_0, \quad u'(0) = v_0,$$

où u_0 et v_0 sont deux nombres fixés. On note U la transformée de Laplace de u , et F la transformée de Laplace de f . Les hypothèses faites assurent l'existence (et l'analyticité) de F dans un domaine $\{p \in \mathbb{C}, \Re(p) > s_u\}$, avec $s_u < \infty$. on se ramène alors à

$$p^2 U(p) - pu_0 - v_0 = F(p),$$

soit encore, pour $p \neq 0$,

$$U(p) = \frac{F(p) + v_0}{p^2} + \frac{u_0}{p},$$

dans un domaine de valeurs de p bien choisi : la fonction U ainsi obtenue est analytique dans le domaine $\{p \in \mathbb{C}, \Re(p) > \max(s_f, 0)\}$. Cette équation permet d'obtenir une solution u à partir de f , par transformation de Laplace inverse (que nous avons déjà vue).

On peut toutefois utiliser ce que l'on sait déjà, c'est à dire les propriétés de linéarité de la transformation de Laplace, sa relation avec le produit de convolution ainsi que les quelques transformées de Laplace déjà vues.

Ainsi, on sait que l'original de Laplace de la fonction $p \rightarrow 1/p$ est la fonction de Heaviside Θ , et que l'original de Laplace de $p \rightarrow 1/p^2$ est la fonction $t \rightarrow t\Theta(t)$. Par ailleurs, l'original de Laplace de la fonction $p \rightarrow F(p)/p^2$ est le produit de convolution de f par la fonction $t \rightarrow t\Theta(t)$. On en déduit donc

$$u(t) = \Theta(t) \left(u_0 + v_0 t + \int_0^t s f(t-s) ds \right).$$

Equations aux dérivées partielles

Les équations aux dérivées partielles (qu'on notera EDP) jouent un rôle fondamental en physique. En particulier les EDP du second ordre permettent de décrire un certain nombre de phénomènes physiques de base, tels que les phénomènes de diffusion, ou de propagation.

L'objectif de ce chapitre est de décrire les principales classes d'EDP du second ordre, et de donner les principales approches pour leur résolution. On verra en particulier que les transformations (notamment la transformation de Fourier) jouent un rôle important pour les problèmes "sans bord" (c'est à dire en espace infini), alors que les méthodes de type "Sturm-Liouville" que nous avons décrites dans le chapitre 6 sont particulièrement appropriées pour ce que l'on appelle les problèmes "aux bords".

7.1 EDP du second ordre, classification

La majorité des équations différentielles qui apparaissent en physique font intervenir des dérivées partielles par rapport aux variables spatiales et temporelle, et sont donc des équations aux dérivées partielles. On se limitera ici aux équations aux dérivées partielles du second ordre, c'est à dire aux équations de la forme

$$\sum_{k,\ell} \alpha_{k\ell} \frac{\partial^2}{\partial x_k \partial x_\ell} \varphi(\underline{x}) + \sum_k \beta_k \frac{\partial}{\partial x_k} \varphi(\underline{x}) + \gamma \varphi(\underline{x}) = 0,$$

où les $\alpha_{k\ell}$, β_k et γ sont des fonctions fixées.

Dans le cas d'équations dépendant de deux variables, on distingue les trois cas suivants :

1. **Equations hyperboliques** : l'exemple classique est l'équation des ondes

$$\frac{\partial^2 \varphi}{\partial x^2}(x, t) - \frac{1}{c^2} \frac{\partial^2 \varphi}{\partial t^2}(x, t) = f(x, t).$$

2. **Equations paraboliques** : l'exemple classique est l'équation de la chaleur

$$\frac{\partial^2 \varphi}{\partial x^2}(x, t) - \alpha \frac{\partial \varphi}{\partial t}(x, t) = f(x, t).$$

3. **Equations elliptiques** : l'exemple classique est l'équation de Laplace

$$\frac{\partial^2 \varphi}{\partial x^2}(x, y) + \frac{\partial^2 \varphi}{\partial y^2}(x, y) = f(x, y).$$

Les solutions de ces différentes équations aux dérivées partielles ont des comportements radicalement différents, qui correspondent à des situations physiques très différentes elles aussi.

La classification est basée sur une analogie avec les cônes, et les considérations suivantes. Considérons une EDP du second ordre, faisant intervenir deux variables t et x :

$$A \frac{\partial^2 f(x,t)}{\partial t^2} + B \frac{\partial^2 f(x,t)}{\partial t \partial x} + C \frac{\partial^2 f(x,t)}{\partial x^2} + D \frac{\partial f(x,t)}{\partial t} + E \frac{\partial f(x,t)}{\partial x} = g(x,t). \quad (7.1)$$

Par analogie avec les cônes, d'équation

$$ax^2 + bxt + ct^2 + dt + ex = g,$$

cette équation est dite

- **Parabolique** si $B^2 - 4AC = 0$.
- **Hyperbolique** si $B^2 - 4AC > 0$.
- **Elliptiques** si $B^2 - 4AC < 0$.

REMARQUE 7.1 En général, les paramètres A, B, C, D et E peuvent être des fonctions de x et t , de sorte que le signe du discriminant $B^2 - 4AC$ peut varier d'un point (x, t) à l'autre, et par là l'équation changer de régime.

7.2 Problèmes elliptiques : le Laplacien

Avant d'aborder les problèmes elliptiques, il est utile de s'attarder un peu sur le Laplacien.

7.2.1 Le Laplacien

L'opérateur Laplacien joue un rôle majeur en Physique. En effet, il intervient par exemple dans le calcul des variations de fonctionnelles faisant intervenir la norme carrée du gradient d'une fonction, ce qui est une situation courante.

Le Laplacien peut prendre différentes formes, suivant le système de coordonnées utilisé. Par exemple, dans le **cas bidimensionnel**, il s'écrit soit en coordonnées Cartésiennes, soit en coordonnées polaires (voir la Figure 7.1) :

$$\Delta f(x, y) = \frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2} \quad (7.2)$$

$$\Delta g(r, \theta) = \frac{\partial^2 g(r, \theta)}{\partial r^2} + \frac{1}{r} \frac{\partial g(r, \theta)}{\partial r} + \frac{1}{r^2} \frac{\partial^2 g(r, \theta)}{\partial \theta^2}. \quad (7.3)$$

Dans le **cas tridimensionnel**, on peut l'exprimer en coordonnées Cartésiennes, cylindriques ou sphériques (voir la Figure 7.2) :

$$\Delta f(x, y, z) = \frac{\partial^2 f(x, y, z)}{\partial x^2} + \frac{\partial^2 f(x, y, z)}{\partial y^2} + \frac{\partial^2 f(x, y, z)}{\partial z^2} \quad (7.4)$$

$$\Delta g(r, \theta, z) = \frac{\partial^2 g(r, \theta, z)}{\partial r^2} + \frac{1}{r} \frac{\partial g(r, \theta, z)}{\partial r} + \frac{1}{r^2} \frac{\partial^2 g(r, \theta, z)}{\partial \theta^2} + \frac{\partial^2 g(r, \theta, z)}{\partial z^2} \quad (7.5)$$

$$\begin{aligned} \Delta h(r, \theta, \varphi) = & \frac{\partial^2 h(r, \theta, \varphi)}{\partial r^2} + \frac{2}{r} \frac{\partial h(r, \theta, \varphi)}{\partial r} + \frac{1}{r^2 \sin^2 \varphi} \frac{\partial^2 h(r, \theta, \varphi)}{\partial \theta^2} \\ & + \frac{1}{r^2} \frac{\partial^2 h(r, \theta, \varphi)}{\partial \varphi^2} + \frac{1}{r^2 \tan \varphi} \frac{\partial h(r, \theta, \varphi)}{\partial \varphi} \end{aligned} \quad (7.6)$$

Il est important de bien comprendre la signification physique du Laplacien, nous allons l'illustrer dans le cas bidimensionnel. L'information fondamentale est que le Laplacien d'une fonction de deux variables F mesure la **concavité** de F au point considéré. Ainsi :

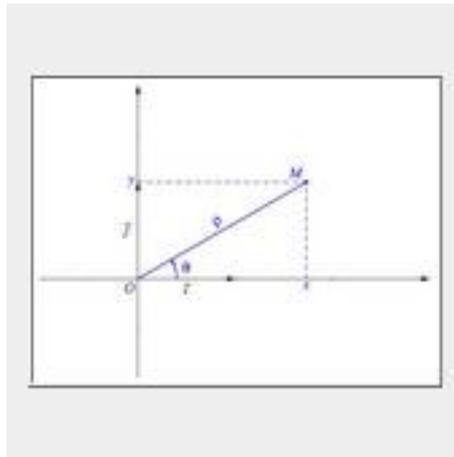


FIG. 7.1: Systèmes de coordonnées polaires dans le plan.

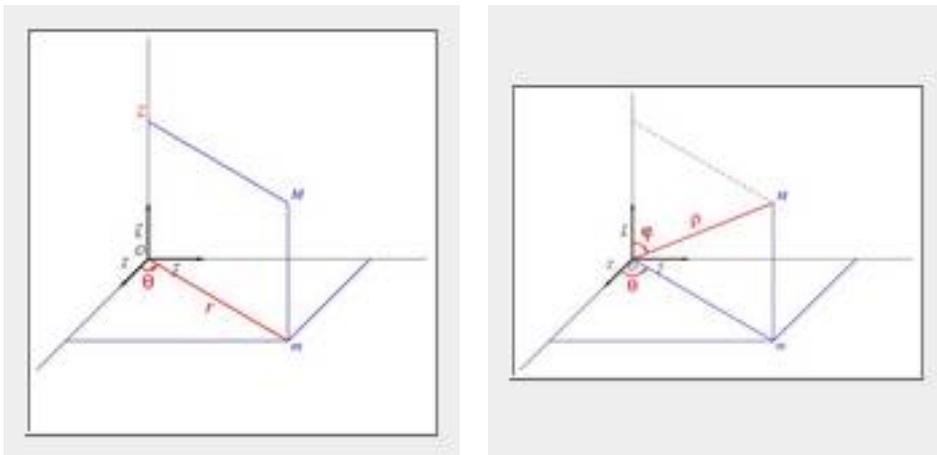


FIG. 7.2: Systèmes de coordonnées cylindriques et sphériques dans l'espace 3D.

- Si $\Delta F(x_0, y_0) > 0$, la fonction F est concave au point (x_0, y_0) . Donc $F(x_0, y_0)$ est automatiquement inférieur à une moyenne de F sur un voisinage de (x_0, y_0) , par exemple un petit cercle centré sur le point (x_0, y_0) .
- Si $\Delta F(x_0, y_0) < 0$, la fonction F est convexe au point (x_0, y_0) . Donc $F(x_0, y_0)$ est automatiquement supérieur à une moyenne de F sur un voisinage de (x_0, y_0) , par exemple un petit cercle centré sur le point (x_0, y_0) .
- Si $\Delta F(x_0, y_0) = 0$, la fonction F est “plate” au point (x_0, y_0) . Donc $F(x_0, y_0)$ est automatiquement égal à une moyenne de F sur un voisinage de (x_0, y_0) , par exemple un petit cercle centré sur le point (x_0, y_0) .

De là, nous pouvons donner une interprétation simple à un certain nombre d'équations classiques de la physique.

1. L'équation de Laplace $\Delta u = 0$ traduit le fait que la solution u est toujours égale à sa moyenne prise sur un voisinage. Par exemple, la hauteur d'une membrane attachée par son bord satisfait l'équation de Laplace. Ceci traduit le fait que la hauteur de la membrane en un point est toujours égale à la moyenne des hauteurs sur un petit cercle centré en ce point.
2. L'équation de la chaleur $u'_t = \alpha^2 \Delta u$ décrit (entre autres) l'évolution d'un champ de température (ou de concentration), et peut s'interpréter comme le fait que la variation de température $u'_t(x, y)$ au point de coordonnées (x, y) est proportionnelle $\Delta u(x, y)$, c'est à dire à la concavité de la

membrane au point de coordonnées (x, y) . Ainsi, si $u(x, y)$ est inférieur à la moyenne de u dans un voisinage de (x, y) , alors la température en ce point va augmenter. Inversement, si $u(x, y)$ est supérieur à la moyenne de u dans un voisinage de (x, y) , alors la température en ce point va diminuer.

3. L'équation des ondes $u''_{tt} = \alpha^2 \Delta u$ décrit (entre autres) le déplacement vertical de la membrane d'un tambour, et peut s'interpréter comme le fait que l'accélération (ou la force) $u''_{tt}(x, y)$ au point de coordonnées (x, y) est proportionnelle $\Delta u(x, y)$, c'est à dire à la concavité de la membrane au point de coordonnées (x, y) . Ainsi, si $u(x, y)$ est inférieur à la moyenne de u dans un voisinage de (x, y) , alors l'accélération (et la force) est positive.

7.2.2 Conditions aux bords

On s'intéresse ici à ce que l'on appelle des **problèmes aux bords**, c'est à dire des équations aux dérivées partielles faisant intervenir des dérivées spatiales, définies dans un domaine borné, et dans lesquels des conditions aux bords sont spécifiées. Ce type de problème peut se rencontrer dans de nombreuses circonstances, par exemple après avoir éliminé une variable temporelle d'un problème parabolique ou hyperbolique par séparation des variables.

Conditions de Dirichlet

Les conditions aux bords de Dirichlet spécifient les valeurs de la solution sur le bord du domaine. L'exemple le plus simple, en deux dimensions, est fourni par l'équation de Laplace

$$\Delta u(x, y) = 0, \quad (7.7)$$

dans un disque Ω de rayon R fixé, avec une condition aux bords

$$u(x, y) = g(x, y), \quad \forall (x, y) \in \partial\Omega, \quad (7.8)$$

g étant une fonction fixée sur $\partial\Omega$. On parle alors de **problème de Dirichlet intérieur**, par opposition au **problème de Dirichlet extérieur**, dans lequel on résout (7.7) dans l'espace extérieur à Ω (son complémentaire dans le plan), toujours avec la condition aux bords (7.8).

Conditions de Neumann

Ici, contrairement au cas des conditions aux bords de Dirichlet, on fixe la valeur de la dérivée normale de la solution à la frontière du domaine considéré. La dérivée normale correspond généralement à une quantité physique appelée le **flux** : ce qui entre (ou qui sort) du domaine considéré.

Etant donné le vecteur $\underline{n}(x, y)$, normal au bord $\partial\Omega$ du domaine Ω au point $(x, y) \in \partial\Omega$ on note

$$\frac{\partial u(x, y)}{\partial n} = \underline{n}(x, y) \cdot \nabla u(x, y)$$

la dérivée normale de u au point (x, y) de $\partial\Omega$.

Les conditions de Neumann prennent la forme

$$\frac{\partial u(x, y)}{\partial n} = g(x, y), \quad \forall (x, y) \in \partial\Omega, \quad (7.9)$$

g étant une fonction fixée sur $\partial\Omega$. On parle alors de **problème de Neumann intérieur**, et de **problème de Neumann extérieur** pour le cas complémentaire.

Conditions mixtes

On verra également apparaître des conditions au bord “mixtes”, faisant intervenir à la fois la valeur de la solution sur le bord et sa dérivée normale. Ce cas de figure se présentera notamment dans le cas de l'équation de la chaleur ci dessous.

$$\frac{\partial u(x, y)}{\partial n} + \lambda u(x, y) = g(x, y), \quad \forall (x, y) \in \partial\Omega, \quad (7.10)$$

7.2.3 Le problème de Dirichlet intérieur en dimension 2 pour l'équation de Laplace

On s'intéresse au problème de Dirichlet intérieur pour l'équation de Laplace dans un disque de rayon unité. En coordonnées polaires, le problème est le suivant :

$$\frac{\partial^2 u(r, \theta)}{\partial r^2} + \frac{1}{r} \frac{\partial u(r, \theta)}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u(r, \theta)}{\partial \theta^2} = 0, \quad r \leq 1 \quad (7.11)$$

$$u(1, \theta) = g(\theta), \quad (7.12)$$

g étant une fonction fixée.

On utilise la méthode de séparation des variables, et on recherche des solutions sous la forme

$$u(r, \theta) = R(r)\Theta(\theta), \quad (7.13)$$

où R et Θ sont des fonctions supposées de classe C^2 .

En reportant cette forme particulière dans l'équation de Laplace, et en divisant membre à membre par $R\Theta$, on obtient

$$\frac{R''(r)}{R(r)} + \frac{1}{r} \frac{R'(r)}{R(r)} + \frac{1}{r^2} \frac{\Theta''(\theta)}{\Theta(\theta)} = 0,$$

qui s'écrit aussi

$$r^2 \frac{R''(r)}{R(r)} + r \frac{R'(r)}{R(r)} = - \frac{\Theta''(\theta)}{\Theta(\theta)}.$$

Le premier membre étant indépendant de θ , et le second étant indépendant de r , ils sont tous deux égaux à une constante, appelée *constante de séparation*, que l'on pose égale à λ^2 , avec $\lambda \in \mathbb{C}$. Ceci nous conduit au système de deux équations découplées

$$r^2 R''(r) + rR'(r) - \lambda^2 R(r) = 0 \quad (7.14)$$

$$\Theta''(\theta) + \lambda^2 \Theta(\theta) = 0. \quad (7.15)$$

L'équation angulaire se résout facilement, et donne des solutions de la forme

$$\Theta(\theta) = Ae^{i\lambda\theta} + Be^{-i\lambda\theta}, \quad (7.16)$$

et les contraintes de régularité sur Θ impliquent que λ doit être un entier, que l'on peut choisir positif sans perte de généralité.

$$\lambda = n \in \mathbb{Z}^+.$$

Passons maintenant à l'équation radiale. Il faut distinguer deux cas, selon que $n = 0$ ou $n \neq 0$.

Si $n = 0$, on a l'équation

$$rR''(r) + R'(r) = 0,$$

qui équivaut à l'équation de Sturm-Liouville simple $(rR'(r))' = 0$, d'où $rR'(r) = \beta$ et donc

$$R_0(r) = \beta \ln(r) + \alpha_0. \quad (7.17)$$

Si l'on veut que la solution soit bornée en $r = 0$, on doit nécessairement avoir $\beta = 0$, d'où

$$R_0(r) = \alpha_0 .$$

Si $n \neq 0$: l'équation prend la forme

$$r^2 R''(r) + rR'(r) - n^2 R(r) = 0 ,$$

qui est une équation à coefficients non constants, qui peut être résolue par la méthode de Fröbenius. Il y a en fait plus simple, en observant que l'on peut rechercher des solutions sous la forme $R(r) = r^\lambda$; en reportant cette forme particulière, on obtient

$$\lambda(\lambda - 1) + \lambda - n^2 = 0 ,$$

d'où la solution $\lambda = \pm n$. On vérifie facilement que

$$R_n(r) = \alpha_n r^n + \beta_n r^{-n} , \quad (7.18)$$

et de nouveau la contrainte d'avoir une solution bornée implique $\beta_n = 0$, d'où la solution

$$R_n(r) = \alpha_n r^n .$$

Les solutions à variables séparées conduisent à une solution de la forme

$$u(r, \theta) = a_0 + \sum_{n=1}^{\infty} r^n \left(a_n e^{in\theta} + b_n e^{-in\theta} \right)$$

pour certaines constantes $a_n, b_n \in \mathbb{C}$, à déterminer.

Il est maintenant temps d'imposer les conditions aux bords :

$$u(1, \theta) = \sum_{n=0}^{\infty} \left(a_n e^{in\theta} + b_n e^{-in\theta} \right) = g(\theta) , \quad \forall \theta .$$

En posant

$$a_{-n} = b_n , \quad \forall n \in \mathbb{Z}^- ,$$

on peut remarquer que les coefficients a_n ne sont autres que les coefficients de Fourier de la fonction g

$$a_n = \frac{1}{2\pi} \int_0^{2\pi} g(\theta) e^{-in\theta} d\theta .$$

Ceci conclut la solution du problème. Dans le cas un peu plus général d'un disque de rayon fixé R , on montre de même

THÉORÈME 7.1 Soit \mathcal{D} un disque de rayon R , centré sur l'origine, et soit g une fonction périodique continue sur $[0, 2\pi]$. La solution du problème de Dirichlet intérieur dans \mathcal{D}

$$\Delta u(r, \theta) = 0 \quad \text{dans } \mathcal{D} \quad (7.19)$$

$$u(R, \theta) = g(\theta) , \quad \text{sur } \partial\mathcal{D} \quad (7.20)$$

est donnée par

$$u(r, \theta) = a_0 + \sum_{n=1}^{\infty} \left(\frac{r}{R} \right)^n \left(a_n e^{in\theta} + b_n e^{-in\theta} \right) \quad (7.21)$$

où les coefficients sont donnés par

$$a_n = \frac{1}{2\pi} \int_0^{2\pi} g(\theta) e^{-in\theta} d\theta , \quad b_n = a_{-n} . \quad (7.22)$$

EXEMPLE 7.1 Considérons le cas

$$g(\theta) = 1.$$

Alors on voit facilement que

$$a_n = \delta_{n,0},$$

et que la solution est donc donnée par

$$u(x, y) = 1, \quad \forall (x, y) \in \mathcal{D}.$$

EXEMPLE 7.2 Considérons le cas

$$g(\theta) = \cos(k\theta), \quad k \in \mathbb{Z}.$$

On voit facilement que

$$a_n = \frac{1}{2} (\delta_{n,k} + \delta_{n,-k}),$$

et la solution est donc de la forme

$$u(x, y) = r^k \cos(\theta), \quad \text{avec } r = \sqrt{x^2 + y^2}, \theta = \arctan(y/x).$$

On voit apparaître dans ce cas particulier une caractéristique générale des solutions : plus la condition au bord varie rapidement, plus la solution décroît vite quand $r \rightarrow 0$. Ceci apparaîtra clairement dans les représentations graphiques de l'exemple suivant.

EXEMPLE 7.3 On a représenté ici la solution (obtenue numériquement) correspondant à la condition au bord

$$g(\theta) = \cos(k\theta)^2,$$

pour différentes valeurs de k . Les solutions pour les cas $k = 1$, $k = 3$ et $k = 5$ se trouvent sur les figures de gauche et droite. On y voit en particulier que la solution « suit » bien la condition au bord, et que la remarque de l'exemple précédent est toujours valide : plus k est grand, plus la solution tend vers zéro rapidement quand $r \rightarrow 0$.

Une autre forme de cette solution est donnée par la **formule intégrale de Poisson**

COROLLAIRE 7.1 La solution du problème de Dirichlet intérieur sur le cercle peut s'écrire

$$u(r, \theta) = \frac{1}{2\pi} \int_0^{2\pi} \frac{R^2 - r^2}{R^2 + r^2 - 2rR \cos(\theta - \alpha)} g(\alpha) d\alpha \quad (7.23)$$

Preuve : Il suffit de reporter les coefficients de Fourier à l'intérieur de la solution trouvée :

$$\begin{aligned} u(r, \theta) &= \frac{1}{2\pi} \left[\int_0^{2\pi} g(\alpha) d\alpha + \sum_{n=1}^{\infty} \left(\frac{r}{R}\right)^n \left(\int_0^{2\pi} e^{in(\theta-\alpha)} g(\alpha) d\alpha + \int_0^{2\pi} e^{-in(\theta-\alpha)} g(\alpha) d\alpha \right) \right] \\ &= \frac{1}{2\pi} \int_0^{2\pi} \left[1 + \sum_{n=1}^{\infty} \left(\frac{r}{R}\right)^n \left(e^{in(\theta-\alpha)} + e^{-in(\theta-\alpha)} \right) \right] g(\alpha) d\alpha \\ &= \frac{1}{2\pi} \int_0^{2\pi} \left[1 + \frac{1}{1 - re^{i(\theta-\alpha)}/R} + \frac{1}{1 - re^{-i(\theta-\alpha)}/R} - 2 \right] g(\alpha) d\alpha, \end{aligned}$$

d'où on déduit le résultat par simplification. ♠

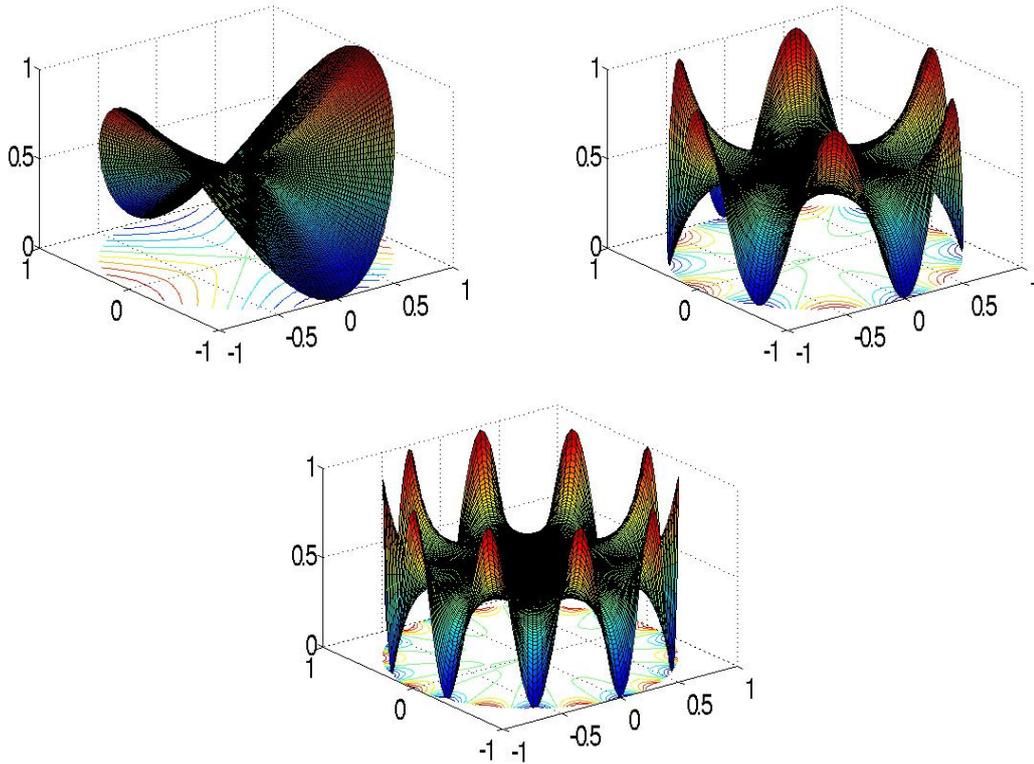


FIG. 7.3: Solution de l'équation de Laplace dans un disque avec condition au bord de Dirichlet pour des conditions au bord oscillantes.

Cette expression est assez remarquable, car elle exprime la solution directement en fonction de la condition au bord, sous la forme d'un produit de convolution

$$u(r, \theta) = (P_r * g)(\theta),$$

où la fonction P_r , appelée *noyau de Poisson* et donnée par

$$P_r(\theta) = \frac{1}{2\pi} \frac{R^2 - r^2}{R^2 + r^2 - 2rR \cos(\theta)}$$

joue donc un rôle fondamental. Notons que ce noyau est singulier en $r = R, \theta = 0$ (mais la singularité est intégrable), ce qui montre que pour $r \approx R$, la contribution principale à l'intégrale (7.23) est donnée par $\theta \approx \alpha$. Le noyau de Poisson est représenté dans la figure 7.4 (où la singularité a été « gommée », pour que la figure soit interprétable); on y voit clairement le comportement mentionné ci-dessus.

7.2.4 Problème extérieur, problème dans une couronne

Le problème extérieur

$$\frac{\partial^2 u(r, \theta)}{\partial r^2} + \frac{1}{r} \frac{\partial u(r, \theta)}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u(r, \theta)}{\partial \theta^2} = 0, \quad r \geq R \tag{7.24}$$

$$u(R, \theta) = g(\theta), \tag{7.25}$$

g étant une fonction fixée. peut être résolu de façon tout à fait similaire. Les solutions générales angulaire (7.16) et radiale (7.17) et (7.18) sont toujours valides. L'exigence d'une solution bornée impose

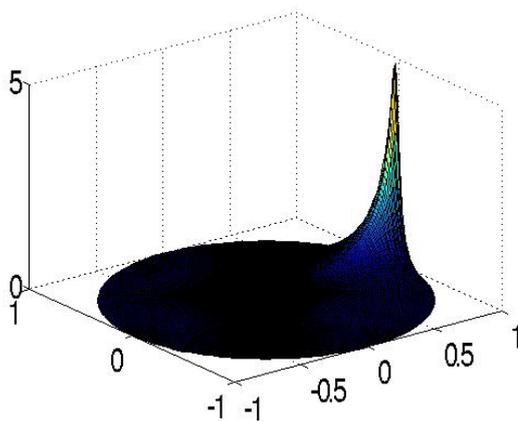


FIG. 7.4: Le noyau de Poisson

encore la nullité du coefficient du terme logarithmique dans (7.17), mais impose cette fois la nullité des coefficients des termes en r^n dans (7.17), ce qui conduit à des solutions de la forme

$$u(r, \theta) = a_0 + \sum_{n=1}^{\infty} \left(\frac{r}{R}\right)^{-n} (a_n e^{in\theta} + b_n e^{-in\theta}) \quad (7.26)$$

où les coefficients sont toujours donnés par (7.22).

Le cas du problème de Dirichlet dans une couronne $R_1 \leq r \leq R_2$, donné par

$$\frac{\partial^2 u(r, \theta)}{\partial r^2} + \frac{1}{r} \frac{\partial u(r, \theta)}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u(r, \theta)}{\partial \theta^2} = 0, \quad R_1 \leq r \leq R_2 \quad (7.27)$$

$$u(R_1, \theta) = g_1(\theta), \quad (7.28)$$

$$u(R_2, \theta) = g_2(\theta), \quad (7.29)$$

où g_1 et g_2 sont deux fonctions fixées est un peu plus complexe, mais se résout par des techniques tout à fait similaires. La solution est de la forme

$$u(r, \theta) = a_0 + b_1 \ln(r) + \sum_{n=1}^{\infty} \left[r^{-n} (a_n e^{in\theta} + b_n e^{-in\theta}) + r^n (c_n e^{in\theta} + d_n e^{-in\theta}) \right],$$

et les constantes sont cette fois encore déterminées à partir des deux conditions aux bords, en calculant leurs coefficients de Fourier.

7.3 Problèmes paraboliques : phénomènes de diffusion

7.3.1 Exemples

L'exemple le plus classique est celui de l'équation de la chaleur (ou équation de Fourier)

$$\frac{\partial u}{\partial t} = \alpha^2 \Delta u \quad (7.30)$$

où $\alpha \in \mathbb{R}$ est un paramètre, et Δ représente le Laplacien en espace. L'équation de la chaleur est utilisée dans de nombreux contextes, par exemple pour décrire l'évolution d'un champ de température dans un espace donné, avec une condition initiale donnée. Elle est également utilisée dans de nombreux autres contextes, qui exhibent des comportements de type **diffusif**.

Comme on l'a vu, le Laplacien mesure la convexité d'une fonction. Si $\Delta u(\underline{x}, t) > 0$, alors $u(\underline{x}, t)$ est inférieure à sa moyenne sur un petit voisinage de \underline{x} ; alors $u'_t(\underline{x}, t) > 0$, ce qui va faire augmenter la valeur de $u(\underline{x}, t)$. Inversement, si $\Delta u(\underline{x}, t) < 0$, alors la valeur de $u(\underline{x}, t)$ va diminuer. Ce type d'équation a donc tendance à moyenniser, ou "lisser" la solution au cours du temps.

Comme autre exemple d'équation parabolique, on peut notamment mentionner l'équation de convection-diffusion, que l'on obtient en ajoutant un terme de convection à l'équation de la chaleur

$$\frac{\partial u}{\partial t} = \alpha^2 \Delta u - \underline{V} \cdot \nabla u ,$$

\underline{V} étant un vecteur (vitesse) fixé.

Les EDP de cette classe peuvent être considérées dans l'espace tout entier, ou dans un domaine borné de l'espace, ce qui donne lieu soit à des conditions au bord, soit des conditions aux limites (dans ce cas on impose en général que la solution tende vers zéro à l'infini). La variable temporelle est en général prise dans \mathbb{R}^+ , et on complète le problème par des conditions initiales.

7.3.2 Le cas de domaines infinis : transformations

Les équations du type "équation de la chaleur" en domaine infini ou semi-infini sont souvent traitées en utilisant des transformations intégrales, comme la transformation de Fourier dans le cas de domaine infini, ou la transformation de Laplace dans le cas semi-infini. L'idée essentielle est que ces transformations transforment la dérivation en multiplication, et donc une EDP en une famille d'équations différentielles ordinaires, plus faciles à traiter.

Comme on a vu quelques exemples auparavant, on ne discutera pas ce cas davantage ici.

7.3.3 Domaines bornés : conditions aux bords

Les conditions aux bords les plus faciles à exploiter sont comme souvent les **conditions de Dirichlet**

$$u(\underline{x}, t) = g(t) , \quad \forall \underline{x} \in \partial\Omega , \quad (7.31)$$

g étant une fonction fixée. Lorsque $g = 0$, on parle de condition au bord homogène.

Il arrive toutefois que les conditions aux bords portent sur le flux, c'est à dire la dérivée normale de la solution sur le bord. On parle alors de **condition de Neumann**

$$\frac{\partial u}{\partial n}(\underline{x}, t) = g(t) , \quad \forall \underline{x} \in \partial\Omega , \quad (7.32)$$

homogène si $g = 0$.

Les conditions au bord mixtes, ou **conditions de Robin** apparaissent assez couramment, notamment dans les problèmes de conduction de la chaleur, pour lesquels la dérivée normale est proportionnelle à la différence entre la valeur de la solution à l'intérieur du domaine et sa valeur à l'extérieur.

$$\frac{\partial u}{\partial n}(\underline{x}, t) - \lambda u(\underline{x}, t) = g(t) , \quad \forall \underline{x} \in \partial\Omega . \quad (7.33)$$

Là encore, la condition au bord est homogène si $g = 0$.

7.3.4 Equation de la chaleur 1D sur un domaine borné

Le champ de température d'une barre conductrice de longueur L fixée, isolée latéralement, est décrit par l'équation de la chaleur (ou équation de Fourier)

$$\frac{\partial u(x, t)}{\partial t} = \alpha^2 \frac{\partial^2 u(x, t)}{\partial x^2} , \quad \alpha \in \mathbb{R} \quad (7.34)$$

à laquelle on ajoute une condition initiale

$$u(x, 0) = \phi(x), \quad (7.35)$$

ϕ étant une fonction fixée, et des conditions aux bords, par exemple de type Dirichlet, Neumann ou Robin.

On recherche en général des solutions par la méthode de séparation des variables, c'est à dire sous la forme

$$u(x, t) = X(x)T(t).$$

Sous cette forme, on se ramène alors à une équation de la forme

$$X(x)T'(t) = \alpha^2 X''(x)T(t),$$

où encore

$$\frac{T'(t)}{T(t)} = \alpha^2 \frac{X''(x)}{X(x)} = -\lambda\alpha^2,$$

λ étant une constante de séparation, a priori complexe.

L'équation temporelle conduit directement à

$$T(t) = Ce^{-\lambda\alpha^2 t},$$

ce qui force $\Re(\lambda) > 0$ si on se limite à des solutions bornées. Pour ce qui est de l'équation spatiale

$$X''(x) + \lambda X(x) = 0,$$

elle conduit à des solutions de la forme

$$X(x) = Ae^{i\lambda^{1/2}x} + Be^{-i\lambda^{1/2}x}. \quad (7.36)$$

Pour aller plus loin, il faut imposer les conditions aux bords. On se limitera ici aux deux types de conditions aux bords les plus courantes, c'est à dire les conditions de Dirichlet et les conditions mixtes de Robin.

Conditions de Dirichlet homogènes

On impose les deux conditions

$$X(0, t) = X(L, t) = 0.$$

Ceci correspond à une situation physique décrite dans la figure 7.5 : une barre conductrice de la chaleur, dont les deux extrémités sont maintenues à température nulle par un thermostat, et isolée latéralement.



FIG. 7.5: Barre conductrice, isolée latéralement, dont les deux extrémités sont à température nulle.

Cette condition au bord conduit à

$$A + B = 0, \quad Ae^{i\lambda^{1/2}L} + Be^{-i\lambda^{1/2}L} = 0.$$

La première équation donne $B = -A$, et la seconde conduit à

$$A \sin(\lambda^{1/2}L) = 0.$$

Si on évite la solution nulle, ceci implique que les seules solutions pour λ sont de la forme

$$\lambda = \lambda_n = \left(\frac{n\pi}{L}\right)^2. \quad (7.37)$$

On obtient donc des solutions à variables séparées de la forme

$$u_n(x, t) = A_n \exp\left\{-\left(\frac{n\pi\alpha}{L}\right)^2 t\right\} \sin\left(\frac{n\pi x}{L}\right),$$

et plus généralement, l'équation et les conditions aux bords étant homogènes, une solution générale

$$u(x, t) = \sum_{n=1}^{\infty} A_n \exp\left\{-\left(\frac{n\pi\alpha}{L}\right)^2 t\right\} \sin\left(\frac{n\pi x}{L}\right).$$

La dernière étape consiste à trouver les valeurs des constantes A_n . On utilise pour cela la condition initiale, qui prend donc la forme

$$\phi(x) = u(x, 0) = \sum_{n=1}^{\infty} A_n \sin\left(\frac{n\pi x}{L}\right).$$

Restent à trouver les constantes A_n . Pour cela, remarquons que

$$\begin{aligned} \int_0^L \phi(x) \sin\left(\frac{m\pi x}{L}\right) dx &= \sum_{n=1}^{\infty} A_n \int_0^L \sin\left(\frac{m\pi x}{L}\right) \sin\left(\frac{n\pi x}{L}\right) dx \\ &= \sum_{n=1}^{\infty} \frac{A_n}{2} \int_0^L \phi(x) \left(\cos\left(\frac{(m-n)\pi x}{L}\right) - \cos\left(\frac{(m+n)\pi x}{L}\right) \right) dx \\ &= \frac{L}{2} A_m \end{aligned}$$

Par conséquent, les coefficients A_n du développement de la solution s'expriment directement à partir de la condition initiale, et on a

THÉORÈME 7.2 *La solution de problème de Dirichlet pour l'équation de la chaleur unidimensionnelle sur l'intervalle $[0, L]$ prend la forme*

$$u(x, t) = \sum_{n=1}^{\infty} A_n \exp\left\{-\left(\frac{n\pi\alpha}{L}\right)^2 t\right\} \sin\left(\frac{n\pi x}{L}\right), \quad (7.38)$$

où les constantes A_n sont données à partir de la condition initiale ϕ par

$$A_n = \frac{2}{L} \int_0^L \phi(x) \sin\left(\frac{n\pi x}{L}\right) dx. \quad (7.39)$$

La manière dont on a retrouvé les coefficients A_n peut sembler un peu arbitraire. En fait, elle provient du résultat suivant, qui suggère donc le calcul fait pour retrouver les A_n :

LEMME 7.1 *La famille de fonctions*

$$e_n : x \in [0, L] \rightarrow e_n(x) = \frac{1}{\sqrt{L}} \sin\left(\frac{\pi n x}{L}\right), \quad n = 1, \dots, \infty \quad (7.40)$$

est une base orthonormée de l'espace $L_0^2([0, L])$ des fonctions de carré intégrable sur $[0, L]$ qui s'annulent en $x = 0$ et en $x = L$.

REMARQUE 7.2 Il est utile de revenir en arrière et retracer l'origine de cette base de fonctions. Nous les avons obtenues comme solutions du problème de type Sturm Liouville

$$X''(x) + \lambda X(x) = 0,$$

en imposant des conditions au bord particulières. Or nous avons vu au chapitre 6 que ces problèmes de Sturm-Liouville fournissent des bases orthonormées d'espaces de Hilbert adaptés aux conditions aux bords choisies. Le lemme ci-dessus se place donc précisément dans ce cadre.

Il est important de donner une interprétation physique à la solution obtenue en (7.38). Le cas $t = 0$ redonne la condition initiale, sous forme d'une série de Fourier particulière (ici une série de sinus). On peut remarquer que l'évolution temporelle de la solution se caractérise par une atténuation de ces coefficients de Fourier, atténuation d'autant plus forte que l'indice fréquentiel (le n du coefficient A_n). Ainsi, les composantes sinusoïdales les plus rapides sont les plus rapidement atténuées. La solution va donc devenir de plus en plus « lisse » au cours du temps.

On peut en voir un exemple dans les représentations graphiques de la figure 7.6, où on a représenté les solutions à différents temps, pour une condition initiale aléatoire.

Conditions de Robin homogènes

On considère maintenant le cas des conditions aux bords mixtes. Ce choix se justifie par des considérations physiques. En effet, la loi de Newton précise que le flux de température à l'interface entre deux milieux est proportionnel à la différence de température entre ces deux milieux. De plus, la loi de Fourier spécifie que ce flux est également proportionnel à la valeur du gradient de température à l'interface. On considère ici la situation décrite dans la figure 7.7, d'une barre conductrice isolée latéralement, dont le côté gauche est maintenu à température nulle, et le droit est en contact avec un milieu se trouvant lui aussi à température nulle.

Le calcul de la solution générale est le même que plus haut, jusqu'à l'équation (7.36). On doit maintenant utiliser la condition au bord, que l'on doit cette fois choisir de type Dirichlet à gauche, et mixte à droite

$$u(0, t) = 0, \quad u'_x(L, t) + hu(L, t) = 0$$

La première condition implique comme précédemment $B = -A$, d'où des solutions de la forme

$$X(x) = A \sin(\lambda^{1/2}x).$$

Comme précédemment, les valeurs possibles de λ sont déterminées par la seconde condition au bord, à savoir

$$\lambda^{1/2} \cos(\lambda^{1/2}L) + h \sin(\lambda^{1/2}L) = 0,$$

d'où les valeurs possibles de $\lambda^{1/2}$ sont les solutions de l'équation

$$\operatorname{tg}(\lambda^{1/2}L) = -\frac{\lambda^{1/2}}{h}, \quad (7.41)$$

équation qui doit être résolue numériquement. On peut toutefois voir que les solutions $\lambda^{1/2}$ sont les points d'intersection de la droite $t \rightarrow -t/h$ (en rouge... pour ceux qui ont la couleur !) avec la courbe $t \rightarrow \operatorname{tg}(tL)$ (en bleu) dans la figure 7.8. Les valeurs admissibles de λ forment donc un ensemble discret $\{\lambda_n, n \in \mathbb{Z}^+\}$.

La solution générale du problème considéré est donc de la forme

$$u(x, t) = \sum_{n=1}^{\infty} a_n e^{-\lambda_n a^2 t} \sin(\lambda_n^{1/2} x), \quad (7.42)$$

(le cas $\lambda_0 = 0$ ne contribue pas), les coefficients a_n restant à déterminer via la condition initiale. On va pour cela utiliser le lemme suivant

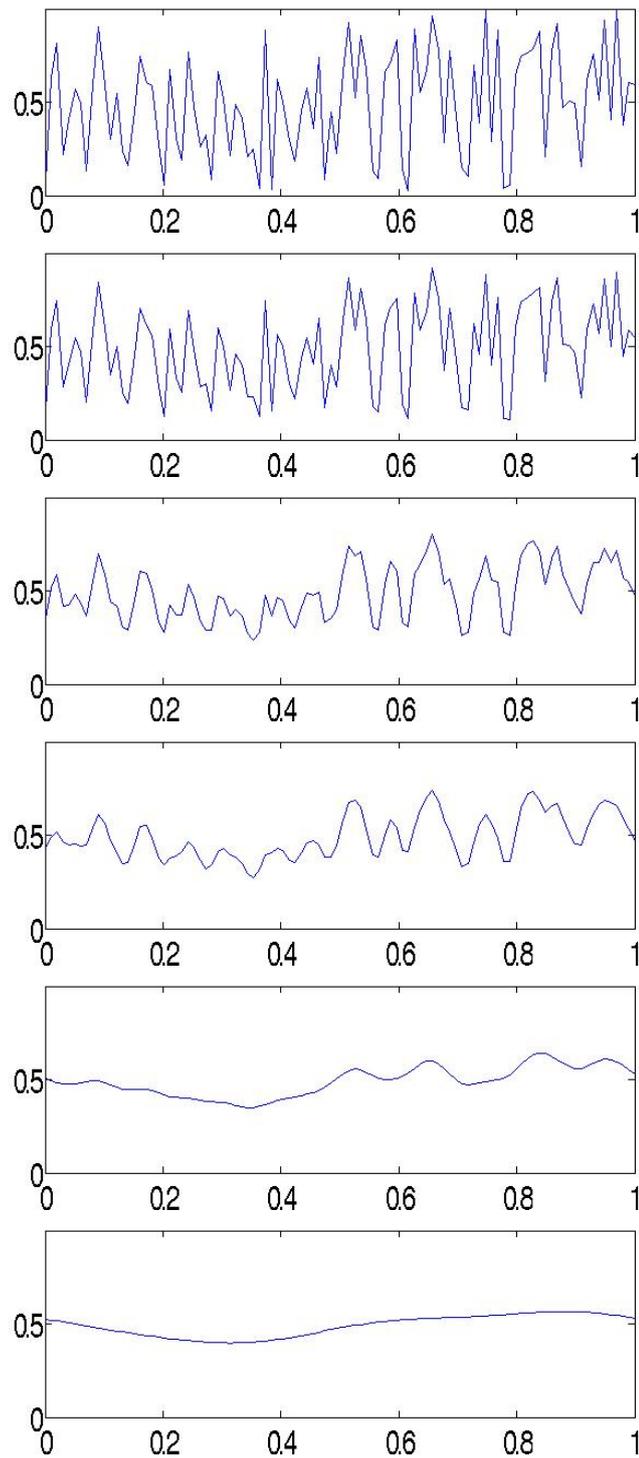


FIG. 7.6: Solutions de l'équation de la chaleur avec une condition initiale aléatoire, et des conditions aux bords de Dirichlet homogènes, pour $t = 0$ (condition initiale), $t = 0,1$, $t = 0,5$, $t = 1$, $t = 5$ et $t = 50$

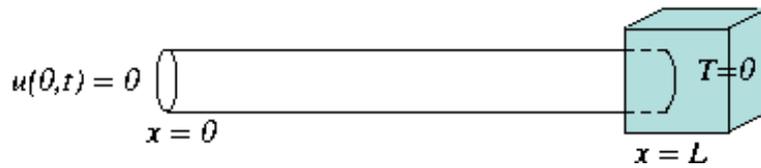


FIG. 7.7: Barre conductrice, isolée latéralement, dont l'extrémité gauche est maintenue à température nulle, et la droite se trouve en contact avec un milieu à température nulle

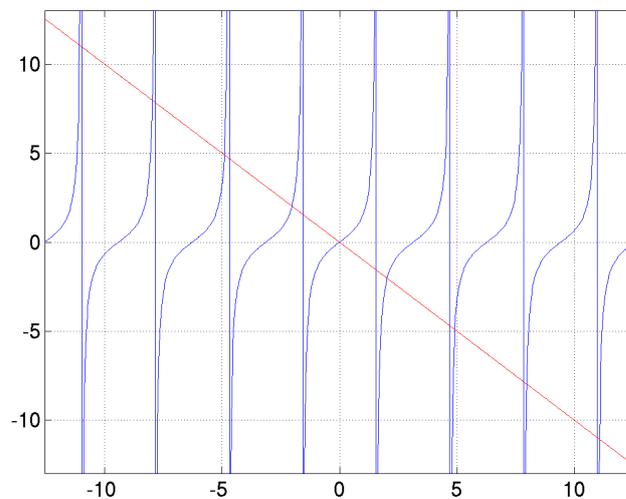


FIG. 7.8: Valeurs propres pour l'équation de la chaleur avec condition au bord mixte : les valeurs de $\lambda^{1/2}$ sont les intersections des deux courbes.

LEMME 7.2 La famille de fonctions

$$f_n : x \in [0, L] \rightarrow f_n(x) = \frac{2}{\sqrt{2L - \sin(2\sqrt{\lambda_n}L)}} \sin(\sqrt{\lambda_n}x), \quad n = 1, \dots, \infty \quad (7.43)$$

est une base orthonormée de l'espace

$$\mathcal{H} = \{f \in C^2([0, L]), f(0) = 0 \text{ et } f'(L) + hf(L) = 0\} .$$

Preuve : pour simplifier, posons $\mu_n = \sqrt{\lambda_n}$, et calculons

$$\begin{aligned}
 \int_0^L \sin(\mu_n x) \sin(\mu_m x) dx &= \left[-\frac{1}{\mu_m} \sin(\mu_n x) \cos(\mu_m x) \right]_0^L + \frac{\mu_n}{\mu_m} \int_0^L \cos(\mu_n x) \cos(\mu_m x) dx \\
 &= -\frac{1}{\mu_m} \sin(\mu_n L) \cos(\mu_m L) + \frac{\mu_n}{\mu_m} \left\{ \left[\frac{\sin(\mu_m x)}{\mu_m} \cos(\mu_n x) \right] \right. \\
 &\quad \left. + \frac{\mu_n}{\mu_m} \int_0^L \sin(\mu_n x) \sin(\mu_m x) dx \right\} \\
 &= -\frac{1}{\mu_m} \left\{ \sin(\mu_n L) \cos(\mu_m L) - [-h \sin(\mu_n L)] [-\cos(\mu_m L)/h] \right\} \\
 &\quad + \left(\frac{\mu_n}{\mu_m} \right)^2 \int_0^L \sin(\mu_n x) \sin(\mu_m x) dx \\
 &= \left(\frac{\mu_n}{\mu_m} \right)^2 \int_0^L \sin(\mu_n x) \sin(\mu_m x) dx ,
 \end{aligned}$$

d'où il s'ensuit que l'intégrale est nulle pour $\mu_n \neq \mu_m$. Pour ce qui est de l'orthogonalité, il suffit de remarquer que

$$\int_0^L \sin(\mu_n x)^2 dx = \frac{L}{2} - \frac{\sin(2\mu_n L)}{4} ,$$

ce qui conclut la preuve. ♠

7.3.5 Problèmes inhomogènes : utilisation des bases "Sturm-Liouville"

Dans le cas où l'équation considérée est inhomogène, il est encore possible d'utiliser des techniques telles que celle que nous venons de voir. Par exemple, prenons le cas de l'équation de la chaleur inhomogène

$$\frac{\partial u}{\partial t}(x, t) = \alpha^2 \frac{\partial^2 u}{\partial x^2}(x, t) + f(x, t) , \quad (7.44)$$

f étant une fonction fixée, jouant le rôle de "source de chaleur", avec condition au bord de Dirichlet. On peut alors utiliser la base de Sturm-Liouville

$$x \rightarrow X_n(x) = C_n \sin(\mu_n x) ,$$

avec $\mu_n = \pi n/L$ et $C_n = 1/\sqrt{L}$, pour décomposer le terme source

$$f(x, t) = \sum_{n=1}^{\infty} F_n(t) X_n(x) ,$$

où

$$F_n(t) = \int_0^L f(x, t) X_n(x) dx ,$$

et de là développer un calcul similaire au précédent.

En effet, on sait que la solution se décompose sous la forme

$$u(x, t) = \sum_{n=1}^{\infty} T_n(t) X_n(x) ,$$

où les fonctions $T_n(t)$ sont à déterminer. L'équation (7.44) se met sous la forme

$$\sum_{n=1}^{\infty} T_n'(t) X_n(x) = \alpha^2 \sum_{n=1}^{\infty} (\lambda_n^2 T_n(t) X_n(x) + F_n(t) X_n(x)) ,$$

d'où on déduit la famille (infinie) d'équations différentielles inhomogènes

$$T_n'(t) - \alpha^2 \lambda_n^2 T_n(t) = F_n(t).$$

En résolvant ces dernières, on obtient les fonctions T_n , qui fournissent la solution du problème inhomogène.

7.4 Problèmes hyperboliques : phénomènes de propagation

Terminons ce chapitre par une rapide discussion des problèmes hyperboliques, qui décrivent les phénomènes de propagation d'ondes par exemple.

7.4.1 Exemples : corde et poutre

L'exemple le plus simple et classique est celui de l'équation aux dérivées partielles décrivant les vibrations transversales d'une corde de violon (ou piano); si on note $u(x, t)$ la position du point x de la corde à l'instant t , on a

$$\frac{\partial^2 u(x, t)}{\partial t^2} = c^2 \frac{\partial^2 u(x, t)}{\partial x^2}. \quad (7.45)$$

Comme on l'a vu plus haut, cette équation signifie qu'à tout instant t s'exerce au point x une force de rappel proportionnelle à la concavité $\Delta u(x, t)$ de u en x .

Bien qu'il ne s'agisse pas d'une équation du second ordre, l'équation régissant les vibrations d'une poutre, appelée équation de Cantilever

$$\frac{\partial^2 u(x, t)}{\partial t^2} = -\alpha^2 \frac{\partial^4 u(x, t)}{\partial x^4} \quad (7.46)$$

conduit à des solutions possédant des propriétés qualitatives similaires.

7.4.2 Problèmes sans bord : la solution de d'Alembert

Le cas unidimensionnel

Commençons par le cas unidimensionnel en espace. L'équation des ondes (7.45) possède des solutions simples. Il s'avère avantageux dans ce cas d'introduire les variables $\xi = x - ct$ et $\zeta = x + ct$. On a alors

$$\frac{\partial}{\partial x} = \frac{\partial}{\partial \xi} + \frac{\partial}{\partial \zeta}, \quad \frac{\partial}{\partial t} = -c \frac{\partial}{\partial \xi} + c \frac{\partial}{\partial \zeta}$$

et

$$\frac{\partial^2}{\partial x^2} = \frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial \zeta^2} + 2 \frac{\partial^2}{\partial \xi \partial \zeta}, \quad \frac{\partial^2}{\partial t^2} = c^2 \left(\frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial \zeta^2} - 2 \frac{\partial^2}{\partial \xi \partial \zeta} \right),$$

de sorte que (7.45) est équivalente à l'équation

$$\frac{\partial^2 u}{\partial \xi \partial \zeta} = 0.$$

En intégrant une première fois par rapport à ζ , on obtient

$$\frac{\partial v}{\partial \xi} = a(\xi),$$

une fonction ne dépendant que de ξ . En intégrant cette fois par rapport à ξ , on obtient finalement

$$v(\xi, \zeta) = f(\xi) + g(\zeta),$$

d'où la solution de d'Alembert

$$u(x, t) = f(x - ct) + g(x + ct). \quad (7.47)$$

On obtient ainsi deux ondes f et g , se propageant respectivement dans la direction des x positifs (onde progressive) et négatifs (onde régressive). Les fonctions f et g sont déterminées à partir des conditions initiales

$$u(x, 0) = \phi(x), \quad \frac{\partial u}{\partial t}(x, 0) = \psi(x),$$

qui conduit au système

$$\begin{aligned} f(x) + g(x) &= \phi(x) \\ -c(f'(x) - g'(x)) &= \psi(x) \end{aligned}$$

qui peut ensuite être résolu directement. On obtient

$$\begin{aligned} f'(x) &= \frac{1}{2} \left(\phi'(x) - \frac{1}{c} \psi(x) \right) \\ g'(x) &= \frac{1}{2} \left(\phi'(x) + \frac{1}{c} \psi(x) \right) \end{aligned}$$

d'où

$$\begin{aligned} f(x) &= \frac{1}{2} \left(\phi(x) - \frac{1}{c} \int_a^x \psi(y) dy \right) \\ g(x) &= \frac{1}{2} \left(\phi(x) + \frac{1}{c} \int_a^x \psi(y) dy \right) \end{aligned}$$

et finalement

$$u(x, t) = \frac{1}{2} (\phi(x - ct) + \phi(x + ct)) + \frac{1}{2c} \int_{x-ct}^{x+ct} \psi(s) ds. \quad (7.48)$$

Le cas tridimensionnel

Le cas tridimensionnel est plus complexe, mais conduit à une solution de nature similaire. On considère l'équation des ondes dans l'espace \mathbb{R}^3

$$\Delta u(\underline{x}, t) = \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2}(\underline{x}, t). \quad (7.49)$$

Considérons le problème défini par l'équation des ondes (7.49) dans l'espace \mathbb{R}^3 , complété par les conditions initiales

$$u(\underline{x}, 0) = \phi(\underline{x}) \quad (7.50)$$

$$u'_t(\underline{x}, 0) = \psi(\underline{x}) \quad (7.51)$$

pour un couple de fonctions ϕ, ψ donné. Il est facile de voir que la solution peut s'exprimer comme la somme d'une solution avec conditions aux bords $(0, \psi)$ et d'une solution avec conditions aux bords $(\phi, 0)$.

Le problème peut se résoudre et utilisant une transformation de Fourier tridimensionnelle, et conduit au résultat suivant. On se place dans un système de coordonnées sphériques (r, θ, φ) , $\varphi \in [0, \pi]$ étant l'angle azimutal et $\theta \in [0, 2\pi]$ l'angle fait par la projection du vecteur dans le plan (xOy) avec l'axe Ox (voir Figure 7.2).

THÉORÈME 7.3 La solution du problème défini par l'équation des ondes (7.49) dans l'espace, complété par les conditions initiales (7.50) et (7.51) est donné par les ondes sphériques

$$u(\underline{x}, t) = t\Psi(\underline{x}, t) + \frac{\partial}{\partial t} [t\Phi(\underline{x}, t)] , \quad (7.52)$$

où Φ et Ψ sont les moyennes des conditions initiales ϕ et ψ , prises sur une sphère de centre $\underline{x} = (x, y, z)$ et de rayon ct :

$$\Psi(\underline{x}, t) = \frac{1}{4\pi c^2 t^2} \int_0^\pi \int_0^{2\pi} \psi(x + ct \sin \varphi \cos \theta, y + ct \sin \varphi \sin \theta, z + ct \cos \varphi) c^2 t^2 \sin \varphi d\varphi d\theta$$

$$\Phi(\underline{x}, t) = \frac{1}{4\pi c^2 t^2} \int_0^\pi \int_0^{2\pi} \phi(x + ct \sin \varphi \cos \theta, y + ct \sin \varphi \sin \theta, z + ct \cos \varphi) c^2 t^2 \sin \varphi d\varphi d\theta$$

Ce résultat est comme annoncé la somme de la solution du problème avec $\phi = 0$ et de la solution avec $\psi = 0$. La première est assez simple à interpréter : on voit que la solution en (x, y, z) à l'instant t est la moyenne de la fonction ψ dans une sphère de rayon ct centrée en (x, y, z) .

7.4.3 Corde vibrante : condition aux bords et résolution

On se place dans le cas unidimensionnel, et on s'intéresse à l'équation

$$\frac{\partial^2}{\partial t^2} u(x, t) = c^2 \frac{\partial^2}{\partial x^2} u(x, t), \quad t \in \mathbb{R}^+, x \in [0, L]. \quad (7.53)$$

La solution de cette équation peut être utilisée pour décrire un grand nombre de situations physiques différentes, comme par exemple

- des ondes acoustiques (transverses),
- des vibrations dans les solides (longitudinales, transverses ou ondes de torsion),
- des ondes de probabilité en mécanique quantique,
- les vibrations (transverses) d'une corde
- ...

C'est ce dernier exemple qu'on prendra comme illustration.

Conditions au bord

On considère généralement trois types de conditions au bord.

1. Une condition de type *Dirichlet*, qu'on appelle dans ce cas condition de bords contrôlés

$$u(0, t) = g_1(t), \quad u(L, t) = g_2(t).$$

2. Une condition de type *Neumann*, qui revient à exercer sur les bords des forces fixées

$$u'_x(0, t) = g_1(t), \quad u'_x(L, t) = g_2(t).$$

3. Une condition mixte, qui revient à supposer un lien élastique sur les bords

$$u'_x(0, t) + h_1 u(0, t) = g_1(t), \quad u'_x(L, t) + h_2 u(L, t) = g_2(t).$$

Résolution du problème de Dirichlet homogène

La méthode de résolution suit fidèlement celle employée pour l'équation de la chaleur. On recherche des solutions sous forme de fonctions à variables séparées

$$u(x, t) = X(x)T(t),$$

ce qui conduit à deux équations pour les deux fonctions

$$T''(t) = c^2 \lambda T(t) \quad (7.54)$$

$$X''(x) = \lambda X(x) \quad (7.55)$$

Suivant les valeurs de λ , on sera confronté à divers types de solutions. Dans la mesure où $u(x, t)$ représente la position d'un point d'une corde, il est naturel de se limiter à des valeurs réelles de λ . On a donc les trois situations

– Si $\lambda = 0$, les solutions sont de la forme

$$u(x, t) = (Ax + B)(Ct + D).$$

La solution ne sera bornée que si $A = C = 0$, et ne satisfera les conditions au bord que si $B = D = 0$.

Reste donc la seule solution nulle $u(x, t) = 0$.

– Si $\lambda = \beta^2 > 0$, les solutions sont de la forme

$$u(x, t) = (Ae^{c\beta t} + Be^{-c\beta t}) (Ce^{\beta x} + De^{-\beta x}),$$

et là encore, seule la solution nulle est admissible.

– Si $\lambda = -\beta^2 < 0$, les solutions sont de la forme

$$u(x, t) = (Ae^{i\beta t} + Be^{-i\beta t}) (Ce^{i\beta x} + De^{-i\beta x}).$$

Seul le dernier cas de figure convient. En imposant les conditions aux bords, on voit facilement que $D = -C$, de sorte que la solution prend la forme

$$u_n(x, t) = (Ae^{i\beta_n t} + Be^{-i\beta_n t}) \sin(\beta_n x),$$

avec

$$\beta_n = \frac{n\pi}{L}.$$

La solution générale est donc de la forme

$$u(x, t) = \sum_{n=1}^{\infty} (A_n e^{i\beta_n t} + B_n e^{-i\beta_n t}) \sin(\beta_n x) \quad (7.56)$$

Pour conclure, il faut maintenant imposer les conditions initiales, que l'on prend de la forme

$$u(x, 0) = \phi(x), \quad u'_t(x, 0) = \psi(x).$$

On aboutit alors au système

$$\begin{aligned} \phi(x) &= \sum_{n=1}^{\infty} (A_n + B_n) \sin(\beta_n x) \\ \psi(x) &= \sum_{n=1}^{\infty} i\beta_n (A_n - B_n) \sin(\beta_n x) \end{aligned}$$

d'où on déduit, grâce au lemme 7.1, le système

$$\begin{aligned} A_n + B_n &= \frac{2}{L} \int_0^L \phi(x) \sin(\beta_n x) dx \\ A_n - B_n &= \frac{2}{icL} \int_0^L \psi(x) \sin(\beta_n x) dx, \end{aligned}$$

de sorte que nous avons donc montré le résultat suivant :

THÉORÈME 7.4 La solution $u : (x, t) \in [0, L] \times \mathbb{R}^+ \rightarrow u(x, t) \in \mathbb{R}$ du problème formé de l'équation des ondes unidimensionnelle (7.53), complétée par les conditions aux bord de Dirichlet $u(0, t) = u(L, t) = 0$ et des conditions initiales $u(x, 0) = \Phi(x)$ et $\partial_x u(0, t) = \psi(t)$ est donnée par la série

$$u(x, t) = \sum_{n=1}^{\infty} \left(A_n e^{in\pi ct/L} + B_n e^{-in\pi ct/L} \right) \sin\left(\frac{n\pi x}{L}\right),$$

où les coefficients A_n et B_n s'expriment en fonction des conditions initiales sous la forme

$$\begin{aligned} A_n &= \frac{1}{L} \int_0^L \left(\phi(x) + \frac{1}{ic} \psi(x) \right) \sin\left(\frac{n\pi x}{L}\right) dx \\ B_n &= \frac{1}{L} \int_0^L \left(\phi(x) - \frac{1}{ic} \psi(x) \right) \sin\left(\frac{n\pi x}{L}\right) dx. \end{aligned}$$

Bien entendu, si d'autres types de conditions aux bords avaient été requises, nous aurions obtenu un résultat similaire, la famille des fonctions $\sin(\beta_n x)$ étant alors remplacée par la base de Sturm-Liouville associée au problème correspondant.

7.4.4 Cas bi-dimensionnel : la membrane d'un tambour

Passons pour finir à un exemple en dimension supérieure, qui nous donnera l'occasion de revoir nos vieilles amies les fonctions de Bessel.

On se propose ici d'étudier les vibrations d'une membrane de tambour, décrites par une équation des ondes bidimensionnelle. On suppose que le bord de la membrane est un cercle de rayon 1, et on note $u(\underline{x}, t) = u(r, \theta, t)$ la hauteur du point de coordonnées polaires $\underline{x} = (r, \theta)$ de la membrane à l'instant $t \in \mathbb{R}^+$. En écrivant le Laplacien en coordonnées polaires, l'évolution de u est décrite par l'équation des ondes

$$\frac{\partial^2 u}{\partial t^2}(r, \theta, t) = c^2 \left(\frac{\partial^2 u}{\partial r^2}(r, \theta, t) + \frac{1}{r} \frac{\partial u}{\partial r}(r, \theta, t) + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2}(r, \theta, t) \right). \quad (7.57)$$

Cette équation est complétée par la condition au bord

$$u(1, \theta, t) = 0,$$

ainsi que les conditions initiales

$$u(r, \theta, 0) = g_1(r, \theta), \quad u'_t(r, \theta, 0) = g_2(r, \theta).$$

On recherche des solutions à variables séparées

$$u(r, \theta, t) = U(r, \theta)T(t),$$

ce qui conduit au système

$$\Delta U(r, \theta) + \lambda^2 U(r, \theta) = 0 \quad (7.58)$$

$$T''(t) + c^2 \lambda^2 T(t) = 0 \quad (7.59)$$

où Δ est le Laplacien spatial, exprimé en coordonnées polaires.

L'équation temporelle est facile à résoudre, et conduit à des solutions de la forme

$$T(t) = Ae^{i\lambda ct} + Be^{-i\lambda ct} .$$

Considérons maintenant l'équation d'Helmholtz (7.58), complétée par la condition au bord

$$U(1, \theta) = 0 .$$

On recherche encore une fois des solutions sous forme de fonction à variables séparées

$$U(r, \theta) = R(r)\Theta(\theta) ,$$

ce qui conduit aux deux équations

$$r^2 R''(r) + rR'(r) + (\lambda^2 r^2 - n^2)R(r) = 0 \quad (7.60)$$

$$\Theta''(\theta) + n^2 \Theta(\theta) = 0 \quad (7.61)$$

où on a pris la constante de séparation égale à n^2 (où n est a priori un nombre complexe quelconque). Il s'agit d'un problème que nous avons déjà rencontré au chapitre 6. Encore une fois, l'équation angulaire est facile à résoudre, et conduit à des solutions

$$\Theta_n(\theta) = Ae^{in\theta} + Be^{-in\theta} ,$$

et la périodicité et la continuité de Θ_n et Θ'_n imposent $n \in \mathbb{Z}^+$.

De plus on sait que la solution de l'équation radiale est donnée par les fonctions de Bessel de première espèce et deuxième espèce $J_n(\lambda r)$ et $Y_n(\lambda r)$. Ces dernières étant non bornées, on a donc des solutions de la forme

$$R(r) = J_n(\lambda r) .$$

On doit maintenant imposer la condition au bord $U(1, \theta) = 0$. Ceci impose

$$J_n(\lambda) = 0 ,$$

c'est à dire que la constante de séparation λ doit être un zéro α_{nm} de la fonction de Bessel J_n . Par conséquent

$$U_{nm}(r, \theta) = \left(A_{nm}e^{in\theta} + B_{nm}e^{-in\theta} \right) J_n(\alpha_{nm}r) .$$

Ainsi, la solution générale de notre problème est de la forme

$$u(r, \theta, t) = \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} J_n(\alpha_{nm}r) \left(A_{nm}e^{in\theta} + B_{nm}e^{-in\theta} \right) (C_{nm} \sin(\alpha_{nm}ct) + D_{nm} \cos(\alpha_{nm}ct)) \quad (7.62)$$

La dernière étape consiste à calculer les constantes à partir des conditions initiales. Le cas général est assez difficile à traiter, on se limitera ici au cas particulier où la vitesse initiale $u'_t(r, \theta, 0)$ est nulle, et où la condition initiale $u(r, \theta, 0)$ dépend uniquement de la variable radiale r . En d'autres termes, on considère le cas

$$u(r, \theta, 0) = f(r) , \quad u'_t(r, \theta, 0) = 0 .$$

On a alors

$$u(r, \theta, 0) = f(r) = \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} J_n(\alpha_{nm}r) \left(A_{nm}e^{in\theta} + B_{nm}e^{-in\theta} \right) D_{nm} ,$$

et on en déduit par exemple pour $k \in \mathbb{N}$,

$$\frac{1}{2\pi} \int_0^{2\pi} u(r, \theta, 0) e^{-ik\theta} d\theta = f(r) \delta_{k,0} = \sum_{m=1}^{\infty} \sum_n J_n(\alpha_{nm}r) A_{nm} \delta_{n,k} = \sum_{m=1}^{\infty} J_k(\alpha_{km}r) A_{km} D_{km} ,$$

et un résultat similaire pour $k \leq 0$, faisant intervenir les coefficients B_{nm} .

Ainsi, pour tout $k \neq 0$ on a

$$\sum_{m=1}^{\infty} J_k(\alpha_{km}r) A_{km} D_{km} = \sum_{m=1}^{\infty} J_k(\alpha_{km}r) B_{km} D_{km} , \quad k \neq 0$$

qui implique $A_{km} D_{km} = B_{km} D_{km} = 0$, et pour $k = 0$,

$$\sum_{m=1}^{\infty} J_0(\alpha_{0m}r) (A_{0m} + B_{0m}) D_{0m} = f(r) ,$$

de sorte que les coefficients $d_m = (A_{0m} + B_{0m}) D_{0m}$ sont les coefficients du développement de f sur la base de Bessel \mathcal{J}_{0m} .

En raisonnant similairement, on montre que les coefficients $A_{mn} C_{mn}$ et $B_{mn} C_{mn}$ sont tous nuls, comme conséquence de la seconde condition initiale.

Ainsi, en posant $k_m = \alpha_{0m}$

$$u(r, \theta, t) = \sum_{m=1}^{\infty} d_m J_0(k_m r) \cos(k_m c t) .$$

Il reste donc à déterminer les constantes d_m . On écrit pour cela

$$f(r) = \sum_{m=1}^{\infty} d_m J_0(k_m r) ,$$

et on utilise maintenant un résultat obtenu au chapitre précédent, à savoir l'orthogonalité des fonctions de Bessel

$$\int_0^1 r J_n(\alpha_{nm}r) J_n(\alpha_{nm'}r) dr = 0 \quad \text{si } m' \neq m .$$

Dans notre cas, on obtient, après calcul

$$\int_0^1 r J_n(k_m r) J_n(k_{m'} r) dr = \frac{1}{2} \delta_{mm'} J_1^2(k_m) ,$$

d'où on déduit les constantes cherchées

$$d_m = \frac{2}{J_1^2(k_m)} \int_0^1 r f(r) J_0(k_m r) dr .$$

REMARQUE 7.3 Notons que nous avons travaillé ici dans un cas particulier, tant pour la condition au bord (indépendante de θ), que pour les conditions initiales. Dans le cas général, on obtient un développement sur les fonctions de Bessel (7.62), et on peut vérifier que les constantes peuvent être déterminées par les conditions initiales et les conditions aux bords.

Quatrième partie

Distributions, fonctions de Green

8 Distributions et fonctions de Green

La théorie des distributions, développée en mathématiques par I.M. Gelfand, puis par L. Schwartz, trouve ses origines dans certains problèmes de physiques. Comme motivation on cite généralement le problème du calcul du potentiel généré par une “distribution” de charge. Lorsque ces charges sont des charges ponctuelles, situées en des points $x_1, \dots, x_N \in \mathbb{R}^3$, le potentiel créé en un point $x \in \mathbb{R}^3$ s’écrit

$$V(x) = \frac{1}{4\pi\epsilon_0} \sum_{i=1}^n \frac{q_i}{|x - x_i|},$$

où ϵ_0 est la permittivité du vide, q_i représente la charge au point x_i , et on note $|x|$ la norme Euclidienne du vecteur $x \in \mathbb{R}^3$. Dans le cas d’une distribution volumique, de densité ρ à support dans un domaine Ω , on écrit alors

$$V(x) = \frac{1}{4\pi\epsilon_0} \int_{\Omega} \frac{\rho(x')}{|x - x'|} dx'.$$

Dans le cas le plus général, on est amené à considérer simultanément ces différents cas, ce qui complique le formalisme. La théorie des distributions permet de donner une approche unifiée pour ces situations.

Le but de la théorie des distributions est de permettre d’étendre des opérations “classiques” sur les fonctions, telles que dérivation, intégration, convolution, transformations de Fourier, à des cadres dans lesquels ces opérations ne sont *a priori* pas définies.

L’exemple le plus simple est celui de la fonction “saut”, ou fonction de Heaviside $t \rightarrow \Theta(t)$. Cette fonction, bien que très simple, est difficile à appréhender dans un cadre de calcul intégral, dans la mesure où elle ne tend pas vers zéro à l’infini. Elle est également difficile à manipuler dans un cadre de calcul différentiel, car elle est discontinue en $t = 0$.

Cependant, l’idée intuitive que l’on peut s’en faire est que sa dérivée, si elle avait un sens, devrait être nulle pour tout $t \neq 0$, et “infinie” en $t = 0$. Dirac a donc proposé de lui associer une “pseudo-fonction” dérivée, notée δ , définie par un passage à la limite

$$\delta(x) = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} g\left(\frac{t}{\epsilon}\right),$$

qui possède bien le comportement désiré. Ici, g est une fonction régulière (par exemple une Gaussienne), d’intégrale égale à 1, possédant un maximum en 0.

On va voir que cette limite n’a pas de sens en tant que fonction usuelle ; il s’agit d’une limite d’une suite de fonctions continues, qui ne définit néanmoins pas une fonction continue (puisque’elle diverge en 0).

La théorie des distributions permet de donner un sens à ce type de limites, via l’introduction d’un concept assez élaboré, le concept de **dual** d’un espace fonctionnel.

8.1 Distributions : définition, propriétés et exemples

8.1.1 Fonctions test et dualité

La notion essentielle pour la théorie des distributions est la notion de dualité.

DÉFINITION 8.1 Etant donné un espace fonctionnel E , on définit l'**espace dual** de E , noté E' , comme l'ensemble des **formes linéaires continues** sur E , c'est à dire l'ensemble des applications linéaires

$$T : E \rightarrow \mathbb{C}, \quad f \in E \rightarrow \langle T, f \rangle \in \mathbb{C}$$

satisfaisant les propriétés suivantes :

1. *Linéarité* : pour toutes $f, g \in E$ et $\lambda, \mu \in \mathbb{C}$,

$$\langle T, \lambda f + \mu g \rangle = \lambda \langle T, f \rangle + \mu \langle T, g \rangle, \quad (8.1)$$

2. *Continuité* :

$$f_n \rightarrow f \text{ dans } E \implies \langle T, f_n \rangle \rightarrow \langle T, f \rangle \text{ dans } \mathbb{C} \quad (8.2)$$

Un espace des distributions sera le dual, dans un sens que nous allons voir, d'un espace vectoriel de fonctions bien choisi, que l'on appellera espace de **fonctions test**. On verra dans ce qui suit différents espaces de fonctions test.

8.1.2 Distributions sur \mathbb{R}

Les distributions sont donc définies par dualité, en considérant pour espace fonctionnel E l'espace des fonctions C^∞ à support compact.

$$\mathcal{D}(\mathbb{R}) = \{f \in C^\infty(\mathbb{R}), \text{ et } \exists K \subset \mathbb{R}, \text{ compact, tel que } f(t) = 0 \forall t \notin K\}. \quad (8.3)$$

Pour construire le dual $\mathcal{D}'(\mathbb{R})$ de $\mathcal{D}(\mathbb{R})$, il faut définir ce que l'on entend par convergence au sens de $\mathcal{D}(\mathbb{R})$. On dira qu'une suite de fonctions $f_n \in \mathcal{D}(\mathbb{R})$ converge vers $f \in \mathcal{D}(\mathbb{R})$ si

- il existe un compact K tel que le support de f_n soit contenu dans K pour tout n .
- Pour tout $k \in \mathbb{Z}^+$, la suite des dérivées $f_n^{(k)}$ converge uniformément vers $f^{(k)}$.

DÉFINITION 8.2 Une distribution sur \mathbb{R} est une forme linéaire continue

$$T : \mathcal{D}(\mathbb{R}) \longrightarrow \mathbb{C}.$$

On note $\mathcal{D}'(\mathbb{R})$ l'espace vectoriel des distributions sur \mathbb{R} .

EXEMPLE 8.1 L'exemple le plus simple est celui de la distribution nulle :

$$0 : f \in \mathcal{D}(\mathbb{R}) \rightarrow 0.$$

Les propriétés de linéarité et continuité sont vérifiées facilement.

EXEMPLE 8.2 La **distribution de Dirac** δ_a , où a est un réel quelconque, est définie de la façon suivante : pour toute fonction de test $f \in \mathcal{D}(\mathbb{R})$,

$$\langle \delta_a, f \rangle = f(a). \quad (8.4)$$

EXEMPLE 8.3 Le **peigne de Dirac**, noté \sqcup_h (lettre de l'alphabet cyrillique), associe à toute fonction test la somme de la série de ses échantillons (ou valeurs ponctuelles) régulièrement espacés.

$$\langle \sqcup_h, f \rangle = \sum_{k=-\infty}^{\infty} f(kh). \quad (8.5)$$

Comme f est supposée à support compact, cette série est automatiquement convergente (car finie). Le peigne de Dirac s'exprime comme

$$\sqcup_h = \sum_{k=-\infty}^{\infty} \delta_{kh}.$$

En quoi les distributions sont elles des généralisations des fonctions ? on va voir que l'espace $\mathcal{D}'(\mathbb{R})$ contient l'espace $L^1_{loc}(\mathbb{R})$... dans un certain sens, qu'on va maintenant préciser. Rappelons tout d'abord qu'en notant χ_I l'indicatrice d'un domaine $I \subset \mathbb{R}$:

$$\chi_I(t) = \begin{cases} 1 & \text{si } t \in I \\ 0 & \text{sinon,} \end{cases}$$

on définit

$$L^1_{loc}(\mathbb{R}) = \left\{ f : \mathbb{R} \rightarrow \mathbb{C}, \quad f\chi_I \in L^1(\mathbb{R}) \forall I \subset \mathbb{R} \text{ borné} \right\}. \quad (8.6)$$

A toute fonction de $L^1_{loc}(\mathbb{R})$ on peut associer une distribution de la façon suivante. Soit $\varphi \in \mathcal{D}(\mathbb{R})$. Soit T_φ l'application définie par

$$\langle T_\varphi, f \rangle = \int \varphi(t)f(t) dt, \quad \forall f \in \mathcal{D}(\mathbb{R}). \quad (8.7)$$

Comme $\varphi, f \in \mathcal{D}(\mathbb{R})$, cette intégrale est absolument convergente. De plus, il est immédiat de vérifier que T_φ est une forme linéaire sur $\mathcal{D}'(\mathbb{R})$. Le point le plus délicat à vérifier est la continuité de T_φ . Ainsi, à toute fonction localement intégrable f , on peut associer une distribution, appelée **distribution régulière** $T_f \in \mathcal{D}'$.

REMARQUE 8.1 En pratique, on note souvent (abusivement) f la distribution régulière T_f associée à f (et on dit donc $L^1_{loc}(\mathbb{R}) \subset \mathcal{D}'(\mathbb{R})$, par abus de notation, au lieu de $T(L^1_{loc}(\mathbb{R})) \subset \mathcal{D}'(\mathbb{R})$). Ainsi, on note parfois une distribution $T \in \mathcal{D}'(\mathbb{R})$ sous la même forme qu'une fonction $t \rightarrow T(t)$, et on écrit alors

$$\langle T, f \rangle = \int_{-\infty}^{\infty} T(t)f(t) dt.$$

Par exemple, on associe à la distribution de Dirac une "fonction généralisée", notée $t \rightarrow \delta(t)$, qui a la propriété que pour toute $f \in \mathcal{D}(\mathbb{R})$

$$\langle \delta, f \rangle = \int_{-\infty}^{\infty} f(t)\delta(t) dt = f(0).$$

Avec ces notations, on peut donc noter

$$\delta_a(t) = \delta(t - a),$$

de sorte que l'on écrit

$$\langle \delta_a, f \rangle = \int_{-\infty}^{\infty} \delta(t - a)f(t) dt = f(a).$$

Dans ce qui suit, on essaie au maximum d'éviter de recourir à ce type de notation, qui est cependant parfois plus pratique pour effectuer des calculs.

8.1.3 Propriétés élémentaires

Produit d'une fonction par une distribution

Soit $u \in C^\infty$. Le produit de u par une fonction $f \in \mathcal{D}(\mathbb{R})$ est toujours une fonction $uf \in \mathcal{D}(\mathbb{R})$. On a donc

PROPOSITION 8.1 *Etant donnée une fonction $u \in C^\infty(\mathbb{R})$ et une distribution $\varphi \in \mathcal{D}'(\mathbb{R})$, la forme linéaire*

$$u\varphi : f \in \mathcal{D}(\mathbb{R}) \rightarrow \langle u\varphi, f \rangle = \langle \varphi, uf \rangle$$

définit une distribution.

EXEMPLE 8.4 Etant donnée la distribution de Dirac δ_a , et $u \in C^\infty(\mathbb{R})$, on a

$$\langle u\delta_a, f \rangle = \langle \delta_a, uf \rangle = u(a)f(a),$$

de sorte que l'on a l'égalité suivante, au sens des distributions

$$u\delta_a = u(a)\delta_a. \quad (8.8)$$

Notons en particulier que l'on a nécessairement $t\delta = 0$. Inversement, on a le résultat fondamental suivant

PROPOSITION 8.2 *Soit T une distribution telle que*

$$tT = 0$$

(c'est à dire le produit de T par la fonction linéaire $t \rightarrow t$ est la distribution nulle). Alors T est multiple de la distribution de Dirac, i.e.

$$T = \alpha\delta, \quad \alpha \in \mathbb{C}.$$

Plus généralement, en appliquant ce résultat plusieurs fois successivement, on aboutit à

PROPOSITION 8.3 *Soit T une distribution telle que*

$$t^n T = 0$$

Alors T est combinaison linéaire de la distribution de Dirac et de ses $n - 1$ premières dérivées

$$T = \sum_{k=0}^{n-1} \alpha_k \delta^{(k)}, \quad \alpha_k \in \mathbb{C}.$$

REMARQUE 8.2 Alors que le produit d'une distribution $T \in \mathcal{D}'(\mathbb{R})$ par une fonction C^∞ est bien définie, le produit d'une distribution par une autre distribution n'a généralement pas de sens, c'est à dire ne définit pas une nouvelle distribution. Par exemple, le produit $\delta\delta$ n'a pas de sens.

Distributions paires et impaires

On considère l'opérateur parité, défini par

$$[\Pi f](t) = f(-t).$$

L'action de cet opérateur sur des distributions se définit de la façon suivante. Étant donnée $T \in \mathcal{D}'(\mathbb{R})$, on note $T^- \in \mathcal{D}'(\mathbb{R})$, la distribution définie par

$$\langle T^- f \rangle = \langle T, \Pi f \rangle, \quad \forall f \in \mathcal{D}(\mathbb{R}). \quad (8.9)$$

On rappelle qu'une fonction f est paire si $\Pi f = f$, et impaire si $\Pi f = -f$.

DÉFINITION 8.3 Une distribution $T \in \mathcal{D}'(\mathbb{R})$ est paire si $T^- = T$, et **impaire** si $T^- = -T$.

Il est clair que pour toute fonction paire, la distribution régulière associée est paire :

$$\langle T_\varphi^-, f \rangle = \int_{-\infty}^{\infty} \varphi(-t)f(t) dt = \int_{-\infty}^{\infty} \varphi(t)f(-t) dt = \int_{-\infty}^{\infty} \varphi(t)f(t) dt = \langle T_\varphi^-, f \rangle.$$

De même la distribution associée à une fonction impaire est impaire.

Par exemple, on vérifie que la distribution de Dirac δ est paire, alors que sa dérivée δ' est impaire.

Distributions périodiques

Soit τ_a l'opérateur de translation par $a \in \mathbb{R}$, défini par

$$[\tau_a]f(t) = f(t - a).$$

Une fonction f est périodique de période a si $\tau_a f = f$. On définit la translatée $\tau_a T$ d'une distribution $T \in \mathcal{D}'(\mathbb{R})$ par

$$\langle \tau_a T, f \rangle = \langle T, \tau_{-a} f \rangle, \quad \forall f \in \mathcal{D}(\mathbb{R}).$$

DÉFINITION 8.4 La distribution $T \in \mathcal{D}'(\mathbb{R})$ est périodique de période $a \in \mathbb{R}$ si $\tau_a T = T$.

On vérifie facilement que le peigne de Dirac \sqcup_h est périodique de période h .

8.2 Distributions sur \mathbb{R}^n

On construit de même les distributions sur \mathbb{R}^n . Il suffit d'introduire les espaces $\mathcal{E}(\mathbb{R}^n)$ des fonctions dont toutes les dérivées sont continues, et

$$\mathcal{D}(\mathbb{R}^n) = \{f \in \mathcal{E}(\mathbb{R}^n), \text{supp}(f) \text{ borné}\}.$$

DÉFINITION 8.5 Une distribution sur \mathbb{R}^n est une forme linéaire continue

$$T : \mathcal{D}(\mathbb{R}^n) \longrightarrow \mathbb{C}.$$

On note $\mathcal{D}'(\mathbb{R}^n)$ l'espace vectoriel des distributions sur \mathbb{R}^n .

Comme en dimension 1, les exemples ne manquent pas. On définit par exemple les distributions régulières de la même façon qu'en dimension 1 : en définissant l'espace des fonction **localement intégrables**

$$L_{loc}^1(\mathbb{R}^n) = \left\{ f : \mathbb{R}^n \rightarrow \mathbb{C}, \quad f \chi_\Omega \in L^1(\mathbb{R}) \quad \forall \Omega \subset \mathbb{R}^n \text{ compact} \right\}. \quad (8.10)$$

on associe à toute $\varphi \in L_{loc}^1(\mathbb{R}^n)$ la **distribution régulière** T_φ , définie par

$$\langle T_\varphi, f \rangle = \int_{\mathbb{R}^n} \varphi(x) f(x) dx. \quad (8.11)$$

Avec les mêmes précautions qu'en dimension 1, on a parfois tendance à identifier une fonction localement intégrable φ avec la distribution régulière associée T_φ ... et à associer à des distributions non régulières $T \in \mathcal{D}'(\mathbb{R}^n)$ une "pseudo-fonction", ou fonction généralisée, notée $x \rightarrow T(x)$, de sorte que l'on écrit

$$\langle T, f \rangle = \int_{\mathbb{R}^n} T(x) f(x) dx.$$

Là encore, il convient d'être très précautionneux avec ce type de notation.

EXEMPLE 8.5 On considère la fonction de Heaviside unidimensionnelle Θ , et on construit la fonction $u : \mathbb{R}^3 \rightarrow \mathbb{C}$ suivante

$$u(x, y, z) = \Theta(x).$$

Il s'agit d'une variante 3D de la fonction de Heaviside. Cette fonction est localement intégrable, et permet donc de définir une distribution régulière.

Soit maintenant $v \in L_{loc}^1(\mathbb{R}^3)$, définie par

$$v(x, y, z) = \Theta(x)\Theta(y)\Theta(z).$$

La fonction v (que l'on note aussi $\Theta \otimes \Theta \otimes \Theta$) est l'indicatrice d'un quadrant de l'espace tridimensionnel, et définit une distribution régulière.

EXEMPLE 8.6 Toujours dans \mathbb{R}^3 , la distribution de Dirac δ_a , où $a \in \mathbb{R}^3$ est définie de façon similaire à la distribution de Dirac unidimensionnelle : pour $f \in \mathcal{D}(\mathbb{R}^3)$,

$$\langle \delta_a, f \rangle = f(a). \quad (8.12)$$

Cependant, il est aussi possible de définir des variantes à la distribution de Dirac. Par exemple, étant donnée une surface $S \subset \mathbb{R}^3$, la distribution de Dirac associée δ_S est définie par

$$\langle \delta_S, f \rangle = \int_S f(s) ds, \quad (8.13)$$

c'est à dire associe à toute fonction test son intégrale sur la surface. Ce type d'exemple est courant en physique, par exemple en électrostatique, où on est couramment amené à considérer des distributions de charge dite **superficielles**.

On définit de façon similaire une distribution de Dirac δ_C associée à une courbe $C \subset \mathbb{R}^3$:

$$\langle \delta_C, f \rangle = \int_C f(u) du. \quad (8.14)$$

Plus généralement, on peut utiliser des distributions unidimensionnelles pour construire des distributions multidimensionnelles, par la méthode du **produit tensoriel**. Commençons par le cas des distributions régulières.

Dans le cas bidimensionnel, soient $u, v \in L_{loc}^1(\mathbb{R})$. On définit $u \otimes v \in \mathcal{D}(\mathbb{R}^2)$ par

$$(u \otimes v)(x, y) = u(x)v(y).$$

Soit $T_{u \otimes v}$ la distribution régulière associée. On a alors, pour $f \in \mathcal{D}(\mathbb{R}^2)$

$$\begin{aligned} \langle T_{u \otimes v}, f \rangle &= \int_{\mathbb{R}^2} u(x)v(y)f(x,y) dx dy \\ &= \int_{-\infty}^{\infty} u(x) \left(\int_{-\infty}^{\infty} v(y)f(x,y) dy \right) dx \\ &= \int_{-\infty}^{\infty} v(x) \left(\int_{-\infty}^{\infty} u(y)f(x,y) dx \right) dy. \end{aligned}$$

De façon générale, étant données $T, S \in \mathcal{D}'(\mathbb{R})$, T agissant sur la variable x et S sur la variable y , on définira $U = T \otimes S$ de la façon suivante : pour toute $f \in \mathcal{D}'(\mathbb{R}^2)$,

$$\langle T \otimes S, \varphi \rangle = \langle T, \langle S, \varphi \rangle \rangle = \langle T, \langle S, \varphi \rangle \rangle,$$

où par exemple $\langle T, \varphi \rangle$ représente la dualité par rapport à la variable x , et produit une fonction de y , et $\langle S, \langle T, \varphi \rangle \rangle$ représente cette fois la dualité par rapport à la variable y .

8.3 Dérivation des distributions

8.3.1 Le cas unidimensionnel

Les distributions constituant une généralisation des fonctions, la dérivation des fonctions, lorsqu'elle est bien définie (c'est à dire pour les fonctions à dérivée continue), doit être un cas particulier de la dérivation des distributions. C'est pourquoi il est utile de se placer tout d'abord dans le cas des distributions régulières, associées aux fonctions de $C^1(\mathbb{R})$.

Soit $\varphi \in C^1(\mathbb{R})$, et soit T_φ la distribution régulière associée. Soit $T_{\varphi'}$ la distribution régulière associée à φ' . On a alors pour tout $f \in \mathcal{D}(\mathbb{R})$

$$\langle T_{\varphi'}, f \rangle = \int_{-\infty}^{\infty} \varphi'(t)f(t) dt = - \int_{-\infty}^{\infty} \varphi(t)f'(t) dt = - \langle T_\varphi, f \rangle,$$

où le terme tout intégré dans l'intégration par parties est nul car f est à support borné. Ceci suggère de définir la dérivée d'une distribution de la façon suivante :

DÉFINITION 8.6 Soit $T \in \mathcal{D}'(\mathbb{R})$. Sa dérivée T' est définie par

$$\langle T', f \rangle = - \langle T, f' \rangle, \quad \forall f \in \mathcal{D}(\mathbb{R}). \quad (8.15)$$

De même, la dérivée p -ième est définie par

$$\langle T^{(p)}, f \rangle = (-1)^p \langle T, f^{(p)} \rangle, \quad \forall f \in \mathcal{D}(\mathbb{R}). \quad (8.16)$$

On vient de voir que la dérivée de la distribution régulière associée à une fonction $\varphi \in C^1$ coïncide avec la distribution régulière de la dérivée de cette fonction. Qu'en est-il lorsque φ n'est plus régulière ? supposons par exemple que φ soit continûment différentiable, sauf en un point t_0 où elle admet une discontinuité de première espèce (c'est à dire que les limites à droite et à gauche de φ en t_0 existent, mais diffèrent). On notera

$$\sigma_{t_0} = \lim_{t \rightarrow t_0^+} \varphi(t) - \lim_{t \rightarrow t_0^-} \varphi(t).$$

Calculons alors

$$\begin{aligned}
 \langle T'_{\varphi}, f \rangle &= - \int_{-\infty}^{\infty} \varphi(t) f'(t) dt \\
 &= - \int_{-\infty}^{t_0} \varphi(t) f'(t) dt - \int_{t_0}^{\infty} \varphi(t) f'(t) dt \\
 &= -[\varphi(t)f(t)]_{-\infty}^{t_0} - [\varphi(t)f(t)]_{t_0}^{\infty} + \int_{-\infty}^{t_0} \varphi'(t)f(t) dt + \int_{t_0}^{\infty} \varphi'(t)f(t) dt \\
 &= \sigma_{t_0}f(t_0) + \int_{-\infty}^{\infty} \varphi'(t)f'(t) dt \\
 &= \langle T_{\varphi'} + \sigma_{t_0} \delta_{t_0}, f \rangle,
 \end{aligned}$$

de sorte que l'on peut écrire

$$T'_{\varphi} = T_{\varphi'} + \sigma_{t_0} \delta_{t_0}. \quad (8.17)$$

On dit alors que **la dérivée au sens des distributions est égale à la dérivée au sens des fonctions, additionnée d'une distribution de Dirac localisée sur le saut, multipliée par la hauteur du saut**. Plus généralement, on montre le théorème suivant, appelé **règle d'or**

THÉORÈME 8.1 (RÈGLE D'OR) Soit φ une fonction de classe $C^1(\mathbb{R})$, sauf en t_0, t_1, \dots , où elle admet des discontinuités de hauteur

$$\sigma_k = \lim_{t \rightarrow t_k^+} \varphi(t) - \lim_{t \rightarrow t_k^-} \varphi(t).$$

Alors, on a

$$T'_{\varphi} = T_{\varphi'} + \sum_k \sigma_{t_k} \delta_{t_k}. \quad (8.18)$$

EXEMPLE 8.7 On considère la fonction "escalier", ou partie entière

$$t \rightarrow E(t) = t [\text{mod } 1].$$

Cette fonction admet une discontinuité en tous les points entiers. Elle est localement intégrable, et définit donc une distribution régulière. Sa dérivée au sens des distributions vaut

$$E' = \sum_{k=-\infty}^{\infty} \delta_k = \sqcup.$$

Similairement, la fonction "dents de scie" $t - E(t)$ admet comme dérivée au sens des distributions la fonction $1 + \sqcup$.

Dans d'autres situations, la dérivée d'une distribution régulière peut prendre une forme plus complexe. C'est le cas de la valeur principale, que nous allons étudier plus en détails un peu plus bas. Auparavant, notons la propriété importante suivante, qui généralise au cas des distributions une propriété utile dans le cas des fonctions.

PROPOSITION 8.4 Soit $T \in \mathcal{D}'(\mathbb{R})$ et soit $u \in C^\infty(\mathbb{R})$. Alors on sait que $uT \in \mathcal{D}'(\mathbb{R})$ est aussi une distribution, et on a alors

$$(uT)' = u'T + uT'. \quad (8.19)$$

Preuve : pour toute $f \in \mathcal{D}(\mathbb{R})$, on a

$$\langle (uT)', f \rangle = -\langle uT, f' \rangle = -\langle T, uf' \rangle = -\langle T, (uf)' - u'f \rangle = \langle u'T + uT', f \rangle,$$

ce qui démontre la proposition. ♠

8.3.2 Un exemple : la valeur principale de $1/t$

Considérons la fonction $\varphi(t) = \ln|t|$. Cette fonction est localement intégrable, et admet donc une distribution régulière associée $T_\varphi \in \mathcal{D}'(\mathbb{R})$. Cependant, sa dérivée usuelle $\varphi'(t) = 1/t$ n'est pas localement intégrable, et ne définit donc pas de distribution régulière. On peut toutefois calculer la dérivée de T_φ au sens des distributions.

Calculons, pour $f \in \mathcal{D}(\mathbb{R})$

$$\langle T'_{\varphi}, f \rangle = -\int_{-\infty}^{\infty} \varphi(t) f'(t) dt = -\int_{-\infty}^{\infty} \ln|t| f'(t) dt.$$

Le point $t = 0$ nécessite un traitement particulier, on écrit donc

$$\langle T'_{\varphi}, f \rangle = -\lim_{\epsilon \rightarrow 0} \int_{|t| \geq \epsilon} \ln|t| f'(t) dt,$$

et on calcule, par exemple

$$\begin{aligned} \int_{\epsilon}^{\infty} \ln|t| f'(t) dt &= [\ln(t)f(t)]_{\epsilon}^{\infty} - \int_{\epsilon}^{\infty} \frac{f(t)}{t} dt \\ &= \ln(\epsilon)f(\epsilon) - \int_{\epsilon}^{\infty} \frac{f(t)}{t} dt. \end{aligned}$$

De même,

$$\begin{aligned} \int_{-\infty}^{-\epsilon} \ln|t| f'(t) dt &= \int_{\epsilon}^{\infty} \ln(t) f'(-t) dt \\ &= -[\ln(t)f(-t)]_{\epsilon}^{\infty} + \int_{\epsilon}^{\infty} \frac{f(-t)}{t} dt, \\ &= -\ln(\epsilon)f(\epsilon) - \int_{-\infty}^{-\epsilon} \frac{f(t)}{t} dt, \end{aligned}$$

d'où on déduit

$$\langle T'_{\varphi}, f \rangle = \lim_{\epsilon \rightarrow 0} \int_{|t| \geq \epsilon} \frac{f(t)}{t} dt. \quad (8.20)$$

La distribution ainsi définie (dûe à Cauchy) est appelée **valeur principale de $1/t$** , et notée

$$T'_{\ln|t|} = \text{v.p.} \left(\frac{1}{t} \right).$$

Il s'agit d'une distribution non régulière (ce qui montre au passage que la dérivée d'une distribution régulière n'est pas nécessairement régulière).

Un calcul similaire permet de montrer que la dérivée de la valeur principale est une autre distribution non régulière, appelée **partie finie de $1/t^2$** (dûe au mathématicien français Hadamard), donnée par

$$\begin{aligned} \frac{d}{dt} \text{v.p.} \left(\frac{1}{t} \right) &= -\text{p.f.} \left(\frac{1}{t^2} \right), \\ \left\langle \text{p.f.} \left(\frac{1}{t^2} \right), f \right\rangle &= \lim_{\epsilon \rightarrow 0} \int_{|t| \geq \epsilon} \frac{f(t) - f(0)}{t^2} dt. \end{aligned}$$

On définit de façon similaire les parties finies de t^{-n} pour tout n , par

$$\text{p.f.} \left(\frac{1}{t^n} \right) = \frac{(-1)^{n-1}}{(n-1)!} \frac{d^{n-1}}{dt^{n-1}} \text{v.p.} \left(\frac{1}{t} \right)$$

On peut montrer facilement que

$$t \text{ v.p.} \left(\frac{1}{t} \right) = 1$$

(1 représentant ici la distribution régulière associée à la fonction constante égale à 1). Inversement, on a aussi

PROPOSITION 8.5 Soit T une distribution solution de

$$tT = 1$$

(le produit de T par la fonction linéaire $t \rightarrow t$ est la distribution régulière associée à la fonction constante 1). Alors T est égale à la valeur principale de $1/t$, plus un multiple de la distribution de Dirac

$$T = \text{v.p.} \left(\frac{1}{t} \right) + \alpha \delta, \quad \alpha \in \mathbb{C}.$$

Preuve : La preuve est une conséquence de la Proposition 8.2. Notons $S = T - \text{v.p.}(1/t)$. On a alors $tS = 0$, et il suffit d'appliquer la Proposition 8.2 pour conclure. ♠

On montre également le résultat suivant

PROPOSITION 8.6 Soit $T \in \mathcal{D}'$, solution de

$$(t^2 - a^2)T = 1.$$

Alors T est de la forme

$$T = \text{v.p.} \left(\frac{1}{t^2 - a^2} \right) + \alpha \delta_a + \beta \delta_{-a},$$

où $\alpha, \beta \in \mathbb{C}$, et où on a noté

$$\text{v.p.} \left(\frac{1}{t^2 - a^2} \right) = \frac{1}{2a} \left(\text{v.p.} \left(\frac{1}{t - a} \right) - \text{v.p.} \left(\frac{1}{t + a} \right) \right).$$

8.3.3 Le cas multidimensionnel

Dans le cas multidimensionnel, les opérations de dérivation partielle des distributions sont définies similairement au cas unidimensionnel, comme généralisations de la dérivation des distributions régulières.

DÉFINITION 8.7 Soit $T \in \mathcal{D}'(\mathbb{R}^n)$.

1. La dérivée partielle k -ième de T par rapport à la variable x_ℓ est définie par

$$\left\langle \frac{\partial^k T}{\partial x_\ell^k}, f \right\rangle = (-1)^k \left\langle T, \frac{\partial^k f}{\partial x_\ell^k} \right\rangle, \quad \forall f \in \mathcal{D}(\mathbb{R}^n). \quad (8.21)$$

2. Soit $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$, où les α_k sont des entiers non-négatifs, un multi-indice. On note

$$D^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}}$$

l'opérateur de dérivation partielle, où on a noté

$$|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n,$$

appelé **degré** de D^α . Alors $D^\alpha T$ est définie par

$$\langle D^\alpha T, f \rangle = (-1)^{|\alpha|} \langle T, D^\alpha f \rangle, \quad \forall f \in \mathcal{D}(\mathbb{R}^n). \quad (8.22)$$

Comme on l'a signalé, la dérivation des distributions régulières associées aux fonctions de classe $C^1(\mathbb{R}^n)$ coïncide avec la notion usuelle : soit $\varphi \in C^1(\mathbb{R}^n)$. On voit facilement, par intégrations par parties, que dans ce cas

$$D^\alpha T_\varphi = T_{D^\alpha \varphi}. \quad (8.23)$$

Le cas des fonctions présentant des discontinuités est quant à lui plus complexe, comme on va le voir sur l'exemple suivant.

EXEMPLE 8.8 Soit $u = \Theta \otimes \Theta \otimes \Theta$ la fonction localement intégrable définie par

$$u(x, y, z) = \Theta(x)\Theta(y)\Theta(z).$$

Calculons par exemple, pour $f \in \mathcal{D}(\mathbb{R}^3)$

$$\begin{aligned} \left\langle \frac{\partial^3 T_u}{\partial x \partial y \partial z}, f \right\rangle &= - \left\langle \Theta \otimes \Theta \otimes \Theta, \frac{\partial^3 f}{\partial x \partial y \partial z} \right\rangle \\ &= \int_0^\infty \int_0^\infty \int_0^\infty \frac{\partial^3 f}{\partial x \partial y \partial z}(x, y, z) dx dy dz \\ &= f(0, 0, 0), \end{aligned}$$

car f est à support compact. Ainsi, on a

$$T_{\Theta \otimes \Theta \otimes \Theta} = \delta,$$

on retrouve ainsi la distribution de Dirac "ponctuelle" à l'origine.

EXEMPLE 8.9 Si nous considérons maintenant $u(x, y, z) = \Theta(x)$, un calcul similaire montre que pour toute fonction test $f \in \mathcal{D}(\mathbb{R}^3)$,

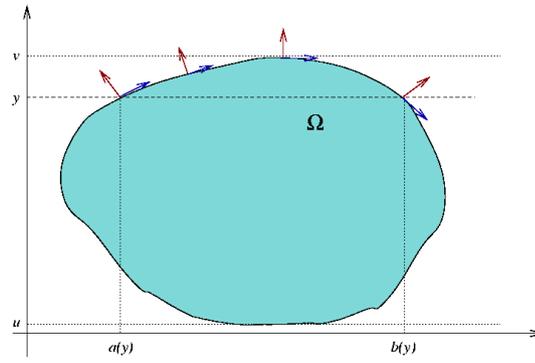
$$\left\langle \frac{\partial T_u}{\partial x}, f \right\rangle = f(0, y, z),$$

de sorte que

$$\frac{\partial T_u}{\partial x} = \delta_{P_{yOz}},$$

la distribution de Dirac associée au plan d'équation $x = 0$, c'est à dire le plan yOz .

On va maintenant s'intéresser de plus près aux cas bi et tridimensionnels.

FIG. 8.1: Un exemple de surface bornée convexe dans \mathbb{R}^2 , à bord différentiable.

8.3.4 Surfaces fermées dans \mathbb{R}^2

Considérons une courbe fermée différentiable \mathcal{C} du plan, définie par la paramétrisation

$$\gamma : s \in [0, 1] \rightarrow \gamma(s) \in \mathbb{R}^2.$$

Cette courbe définit une surface fermée $\Omega \subset \mathbb{R}^2$, de bord différentiable $\partial\Omega = \mathcal{C}$ (voir par exemple la FIG. 8.1, où un exemple de surface fermée convexe du plan est représenté).

La dérivée en tous points de γ permet de définir un vecteur tangent à la courbe en tous points ; on note \underline{t} la tangente à la courbe, normalisée de sorte que $|\underline{t}| = 1$:

$$\underline{t}(s) = K \frac{d\gamma(s)}{ds}, \quad K = \left| \frac{d\gamma(s)}{ds} \right|^{-1}.$$

Dans le plan, la tangente suffit à définir le vecteur normal à la courbe, noté \underline{n} , que l'on choisit orienté vers l'extérieur.

Soit maintenant $\varphi \in L^1_{loc}(\mathbb{R}^2)$, telle que φ soit continûment différentiable partout sauf sur $\partial\Omega$, où elle admet une discontinuité de première espèce. On notera, pour $\omega \in \partial\Omega$

$$\sigma(\omega) = \varphi_{ext}(\omega) - \varphi_{in}(\omega),$$

où $\varphi_{ext}(\omega)$ est la limite de $\varphi(x)$ quand $x \rightarrow \omega$ par l'extérieur, et une définition similaire pour $\varphi_{in}(\omega)$.

Considérons par exemple la dérivation partielle par rapport à la coordonnée x , et étudions ce qui se passe le long d'une droite d'ordonnée y . Lorsque cette droite n'intersecte pas le domaine Ω , la dérivée au sens des distributions est égale à la dérivée au sens des fonctions. Lorsqu'elle la droite d'ordonnée y (représentée en petits tirets sur la FIG. 8.1) intersecte Ω , on notera $a(y)$ et $b(y)$ les abscisses respectives des deux points où elle intersecte $\partial\Omega$ (on a choisi un exemple convexe, il n'y a donc que deux points d'intersection). Le long de cette droite, φ est continûment différentiable par rapport à x partout sauf en $x = a(y)$ et $x = b(y)$.

Calculons pour tout $f \in \mathcal{D}(\mathbb{R}^2)$,

$$\begin{aligned}
\left\langle \frac{\partial T_\varphi}{\partial x}, f \right\rangle &= - \int_{-\infty}^u \int_{-\infty}^{\infty} \frac{\partial f(x, y)}{\partial x} \varphi(x, y) dx - \int_v^{\infty} \int_{-\infty}^{\infty} \frac{\partial f(x, y)}{\partial x} \varphi(x, y) dx \\
&\quad - \int_u^v \left(\int_{-\infty}^{a(x)} \frac{\partial f(x, y)}{\partial x} \varphi(x, y) dx + \int_{a(x)}^{b(x)} \frac{\partial f(x, y)}{\partial x} \varphi(x, y) dx \right. \\
&\quad \quad \left. + \int_{b(x)}^{\infty} \frac{\partial f(x, y)}{\partial x} \varphi(x, y) dx \right) dy \\
&= - \left\langle T_\varphi, \frac{\partial f}{\partial x} \right\rangle + \int_{-\infty}^{\infty} f(a(y), y) [\varphi(a_+(y), y) - \varphi(a_-(y), y)] dy \\
&\quad + \int_{-\infty}^{\infty} f(b(y), y) [\varphi(b_+(y), y) - \varphi(b_-(y), y)] dy \\
&= \left\langle T_{\frac{\partial \varphi}{\partial x}}, f \right\rangle - \int_{\partial \Omega} f(\omega) \sigma_{\partial \Omega}(\omega) n_x(\omega) d\omega,
\end{aligned}$$

où on rappelle que la distribution de Dirac sur une courbe C du plan est définie par

$$\langle \delta_C, f \rangle = \int_C f(\omega) d\omega.$$

Ce calcul ne fonctionne que dans le cas particulier d'une surface convexe. Plus généralement, on montre le théorème suivant

THÉORÈME 8.2 (RÈGLE D'OR EN DEUX DIMENSIONS) Soit $\Omega \subset \mathbb{R}^2$ une surface du plan, de bord différentiable $\partial\Omega$. Soit $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$, continûment différentiable partout sauf sur la courbe $\partial\Omega$ où elle admet une discontinuité de première espèce. Soit $\underline{n}(\omega) = (n_x(\omega), n_y(\omega))^t$ la normale extérieure à $\partial\Omega$ en $\omega \in \partial\Omega$. Alors on a

$$\begin{cases} \frac{\partial T_\varphi}{\partial x} = T_{\frac{\partial \varphi}{\partial x}} + \sigma_{\partial \Omega} \delta_{\partial \Omega} n_x \\ \frac{\partial T_\varphi}{\partial y} = T_{\frac{\partial \varphi}{\partial y}} + \sigma_{\partial \Omega} \delta_{\partial \Omega} n_y. \end{cases} \quad (8.24)$$

Pour une fonction localement intégrable vectorielle $\underline{\psi} = (\psi_1, \psi_2) \in \mathcal{D}(\mathbb{R}^2)^2$, notons $T_{\underline{\psi}}$ la distribution régulière (vectorielle) correspondante.

On peut appliquer le théorème précédent au gradient d'une distribution : étant donnée $\varphi \in L_{loc}^2(\mathbb{R}^2)$, continûment différentiable partout sauf sur $\partial\Omega$ où elle admet une discontinuité de première espèce, on a

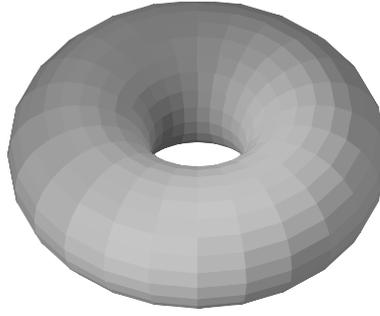
$$\text{Grad } T_\varphi = T_{\text{Grad } \varphi} + \underline{n} \sigma_{\partial \Omega} \delta_{\partial \Omega}. \quad (8.25)$$

Dans cette équation, T_φ est une distribution régulière scalaire, son gradient est une distribution vectorielle, et $T_{\text{Grad } \varphi}$ est également vectorielle.

8.3.5 Domaines fermés dans \mathbb{R}^3

Passons maintenant au cas tridimensionnel. On considère une surface différentiable \mathcal{S} de l'espace, définie par une application différentiable

$$\gamma : (u, v) \in [0, 1] \times [0, 1] \rightarrow \gamma(u, v) \in \mathbb{R}^3.$$

FIG. 8.2: Un exemple de domaine borné de \mathbb{R}^3 , à bord différentiable : le tore

Un exemple de telle surface se trouve en FIG. 8.2. Comme dans le cas bidimensionnel, la différentiabilité de la surface permet de définir en tout point de la surface $\omega \in \mathcal{S}$ non plus une tangente, mais un plan tangent, caractérisé par les deux vecteurs unitaires

$$\begin{cases} \underline{t}(\omega) &= K \frac{\partial \gamma(u,v)}{\partial u}, & K &= \left| \frac{\partial \gamma(u,v)}{\partial u} \right|^{-1} \\ \underline{t}'(\omega) &= K' \frac{\partial \gamma(u,v)}{\partial v}, & K' &= \left| \frac{\partial \gamma(u,v)}{\partial v} \right|^{-1}, \end{cases}$$

et une normale extérieure unitaire à la surface \underline{n} , perpendiculaire à \underline{t} et \underline{t}' .

On définit la distribution de Dirac associée à la surface \mathcal{S} de la façon suivante : étant donnée $f \in \mathcal{D}(\mathbb{R}^3)$,

$$\langle \delta_{\mathcal{S}}, f \rangle = \int_{\mathcal{S}} f(\omega) d\omega.$$

On peut alors montrer le résultat suivant :

THÉORÈME 8.3 (RÈGLE D'OR EN TROIS DIMENSIONS) Soit Ω le domaine de \mathbb{R}^3 , intérieur à \mathcal{S} (donc $\partial\Omega = \mathcal{S}$), et soit φ une fonction localement intégrable dans l'espace \mathbb{R}^3 , de classe C^1 partout sauf sur la surface $\mathcal{S} = \partial\Omega$, où elle admet une discontinuité de première espèce. Soit $\underline{n}(\omega) = (n_x(\omega), n_y(\omega), n_z(\omega))^t$ la normale extérieure à $\partial\Omega$ en $\omega \in \partial\Omega$. Alors on a

$$\begin{cases} \frac{\partial T_\varphi}{\partial x} &= T_{\frac{\partial \varphi}{\partial x}} + \sigma_{\partial\Omega} \delta_{\partial\Omega} n_x \\ \frac{\partial T_\varphi}{\partial y} &= T_{\frac{\partial \varphi}{\partial y}} + \sigma_{\partial\Omega} \delta_{\partial\Omega} n_y \\ \frac{\partial T_\varphi}{\partial z} &= T_{\frac{\partial \varphi}{\partial z}} + \sigma_{\partial\Omega} \delta_{\partial\Omega} n_z. \end{cases} \quad (8.26)$$

Comme application, on peut s'intéresser aux trois opérateurs de dérivation classiques : gradient, divergence et rotationnel. Avec les mêmes notations que ci-dessus, le théorème peut s'écrire

$$\text{Grad } T_\varphi = T_{\text{Grad } \varphi} + \underline{n} \sigma \delta_{\partial\Omega}, \quad (8.27)$$

où on a noté $T_{\text{Grad } \varphi}$ la distribution vectorielle dont les trois composantes sont les dérivées partielles de T_φ .

Soit maintenant une fonction vectorielle $\underline{\varphi} = (\varphi_1, \varphi_2, \varphi_3)$ où $\varphi_i \in L^1_{loc}(\mathbb{R}^3)$ pour $i = 1, 2, 3$ satisfont les conditions ci-dessus. On note $\underline{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ la fonction vectorielle décrivant les sauts des trois composantes de $\underline{\varphi}$ sur le bord $\partial\Omega$.

Alors

$$\begin{cases} \text{Div } T_\varphi &= T_{\text{Div } \underline{\varphi}} + \underline{n} \cdot \underline{\sigma} \delta_{\partial\Omega} \\ \text{Rot } T_\varphi &= T_{\text{Rot } \underline{\varphi}} + \underline{n} \wedge \underline{\sigma} \delta_{\partial\Omega}. \end{cases} \quad (8.28)$$

EXEMPLE 8.10 Les équations de Maxwell au sens des distributions : Dans la même situation que ci-dessus, c'est à dire en présence d'un domaine Ω de bord différentiable $\partial\Omega$, considérons l'équation de Maxwell

$$\text{Rot } \underline{E} = -\frac{\partial \underline{B}}{\partial t}.$$

Supposons l'existence de densités de charge et de courant superficielles sur $\partial\Omega$ seulement (pas de densité volumique). En considérant les champs \underline{B} et \underline{E} comme des distributions, écrivons ces équations au sens des distributions :

$$\text{Rot } T_{\underline{E}} = -\frac{\partial}{\partial t} T_{\underline{B}},$$

ce qui en appliquant la règle d'or conduit à

$$T_{\text{Rot } (\underline{E})} + \underline{n} \wedge \underline{\sigma}_{\underline{E}} \delta_{\partial\Omega} = -T_{\frac{\partial \underline{B}}{\partial t}}.$$

La composante "volumique" de cette équation (c'est à dire ne faisant pas apparaître de distribution superficielle $\delta_{\partial\Omega}$) redonne l'équation de Maxwell précédente. Par contre, l'équation superficielle conduit à

$$\underline{n} \wedge \underline{\sigma}_{\underline{E}} = 0,$$

c'est à dire à l'égalité des composantes tangentielles du champ électrique à la frontière du domaine. De même, supposant l'existence d'une densité superficielle de courant \underline{j} sur $\partial\Omega$, et écrivant au sens des distributions

$$\text{Rot } T_{\underline{H}} = T_{\underline{j}} + \frac{\partial}{\partial t} T_{\underline{D}},$$

ce qui en séparant la partie volumique de la partie superficielle, donne l'équation de Maxwell classique

$$T_{\text{Rot } \underline{H}} = T_{\frac{\partial \underline{D}}{\partial t}},$$

ainsi que la condition de raccordement du champ magnétique à la frontière, c'est à dire la discontinuité des composantes tangentielles de \underline{H} :

$$\underline{n} \wedge \underline{\sigma}_{\underline{H}} = \underline{j}.$$

Passons aux équations à la divergence, et supposons l'existence d'une densité superficielle de charge ρ sur $\partial\Omega$. En écrivant au sens des distributions

$$\text{Div } T_{\underline{D}} = T_{\rho},$$

on aboutit à l'équation classique pour la partie volumique

$$T_{\text{Div } \underline{D}} = T_{\rho},$$

ainsi qu'à la condition de raccordement

$$\underline{n} \cdot \underline{\sigma}_{\underline{D}} = \rho,$$

qui spécifie donc la discontinuité de la composante normale de l'induction électrique. De même, en utilisant la dernière équation

$$\text{Div } T_{\underline{B}} = 0,$$

on obtient la continuité de la composante normale de l'induction magnétique.

Ainsi, l'écriture des équations de Maxwell au sens des distributions permet d'unifier le traitement des équations elles-mêmes et des conditions de raccordement.

8.4 Suites de distributions

De nombreuses distributions sont définies via un passage à la limite d'une suite de distributions régulières. La convergence au sens des distributions est définie comme suit.

DÉFINITION 8.8 Une suite de distributions $T_n \in \mathcal{D}'(\mathbb{R})$, $n \in \mathbb{Z}^+$ converge vers une distribution limite $T \in \mathcal{D}'(\mathbb{R})$ si pour toute fonction test $f \in \mathcal{D}(\mathbb{R})$, on a

$$\lim_{n \rightarrow \infty} \langle T_n, f \rangle = \langle T, f \rangle. \quad (8.29)$$

Supposons que $T_n \rightarrow T$ dans $\mathcal{D}'(\mathbb{R})$, et considérons la suite des dérivées p -ièmes $T_n^{(p)}$ de T_n . On a pour tout $f \in \mathcal{D}(\mathbb{R})$

$$\lim_{n \rightarrow \infty} \langle T_n^{(p)}, f \rangle = (-1)^p \lim_{n \rightarrow \infty} \langle T_n, f^{(p)} \rangle = (-1)^p \langle T, f^{(p)} \rangle = \langle T^{(p)}, f \rangle.$$

Ainsi, on a montré

PROPOSITION 8.7 Soit T_n une suite de distributions convergeant vers T dans $\mathcal{D}'(\mathbb{R})$. Alors

$$T_n^{(p)} \longmapsto T^{(p)}$$

quand $n \rightarrow \infty$.

Les distributions permettent de donner un sens à des limites de suites de fonctions qui n'existent pas au sens des fonctions, comme le montre l'exemple suivant.

EXEMPLE 8.11 Soit φ_n la suite fonctions localement intégrables définies par

$$\varphi_n(t) = \sin(2\pi nt) \chi_{[-1/2, 1/2]}(t).$$

La suite φ ne converge pas au sens usuel quand $n \rightarrow \infty$; en fait, à l'exception des points $t = 0$ et $t = \pm 1/2$ où φ_n s'annule pour tout n , on n'a jamais convergence simple. Par contre, il est possible de montrer que

$$\lim_{n \rightarrow \infty} \langle T_{\varphi_n}, f \rangle = 0 \quad \forall f \in \mathcal{D}(\mathbb{R}),$$

d'où

$$\lim_{n \rightarrow \infty} T_{\varphi_n} = 0$$

au sens des distributions.

Il est tentant de penser que toute suite de fonctions localement intégrables convergeant presque partout vers une fonction localement intégrable converge au sens des distributions vers la distribution régulière correspondante. Ceci n'est pas tout à fait vrai, mais le devient via une hypothèse supplémentaire.

PROPOSITION 8.8 Soit $\{\varphi_n \in L^1_{loc}(\mathbb{R}), n \in \mathbb{Z}^+\}$ une suite convergeant presque partout vers $\varphi \in L^1_{loc}(\mathbb{R})$, et telle qu'il existe $g \in L^1_{loc}(\mathbb{R})$ vérifiant

$$|\varphi_n(t)| \leq g(t) \quad \forall t \in \mathbb{R}.$$

Alors

$$T_{\varphi_n} \longmapsto T_\varphi \quad \text{quand } n \rightarrow \infty.$$

EXEMPLE 8.12 Soit φ_n la suite fonctions localement intégrables définies par

$$\varphi_n(t) = \begin{cases} n & \text{si } |t| \leq 1/2n \\ 0 & \text{sinon.} \end{cases}$$

On voit que φ_n converge presque partout vers la fonction nulle (en fait, partout sauf en 0, où la limite n'est pas définie). Par contre, la limite de la suite φ_n (ou plus précisément de la suite des distributions régulières associées) est bien définie au sens des distributions. Soit en effet $f \in \mathcal{D}(\mathbb{R})$.

$$\langle T_{\varphi_n}, f \rangle = n \int_{-1/2n}^{1/2n} f(t) dt \rightarrow f(0),$$

d'où on déduit

$$T_{\varphi_n} \longmapsto \delta$$

au sens des distributions. Ainsi, la limite de cette suite de distributions régulières est dans ce cas une distribution non régulière.

Un cas particulièrement important est celui de la convergence vers la distribution de Dirac, qu'on vient de voir. Plus généralement, on a

THÉORÈME 8.4 Soit $\{\varphi_n\}$ une suite de fonctions localement intégrables, telles que

1. Il existe $A > 0$ tel que pour tout $t \in \mathbb{R}$, $|t| \leq A$, $\varphi_n(t) \geq 0$.
2. Pour tout $\epsilon \in]0, 1[$, $\int_{|t| \leq \epsilon} \varphi_n(t) dt \rightarrow 1$ quand $n \rightarrow \infty$.
3. Pour tout $\epsilon \geq 0$, $\varphi_n \rightarrow 0$ uniformément dans le domaine $|t| \in]\epsilon, 1/\epsilon[$.

Alors

$$T_{\varphi_n} \rightarrow \delta \quad \text{quand } n \rightarrow \infty$$

8.5 Produit de convolution des fonctions et des distributions

Le produit de convolution des fonctions (et des distributions) joue un rôle fondamental en Physique. On va se limiter ici au cas des fonctions (et distributions) d'une variable; la généralisation des définitions au cas de plus d'une variable se fait assez simplement, mais certaines propriétés et applications sont parfois bien plus complexes en dimensions supérieures.

8.5.1 Convolution des fonctions

Rappelons tout d'abord la définition du produit de convolution des fonctions. Etant données deux fonctions f, g d'une variable réelle, leur produit de convolution est la fonction $h = f * g$ définie par

$$(f * g)(t) = \int_{-\infty}^{\infty} f(s)g(t-s) ds, \quad (8.30)$$

pour tout t tel que l'intégrale converge. Le produit de convolution possède des propriétés simples : il est commutatif, et distributif par rapport à l'addition.

La notion de convolution joue un rôle fondamental en physique. Elle intervient en particulier lorsqu'une quantité physique, représentée par une fonction f , est mesurée à l'aide d'un dispositif expérimental. Ce dernier est généralement incapable de reproduire les fluctuations les plus rapides de f , et ne donne pour approximation de la valeur $f(t)$ en t qu'une moyenne locale des valeurs de f dans un voisinage de t . On représente ceci via un produit de convolution comme en (8.30), où la fonction g est une fonction localisée autour de 0 (de sorte que $g(t-s)$ est localisée autour de $s=t$), qui représente en quelque sorte la réponse du dispositif expérimental.

REMARQUE 8.3 Un dispositif expérimental parfait admettrait comme réponse une fonction g telle que $f * g = f$. Or il s'avère que la réponse g permettant ceci n'est autre que la distribution de Dirac :

$$\int_{-\infty}^{\infty} f(s)\delta(t-s) ds = f(t).$$

8.5.2 Support d'une distribution

Rappelons tout d'abord la notion de support d'une fonction. Par définition, le support d'une fonction d'une variable réelle f est le plus petit sous-ensemble fermé de \mathbb{R} tel que f s'annule en dehors de ce domaine. Plus généralement, en dimension quelconque, étant donnée $f : \mathbb{R}^n \rightarrow \mathbb{C}$,

$$\text{Supp}(f) = \overline{\{x \in \mathbb{R}^n, f(x) \neq 0\}},$$

la barre représentant la fermeture.

Ceci nous permet de donner un sens à la notion de support d'une distribution.

DÉFINITION 8.9 Soit $T \in \mathcal{D}'(\mathbb{R}^n)$.

1. T est nulle sur un domaine $\Omega \subset \mathbb{R}^n$ si

$$\langle T, f \rangle = 0 \quad \forall f \in \mathcal{D}(\mathbb{R}^n), \text{ Supp}(f) \subset \Omega.$$

2. Le support de T est le plus petit domaine fermé $\Omega \subset \mathbb{R}^n$ tel que T soit nulle en dehors de ce domaine.

Bien évidemment, le support d'une distribution régulière T_φ coïncide avec le support de φ .

EXEMPLE 8.13 Il est facile de vérifier que le support de la distribution de Dirac δ est le singleton $\{0\}$. De même, le support de δ_a est le singleton $\{a\}$. En dimension supérieure, le support de la distribution δ_C sur une courbe C de \mathbb{R}^2 est la courbe C elle-même.

Muni de cette définition, on peut alors introduire aussi les notions suivantes

DÉFINITION 8.10 1. Une distribution $T \in \mathcal{D}'(\mathbb{R})$ est à support limité à gauche si $\text{Supp}(T) \subset [a, \infty[$, pour un certain $a \in \mathbb{R}$. On note en particulier $\mathcal{D}'_+(\mathbb{R})$ les distributions à support limité à gauche telles que $a \geq 0$.

2. Une distribution $T \in \mathcal{D}'(\mathbb{R})$ est à support limité à droite si $\text{Supp}(T) \subset]-\infty, b[$, pour un certain $b \in \mathbb{R}$. On note en particulier $\mathcal{D}'_-(\mathbb{R})$ les distributions à support limité à droite telles que $b \leq 0$.

8.5.3 Convolution des distributions

Comme d'habitude, la convolution des distributions s'obtient comme généralisation de la convolution des distributions régulières. Considérons donc $\varphi, \psi \in \mathcal{D}(\mathbb{R})$ et les distributions régulières associées. Alors, $\varphi * \psi \in \mathcal{D}(\mathbb{R})$, et on peut considérer $T_{\varphi * \psi}$: soit $f \in \mathcal{D}(\mathbb{R})$,

$$\langle T_{\varphi * \psi}, f \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varphi(s) \psi(t-s) f(t) ds dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varphi(s) [\tau_s \psi](t) f(t) dt ds ,$$

où τ_s est l'opérateur de translation

$$[\tau_s f](t) = f(t-s) .$$

On reconnaît dans l'intégrale sur t l'action de la distribution régulière $T_{\tau_s \psi} = \tau_s T_\psi$ sur f

$$\int_{-\infty}^{\infty} [\tau_s \psi](t) f(t) dt = \langle \tau_s T_\psi, f \rangle ,$$

qui produit donc une fonction de s ; l'intégrale par rapport à t représente alors l'action de T_φ sur cette fonction de s , ce que l'on note

$$\langle T_{\varphi * \psi}, f \rangle = \langle T_\varphi, \langle \tau_\bullet T_\psi, f \rangle \rangle ,$$

où on a aussi noté $\langle \tau_\bullet T_\psi, f \rangle$ la fonction d'une variable réelle

$$\langle \tau_\bullet T_\psi, f \rangle : t \longrightarrow \langle \tau_t T_\psi, f \rangle = \langle T_{\tau_t \psi}, f \rangle .$$

Cette fonction possède la propriété suivante

LEMME 8.1 Soit $T \in \mathcal{D}'(\mathbb{R})$. Pour tout $f \in \mathcal{D}(\mathbb{R})$, la fonction $\langle \tau_\bullet T, f \rangle$ est C^∞ . Cette fonction est à support limité à gauche (resp. à droite) si T est à support limité à droite (resp. à gauche).

Dans un cadre plus général, en définissant l'action d'un opérateur de translation sur une distribution $T \in \mathcal{D}'(\mathbb{R})$ par

$$\langle \tau_s T, f \rangle = \langle T, \tau_{-s} f \rangle ,$$

on est conduit à définir de même la fonction $\langle \tau_\bullet T, f \rangle$ par

$$\langle \tau_\bullet T, f \rangle : t \longrightarrow \langle \tau_t T, f \rangle = \langle T, \tau_{-t} f \rangle . \quad (8.31)$$

On peut alors introduire le produit de convolution de deux distributions :

DÉFINITION 8.11 Soient $T, S \in \mathcal{D}'(\mathbb{R})$. Le produit de convolution $T * S$ est défini par

$$\langle T * S, f \rangle = \langle T, \langle \tau_\bullet S, f \rangle \rangle , \quad (8.32)$$

lorsque ces opérations ont un sens.

Malgré cette définition quelque peu formelle, le produit de convolution est relativement simple à utiliser, comme on le verra. En pratique, on utilise souvent la notation des distributions comme fonctions généralisées, ce qui permet de se débarrasser des notations trop lourdes.

En fait, le produit de convolution de deux distributions données n'est pas toujours bien défini. Il l'est toutefois au prix d'hypothèses supplémentaires.

PROPOSITION 8.9 Soient $T, S \in \mathcal{D}'(\mathbb{R})$. Le produit de convolution est bien défini en tant que distribution de $\mathcal{D}'(\mathbb{R})$ quand

1. T et S sont à support limité à gauche. $T * S$ est alors à support limité à gauche. En particulier, le produit de deux distributions de $\mathcal{D}'_+(\mathbb{R})$ est toujours une distribution de $\mathcal{D}'_+(\mathbb{R})$.
2. T et S sont à support limité à droite. $T * S$ est alors à support limité à droite. En particulier, le produit de deux distributions de $\mathcal{D}'_-(\mathbb{R})$ est toujours une distribution de $\mathcal{D}'_-(\mathbb{R})$.
3. T et S appartiennent à l'espace $\mathcal{E}'(\mathbb{R})$ des distributions à support compact, dual de l'espace $\mathcal{E}(\mathbb{R})$ des fonctions indéfiniment différentiables.

Alors, le produit de convolution possède les propriétés suivantes :

1. Il est commutatif :

$$T * S = S * T, \quad S, T \in \mathcal{D}'(\mathbb{R}).$$

2. Il est distributif par rapport à l'addition :

$$T * (U + V) = T * U + T * V, \quad T, U, V \in \mathcal{D}'(\mathbb{R}).$$

3. Il est associatif,

$$T * (U * V) = (T * U) * V, \quad T, U, V \in \mathcal{D}'(\mathbb{R}),$$

lorsque toutes ces opérations sont bien définies.

EXEMPLE 8.14 La distribution de Dirac est élément neutre pour le produit de convolution :

$$\delta * T = T, \quad \forall T \in \mathcal{D}'(\mathbb{R}).$$

Similairement, pour tout réel a , on montre que

$$\delta_a * T = \tau_a T,$$

d'où, en appliquant cela au peigne de Dirac, on obtient pour tout $T \in \mathcal{D}'(\mathbb{R})$

$$\sqcup_h * T = \sum_{k \in \mathbb{Z}} \tau_{kh} T.$$

Cette distribution est appelée périodisée de T de période h .

8.5.4 Equations de convolution

L'utilisation des distributions pour la résolution d'équations différentielles et équations aux dérivées partielles repose essentiellement sur le résultat suivant :

PROPOSITION 8.10 1. Soient $S, T \in \mathcal{D}'(\mathbb{R})$, telles que $S * T$ existe dans $\mathcal{D}'(\mathbb{R})$. Alors

$$(S * T)' = S' * T = S * T',$$

et plus généralement, pour tout entier positif n ,

$$(S * T)^{(n)} = S^{(n)} * T = S * T^{(n)}. \quad (8.33)$$

2. Pour tout $T \in \mathcal{D}'(\mathbb{R})$, et tout entier positif n ,

$$T^{(n)} = \delta^{(n)} * T. \quad (8.34)$$

Preuve : Commençons par la première partie, et supposons que $S * T$ existe. Alors $(S * T)' \in \mathcal{D}'(\mathbb{R})$ aussi, et

$$\begin{aligned} \langle (S * T)', f \rangle &= -\langle S * T, f' \rangle \\ &= -\langle S, \langle \tau \bullet, T, f' \rangle \rangle \\ &= \langle S, \langle (\tau \bullet T)', f \rangle \rangle \\ &= \langle S, \langle \tau \bullet, T', f \rangle \rangle \\ &= \langle S * T', f \rangle \end{aligned}$$

car la dérivation commute avec les translations. L'autre égalité se montre de même. La preuve pour les dérivées de tous ordres est identique.

Pour la seconde partie, il suffit de remarquer que

$$T' = T' * \delta = T * \delta',$$

cette démonstration se généralisant elle aussi à tous les ordres. ♠

La conséquence immédiate de cette proposition est que les équations différentielles au sens des distributions peuvent se mettre sous la forme d'équations de convolution. Plus précisément, considérons une équation différentielle d'ordre n , de la forme

$$\alpha_n X^{(n)} + \alpha_{n-1} X^{(n-1)} + \dots + \alpha_1 X' + \alpha_0 X = Y,$$

où les α_k sont des coefficients fixés, et le membre de droite $Y \in \mathcal{D}'(\mathbb{R})$ est lui aussi supposé connu. Cette équation se met sous la forme

$$T * X = Y,$$

où la distribution T est donnée par

$$T = \alpha_n \delta^{(n)} + \alpha_{n-1} \delta^{(n-1)} + \dots + \alpha_1 \delta' + \alpha_0 \delta.$$

La question est alors la suivante : peut-on trouver un "inverse" à T pour le produit de convolution. En d'autres termes, puisque comme on l'a vu, la distribution de Dirac est élément neutre pour le produit de convolution, existe-t-il une distribution, que l'on note T^{*-1} , telle que

$$T^{*-1} * T = \delta ?$$

Si tel est le cas, on pourra alors écrire

$$X = T^{*-1} * T * X = T^{*-1} * Y,$$

ce qui résout l'équation différentielle.

La question est : dans quel cadre ce calcul a-t-il un sens ?

Un cadre possible est donné par les algèbres de convolution.

DÉFINITION 8.12 Une algèbre de convolution dans $\mathcal{D}'(\mathbb{R})$ est un sous-espace vectoriel \mathcal{A} de $\mathcal{D}'(\mathbb{R})$, tel que

1. \mathcal{A} est stable par convolution : $\forall T, S \in \mathcal{A}, T * S \in \mathcal{A}$, et le produit de convolution est commutatif et associatif dans \mathcal{A} .
2. $\delta \in \mathcal{A}$.

EXEMPLE 8.15 Les sous-espaces de $\mathcal{D}'(\mathbb{R})$ suivants sont des algèbres de convolution :

- L'espace $\mathcal{E}'(\mathbb{R})$ des distributions à support compact.
- L'espace des distributions à support limité à gauche.
- L'espace des distributions à support limité à droite.
- L'espace $\mathcal{D}'_+(\mathbb{R})$ des distributions à support dans $[0, \infty[$.
- L'espace $\mathcal{D}'_-(\mathbb{R})$ des distributions à support dans $] - \infty, 0]$.

Considérons l'équation de convolution

$$T * X = Y \quad (8.35)$$

Etant donnée une algèbre de convolution \mathcal{A} , et une distribution $T \in \mathcal{A}$, si il existe $U \in \mathcal{A}$ telle que

$$U * T = \delta, \quad (8.36)$$

alors il est possible de montrer que U est unique dans \mathcal{A} .

DÉFINITION 8.13 Etant donnée $T \in \mathcal{A}$, l'**inverse de convolution** $U \in \mathcal{A}$ de $T \in \mathcal{A}$, vérifiant (8.36), si il existe, est appelé **solution élémentaire** de (8.35).

EXEMPLE 8.16 Dans $\mathcal{D}'_+(\mathbb{R})$, soit $T_\Theta \in \mathcal{D}'_+(\mathbb{R})$ la distribution de Heaviside (distribution régulière associée à la fonction de Heaviside). On considère l'équation

$$T_\Theta * X = Y.$$

On a déjà vu que $\delta = T'_\Theta$, et comme $\delta \in \mathcal{D}'_+(\mathbb{R})$, on a

$$\delta' * T_\Theta = \delta * T'_\Theta = \delta * \delta = \delta.$$

δ' est l'inverse de convolution de T_Θ , et est donc la solution élémentaire dans $\mathcal{D}'_+(\mathbb{R})$ de l'équation ci-dessus.

EXEMPLE 8.17 La primitive peut aussi s'interpréter en termes de solution fondamentale. On considère cette fois l'équation, toujours dans $\mathcal{D}'_+(\mathbb{R})$

$$X' = \delta' * X = Y.$$

Comme on vient de le voir, l'inverse de convolution de δ' est T_Θ , qui est donc solution fondamentale de l'équation ci-dessus.

8.5.5 Application aux problèmes de valeur initiale

On vient de voir que la dérivation des distributions est équivalente à la convolution avec une dérivée de la distribution de Dirac. Un domaine d'applications immédiat est constitué par les équations différentielles. On se limite ici à un cadre relativement simple, en se limitant à l'algèbre de convolution \mathcal{D}'_+ , qui permet toutefois de cerner les points importants.

Considérons l'équation différentielle au sens des distributions

$$X^{(n)} + a_{n-1}X^{(n-1)} + \dots + a_1X' + a_0X = Y, \quad (8.37)$$

où les a_k sont des coefficients fixés, et le membre de droite $Y \in \mathcal{D}'_+(\mathbb{R})$ est lui aussi fixé. Comme on l'a vu, cette équation se met sous la forme

$$D * X = Y,$$

où la distribution D est donnée par

$$D = \delta^{(n)} + a_{n-1}\delta^{(n-1)} + \dots + a_1\delta' + a_0\delta. \quad (8.38)$$

THÉORÈME 8.5 Soit f une fonction, solution de l'équation différentielle

$$f^{(n)}(t) + a_{n-1}f^{(n-1)}(t) + \dots + a_1f'(t) + a_0f(t) = 0, \quad (8.39)$$

avec les conditions initiales

$$f(0) = f'(0) = \dots = f^{(n-2)}(0) = 0, \quad f^{(n-1)}(0) = 1. \quad (8.40)$$

Alors la solution fondamentale $U = D^{*-1} \in \mathcal{D}'_+(\mathbb{R})$ de l'équation

$$D * X = Y,$$

où D est donné en (8.38) est la distribution régulière

$$U = T_{f\ominus}, \quad (8.41)$$

et (8.37) admet donc l'unique solution dans $\mathcal{D}'_+(\mathbb{R})$

$$X = T_{f\ominus} * Y \in \mathcal{D}'_+(\mathbb{R}). \quad (8.42)$$

Preuve : soit donc $U = T_{f\ominus}$. Un calcul direct montre que

$$U' = T_{f'\ominus}, \quad \dots, U^{(n-1)} = T_{f^{(n-1)}\ominus}, \quad U^{(n)} = T_{f^{(n)}\ominus} + \delta.$$

Par conséquent, on a bien

$$U^{(n)} + a_{n-1}U^{(n-1)} + \dots + a_1U' + a_0U(t) = T_{f^{(n)} + a_{n-1}f^{(n-1)} + \dots + a_1f' + a_0f} + \delta = \delta,$$

ce qui montre que U est une solution fondamentale. $U \in \mathcal{D}'_+$, et est donc l'unique solution fondamentale dans \mathcal{D}'_+ . Finalement, on voit facilement que $U * Y$ est solution de l'équation différentielle. Comme $U, Y \in \mathcal{D}'_+$, qui est une algèbre de convolution, on en déduit le résultat. ♠

EXEMPLE 8.18 Considérons l'exemple simple d'un circuit RC ; on note V_1 la tension générée par un générateur, et V_2 la tension aux bornes du condensateur. Comme on l'a vu, V_1 et V_2 sont liés par l'équation

$$\tau V_2' + V_2 = V_1,$$

où $\tau = RC$. Il s'agit donc d'un problème de la forme

$$X' + \lambda X = Y,$$

avec $X = V_2$, $\lambda = 1/\tau$ et $Y = V_1/\tau$. La solution élémentaire prend donc la forme d'une distribution régulière

$$U = T_{u\theta},$$

où u , solution de $u' + \lambda u$ avec $u(0) = 1$ est donné par

$$u(t) = e^{-\lambda t}.$$

Finalement, la solution est donc

$$V_2 = X = T_{u\theta} * Y = \frac{1}{\tau} T_{u\theta} * V_1.$$

La solution élémentaire possède une interprétation physique simple : elle donne la réponse du système à une "impulsion élémentaire", représentée par une distribution de Dirac :

$$T_{u\Theta} = T_{u\Theta} * \delta.$$

C'est pourquoi on appelle parfois **réponse impulsionnelle** la solution élémentaire.

Dans le cas où le système étudié (ici le circuit RC) est excité par une distribution Y plus complexe, la solution V_2 apparaît comme une superposition de ces réponses impulsionnelles.

Supposons maintenant que Y soit une distribution régulière, $Y = T_\varphi$ pour une certaine fonction $\varphi \in L^1_{loc}(\mathbb{R})$. On a alors

$$V_1 = T_{u\Theta} * T_\varphi = T_{(u\Theta)*\varphi}.$$

La fonction $v = (u\Theta) * \varphi$ prend la forme

$$v(t) = \int_{-\infty}^{\infty} u(s)\Theta(s)\varphi(t-s) ds = \int_0^{\infty} u(s)\varphi(t-s) ds = \int_0^{\infty} e^{-s/\tau}\varphi(t-s) ds.$$

8.6 Transformations de Laplace et Fourier des distributions

La transformation de Fourier joue elle aussi un rôle fondamental, et la théorie des distributions permet de lui donner un sens dans des cas où la définition usuelle (au sens des fonctions) n'en a plus.

8.6.1 Espace de Schwartz et distributions tempérées

DÉFINITION 8.14 1. Une fonction $f \in C^\infty(\mathbb{R})$ est dite à décroissance rapide si pour tout entier positif k , on a

$$\lim_{t \rightarrow \pm\infty} t^k |f(t)| = 0. \quad (8.43)$$

L'espace des fonctions C^∞ dont toutes les dérivées sont à décroissance rapide est appelé **Espace de Schwartz**, et noté $\mathcal{S}(\mathbb{R})$.

On peut facilement voir que toute fonction de $\mathcal{D}(\mathbb{R})$ appartient automatiquement à l'espace de Schwartz. De plus, toute fonction de $\mathcal{S}(\mathbb{R})$ est de carré intégrable. On a donc

$$\mathcal{D}(\mathbb{R}) \subset \mathcal{S}(\mathbb{R}) \subset L^2(\mathbb{R}).$$

Pour préciser la topologie de $\mathcal{S}(\mathbb{R})$, il est nécessaire de définir la convergence dans $\mathcal{S}(\mathbb{R})$. On dira qu'une suite de fonctions $\{f_n \in C^\infty(\mathbb{R}), n \in \mathbb{Z}^+\}$ converge vers f si pour tout k , la suite des dérivées d'ordre k de f_n converge uniformément vers $f^{(k)}$.

On peut alors définir les distributions tempérées :

DÉFINITION 8.15 Une distribution tempérée est une forme linéaire continue sur l'espace de Schwartz. Les distributions tempérées forment un espace vectoriel, noté $\mathcal{S}'(\mathbb{R})$.

Comme conséquence des inclusions ci-dessus, on montre également les inclusions inverses au niveau des espaces duaux :

$$L^2(\mathbb{R}) \subset \mathcal{S}'(\mathbb{R}) \subset \mathcal{D}'(\mathbb{R}).$$

EXEMPLE 8.19 Les exemples les plus simples de distributions tempérées sont les distributions régulières. Cependant, toute distribution régulière n'est pas nécessairement tempérée. En effet, étant donnée $\varphi \in L^1_{loc}(\mathbb{R})$, $\langle T_\varphi, f \rangle$ peut être bien définie pour toute fonction test $f \in \mathcal{D}(\mathbb{R})$, mais pas pour $f \in \mathcal{S}(\mathbb{R})$. Par exemple, en prenant $\varphi(t) = e^{t^2}$ pour tout t , $\langle T_\varphi, f \rangle$ n'est pas bien défini pour une large classe de fonctions $f \in \mathcal{S}(\mathbb{R})$ (par exemple $f(t) = e^{-t^2}$). Il est nécessaire de faire des hypothèses supplémentaires sur φ .

On dit que $\varphi \in L^1_{loc}(\mathbb{R})$ est à croissance lente si il existe A, M et k tels que pour tout $|t| \geq M$, on ait $|\varphi(t)| \leq A|t|^k$. On montre alors que les distributions régulières associées à de telles fonctions à croissance lente sont tempérées.

Par exemple, T_Θ est une distribution tempérée.

EXEMPLE 8.20 On montre facilement que toutes les distributions de Dirac et leurs dérivées sont des distributions tempérées. En effet, leur action sur les fonctions test de l'espace de Schwartz est bien définie. Par exemple, $\langle \delta^{(n)}, f \rangle = (-1)^n f^{(n)}(0)$ est bien défini pour tout entier n si $f \in \mathcal{S}(\mathbb{R})$.

Il est possible de montrer pour les distributions tempérées un certain nombre de propriétés similaires à celles que nous avons rencontrées dans $\mathcal{D}'(\mathbb{R})$. Notamment

PROPOSITION 8.11 Soit $T \in \mathcal{S}'(\mathbb{R})$. Alors

1. $T^{(p)} \in \mathcal{S}'(\mathbb{R})$ pour tout $p \in \mathbb{Z}^+$.
2. $PT \in \mathcal{S}'(\mathbb{R})$ pour tout polynôme P .

8.6.2 Transformation de Laplace des distributions

Commençons par aborder brièvement la transformation de Laplace. Comme dans le cas des fonctions, on se limitera ici au cas de distributions à support borné à gauche, plus précisément aux distributions de $\mathcal{D}'_+(\mathbb{R})$.

DÉFINITION 8.16 Soit T une distribution à support borné à gauche, telle qu'il existe α tel que pour tout $x \geq \alpha$, la distribution $e^{-xt}T(t)$ soit tempérée. La transformée de Laplace unilatérale de T est la **distribution régulière**

$$[\mathcal{L}T](p) = \langle T, e^{-p\bullet} \rangle. \quad (8.44)$$

La borne inférieure des α satisfaisant cette propriété est appelée abscisse de sommabilité.

Il est important d'insister sur le fait que la transformée de Laplace ainsi définie est une distribution régulière, c'est à dire une distribution de type « fonction ». Ceci simplifie beaucoup les choses dans certains calculs.

Par exemple, la distribution de Dirac est elle même une distribution tempérée, à support borné à gauche, et admet une transformée de Laplace

$$[\mathcal{L}\delta](p) = 1, \quad (8.45)$$

avec abscisse de sommabilité égal à $-\infty$.

Le lien entre transformation de Laplace et dérivation reste assez simple, voire plus simple que dans le cas des fonctions.

PROPOSITION 8.12 Soit $T \in \mathcal{D}'_+$, et soit T' sa dérivée. Alors si T admet une transformée de Laplace $\mathcal{L}T$, T' admet également une transformée de Laplace $\mathcal{L}T'$, qui satisfait

$$[\mathcal{L}T'](p) = p[\mathcal{L}T](p).$$

Il est à noter que ce résultat est compatible avec celui que nous avons vu dans le cas des fonctions. En effet, considérons $\varphi \in L^1_{loc}(\mathbb{R}^+)$, continûment différentiable sur \mathbb{R}^+ , et soit $T = T_{\Theta\varphi}$. On a alors

$$[T_{\Theta\varphi}]' = T_{\Theta\varphi'} + \varphi(0)\delta,$$

et par application de la transformation de Laplace,

$$[\mathcal{L}T_{\Theta\varphi'}](p) = p[\mathcal{L}T_{\Theta\varphi}](p) - \varphi(0),$$

ce qui coïncide avec la règle de transformation de Laplace de la dérivée des fonctions.

EXEMPLE 8.21 On a déjà vu que la transformée de Laplace de la distribution de Dirac est la fonction constante égale à 1. plus généralement, on a en notation « fonction » (rappelons que la transformée de Laplace est une fonction)

$$[\mathcal{L}\delta^{(n)}](p) = p^n, \quad n \in \mathbb{N}. \quad (8.46)$$

Pour finir, il est intéressant de caractériser les distributions qui sont transformées de Laplace d'une distribution :

PROPOSITION 8.13 Toute fonction $p \rightarrow F(p)$, holomorphe et majorée par un polynôme en p dans un demi-plan $\Re(p) > \alpha$ est transformée de Laplace d'une distribution de \mathcal{D}'_{\pm} .

EXEMPLE 8.22 Application au calcul de certaines fonctions de Green dans \mathcal{D}'_+

1. Considérons le cas de la dérivation : la fonction de Green correspondante G dans \mathcal{D}'_+ doit satisfaire $G' = \delta$, d'où par transformation de Laplace

$$pG(p) = 1.$$

On en déduit (rappelons que la transformée de Laplace d'une distribution est une distribution régulière) que

$$[\mathcal{L}G](p) = \frac{1}{p}, \quad \Re(p) > 0,$$

dont on a vu que c'était la transformée de Laplace de la fonction de Heaviside, d'où

$$G = T_{\Theta}.$$

Donc, la solution de l'équation $T' = S$ est de la forme $T = T_{\Theta} * S$, sans surprise quand on se souvient que convoluer avec la fonction de Heaviside est la même chose que calculer la primitive valant 0 à l'origine.

2. Considérons le cas de l'opérateur de Helmholtz : $d^2/dt^2 + \omega_0^2$. Les fonctions de Green satisfont $G'' + \omega_0^2 G = \delta$, soit dans le domaine de Laplace

$$(p^2 + \omega_0^2)[\mathcal{L}G](p) = 1,$$

dont la solution ($\mathcal{L}G$ étant une fonction) est $[\mathcal{L}G](p) = 1/(p^2 + \omega_0^2)$, avec $\Re(p) > 0$. La solution prend donc la forme de la distribution régulière

$$G(t) = T_{\Theta g}, \quad g(t) = \frac{1}{\omega_0} \sin(\omega_0 t).$$

Notons que la présence de la fonction de Heaviside dans cette affaire provient du domaine de définition de $\mathcal{L}G$, à savoir $\Re(p) > 0$.

8.6.3 Transformée de Fourier des distributions tempérées

On rappelle la propriété centrale de la transformation de Fourier des fonctions, appelée **formule d'échange** : pour toutes fonctions $f, g \in L^1(\mathbb{R})$, on a

$$\int_{-\infty}^{\infty} f(t)\hat{g}(t) dt = \int_{-\infty}^{\infty} \hat{f}(s)g(s) ds.$$

La transformation de Fourier des distributions généralise cette propriété aux distributions tempérées.

DÉFINITION 8.17 Soit $T \in \mathcal{S}'(\mathbb{R})$. Sa transformée de Fourier \hat{T} est la distribution tempérée définie par

$$\langle \hat{T}, f \rangle = \langle T, \hat{f} \rangle, \quad \forall f \in \mathcal{S}(\mathbb{R}). \quad (8.47)$$

On note \mathcal{F} l'opérateur de transformation de Fourier :

$$\hat{T} = \mathcal{F}T.$$

EXEMPLE 8.23 Distributions de Dirac. Calculons $\hat{\delta}$: pour toute fonction test $f \in \mathcal{S}(\mathbb{R})$,

$$\langle \hat{\delta}, f \rangle = \langle \delta, \hat{f} \rangle = \hat{f}(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) dt.$$

Par conséquent, la transformée de Fourier de la distribution de Dirac est, à une constante près, la distribution régulière associée à la fonction identiquement égale à 1 :

$$\hat{\delta} = \frac{1}{\sqrt{2\pi}} T_1. \quad (8.48)$$

De même, pour tout $a \in \mathbb{R}$,

$$\langle \hat{\delta}_a, f \rangle = \langle \delta_a, \hat{f} \rangle = \hat{f}(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{-iat} dt,$$

d'où on déduit

$$\hat{\delta}_a = \frac{1}{\sqrt{2\pi}} T_{\epsilon_a}, \quad (8.49)$$

où ϵ_a est la sinusoïde

$$\epsilon_a(t) = e^{-iat}.$$

EXEMPLE 8.24 Sinusoïdes. Calculons tout d'abord

$$\langle \hat{T}_1, f \rangle = \langle T_1, \hat{f} \rangle = \int_{-\infty}^{\infty} \hat{f}(\omega) d\omega = \sqrt{2\pi} f(0),$$

d'après la formule d'inversion de Fourier, qui s'applique (point par point) pour les fonctions de $\mathcal{S}(\mathbb{R})$. On a donc

$$\widehat{T}_1 = \sqrt{2\pi} \delta. \quad (8.50)$$

Si on prend le parti de noter les distributions sous forme de fonction généralisée, on notera alors

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-i\omega t} dt = \sqrt{2\pi} \delta(\omega),$$

d'où la formule suivante, extrêmement utile dans les calculs :

$$\delta(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t} dt. \quad (8.51)$$

De même, on montre facilement que

$$\widehat{T}_{\epsilon_a} = \sqrt{2\pi} \delta_{-a}, \quad (8.52)$$

que l'on interprète aussi en termes de fonctions généralisées comme

$$\delta(\omega + a) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i(\omega+a)t} dt.$$

8.6.4 Propriétés

La transformation de Fourier des distributions tempérées possède la plupart des propriétés "standard" de la transformation de Fourier des fonctions. On en donne une liste (non exhaustive) ci-dessous.

1. **Linéarité.** On vérifie facilement que pour tous $T, S \in \mathcal{S}'(\mathbb{R})$ et $\lambda \in \mathbb{C}$,

$$\widehat{T+S} = \widehat{T} + \widehat{S} \quad (8.53)$$

$$\widehat{\lambda T} = \lambda \widehat{T}. \quad (8.54)$$

2. **Comportement vis à vis des translations et modulations.** Soit $a \in \mathbb{R}$, et soit τ_a l'opérateur de translation par a : $\tau_a f(t) = f(t - a)$. En notant comme plus haut ϵ_a la fonction qui à tout $t \in \mathbb{R}$ associe $\epsilon_a(t) = e^{-iat}$, on a

$$\widehat{\tau_a T} = \epsilon_a \widehat{T}. \quad (8.55)$$

De même, en notant comme d'habitude $\epsilon_a T$ le produit de la distribution (tempérée) par la fonction C^∞ ϵ_a ,

$$\widehat{\epsilon_a T} = \tau_{-a} \widehat{T}. \quad (8.56)$$

8.6.5 Transformation de Fourier et dérivation

Comme on l'a vu plus haut, l'espace de Schwartz ainsi que l'espace des distributions tempérées sont stables par dérivation. Par conséquent, étant donnée une distribution tempérée T , à laquelle on sait associer sa transformée de Fourier \widehat{T} , on sait également associer une transformée de Fourier à toutes ses dérivées, qui sont elles aussi dans $\mathcal{S}'(\mathbb{R})$.

Par ailleurs, on sait que pour toute fonction $f \in \mathcal{S}(\mathbb{R})$,

$$\begin{aligned} (\widehat{f})'(\omega) &= \widehat{(-it)f(\omega)} \\ \widehat{f}'(\omega) &= i\omega \widehat{f}(\omega), \end{aligned}$$

où on a noté it la fonction qui à t associe it .

Ecrivons donc, pour $T \in \mathcal{S}'(\mathbb{R})$ et tout $f \in \mathcal{S}(\mathbb{R})$

$$\langle \widehat{T}', f \rangle = \langle T', \widehat{f} \rangle = -\langle T, (\widehat{f})' \rangle = -\langle T, \widehat{(-it f)} \rangle = \langle \widehat{T}, it f \rangle = \langle it \widehat{T}, f \rangle.$$

Ceci étant vrai pour tout $f \in \mathcal{S}'(\mathbb{R})$, on a donc montré

$$\widehat{T'} = it \widehat{T}. \quad (8.57)$$

Similairement, calculons

$$\langle \widehat{T'}, f \rangle = -\langle \widehat{T}, f' \rangle = -\langle T, \widehat{f'} \rangle = -\langle T, it \widehat{f} \rangle = \langle (-it) T, f \rangle.$$

En étendant ces calculs aux dérivées d'ordres quelconques, on montre similairement

THÉORÈME 8.6 Soit $T \in \mathcal{S}'(\mathbb{R})$. Alors on a, pour tout entier positif n ,

$$\widehat{T^{(n)}} = (it)^n \widehat{T} \quad (8.58)$$

$$\widehat{T}^{(n)} = (-it)^n T \quad (8.59)$$

EXEMPLE 8.25 Dérivées de distributions de Dirac. En appliquant directement les expressions que nous venons de voir, et l'expression de la transformée de Fourier de distributions de Dirac, on obtient directement

$$\widehat{\delta^{(p)}} = (it)^p \widehat{\delta} = \frac{1}{\sqrt{2\pi}} (it)^p T_1. \quad (8.60)$$

EXEMPLE 8.26 Distribution de Heaviside. Pour simplifier les notations, on notera Θ la distribution régulière T_Θ associée à la fonction de Heaviside Θ . De l'égalité $\Theta' = \delta$ on déduit la relation

$$i\omega \widehat{\Theta} = \frac{1}{\sqrt{2\pi}}$$

d'où on déduit, d'après la Proposition 8.5 que $\widehat{\Theta}$ est nécessairement de la forme

$$\widehat{\Theta} = \frac{1}{i\sqrt{2\pi}} \text{v.p.} \left(\frac{1}{\omega} \right) + a\delta$$

pour une certaine constante a . Cette dernière est déterminée grâce à la remarque suivante : la fonction $\Theta - 1/2$ est réelle et impaire, de même que la distribution associée. Par conséquent, la transformée de Fourier de $\Theta - 1/2$ doit être impaire également. Donc, la transformée de Fourier de $1/2$ valant $\delta/2\sqrt{2}$, et la distribution de Dirac étant paire, on a nécessairement $a = 1/2\sqrt{2}$, d'où

$$\widehat{\Theta} = \frac{1}{i\sqrt{2\pi}} \left[\text{v.p.} \left(\frac{1}{\omega} \right) + i\pi\delta \right]. \quad (8.61)$$

De même, en notant Θ_- la distribution régulière associée à $t \rightarrow \Theta(-t)$, on a

$$\widehat{\Theta}_- = \frac{1}{i\sqrt{2\pi}} \left[-\text{v.p.} \left(\frac{1}{\omega} \right) + i\pi\delta \right]. \quad (8.62)$$

Finalement, notons $\text{sgn} = \Theta + \Theta_-$ la distribution régulière associée à la fonction qui à t associe son signe :

$$\text{sgn} = \Theta - \Theta_-.$$

On a alors

$$\widehat{\text{sgn}} = \frac{2}{i\sqrt{2\pi}} \text{v.p.} \left(\frac{1}{\omega} \right)$$

8.6.6 Inversion de la transformation de Fourier

La transformation de Fourier des distributions tempérées hérite de la propriété d'inversibilité de la transformation de Fourier dans l'espace de Schwartz. En effet, si $f \in \mathcal{S}(\mathbb{R})$, alors $\hat{f}(\omega)$ est bien défini pour tout $\omega \in \mathbb{R}$, et on a, pour tout $t \in \mathbb{R}$,

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\omega) e^{i\omega t} d\omega, \quad (8.63)$$

ce que l'on note

$$f = \overline{\mathcal{F}} \hat{f},$$

l'opérateur linéaire $\overline{\mathcal{F}}$ étant appelé **transformée de Fourier conjuguée**, ou **transformée de Fourier adjointe**. On peut donc noter

$$\overline{\mathcal{F}}\mathcal{F} = \mathbf{1}_{\mathcal{S}(\mathbb{R})}, \quad \mathcal{F}\overline{\mathcal{F}} = \mathbf{1}_{\mathcal{S}(\mathbb{R})}. \quad (8.64)$$

Ceci permet de définir la transformation de Fourier conjuguée des distributions tempérées : si $T \in \mathcal{S}'(\mathbb{R})$, on définit sa transformée de Fourier conjuguée $\check{T} = \overline{\mathcal{F}}T$ par dualité :

$$\langle \check{T}, f \rangle = \langle T, \check{f} \rangle. \quad (8.65)$$

On voit alors facilement que pour tout $TR \in \mathcal{S}'(\mathbb{R})$ et tout $f \in \mathcal{S}(\mathbb{R})$,

$$\langle \overline{\mathcal{F}}\mathcal{F}T, f \rangle = \langle \mathcal{F}T, \overline{\mathcal{F}}f \rangle = \langle T, \mathcal{F}\overline{\mathcal{F}}f \rangle = \langle T, f \rangle,$$

d'après (8.64), d'où on déduit

THÉORÈME 8.7 *La transformation de Fourier des distributions tempérées est inversible : on a*

$$\overline{\mathcal{F}}\mathcal{F} = \mathbf{1}_{\mathcal{S}'(\mathbb{R})}, \quad \mathcal{F}\overline{\mathcal{F}} = \mathbf{1}_{\mathcal{S}'(\mathbb{R})}. \quad (8.66)$$

Ce résultat trouve de nombreuses applications, notamment pour la résolution d'équations différentielles, et d'équations aux dérivées partielles en dimensions supérieures.

8.7 Application au calcul des fonctions de Green

On a déjà vu plus haut l'intérêt de la transformation de Fourier pour la résolution des équations différentielles (ou équations aux dérivées partielles) à coefficients constants. La transformation de Fourier permet également le calcul explicite de Fonctions de Green, dans le même contexte, et en prenant en compte des conditions aux bords ou aux limites.

Le schéma général est le suivant : étant donné un opérateur différentiel linéaire D , une fonction de Green G de D satisfait

$$DG = \delta.$$

Par transformation de Fourier, cette équation devient

$$\widehat{DG} = H\widehat{G} = \frac{1}{\sqrt{2\pi}},$$

où H est une distribution, et c'est cette dernière équation qu'il faut ensuite résoudre au sens des distributions.

8.7.1 Exemple simple : fonction de Green de l'opérateur de dérivation dans \mathcal{D}'_+

Reprenons l'exemple de l'opérateur $D = d/dt$ en dimension 1. Par transformation de Fourier, on aboutit à l'équation

$$i\omega \widehat{G}(\omega) = \frac{1}{\sqrt{2\pi}},$$

d'où on déduit

$$\widehat{G}(\omega) = \frac{1}{i\sqrt{2\pi}} \text{v.p.} \left(\frac{1}{\omega} \right) + c\delta,$$

pour une certaine constante c . Or nous savons d'après l'Exemple 8.26 que \widehat{G} est donc égal à la transformée de Fourier de la distribution de Heaviside, à un terme additif de la forme $c'\delta$ près. Donc G est de la forme $G = T_{\Theta} + c''$ pour une certaine constante c'' . Parmi toutes les valeurs possibles de c'' , la seule qui assure $G \in \mathcal{D}'_+$ est $c'' = 0$. Donc l'unique fonction de Green $G \in \mathcal{D}'_+$ du problème est la distribution de Heaviside

$$G = T_{\Theta}. \quad (8.67)$$

8.7.2 Oscillateur harmonique 1D

Considérons le cas de l'oscillateur harmonique, défini comme solution de l'équation différentielle

$$f''(t) + \omega_0^2 f(t) = u(t). \quad (8.68)$$

Par définition, une fonction de Green associée à cette équation doit donc satisfaire l'équation

$$G'' + \omega_0^2 G = (\delta'' + \omega_0^2 \delta) * G = \delta, \quad (8.69)$$

ce qui en passant dans le domaine de Fourier, conduit à

$$(\omega_0^2 - \omega^2) \widehat{G}(\omega) = \frac{1}{\sqrt{2\pi}}. \quad (8.70)$$

A ce point, on pourrait conclure que $\widehat{G}(\omega) = 1/(\omega_0^2 - \omega^2)$, et terminer le calcul par transformation de Fourier inverse, mais c'est aller un peu vite en besogne. En effet, au sens des distributions, l'équation (8.70) n'admet pas une solution unique, comme le montre la Proposition 8.6. La solution est de la forme

$$\begin{aligned} \widehat{G}(\omega) &= \frac{1}{\sqrt{2\pi}} \text{v.p.} \left(\frac{1}{\omega_0^2 - \omega^2} \right) + \alpha \delta_{\omega_0} + \beta \delta_{-\omega_0} \\ &= \frac{1}{2\omega_0 \sqrt{2\pi}} \left(-\text{v.p.} \left(\frac{1}{\omega - \omega_0} \right) + \text{v.p.} \left(\frac{1}{\omega + \omega_0} \right) \right) + \alpha \delta_{\omega_0} + \beta \delta_{-\omega_0}, \end{aligned}$$

pour certaines constantes $\alpha, \beta \in \mathbb{C}$. Notons que ces deux constantes sont précisément les constantes d'intégration qui apparaissent lorsque l'on veut résoudre une équation différentielle du second ordre. On obtient de la sorte une famille paramétrique de fonctions de Green, définies pour l'instant dans le domaine de Fourier. Il faut maintenant en calculer les transformées de Fourier inverses. Pour cela, nous avons déjà vu dans l'Exemple 8.16 que $\text{v.p.}(1/\omega)$ est la transformée de Fourier de la distribution $i\text{sgn}/2$. On a donc

$$G(t) = \frac{1}{2\omega_0} \sin(\omega_0 t) \text{sgn}(t) + \alpha' e^{i\omega_0 t} + \beta' e^{-i\omega_0 t}$$

pour certaines constantes $\alpha', \beta' \in \mathbb{C}$.

Il est maintenant temps de se préoccuper des conditions aux bords. On peut par exemple voir que le choix $\alpha' = -\beta' = 1/2\omega_0$ conduit à une fonction de Green causale $G_{\text{ret}} \in \mathcal{D}'_+$, appelée **fonction de Green retardée**

$$G_{\text{ret}}(t) = \frac{1}{\omega_0} \Theta(t) \sin(\omega_0 t) \quad (8.71)$$

alors que le choix $\alpha' = -\beta' = -1/2\omega_0$ produit la **fonction de Green avancée** $G_{\text{av}} \in \mathcal{D}'_-$, qui est anticausale

$$G_{\text{av}}(t) = -\frac{1}{\omega_0} \Theta(-t) \sin(\omega_0 t) \quad (8.72)$$

La fonction de Green retardée est précisément celle que nous avons obtenue plus haut en utilisant la transformation de Laplace.

REMARQUE 8.4 Nous nous sommes spécialisés ici sur les problèmes de type *problème de Cauchy*, c'est à dire des problèmes dans lesquels l'équation considérée est complétée par des données initiales. Des techniques similaires s'appliquent pour des problèmes aux bords, par exemple des problèmes dans lesquels l'équation différentielle est considérée sur un intervalle, et complétée par des conditions aux bords. Les conditions aux bords peuvent alors être utilisées pour fixer les valeurs des constantes α et β ci-dessus.

8.7.3 Un exemple multidimensionnel : l'opérateur de la chaleur

Nous avons déjà vu l'utilisation de la transformation de Laplace pour résoudre l'équation de la chaleur. Nous allons maintenant étudier l'approche basée sur les distributions.

On considère l'opérateur différentiel L agissant sur les fonctions de deux variables par

$$Lu(x, t) = \frac{\partial u}{\partial t}(x, t) - \alpha \frac{\partial^2 u}{\partial x^2}(x, t) \quad (8.73)$$

Soit G une fonction de Green de L : on a donc

$$\frac{\partial G}{\partial t}(x, t) - \alpha \frac{\partial^2 G}{\partial x^2}(x, t) = \delta(x)\delta(t). \quad (8.74)$$

Introduisons une transformation de Fourier en temps et en espace. En notant respectivement ω et k les fréquences temporelle et spatiale, la transformée de Fourier \hat{G} de G est solution de

$$(i\omega + \alpha k^2)\hat{G}(k, \omega) = \frac{1}{2\pi}. \quad (8.75)$$

Donc, à une distribution solution de $(i\omega + \alpha k^2)\hat{G}(k, \omega) = 0$ près, la solution est de la forme

$$\hat{G}(k, \omega) = \frac{1}{2\pi} \frac{1}{i\omega + \alpha k^2} = \frac{1}{2i\pi} \frac{1}{\omega - i\alpha k^2}. \quad (8.76)$$

La transformation de Fourier inverse par rapport à la variable ω peut se calculer en utilisant le théorème des résidus. Sans entrer dans les détails, notons que pour $t < 0$, la présence d'un terme $e^{i\omega t}$ dans l'intégrand force à prolonger le contour d'intégration par un demi-cercle dans le demi-plan complexe inférieur, à l'intérieur duquel l'intégrand est holomorphe. Ainsi,

$$G(x, t) = 0 \quad \forall t < 0. \quad (8.77)$$

Pour $t > 0$ on doit prolonger par un demi-cercle dans le demi-plan complexe supérieur, à l'intérieur duquel l'intégrand admet un pôle simple en $\omega = i\alpha k^2$. En notant $\tilde{G}(k, t)$ la transformée de Fourier temporelle inverse, le théorème des résidus¹ fournit donc

$$\tilde{G}(k, t) = \frac{1}{2i\pi} 2i\pi e^{-\alpha k^2 t} = e^{-\alpha k^2 t}.$$

¹La vérification du fait que l'intégrale sur le demi-cercle tend vers 0 quand le rayon tend vers l'infini est laissée en exercice au lecteur... qui ne demande que ça.

Il suffit maintenant de calculer la transformée de Fourier inverse de \tilde{G} :

$$G(x, t) = -\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{G}(k, t) e^{ikx} dk = \frac{1}{\sqrt{2\alpha t}} e^{-x^2/4\alpha t}, \quad (8.78)$$

une expression déjà vue quelque part.

La solution de l'équation aux dérivées partielles

$$\frac{\partial u}{\partial t}(x, t) - \alpha \frac{\partial^2 u}{\partial x^2}(x, t) = v$$

prend finalement la forme d'un produit de convolution bidimensionnel

$$u(x, t) = (G * v)(x, t) = \int_{-\infty}^{\infty} \int_0^{\infty} G(x', t') v(x - x', t - t') dt' dx'.$$

Il faut noter que cette solution correspond à une condition initiale nulle : $\lim_{t \rightarrow 0} u(x, t) = 0$. Pour tenir compte d'une condition initiale non-nulle, il faut ajouter une solution u_0 de l'équation homogène, satisfaisant une condition initiale donnée

$$u_0(x, 0) = T_0(x).$$

Ce problème peut lui aussi être résolu grâce aux distributions. Introduisons une solution sous forme de distribution régulière (notée comme fonction pour simplifier) $U_0(x, t) = u_0(x, t)\Theta(t)$. La distribution U_0 satisfait

$$\frac{\partial U_0}{\partial t}(x, t) - \alpha \frac{\partial^2 U_0}{\partial x^2}(x, t) = T_0(x)\delta(t)$$

d'où on déduit la solution

$$U_0(x, t) = \int_{-\infty}^{\infty} G(x', t) T_0(x - x') dx',$$

et finalement

$$v(x, t) = \int_{-\infty}^{\infty} \int_0^{\infty} G(x', t') v(x - x', t - t') dt' dx' + \int_{-\infty}^{\infty} G(x', t) T_0(x - x') dx'. \quad (8.79)$$

Cinquième partie

Opérateurs linéaires

Opérateurs linéaires

9.1 Introduction, rappels

Dans cette section, on revoit les notions de base d'algèbre linéaire, en se plaçant sur le corps des complexes \mathbb{C} , en dimension finie et infinie, avec toutefois l'accent sur la dimension finie.

Commençons par le premier rappel suivant :

Soient E, F deux espaces vectoriels sur \mathbb{C} . Une application $f : E \rightarrow F$ est linéaire si pour tous $x, y \in E$ et tous $\lambda, \mu \in \mathbb{C}$, on a $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$. Une application $f : E \rightarrow F$ est antilinéaire si pour tous $x, y \in E$ et tous $\lambda, \mu \in \mathbb{C}$, on a $f(\lambda x + \mu y) = \bar{\lambda}f(x) + \bar{\mu}f(y)$.

9.1.1 Formes sesquilinéaires, formes Hermitiennes

DÉFINITION 9.1 Soit E un espace vectoriel sur \mathbb{C} .

1. Une forme sesquilinéaire sur E est une application

$$\gamma : E \times E \rightarrow \mathbb{C},$$

qui est antilinéaire par rapport à la première variable, et linéaire par rapport à la seconde.

En d'autres termes, pour tous $x, y, z \in E$ et $\lambda, \mu \in \mathbb{C}$, on a

$$\gamma(\lambda x + \mu y, z) = \bar{\lambda}\gamma(x, z) + \bar{\mu}\gamma(y, z), \quad (9.1)$$

$$\gamma(x, \lambda y + \mu z) = \lambda\gamma(x, y) + \mu\gamma(x, z). \quad (9.2)$$

2. Une forme sesquilinéaire γ sur E est Hermitienne si pour tous $x, y \in E$,

$$\gamma(y, x) = \overline{\gamma(x, y)}. \quad (9.3)$$

EXEMPLE 9.1 les exemples suivants donnent un premier aperçu de situations possibles.

1. Soit $E = \mathbb{C}^2$, muni de la base canonique. Pour tous $x, y \in E$, caractérisés par leurs coordonnées

par rapport à la base : $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ et $\underline{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, on définit

$$\gamma(x, y) = \bar{x}_1 y_1 + \bar{x}_2 y_2.$$

On vérifie facilement qu'il s'agit d'une forme Hermitienne.

2. Avec les mêmes notations, toujours dans $E = \mathbb{C}^2$, on définit

$$\gamma(x, y) = \bar{x}_1 y_1 + 3\bar{x}_2 y_2 + i\bar{x}_1 y_2 - i\bar{x}_2 y_1 = \bar{\underline{x}}^t \begin{pmatrix} 1 & i \\ -i & 3 \end{pmatrix} \underline{y}.$$

De nouveau, on vérifie facilement qu'il s'agit d'une forme Hermitienne.

3. Passons maintenant à un exemple en dimension infinie. On note $C_0(\mathbb{Z})$ l'ensemble des suites (infinies) $\{u_n, n \in \mathbb{Z}\}$ à support fini :

$$C_0(\mathbb{Z}) = \{u : \mathbb{Z} \rightarrow \mathbb{C}, \exists N \in \mathbb{Z}^* | u_n = 0 \text{ si } |n| > N\} .$$

On définit alors $\gamma = C_0(\mathbb{Z}) \times C_0(\mathbb{Z}) \rightarrow \mathbb{C}$ par

$$\gamma(x, y) = \sum_{n=-\infty}^{\infty} \bar{x}_n y_n, \quad x, y \in C_0(\mathbb{Z}) .$$

Il s'agit là encore d'une forme Hermitienne.

4. Considérons maintenant l'espace $C([0, 1])$ des fonctions continues dans l'intervalle $[0, 1]$. On vérifie facilement que l'application γ définie par

$$\gamma(f, g) = \int_0^1 \bar{f}(t)g(t) dt, \quad f, g \in C([0, 1])$$

est une forme Hermitienne.

5. Avec les mêmes notations, étant donnée une fonction de deux variables $a : [0, 1] \rightarrow \mathbb{C}$ bornée, l'application γ définie par

$$\gamma(f, g) = \int_0^1 a(t)\bar{f}(t)g(t) dt, \quad f, g \in C([0, 1])$$

est une forme sesquilinéaire, qui est Hermitienne si et seulement si a est à valeurs réelles.

9.1.2 Bases, matrices

Commençons par considérer le cas d'espaces vectoriels complexes de dimension finie (voir la remarque 9.3 ci dessous pour la dimension infinie). Soit E un tel espace, de dimension $\dim(E) = n$, et soit $\mathcal{B} = \{e_1, \dots, e_n\}$ une base de E . Soit $\gamma : E \times E \rightarrow \mathbb{C}$ une forme sesquilinéaire sur E .

Soient $x, y \in E$, quelconques, et soient $\underline{x} = (x_1, \dots, x_n)^t$ et $\underline{y} = (y_1, \dots, y_n)^t$ leurs vecteurs colonne de coordonnées par rapport à la base \mathcal{B} . Alors on peut écrire

$$\gamma(x, y) = \gamma\left(\sum_{i=1}^n x_i e_i, \sum_{j=1}^n y_j e_j\right) = \sum_{i=1}^n \sum_{j=1}^n \bar{x}_i y_j \gamma(e_i, e_j) = \bar{\underline{x}}^t \Gamma \underline{y}, \quad (9.4)$$

où Γ est la matrice dont les éléments sont donnés par

$$\Gamma_{ij} = \gamma(e_i, e_j). \quad (9.5)$$

REMARQUE 9.1 La forme sesquilinéaire γ est Hermitienne si et seulement si sa matrice dans la base considérée est telle que

$$\bar{\Gamma}^t = \Gamma .$$

On dit que la matrice Γ est *Hermitienne*.

Par exemple, dans les exemples (1) et (2) de l'exemple 9.1, les matrices, respectivement

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{et} \quad \begin{pmatrix} 1 & i \\ -i & 3 \end{pmatrix},$$

sont Hermitiennes.

Il est important de se souvenir que la matrice d'une forme sesquilinéaire dépend de la base choisie. Lorsque l'on passe d'une base à une autre, la matrice change, comme l'exprime la remarque suivante.

REMARQUE 9.2 Soit γ une forme sesquilinéaire sur E , soit $\mathcal{B} = \{e_1, \dots, e_n\}$ une base de E , et soit Γ la matrice de γ par rapport à cette base. Soit $\mathcal{B}' = \{e'_1, \dots, e'_n\}$ une autre base de E . On rappelle que les coordonnées \underline{x} d'un vecteur $x \in E$ par rapport à la base \mathcal{B} se transforment en coordonnées \underline{x}' par rapport à la base \mathcal{B}' comme

$$\underline{x}' = P^{-1}\underline{x}, \quad (9.6)$$

où $P = P_{\mathcal{B} \rightarrow \mathcal{B}'}$ est la matrice de passage de la base \mathcal{B} à la base \mathcal{B}' , dont les colonnes sont les composantes des vecteurs e'_i dans la base \mathcal{B} . En d'autres termes,

$$\begin{cases} e'_1 &= P_{11}e_1 + P_{21}e_2 + \dots + P_{n1}e_n, \\ e'_2 &= P_{12}e_1 + P_{22}e_2 + \dots + P_{n2}e_n, \\ \dots &= \dots \\ e'_n &= P_{1n}e_1 + P_{2n}e_2 + \dots + P_{nn}e_n. \end{cases} \quad (9.7)$$

On peut alors écrire, pour tous $x, y \in E$,

$$\gamma(x, y) = \overline{\underline{x}'}^t \Gamma \underline{y} = \overline{\underline{x}'}^t \overline{P}^t \Gamma P \underline{y}',$$

de sorte que la matrice de γ dans la base \mathcal{B}' prend la forme

$$\Gamma' = \overline{P}^t \Gamma P. \quad (9.8)$$

REMARQUE 9.3 Le cas des espaces de dimension infinie. Si l'espace vectoriel considéré est un espace de dimension infinie, et étant donnée une base de cet espace $\{e_1, e_2, \dots\}$, il est encore possible d'associer une matrice Γ à une forme sesquilinéaire $\gamma : E \times E \rightarrow \mathbb{C}$. Cette matrice est toujours constituée des images des vecteurs de base, comme en (9.5), mais il s'agit cette fois d'une matrice infinie. Les manipulations effectuées dans cette section se transposent facilement à ce nouveau cas, à condition toutefois de vérifier que les opérations sont bien définies.

9.1.3 Rang et noyau

Dans cette section, on se limite de nouveau au cas d'espaces vectoriels complexes de dimension finie.

- DÉFINITION 9.2**
1. Soit γ une forme sesquilinéaire sur E . Le rang de γ , noté $\text{rg}(\gamma)$, est défini comme le rang de la matrice Γ , c'est à dire la dimension de l'espace engendré par les vecteurs-colonne de Γ .
 2. γ est non-dégénérée si son rang est égal à la dimension de E , ou de façon équivalente, si $\det(\Gamma) \neq 0$.

Cette définition peut donner l'impression que le rang d'une forme sesquilinéaire dépend de la base dans laquelle la matrice est calculée. Il n'en est rien, en effet les matrices de passage étant toujours inversibles, l'équation (9.8) montre que pour toute autre base \mathcal{B}' , le rang de la matrice associée Γ' est égal au rang de Γ .

DÉFINITION 9.3 Le noyau de la forme sesquilinéaire $\gamma : E \times E \rightarrow \mathbb{C}$ est défini par

$$N(\gamma) = \{y \in E, \gamma(x, y) = 0 \forall x \in E\}. \quad (9.9)$$

REMARQUE 9.4 Nous avons déjà remarqué que dans une base quelconque de E , on a $\gamma(x, y) = \underline{x}^t \Gamma \underline{y}$. Donc, dire que $\gamma(x, y) = 0$ pour tout $x \in E$ revient à dire que $\Gamma \underline{y} = 0$, de sorte que la définition du noyau de γ coïncide avec la définition du noyau de la matrice associée Γ .

Le dernier résultat de cette section, donné sans démonstration, est un résultat fondamental de l'algèbre linéaire.

THÉORÈME 9.1 Soit γ une forme sesquilinéaire sur l'espace vectoriel complexe E . Alors on a

$$\dim(E) = \text{rg}(\gamma) + \dim(N(\gamma)). \quad (9.10)$$

9.1.4 Réduction d'une forme Hermitienne

Parmi les bases d'un espace vectoriel sur \mathbb{C} , certaines sont particulièrement bien adaptées à une forme Hermitienne donnée.

DÉFINITION 9.4 Une base $\mathcal{B} = \{e_1, e_2, \dots\}$ de l'espace vectoriel complexe E est orthogonale pour la forme Hermitienne γ si pour tous i, j , on a

$$\gamma(e_i, e_j) = 0, \quad \text{si } i \neq j. \quad (9.11)$$

Dans une telle base, la matrice Γ de γ prend une forme diagonale :

$$\Gamma = \begin{pmatrix} a_1 & 0 & 0 & \dots \\ 0 & a_2 & 0 & \dots \\ 0 & 0 & a_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

les nombres a_i étant des réels (car γ est supposée Hermitienne). En d'autres termes, étant donnés $x, y \in E$, en notant $\underline{x} = (x_1, x_2, \dots)^t$ et $\underline{y} = (y_1, y_2, \dots)^t$ leurs coordonnées par rapport à cette base, on peut écrire,

$$\gamma(x, y) = \sum_i a_i \bar{x}_i y_i.$$

Un autre résultat fondamental de l'algèbre linéaire montre qu'en dimension finie, il existe toujours de telles bases. Plus précisément,

9.1.5 Espaces pré-Hilbertiens, espaces Hermitiens

- DÉFINITION 9.5**
1. Une forme Hermitienne γ sur E est positive si $\gamma(x, x) \geq 0$ pour tout $x \in E$.
 2. Elle est définie si $\gamma(x, x) = 0$ implique $x = 0$.
 3. Si γ n'est pas définie, un vecteur $x \in E$ tel que $\gamma(x, x) = 0$ est dit isotrope.

Notons qu'une forme Hermitienne définie est non-dégénérée.

- DÉFINITION 9.6**
1. Un espace pré-Hilbertien est un espace vectoriel sur \mathbb{C} , muni d'une forme Hermitienne définie positive.
 2. Un espace pré-Hilbertien de dimension finie est appelé espace Hermitien.

On utilise généralement une notation spécifique lorsque l'on se place dans un cadre d'espace pré-Hilbertien. La forme Hermitienne correspondante, appelée *produit scalaire Hermitien* est notée $\langle \cdot | \cdot \rangle$ et on note aussi

$$\| \cdot \| : x \rightarrow \|x\| = \sqrt{\langle x | x \rangle}.$$

REMARQUE 9.5 Cette application est appelée *norme* associée au produit Hermitien. Elle possède en effet les propriétés d'une norme :

1. $\|\lambda x\| = |\lambda| \|x\|$, pour tous $x \in E$ et $\lambda \in \mathbb{C}$.
2. $\|x\| = 0$ implique $x = 0$.
3. $\|x + y\| \leq \|x\| + \|y\|$ pour tous $x, y \in E$ (inégalité de Minkowsky).

Par ailleurs, la connaissance de la norme est suffisante pour caractériser le produit scalaire Hermitien, grâce à l'identité de polarisation

$$\langle x | y \rangle = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2 - i\|x + iy\|^2 + i\|x - iy\|^2).$$

Une propriété essentielle des espaces pré-Hilbertiens est l'inégalité de Cauchy-Schwarz :

PROPOSITION 9.1 Soit E un espace pré-Hilbertien. Alors pour tous $x, y \in E$, on a

$$|\langle x, y \rangle| \leq \|x\| \|y\|,$$

l'inégalité devenant une égalité si et seulement si x et y sont proportionnels.

Une grande vertu des espaces Hermitiens est qu'ils possèdent toujours une base orthonormée. En effet, d'après le théorème d'inertie de Sylvester (voir théorème 9.2 plus haut), on peut toujours trouver une base orthogonale pour la forme Hermitienne, telle que les éléments diagonaux de la matrice valent 1, -1 ou 0. La forme Hermitienne étant supposée définie positive, les valeurs -1 et 0 sont impossibles, ce qui prouve ce résultat.

9.2 Opérateurs linéaires dans les espaces Hermitiens

Les notions introduites dans ce qui suit peuvent l'être dans un cadre plus général que ce qui est fait ici. On a plutôt choisi pour simplifier de se limiter au cadre des espaces Hermitiens, dans lesquels on a l'assurance de pouvoir disposer de bases orthonormées par exemple.

9.2.1 Définitions

DÉFINITION 9.7 1. Soit \mathcal{H} un espace Hermitien. Un opérateur linéaire de \mathcal{H} (aussi appelé endomorphisme de \mathcal{H}) est une application linéaire $A : \mathcal{H} \rightarrow \mathcal{H}$.

2. La norme d'un opérateur linéaire A sur \mathcal{H} est définie par

$$\|A\|_{op} = \sup_{x \in \mathcal{H}} \frac{\|Ax\|}{\|x\|}. \quad (9.14)$$

La norme d'un opérateur A sur un espace Hermitien \mathcal{H} peut aussi s'écrire

$$\|A\|_{op} = \sup_{x, y \in \mathcal{H}} \frac{|\langle Ax | y \rangle|}{\|x\| \|y\|}.$$

Etant donné un opérateur linéaire A sur \mathcal{H} , et une base \mathcal{B} de \mathcal{H} , on peut construire la matrice $M = M_A$ de A par rapport à cette base : il s'agit de la matrice dont les colonnes sont les composantes de Ae_1, Ae_2, \dots, Ae_n par rapport à la base :

$$\begin{cases} Ae_1 &= M_{11}e_1 + M_{21}e_2 + \dots + M_{n1}e_n, \\ Ae_2 &= M_{12}e_1 + M_{22}e_2 + \dots + M_{n2}e_n, \\ \dots &= \dots \\ Ae_n &= M_{1n}e_1 + M_{2n}e_2 + \dots + M_{nn}e_n. \end{cases} \quad (9.15)$$

Considérons $x \in \mathcal{H}$, et notons $\underline{x} = (x_1, \dots, x_n)^t$ le vecteur colonne de ses coordonnées par rapport à la base \mathcal{B} . On peut donc écrire

$$x = \sum_{i=1}^n x_i e_i,$$

et donc

$$Ax = \sum_{i=1}^n x_i \sum_{j=1}^n M_{ji} e_j = \sum_{j=1}^n \left(\sum_{i=1}^n M_{ji} x_i \right) e_j = \sum_{j=1}^n (M\underline{x})_j e_j.$$

Par conséquent, les coordonnées de Ax dans la base $\mathcal{B} = \{e_1, \dots, e_n\}$ sont données par le vecteur colonne $M\underline{x}$:

$$\underline{Ax} = M_A \underline{x}.$$

REMARQUE 9.6 En prenant des produits scalaires avec les vecteurs de base, on obtient par exemple

$$\langle e_1 | Ae_1 \rangle = M_{11} \langle e_1 | e_1 \rangle + M_{21} \langle e_2 | e_1 \rangle + \dots + M_{n1} \langle e_n | e_1 \rangle = (\Gamma^t M)_{11},$$

et plus généralement, en notant \tilde{M} la matrice définie par

$$\tilde{M}_{ij} = \langle e_i | Ae_j \rangle,$$

on a

$$\tilde{M} = \Gamma^t M .$$

La situation est évidemment plus simple si la base \mathcal{B} considérée est orthonormée pour la forme Hermitienne $\langle \cdot | \cdot \rangle$. Dans ce cas, Γ est la matrice identité, et $\tilde{M} = M$.

DÉFINITION 9.8 Soit A un opérateur linéaire de \mathcal{H} . Le rang de A est la dimension de l'image $A\mathcal{H}$ de \mathcal{H} par A . Le noyau de A est le sous-espace de \mathcal{H} dont l'image par A est nulle :

$$N(A) = \{x \in \mathcal{H}, Ax = 0\} . \quad (9.16)$$

Le théorème de la dimension est un résultat fondamental :

THÉORÈME 9.3 Soit $A : E \rightarrow E$ un opérateur linéaire sur l'espace vectoriel de dimension finie E . Alors on a

$$\dim(E) = \text{rg}(A) + \dim(N(A)) . \quad (9.17)$$

Les exemples les plus simples d'opérateurs sont les opérateurs de rang 1, c'est à dire les opérateurs dont l'image est de dimension 1. La forme générale de ces opérateurs est la suivante : étant donné $u, v \in \mathcal{H}$, on construit l'opérateur suivant :

$$A(u, v) : x \in \mathcal{H} \rightarrow \langle v|x \rangle u , \quad (9.18)$$

qui sont parfois notés (en mécanique quantique notamment) comme des "dyades"

$$A(u, v) = |u\rangle\langle v| .$$

Un cas particulier important est fourni par les opérateurs de projection orthogonale de rang 1, qui correspondent au cas $v = u$, avec $\|u\| = 1$.

On peut également s'intéresser aux opérateurs de projection orthogonale de rang supérieur à 1. On peut facilement montrer la proposition suivante :

PROPOSITION 9.2 Soit \mathcal{H} un espace Hermitien, soit $F \subset \mathcal{H}$ un sous-espace de \mathcal{H} , et soit $\{f_1, \dots, f_m\}$ une base orthonormée de F . Alors l'opérateur de projection orthogonale Π_F de \mathcal{H} sur F est la somme des projecteurs orthogonaux de rang 1

$$\Pi_F = \sum_{i=1}^m A(f_i, f_i) . \quad (9.19)$$

Preuve : Soit $x \in \mathcal{H}$, quelconque, et notons $x_0 = \Pi_F x = \sum_{k=1}^m \alpha_k f_k$ son projeté orthogonal sur F . Alors on a pour tout $i = 1, \dots, m$, $x - x_0 \perp f_i$, ce qui s'écrit

$$\langle f_i | x - x_0 \rangle = 0 , \quad \forall i = 1, \dots, m .$$

En remplaçant x_0 par son expression, on aboutit à

$$\langle f_i | x \rangle = \sum_{k=1}^m \alpha_k \langle f_i | f_k \rangle = \alpha_i ,$$

d'où

$$x_0 = \Pi_F x = \sum_{k=1}^m \langle f_k | x \rangle f_k ,$$

ce qui est le résultat recherché.

9.2.2 Opérateur adjoint

La notion d'opérateur adjoint est une notion fondamentale.

THÉORÈME 9.4 Soit \mathcal{H} un espace Hermitien. Pour tout opérateur linéaire A de \mathcal{H} , il existe un unique opérateur linéaire A^* , appelé opérateur adjoint de A , tel que pour tous $x, y \in \mathcal{H}$,

$$\langle x|Ay \rangle = \langle A^*x|y \rangle. \quad (9.20)$$

Considérons une base \mathcal{B} , orthonormée pour la forme Hermitienne de \mathcal{H} . Alors la matrice de l'adjoint A^* de A dans cette base n'est autre que la conjuguée Hermitienne (en d'autres termes, le complexe conjugué de la transposée) de la matrice de A . En effet, on peut écrire

$$\langle x|Ay \rangle = \underline{\bar{x}}^t M \underline{y} = \sum_{i=1}^n \bar{x}_i \sum_{j=1}^n M_{ij} y_j = \sum_{j=1}^n \sum_{i=1}^n \overline{M_{ij} x_i} y_j = \overline{M^t \underline{x}}^t \underline{y}$$

Ceci s'écrit donc

$$M_{A^*} = \overline{M_A^t}. \quad (9.21)$$

REMARQUE 9.7 Lorsque la base considérée n'est pas orthonormée, la relation est quelque peu plus complexe : il faut cette fois écrire

$$M_{A^*} = \Gamma^{-1} \overline{M_A^t} \Gamma.$$

L'adjoint possède un certain nombre de propriétés simples, qui sont énoncées ci-après

PROPOSITION 9.3 Pour tous opérateurs linéaires A, B sur \mathcal{H} , et tout $\lambda \in \mathbb{C}$,

1. $A^{**} = A$; $1^* = 1$.
2. $(A + B)^* = A^* + B^*$; $(\lambda A)^* = \bar{\lambda} A^*$; $(AB)^* = B^* A^*$.
3. $rg(A^*) = rg(A)$.

Preuve : La plupart de ces propriétés sont démontrées de façon élémentaire. Par exemple pour la première : pour tous $x, y \in \mathcal{H}$, on a

$$\langle x|A^{**}y \rangle = \langle A^*x|y \rangle = \langle x|Ay \rangle,$$

ce qui montre bien que $A^{**} = A$. De façon similaire,

$$\langle (AB)^*x|y \rangle = \langle x|AB y \rangle = \langle A^*x|B y \rangle = \langle B^* A^*x|y \rangle.$$

EXEMPLE 9.3 1. Dans \mathbb{C}^2 , on considère l'opérateur linéaire déterminé dans la base canonique par la matrice

$$M = \begin{pmatrix} \frac{1-i}{\sqrt{3}} & \frac{-1}{\sqrt{3}} \\ \frac{-1}{\sqrt{3}} & \frac{1+i}{\sqrt{3}} \end{pmatrix}$$

L'adjoint de cet opérateur est déterminé par la matrice

$$M^* = \begin{pmatrix} \frac{1+i}{\sqrt{3}} & \frac{-1}{\sqrt{3}} \\ \frac{-1}{\sqrt{3}} & \frac{1-i}{\sqrt{3}} \end{pmatrix}.$$

2. Dans \mathbb{C}^3 , on considère l'opérateur linéaire déterminé par la matrice

$$M = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ i/\sqrt{2} & -i/\sqrt{2} & 0 \\ 0 & 0 & i \end{pmatrix} .$$

On peut aussi vérifier que M^* est également l'inverse de M , de sorte que l'on a $MM^* = M^*M = 1$. De tels opérateurs sont dits *unitaires*.

3. Soient $u, v \in \mathcal{H}$, espace Hermitien, et soit $A(u, v)$ l'opérateur de rang 1 associé, défini en (9.18). Pour tous $x, y \in \mathcal{H}$, on a

$$\langle x|A(u, v)y \rangle = \langle x|\langle v|y \rangle u \rangle = \langle v|y \rangle \langle x|u \rangle = \langle \langle u|x \rangle v|y \rangle = \langle A(v, u)x|y \rangle ,$$

de sorte que l'adjoint de l'opérateur de rang 1 $A(u, v)$ n'est autre que l'opérateur de rang 1 $A(v, u)$:

$$A(u, v)^* = A(v, u) . \tag{9.22}$$

4. De façon plus générale, étant donnés un espace Hermitien \mathcal{H} et un sous-espace $F \subset \mathcal{H}$, la projection orthogonale Π_F de \mathcal{H} sur F (voir Proposition 9.2) est auto-adjointe :

$$\Pi_F^* = \Pi_F .$$

9.2.3 Diagonalisation des opérateurs autoadjoints

On rappelle que $v \in E$ est vecteur propre de l'opérateur A si il vérifie

$$Av = \lambda v ,$$

pour un certain $\lambda \in \mathbb{C}$. λ est appelé valeur propre de A . Le *sous-espace propre* associé à λ est l'ensemble des vecteurs propres de A de valeur propre λ :

$$E_\lambda = \{v \in E, Av = \lambda v\} .$$

A est diagonalisable s'il existe une base $\{e_1, \dots\}$ telle que la représentation de A dans cette base soit diagonale.

$$M_A = \begin{pmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \end{pmatrix} \tag{9.23}$$

DÉFINITION 9.9 Un opérateur linéaire A sur \mathcal{H} est auto-adjoint si $A^* = A$, c'est à dire si pour tous $x, y \in \mathcal{H}$,

$$\langle x|Ay \rangle = \langle Ax|y \rangle . \tag{9.24}$$

Compte tenu de la forme donnée en (9.21) pour la matrice de l'adjoint dans une base orthonormée, il est clair que A est auto-adjoint si et seulement si sa matrice M_A dans une base orthonormée est Hermitienne, c'est à dire telle que

$$\overline{M_A^t} = M_A . \tag{9.25}$$

Le résultat fondamental concernant les opérateurs linéaires auto-adjoints en dimension finie est le suivant :

THÉORÈME 9.5 Soit A un opérateur auto-adjoint sur un espace Hermitien \mathcal{H} .

1. Les valeurs propres de A sont toutes réelles.
2. A est diagonalisable.
3. Les sous-espaces propres de A sont orthogonaux deux à deux. On peut donc construire une base orthonormée de \mathcal{H} en choisissant une base orthonormée dans chaque sous-espace propre de A .

En termes de matrices et vecteurs, ce résultat est équivalent au résultat suivant.

COROLLAIRE 9.1 Toute matrice Hermitienne est diagonalisable, et possède des valeurs propres réelles. Les sous-espaces propres correspondants sont deux à deux orthogonaux pour le produit scalaire Hermitien canonique dans \mathbb{C}^n : $\langle \underline{x} | \underline{y} \rangle = \sum_1^n \bar{x}_i y_i$.

Partant du théorème 9.5, on peut donner de la formule (9.23) l'interprétation suivante : soit $x \in \mathcal{H}$, sa décomposition sur la base orthonormée qui diagonalise A est

$$x = \sum_{i=1}^n \langle e_i | x \rangle e_i, \quad (9.26)$$

et l'image de x par A s'écrit quant à elle

$$Ax = \sum_{i=1}^n a_i \langle e_i | x \rangle e_i. \quad (9.27)$$

Une telle expression porte le nom de représentation spectrale de l'opérateur A .

REMARQUE 9.8 Partant de leur représentation spectrale, il est possible de construire simplement certaines fonctions d'opérateurs. Commençons par les puissances. Étant donné un opérateur autoadjoint quelconque $A \in \mathcal{L}(\mathcal{H})$, sa puissance k -ième est obtenue en appliquant k fois consécutivement A . En utilisant (9.27), on obtient facilement la représentation spectrale du carré de A :

$$A^2 x = \sum_{i=1}^n a_i \langle e_i | x \rangle A e_i = \sum_{i=1}^n a_i^2 \langle e_i | x \rangle e_i.$$

Plus généralement, on a de même, pour tout $k \in \mathbb{N}$

$$A^k x = \sum_{i=1}^n a_i^k \langle e_i | x \rangle e_i.$$

De là, on peut également construire des fonctions plus élaborées d'opérateurs. L'exemple le plus simple est la fonction exponentielle : en écrivant

$$\exp(A) = \sum_{k=0}^{\infty} \frac{1}{k!} A^k,$$

on obtient de même la représentation spectrale de l'exponentielle de l'opérateur auto-adjoint A : pour tout $x \in \mathcal{H}$,

$$\exp(A)x = \sum_{k=0}^{\infty} \frac{1}{k!} A^k x = \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{i=1}^n a_i^k \langle e_i | x \rangle e_i = \sum_{i=1}^n \exp(a_i) \langle e_i | x \rangle e_i.$$

Partant de là, on obtient également les représentations spectrales d'autres fonctions d'opérateurs, pourvu que ces fonctions puissent être développées en série entière, et que les problèmes de convergence de ces dernières puissent être évités. Nous n'entrerons pas dans les détails de ces problèmes ici.

La norme des opérateurs auto-adjoints prend une forme particulièrement simple. En effet, on a

PROPOSITION 9.4 Soit A un opérateur auto-adjoint sur l'espace Hermitien \mathcal{H} , et soit λ_{max} la plus grande (en valeur absolue) de ses valeurs propres. Alors

$$\|A\|_{op} = |\lambda_{max}|. \quad (9.28)$$

Preuve : Supposons que A ait été diagonalisé, notons a_1, \dots, a_N ses valeurs propres (réelles, car A est auto-adjoint). Partant de la représentation spectrale (9.27), on peut écrire pour tout $x = \sum_{i=1}^n \alpha_i e_i$

$$\frac{\|Ax\|^2}{\|x\|^2} = \frac{\sum_{i=1}^n a_i^2 |\alpha_i|^2}{\sum_{i=1}^n |\alpha_i|^2}.$$

Supposons sans perte de généralité que la plus grande valeur propre (en valeur absolue) soit a_1 , et écrivons

$$\frac{\|Ax\|^2}{\|x\|^2} = a_1^2 \frac{\sum_{i=1}^n (a_i/a_1)^2 |\alpha_i|^2}{\sum_{i=1}^n |\alpha_i|^2} = a_1^2 + \frac{\sum_{i=1}^n ((a_i/a_1)^2 - 1) |\alpha_i|^2}{\sum_{i=1}^n |\alpha_i|^2}.$$

Le second terme est toujours négatif ou nul, de sorte que l'on a

$$\frac{\|Ax\|^2}{\|x\|^2} \leq a_1^2,$$

et l'inégalité devient une égalité si et seulement si $\alpha_i = 0$ pour tout i tel que $|a_i| < |a_1|$. Donc $\|A\|_{op} = |a_1|$, ce qui prouve la proposition.

EXEMPLE 9.4 1. Dans un espace Hermitien \mathcal{H} , soit $u \in \mathcal{H}$, tel que $\|u\| = 1$, et soit $A(u, u)$ le projecteur orthogonal sur $\mathbb{C}u \subset \mathcal{H}$. Si on note \mathcal{K} le complément orthogonal de $\mathbb{C}u$ dans \mathcal{H} , c'est à dire le sous-espace de \mathcal{H} engendré par les vecteurs orthogonaux à u , on a pour tout $v \in \mathbb{C}u$, $A(u, u)v = v$, et pour tout $v \in \mathcal{K}$, $A(u, u)v = 0$. Par conséquent, $A(u, u)$ admet 0 et 1 pour valeurs propres, et la décomposition

$$\mathcal{H} = \mathbb{C}u \oplus \mathcal{K}$$

est la décomposition de \mathcal{H} en sous-espaces propres correspondants. Il est clair que $\|A(u, u)\|_{op} = 1$.

2. Similairement, soit $\mathcal{F} \subset \mathcal{H}$ un sous-espace de \mathcal{H} , et soit $\Pi_{\mathcal{F}}$ le projecteur orthogonal de \mathcal{H} sur \mathcal{F} . Soit \mathcal{K} le complément orthogonal de \mathcal{F} dans \mathcal{H} . Alors pour tout $v \in \mathcal{F}$, $\Pi_{\mathcal{F}}v = v$, et pour tout $v \in \mathcal{K}$, $\Pi_{\mathcal{F}}v = 0$. Le projecteur orthogonal $\Pi_{\mathcal{F}}$ admet 0 et 1 pour valeurs propres, et la décomposition

$$\mathcal{H} = \mathcal{F} \oplus \mathcal{K}$$

est la décomposition de \mathcal{H} en sous-espaces propres correspondants. De nouveau $\|\Pi_{\mathcal{F}}\|_{op} = 1$.

3. Considérons le cas de $\mathcal{H} = \mathbb{C}^2$, et soit $A \in \mathcal{L}(\mathbb{C}^2)$, défini par sa matrice dans la base canonique :

$$M_A = \begin{pmatrix} 2 & 1+i \\ 1-i & 2 \end{pmatrix}.$$

A est clairement auto-adjoint (car M_A possède la symétrie Hermitienne : $M_A^t = \overline{M_A}$), et est donc diagonalisable. Chercher les vecteurs propres et valeurs propres de A revient à diagonaliser M_A , donc à rechercher les racines (en λ) du polynôme caractéristique

$$P(\lambda) = \det(M_A - \lambda \mathbf{1}) = (2 - \lambda)^2 - (1 + i)(1 - i) = \lambda^2 - 4\lambda + 2.$$

Les racines sont $\lambda_{\pm} = 2 \pm \sqrt{2}$, donc $\|A\|_{op} = \lambda_+ = 2 + \sqrt{2}$. Les vecteurs propres correspondants ont pour coordonnées dans la base canonique

$$e_{\pm} = \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{1+i}{\sqrt{2}} \\ \pm 1 \end{pmatrix}$$

On peut écrire

$$M_A = P \begin{pmatrix} 2 + \sqrt{2} & 0 \\ 0 & 2 - \sqrt{2} \end{pmatrix} P^{-1},$$

où P est la matrice de passage, dont les colonnes ne sont autres que les vecteurs e_{\pm} :

$$P = \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{1+i}{\sqrt{2}} & \frac{1+i}{\sqrt{2}} \\ 1 & -1 \end{pmatrix}.$$

Finalement, la représentation spectrale de A s'écrit

$$Ax = (2 + \sqrt{2})\langle e_+ | x \rangle e_+ + (2 - \sqrt{2})\langle e_- | x \rangle e_-.$$

9.2.4 Diagonalisation simultanée d'opérateurs auto-adjoints qui commutent

Etant donnés deux opérateurs sur un espace Hermitien, ils ne sont généralement pas diagonaux dans la même base, sauf dans certaines situations particulières. Considérons tout d'abord le cas de deux opérateurs auto-adjoints A et B sur \mathcal{H} , et soit

$$[A, B] = AB - BA$$

leur commutateur. Supposons que $[A, B] = 0$, soit $a \in \mathbb{R}$ une valeur propre quelconque de A , soit

$$\mathcal{H}_a = \{x \in \mathcal{H}, Ax = ax\}$$

le sous-espace propre associé, et soit $x \in \mathcal{H}_a$ un vecteur propre de A , de valeur propre a . Calculons

$$ABx = BAx = aBx.$$

Nous avons montré que Bx est lui aussi vecteur propre de A avec la même valeur propre. En d'autres termes, le sous-espace propre de A de valeur propre a est *stable* par B . En diagonalisant la restriction de B à ce sous-espace propre, nous obtenons donc une base orthonormée de \mathcal{H}_a , formée de vecteurs propres *simultanés* de A et B . Ceci étant vrai pour toute valeur propre a de A , on a donc montré le résultat important suivant :

THÉORÈME 9.6 Soient A et B deux opérateurs auto-adjoints sur \mathcal{H} , tels que $[A, B] = 0$. Alors il existe une base orthonormée de \mathcal{H} formée de vecteurs propres simultanés de A et B .

Ce résultat, et sa généralisation aux opérateurs auto-adjoints en dimension infinie, prend toute son importance en mécanique quantique, où la diagonalisation simultanée d'observables (opérateurs auto-adjoints) qui commutent permet de spécifier l'état quantique d'un système. Nous considérerons des exemples plus loin.

9.3 Opérateurs linéaires dans les espaces de Hilbert

9.3.1 Généralités

Après avoir analysé le cas de la situation finie, passons maintenant au cas plus général.

DÉFINITION 9.10 *Un espace de Hilbert est un espace pré-Hilbertien, complet pour la norme issue du produit Hermitien.*

En d'autres termes, on impose que toute suite de Cauchy dans un espace de Hilbert converge dans cet espace.

REMARQUE 9.9 1. Le cas de la dimension finie est le plus simple : tout espace Hermitien est un espace de Hilbert.

2. L'exemple classique d'espace de Hilbert est l'exemple des espaces de fonctions de module carré intégrable. Par exemple, l'espace

$$L^2([0, 1]) = \left\{ f : [0, 1] \rightarrow \mathbb{C}, \int_0^1 |f(t)|^2 dt < \infty \right\},$$

muni du produit Hermitien

$$\langle f | g \rangle = \int_0^1 \bar{f}(t)g(t) dt, \quad f, g \in L^2([0, 1])$$

est un espace de Hilbert.

3. Un exemple d'espace pré-Hilbertien qui n'est pas un espace de Hilbert est fourni par l'espace $C([0, 1])$ des fonctions continues sur $[0, 1]$, muni du même produit Hermitien que $L^2([0, 1])$. En effet, considérons la suite de fonctions $f_n \in C([0, 1])$, définies par $f_n(t) = \inf(n, t^{-1/3})$. Un calcul simple montre que pour tout n ,

$$\|f_{n+p} - f_n\|^2 \leq \frac{3}{n},$$

de sorte qu'il s'agit bien d'une suite de Cauchy. Par contre, il est aussi facile de voir que lorsque $n \rightarrow \infty$, f_n tend vers $t \rightarrow t^{-1/3}$, qui n'est pas continue en $t = 0$. Ainsi, cette suite ne converge pas dans $C([0, 1])$, et cet espace n'est donc pas un espace de Hilbert.

REMARQUE 9.10 On se limitera ici au cas des espaces de Hilbert séparables, c'est à dire admettant au moins une base orthonormée. C'est un cadre suffisamment large pour ce dont nous avons besoin. On peut par exemple montrer que tous les espaces de fonctions de module carré intégrable sont des espaces de Hilbert séparables.

Comme plus haut, étant donné un espace de Hilbert \mathcal{H} , nous appellerons opérateur sur \mathcal{H} une application linéaire

$$A : \mathcal{H} \rightarrow \mathcal{H}$$

bornée, au sens où sa norme

$$\|A\|_{op} = \sup_{x \in \mathcal{H}} \frac{\|Ax\|}{\|x\|} = \sup_{x, y \in \mathcal{H}} \frac{|\langle x | Ay \rangle|}{\|x\| \|y\|}$$

est finie. On note $\mathcal{L}(\mathcal{H})$ l'ensemble des opérateurs sur un espace de Hilbert \mathcal{H} . On montre facilement qu'étant donnés deux opérateurs $A, B \in \mathcal{L}(\mathcal{H})$, leur somme et leur produit (ou leur composition) est également un opérateur de $\mathcal{L}(\mathcal{H})$. De plus, on a l'inégalité

$$\|AB\|_{op} \leq \|A\|_{op}\|B\|_{op}.$$

EXEMPLE 9.5 1. Opérateurs scalaires : étant donné un espace de Hilbert \mathcal{H} , on considère l'opérateur $A \in \mathcal{L}(\mathcal{H})$ défini par

$$Ax = \lambda x, \quad \forall x \in \mathcal{H},$$

où $\lambda \in \mathbb{C}$ est une constante. On vérifie facilement que $\|A\|_{op} = |\lambda|$.

2. Opérateurs de rang 1 : ces opérateurs sont définis similairement à leurs analogues en dimension finie. Étant donnés $u, v \in \mathcal{H}$, on note $A(u, v)$ l'opérateur défini par

$$A(u, v)x = \langle v|x \rangle u, \quad \forall x \in \mathcal{H}.$$

Il est clair que si $u \neq 0$ et $v \neq 0$, $\dim(A(u, v)\mathcal{H}) = \dim(\mathbb{C}u) = 1$, il s'agit donc d'un opérateur de rang 1.

3. Opérateurs de rang fini et projecteurs orthogonaux : Les opérateurs de rang fini, qui jouent un rôle important dans la théorie des opérateurs compacts (voir plus bas), sont définis comme sommes d'opérateurs de rang 1 : partant d'une suite de vecteurs $u_1, \dots, u_N, v_1, \dots, v_N \in \mathcal{H}$, on construit $A : \mathcal{H} \rightarrow \mathcal{H}$, défini par

$$A = \sum_{n=1}^N A(u_n, v_n) : x \in \mathcal{H} \rightarrow \sum_{n=1}^N \langle v_n|x \rangle u_n.$$

Un cas particulier est fourni par le cas dans lequel $u_n = v_n$ pour tout n , et les vecteurs u_n forment une base orthonormée de l'espace qu'ils engendrent. On montre alors facilement qu'un tel opérateur A est le projecteur orthogonal sur l'espace en question.

4. Opérateurs définis par une matrice infinie : Supposons que $\mathcal{B} = \{e_1, e_2, \dots\}$ soit une base orthonormée de l'espace de Hilbert \mathcal{H} , et soit $A \in \mathcal{L}(\mathcal{H})$. A est caractérisé par les images des vecteurs de base. Plus précisément, pour tout

$$v = \sum_{i=1}^{\infty} \langle e_i|v \rangle e_i \in \mathcal{H},$$

on a

$$Av = \sum_{i=1}^{\infty} \langle e_i|v \rangle Ae_i,$$

et

$$Av = \sum_{j=1}^{\infty} \langle e_j|Av \rangle e_j = \sum_{j=1}^{\infty} \left(\sum_{i=1}^{\infty} \langle e_j|Ae_i \rangle \langle e_i|v \rangle \right) e_j.$$

Ainsi, A est caractérisé par la (double) suite de nombres complexes

$$M_{ij} = \langle e_j|Ae_i \rangle,$$

qui forme la matrice (infinie) M_A de A dans la base orthonormée choisie.

Inversement, soit $M = \{M_{ij}, i, j = 1, 2, \dots\}$ une famille infinie de nombres complexes. Il est tentant d'introduire un opérateur A sur \mathcal{H} , qui serait défini par $Ae_i = \sum_{j=1}^{\infty} M_{ji}e_j$. Ceci n'est malheureusement pas possible en général, car une telle suite produit généralement un opérateur A qui ne serait pas borné.

5. Opérateurs de translation : Considérons l'espace $L^2(\mathbb{R})$ des fonctions de module carré intégrable sur \mathbb{R} , et soit $a \in \mathbb{R}$. On lui associe l'opérateur T_a de translation par a , qui associe à toute fonction $f \in L^2(\mathbb{R})$ sa tradatée

$$(T_a f)(t) = f(t - a).$$

Il est immédiat de vérifier que pour tout $f \in L^2(\mathbb{R})$, $\|T_a f\| = \|f\|$, (on dit que T_a est une isométrie), et on a donc $\|T_a\|_{op} = 1$.

6. Opérateurs de multiplication : Toujours dans $L^2(\mathbb{R})$, soit $m : t \in \mathbb{R} \rightarrow m(t)$ une fonction bornée (c'est à dire telle qu'il existe $C > 0$ tel que $|m(t)| \leq C$ pour tout t). On associe à cette fonction l'opérateur de multiplication M_m sur $L^2(\mathbb{R})$, défini par

$$(M_m f)(t) = m(t)f(t), \forall f \in L^2(\mathbb{R}).$$

Il est facile de vérifier que la norme d'un tel opérateur vaut

$$\|M_m\|_{op} = \sup_{t \in \mathbb{R}} |m(t)|.$$

7. Opérateurs à noyau : Plaçons nous encore une fois dans le cadre de l'espace de Hilbert $L^2(\mathbb{R})$, et soit k une fonction de deux variables, à valeurs complexes. On peut associer à cette fonction un opérateur, défini comme suit. Pour tout $f \in \mathcal{H}$,

$$\mathcal{K}f(x) = \int_{-\infty}^{\infty} k(x, y)f(y) dy, \quad \forall x \in \mathbb{R}.$$

Pour vérifier que cette expression définit bien un opérateur borné sur $L^2(\mathbb{R})$, des hypothèses supplémentaires sur le noyau k sont nécessaires. Nous verrons quelques exemples plus loin.

Attention : ne pas confondre le noyau d'un opérateur tel que celui-ci, et l'ensemble des fonctions dont l'image par \mathcal{K} est nulle, que l'on appelle aussi noyau. Il y a là une regrettable confusion de terminologie.

9.3.2 Adjoint d'un opérateur sur un espace de Hilbert

L'objectif est de pouvoir construire une théorie spectrale pour les opérateurs linéaires sur un espace de Hilbert, similaire à celle que nous avons construite sur les matrices. Nous avons besoin pour cela d'une version adaptée de l'opérateur adjoint, dont l'existence est donnée par le résultat important suivant, dont nous omettrons la preuve (pour les connaisseurs, c'est une conséquence du théorème de Riesz-Fisher).

PROPOSITION 9.5 Soit $A \in \mathcal{L}(\mathcal{H})$ un opérateur linéaire sur l'espace de Hilbert \mathcal{H} . Alors il existe un unique opérateur $A^* \in \mathcal{L}(\mathcal{H})$, appelé adjoint de A , tel que pour tous $u, v \in \mathcal{H}$

$$\langle u | Av \rangle = \langle A^* u | v \rangle. \quad (9.29)$$

Notons que l'on a aussi, pour tous $u, v \in \mathcal{H}$,

$$\langle Au | v \rangle = \langle u | A^* v \rangle.$$

Les propriétés de l'adjoint généralisent celles que nous avons vues dans le cas des espaces Hermitiens :

PROPOSITION 9.6 Soit \mathcal{H} un espace de Hilbert. Pour tous $A, B \in \mathcal{L}(\mathcal{H})$, et tout $\lambda \in \mathbb{C}$,

1. $A^{**} = A$; $1^* = 1$.
2. $(A + B)^* = A^* + B^*$; $(\lambda A)^* = \bar{\lambda}A^*$; $(AB)^* = B^*A^*$.
3. $\|A^*\|_{op} = \|A\|_{op}$.

Preuve : La preuve des premiers points reprend essentiellement la preuve donnée dans le cas de la dimension finie. Pour ce qui est du dernier point, notons que pour tous $u, v \in \mathcal{H}$, on a

$$|\langle u|A^*v \rangle| = |\langle Au|v \rangle| = |\langle v|Au \rangle|,$$

de sorte que

$$\|A^*\|_{op} = \sup_{u,v \in \mathcal{H}} \frac{|\langle u|A^*v \rangle|}{\|u\| \|v\|} = \sup_{u,v \in \mathcal{H}} \frac{|\langle v|Au \rangle|}{\|u\| \|v\|} = \|A\|_{op}.$$

ce qui conclut la preuve.

DÉFINITION 9.11 Soit $A \in \mathcal{L}(\mathcal{H})$ A est auto-adjoint si $A^* = A$.

Les exemples suivants donnent quelques illustrations pour la détermination de l'adjoint d'un opérateur sur un espace de Hilbert, dans les situations considérées dans l'exemple 9.5

EXEMPLE 9.6 1. Opérateurs scalaires : on voit facilement que dans ce cas, pour tous $x, y \in \mathcal{H}$,

$$\langle x|Ay \rangle = \langle x|\lambda y \rangle = \langle \bar{\lambda}x|y \rangle,$$

de sorte que A^* est l'opérateur scalaire de multiplication par $\bar{\lambda}$, complexe conjugué de λ . Donc A est auto-adjoint si et seulement si $\lambda \in \mathbb{R}$.

2. Opérateurs de rang 1 : soient $u, v \in \mathcal{H}$, et soit $A(u, v)$ l'opérateur de rang 1 correspondant. Pour tous $x, y \in \mathcal{H}$, on a

$$\langle x|A(u, v)y \rangle = \langle x|\langle y|v \rangle u \rangle = \langle \langle x|u \rangle v|y \rangle = \langle A(v, u)x|y \rangle,$$

de sorte que

$$A(u, v)^* = A(v, u).$$

Par conséquent, $A(u, v)$ est autoadjoint si et seulement si $v = u$.

3. Opérateurs de rang fini et projecteurs orthogonaux : Comme corollaire immédiat de l'exemple précédent, l'adjoint de

$$A = \sum_{n=1}^N A(u_n, v_n)$$

est

$$A^* = \sum_{n=1}^N A(v_n, u_n).$$

Les projecteurs orthogonaux sont toujours auto-adjoints.

4. Opérateurs définis par une matrice infinie : Soit $A \in \mathcal{L}(\mathcal{H})$, et soit $M = M_A$ sa matrice (infinie) dans la base orthonormée $\mathcal{B} = \{e_1, e_2, \dots\}$. On vérifie facilement que l'adjoint de A est lui aussi caractérisé par une matrice infinie :

$$\langle u | Av \rangle = \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} \langle e_j | Ae_i \rangle \langle e_i | v \rangle \langle u | e_j \rangle = \sum_{i,j} \langle \langle u | e_i \rangle \overline{\langle Ae_i | e_j \rangle} e_j | v \rangle$$

Ainsi, la matrice de A^* n'est autre que la conjuguée Hermitienne de la matrice de A :

$$M_{A^*} = \overline{M_A^t}.$$

5. Opérateurs de translation : Soit $a \in \mathbb{R}$, et soit T_A l'opérateur de translation par a dans $L^2(\mathbb{R})$. Alors pour tous $f, g \in L^2(\mathbb{R})$, on a

$$\langle f | T_a g \rangle = \int_{-\infty}^{\infty} \bar{f}(t) g(t-a) dt = \int_{-\infty}^{\infty} \bar{f}(t+a) g(t) dt = \langle T_{-a} f | g \rangle.$$

Donc

$$(T_a)^* = T_{-a}.$$

6. Opérateurs de multiplication : Dans $L^2(\mathbb{R})$, soit M_m l'opérateur de multiplication par la fonction bornée m . Alors on voit facilement que l'adjoint de M_m est donné par

$$M_m^* = M_{\bar{m}},$$

c'est à dire un opérateur de multiplication par la complexe conjuguée de m . Evidemment, M_m est auto-adjoint si et seulement si m est à valeurs réelles.

7. Opérateurs à noyau : Toujours dans $L^2(\mathbb{R})$, soit \mathcal{K} l'opérateur à noyau, de noyau k . On vérifie de même que l'adjoint de \mathcal{K} est toujours un opérateur à noyau, de noyau \tilde{k} , défini par

$$\tilde{k}(x, y) = \overline{k(y, x)}.$$

9.3.3 Spectre d'un opérateur sur un espace de Hilbert

Dans le cas d'opérateurs en dimension finie, l'analyse spectrale se ramène au calcul des valeurs propres, et (si possible) à la diagonalisation de l'opérateur étudié. Nous allons voir qu'en dimension infinie, la situation peut malheureusement être bien plus complexe, comme nous allons le voir sur les exemples suivants.

EXEMPLE 9.7 1. Considérons tout d'abord le cas de l'espace $L^2([0, 1])$ des fonctions de module carré intégrable sur l'intervalle $[0, 1]$, et de l'opérateur de "position" de la mécanique quantique, noté X , et défini comme suit. Pour tout $f \in L^2([0, 1])$,

$$(Xf)(x) = xf(x).$$

Manifestement, X est auto-adjoint, et $\|X\|_{op} = 1$. La question est de savoir si X admet des valeurs propres et vecteurs propres. En d'autres termes, existe-t-il $\lambda \in \mathbb{R}$ et $\varphi \in L^2([0, 1])$ tels que $X\varphi = \lambda\varphi$? La réponse est négative... On aurait pu penser que pour $\lambda \in [0, 1]$, la distribution de Dirac δ_λ fasse l'affaire... mais ça n'est pas une fonction de $L^2([0, 1])$.

2. Considérons maintenant le cas de l'espace $L^2(\mathbb{R})$, et de l'opérateur position, défini comme ci-dessus. Dans ce cas, la situation est même plus complexe, dans la mesure où X n'est même plus un opérateur borné sur $L^2(\mathbb{R})$. Par exemple, la fonction $f(x) = 1/(1+x^2) \in L^2(\mathbb{R})$, mais $Xf \notin L^2(\mathbb{R})$. Comme on ne peut pas se passer d'un tel opérateur, il y a là une difficulté supplémentaire à lever.

La notion qui étend la notion de valeur propre dans ce nouveau cas est la notion de *spectre*.

DÉFINITION 9.12 Soit $A \in \mathcal{L}(\mathcal{H})$.

1. Le spectre de A est l'ensemble

$$Sp(A) = \{\lambda \in \mathbb{C} : (A - \lambda \mathbf{1}) \text{ n'est pas inversible} \} . \quad (9.30)$$

2. L'ensemble résolvant $\Omega(A)$ est le complément de $Sp(A)$ dans \mathbb{C} .

3. La famille d'opérateurs

$$R(\cdot, A) : \lambda \in \Omega(A) \rightarrow R(\lambda, A) = (\lambda \mathbf{1} - A)^{-1} \quad (9.31)$$

est appelée la résolvante de A .

4. Le nombre

$$\rho(A) = \left\{ \sup_{\lambda \in Sp(A)} |\lambda| \right\} \quad (9.32)$$

est le rayon spectral de A .

La question suivante est de savoir, pour un opérateur donné $A \in \mathcal{L}(\mathcal{H})$ donné, de quoi son spectre est formé. Nous avons déjà vu que si \mathcal{H} est un opérateur Hermitien, le spectre est formé de valeurs propres. Tel n'est plus le cas dans le cas général.

EXEMPLE 9.8 Reprenons l'exemple de l'opérateur position sur $L^2([0, 1])$, et cherchons les valeurs $\lambda \in \mathbb{C}$ telles que $X - \lambda$ ne soit pas inversible. Etant donnée $\varphi \in L^2([0, 1])$, soit $\psi \in L^2([0, 1])$, telle que $\varphi = (X - \lambda)\psi$. Alors on aurait $\psi(x) = \varphi(x)/(x - \lambda)$. Si $\lambda \notin [0, 1]$, la fonction $x \rightarrow 1/(x - \lambda)$ est bornée, et $\|\psi\| \leq \|\varphi\| \sup_x (1/(x - \lambda))$, d'où $(X - \lambda)^{-1}$ est borné. Donc dans ce cas $\lambda \in \Omega(X)$. Supposons maintenant $\lambda \in [0, 1]$. Nous allons maintenant voir que $(X - \lambda)$ n'est plus inversible. Soit φ définie par $\varphi(x) = 1$ pour tout $x \in [0, 1]$. Alors $(X - \lambda)^{-1}\varphi(x) = 1/(x - \lambda)$, et cette fonction n'appartient pas à $L^2([0, 1])$.

Ainsi, dans ce cas, $Sp(X) = [0, 1]$, mais aucune des valeurs du spectre n'est valeur propre, comme nous l'avons vu.

DÉFINITION 9.13 Soit $A \in \mathcal{L}(\mathcal{H})$.

1. $\lambda \in \mathbb{C}$ est valeur propre de A s'il existe $\psi \in \mathcal{H}$, $\psi \neq 0$, telle que

$$A\psi = \lambda\psi . \quad (9.33)$$

Un tel ψ est appelé vecteur propre de A (ou fonction propre si \mathcal{H} est un espace de fonctions).

2. λ est valeur propre approchée de A si il existe une suite de fonctions $\psi_n \in \mathcal{H}$, telles que $\|\psi_n\| = 1$ pour tout n , et

$$\|A\psi_n - \lambda\psi_n\| \longrightarrow 0 \quad \text{quand } n \rightarrow \infty . \quad (9.34)$$

EXEMPLE 9.9 Revenons une fois encore sur le cas de l'opérateur position sur $L^2([0, 1])$, et montrons que tout $\lambda \in [0, 1]$ est valeur propre approchée. Pour cela, soit π_n la fonction définie par

$$\psi_n(x) = \begin{cases} \sqrt{\frac{\pi}{2}} & \text{si } x \in [\lambda - 1/n, \lambda + 1/n] \\ 0 & \text{sinon} \end{cases}$$

On vérifie facilement que $\|\psi_n\| = 1$ pour tout n . Calculons maintenant

$$\|(X - \lambda)\psi_n\|^2 = \int_{\lambda-1/n}^{\lambda+1/n} (x - \lambda)^2 dx = \frac{n}{2} \int_{-1/n}^{1/n} x^2 dx = \frac{1}{3n^2} \rightarrow 0 \quad \text{quand } n \rightarrow \infty,$$

ce qui est le résultat désiré. Donc,

$$Sp(X) = [0, 1].$$

Il est clair que toute valeur propre de $A \in \mathcal{L}(\mathcal{H})$ appartient au spectre de A . Plus généralement,

PROPOSITION 9.7 *Toute valeur propre approchée d'un opérateur $A \in \mathcal{L}(\mathcal{H})$ appartient à $Sp(A)$.*

Dans une situation quelconque, le spectre d'un opérateur peut ne pas être seulement composé de ses valeurs propres et ses valeurs propres approchées. Toutefois, la situation est plus simple dans certains cas, comme par exemple le cas des opérateurs normaux.

DÉFINITION 9.14 *Un opérateur $A \in \mathcal{L}(\mathcal{H})$ est dit normal s'il commute avec son adjoint, c'est à dire si*

$$AA^* = A^*A.$$

Les opérateurs normaux sont les opérateurs qui permettent d'obtenir des résultats de décomposition spectrale proches de ceux que l'on peut obtenir en dimension finie.

REMARQUE 9.11 Notons que tout opérateur auto-adjoint est normal. Par contre, il existe des opérateurs normaux qui ne sont pas auto-adjoints. Par exemple, si A est auto-adjoint, λA est normal pour tout $\lambda \in \mathbb{C}$, mais n'est pas auto-adjoint dès que la partie imaginaire de λ est non-nulle.

Prenons l'exemple suivant : dans \mathbb{C}^3 , l'opérateur linéaire A défini par sa matrice dans la base canonique

$$M_A = \begin{pmatrix} -i & -i & 0 \\ -i & i & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

est effectivement normal. Il est aussi diagonalisable, et ses vecteurs propres sont orthogonaux deux à deux (le vérifier).

LEMME 9.1 *Soit $A \in \mathcal{L}(\mathcal{H})$ un opérateur normal. Alors, son image est orthogonale à son noyau.*

Preuve : pour tout $x \in \mathcal{H}$, on a

$$\|Ax\|^2 = \langle Ax | Ax \rangle = \langle x | A^* Ax \rangle = \langle x | AA^* x \rangle = \|A^* x\|^2.$$

Donc, $Ax = 0$ implique $A^* x = 0$, autrement dit, pour tout $y \in \mathcal{H}$,

$$0 = \langle A^* x | y \rangle = \langle x | Ay \rangle,$$

ce qui implique que x est orthogonal à Ay , pour tout $y \in \mathcal{H}$. Ceci prouve le résultat.

Une conséquence simple (dont la démonstration est omise ici car elle dépasse le cadre de ce cours) est le résultat suivant, que nous avons déjà rencontré dans le cas de l'opérateur position :

THÉORÈME 9.7 Le spectre d'un opérateur normal est l'ensemble de ses valeurs propres approchées.

Il est souvent relativement facile de trouver les valeurs propres approchées d'un opérateur normal. Par contre, la représentation spectrale d'un opérateur normal quelconque est souvent bien plus complexe, et fait appel à des notions qui dépassent le cadre de ce cours. On se contentera de signaler ici que dans ce cas, il existe tout de même un théorème, appelé *Théorème spectral*, permettant de donner un sens précis à ces notions. Comme on le verra, ce théorème spectral prendra toutefois une forme plus simple dans le cas des opérateurs compacts.

EXEMPLE 9.10 Prenons l'équation de la chaleur unidimensionnelle sur \mathbb{R} . Le champ de température au point x et au temps t , noté $T(x, t)$, est solution de l'équation aux dérivées partielles

$$\frac{\partial T(x, t)}{\partial t} = C \frac{\partial^2 T(x, t)}{\partial x^2},$$

C étant une constante positive. Comme on le verra plus tard, il est possible de montrer que pour une condition initiale fixée $T_0 \in L^2(\mathbb{R})$, il existe une et une seule solution, donnée par

$$T(x, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} e^{-Ck^2 t} \hat{T}_0(k) dk,$$

où \hat{T}_0 est la transformée de Fourier spatiale de la condition initiale :

$$\hat{T}_0(k) = \int_{-\infty}^{\infty} T_0(x) e^{-ikx} dx.$$

Soit $P_t : \varphi \in L^2(\mathbb{R}) \rightarrow P_t \varphi$ l'opérateur linéaire sur $L^2(\mathbb{R})$ qui à la condition initiale associe la solution à l'instant t , donc défini pour tout $t \in \mathbb{R}^+$ par

$$(P_t \varphi)(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} e^{-Ck^2 t} \hat{\varphi}(k) dk$$

Pour tout $t \in \mathbb{R}^+$, P_t est auto-adjoint, donc normal. Il est possible de montrer que tout $\lambda \in \mathbb{R}^+$ est valeur propre approchée de P_t . En effet, pour tout $\lambda \in \mathbb{R}^+$, on peut écrire pour tout $\varphi \in L^2(\mathbb{R})$ (grâce à la formule de Plancherel)

$$\|(P_t - \lambda \mathbf{1})\varphi\|^2 = \frac{1}{2\pi} \|\widehat{(P_t - \lambda \mathbf{1})\varphi}\|^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} |e^{-Ck^2 t} - \lambda|^2 |\hat{\varphi}(k)|^2 dk,$$

et il est facile de construire une suite de fonctions φ_n de norme unité telles que $\|(P_t - \lambda \mathbf{1})\varphi_n\| \rightarrow 0$ quand $n \rightarrow \infty$: par exemple, les fonctions φ_n définies par leur transformée de Fourier

$$\hat{\varphi}_n(k) = \sqrt{2n\pi} \chi_{[\alpha-1/2n, \alpha+1/2n]}(k)$$

font parfaitement l'affaire, et sont associées à la valeur propre approchée $\lambda = \exp(-C\alpha^2 t)$. Par contre, aucune de ces valeurs propres approchées n'est une "vraie" valeur propre. Il n'existe pas dans $L^2(\mathbb{R})$ de fonction propre correspondante. Les candidats naturels seraient les fonctions $x \rightarrow e^{ikx}$, mais ces fonctions ne sont pas de module carré intégrable.

9.4 Théorie spectrale des opérateurs compacts

Le cas des opérateurs compacts constitue un cas particulier assez simple, pour lequel les résultats de l'analyse spectrale généralisent assez simplement les résultats vus en dimension finie.

DÉFINITION 9.15 Un opérateur $A \in \mathcal{L}(\mathcal{H})$ sur l'espace de Hilbert \mathcal{H} est compact si pour toute suite bornée $\{u_n \in \mathcal{H}, n \in \mathbb{Z}\}$, on peut extraire de la suite $\{Au_n, n \in \mathbb{Z}\}$ une sous suite convergente.

Les opérateurs compacts possèdent un certain nombre de propriétés intéressantes, que nous énonçons ci-dessous, sans démonstration.

PROPOSITION 9.8

1. L'ensemble des opérateurs compacts sur un espace de Hilbert \mathcal{H} est un sous-espace vectoriel de $\mathcal{L}(\mathcal{H})$.
2. Si $A, B \in \mathcal{L}(\mathcal{H})$ avec A compact, alors AB et BA sont compacts.
3. L'adjoint A^* d'un opérateur compact A est compact.

Une propriété remarquable des opérateurs compacts est qu'ils peuvent être approchés par des opérateurs de rang fini (c'est à dire par des opérateurs dont l'image est de dimension finie).

THÉORÈME 9.8 Pour tout opérateur compact $A \in \mathcal{L}(\mathcal{H})$ sur un espace de Hilbert séparable \mathcal{H} , il existe une suite d'opérateurs de rang fini $A_n \in \mathcal{L}(\mathcal{H})$, tels que

$$\lim_{n \rightarrow \infty} \|A - A_n\|_{op} = 0. \quad (9.35)$$

Inversement, étant donné un opérateur A tel qu'il puisse être arbitrairement bien approché par des opérateurs de rang fini comme en (9.35); alors A est compact.

Le résultat essentiel de la théorie spectrale des opérateurs compacts tient dans les deux résultats suivants.

THÉORÈME 9.9 Soit $A \in \mathcal{L}(\mathcal{H})$ un opérateur compact sur un espace de Hilbert séparable \mathcal{H} . Alors les éléments non-nuls du spectre de A sont des valeurs propres, de multiplicité finie. De plus, pour tout $\epsilon > 0$, le nombre de valeurs propres supérieures à ϵ en module est fini.

Ainsi, ce théorème montre que le spectre d'un opérateur normal est constitué de valeurs propres de multiplicité finie, qui peuvent être ordonnés de façon à tendre vers 0 (qui est valeur propre approchée). Dans le cas des opérateurs compacts normaux, la représentation spectrale prend une forme simple, réminiscente de la situation rencontrée en dimension finie.

THÉORÈME 9.10 Soit $A \in \mathcal{L}(\mathcal{H})$ un opérateur compact normal sur un espace de Hilbert séparable \mathcal{H} . Alors il existe une suite orthonormée de vecteurs propres ψ_n de A et une suite de nombres $\{\lambda_n \in \mathbb{C}, n = 1, \dots\}$, telles que

$$\lambda_n \longrightarrow 0 \quad \text{quand } n \rightarrow \infty \quad (9.36)$$

et, pour tout $x \in \mathcal{H}$,

$$Ax = \sum_n \lambda_n \langle \psi_n | x \rangle \psi_n . \quad (9.37)$$

EXEMPLE 9.11 Soit $\mathcal{H} = L_p^2([0, 1])$ l'espace des fonctions périodiques de période 1, de module carré intégrable sur $[0, 1]$. Soit $h \in L_p^1([0, 1])$, et soit $T = K_h$ l'opérateur de convolution par h , défini sur $L_p^2([0, 1])$ par

$$K_h f(t) = \int_0^1 h(s) f(t-s) ds .$$

Grâce à la formule de Parseval, on peut écrire

$$K_h f(t) = \sum_{-\infty}^{\infty} c_n(h) c_n(f) e^{2i\pi n t} ,$$

expression qui fait intervenir les coefficients de Fourier de h

$$c_k(h) = \int_0^1 h(t) e^{-2i\pi k t} dt$$

et ceux de f . Il est possible de montrer que cet opérateur est un opérateur compact. Les fonctions propres correspondantes sont les fonctions

$$\epsilon_n : t \in [0, 1] \longrightarrow e^{2i\pi n t} ,$$

et les valeurs propres correspondantes ne sont autres que les coefficients de Fourier $c_k(h)$ de h . Comme $h \in L^2([0, 1])$, le théorème de Riemann-Lebesgue implique que ces coefficients de Fourier tendent vers 0 lorsque $k \rightarrow \infty$, et 0 est valeur propre approchée, sans être nécessairement valeur propre (sauf si l'un des coefficients de Fourier de h est nul).

Parmi les opérateurs compacts, il existe une classe particulièrement intéressante : il s'agit des opérateurs à noyaux (comme on les a déjà rencontrés plus haut) dits *de Hilbert-Schmidt*.

DÉFINITION 9.16 Soit $k \in L^2(\mathbb{R}^2)$, et soit \mathcal{K} l'opérateur sur $L^2(\mathbb{R})$ défini par le noyau k : pour tout $f \in L^2(\mathbb{R})$,

$$\mathcal{K}f(x) = \int_{-\infty}^{\infty} k(x, y) f(y) dy . \quad (9.38)$$

Un tel opérateur est appelé opérateur de Hilbert-Schmidt.

Pour que cette définition définisse réellement un opérateur sur $L^2(\mathbb{R})$, il faut montrer que si $f \in L^2(\mathbb{R})$, alors $\mathcal{K}f \in L^2(\mathbb{R})$ aussi. Pour cela, on a par l'inégalité de Cauchy-Schwarz

$$|\mathcal{K}f(x)| = \left| \int_{-\infty}^{\infty} k(x, y) f(y) dy \right| \leq \sqrt{\int_{-\infty}^{\infty} |k(x, y)|^2 dy} \|f\| ,$$

d'où on déduit

$$\|\mathcal{K}f(x)\|^2 \leq \|k\|_{L^2(\mathbb{R}^2)}^2 \|f\|^2 < \infty.$$

Il est facile de montrer que tout opérateur de Hilbert-Schmidt est automatiquement compact. Ainsi, les propriétés spectrales que nous avons énoncées dans les théorèmes 9.9 et 9.10 sont valables.

9.5 Quelques mots sur les opérateurs non bornés

Jusqu'à présent, nous nous sommes limités aux opérateurs bornés sur un espace de Hilbert. Malheureusement, il existe de multiples exemples d'opérateurs, ayant un intérêt physique indéniable, qui ne sont pas bornés. Par exemple, comme nous l'avons vu, l'opérateur position de la mécanique quantique n'est pas borné sur $L^2(\mathbb{R})$. De même, l'opérateur de dérivation n'est pas borné sur $L^2(\mathbb{R})$, ni sur $L^2([0, 1])$ par exemple. Dans ces situations, il est nécessaire de se limiter à un *domaine* (ou *domaine de définition*) dans lequel l'opérateur en question est bien défini. On note alors (A, \mathcal{D}_A) le couple constitué de l'opérateur A et de son domaine $\mathcal{D}_A \subset \mathcal{H}$.

Dans cette situation, la notion d'opérateur auto-adjoint est elle aussi plus subtile, et on doit recourir à la notion (plus faible) d'opérateur symétrique.

DÉFINITION 9.17 Soit A un opérateur linéaire sur \mathcal{H} , de domaine \mathcal{D}_A . (A, \mathcal{D}_A) est symétrique si pour tous $x, y \in \mathcal{D}_A$, on a

$$\langle x | Ay \rangle = \langle Ax | y \rangle. \quad (9.39)$$

EXEMPLE 9.12 1. Soit $\mathcal{H} = L^2(\mathbb{R})$, et considérons l'opérateur position X vu précédemment. Alors

$$\mathcal{D}_X = \{f \in L^2(\mathbb{R}), \int_{-\infty}^{\infty} x^2 |f(x)|^2 dx < \infty\}.$$

Il est facile de vérifier que (X, \mathcal{D}_X) est symétrique. Pour tout $f, g \in \mathcal{D}_X$, on a en effet

$$\langle f | Xg \rangle = \int_{-\infty}^{\infty} \bar{f}(x) xg(x) dx = \int_{-\infty}^{\infty} x\bar{f}(x) g(x) dx = \langle Xf | g \rangle.$$

2. Soit de nouveau $\mathcal{H} = L^2(\mathbb{R})$, et soit P l'opérateur "impulsion" de la mécanique quantique :

$$(Pf)(x) = if'(x).$$

Le domaine de P est l'espace

$$\mathcal{D}_P = \{f \in L^2(\mathbb{R}), f' \in L^2(\mathbb{R})\}.$$

Etant données $f, g \in \mathcal{D}_P$, on peut calculer

$$\langle f | Pg \rangle = i \int_{-\infty}^{\infty} \bar{f}(t)g'(t) dt = \int_{-\infty}^{\infty} i\bar{f}'(t)g(t) dt = \langle Pf | g \rangle.$$

Donc P est symétrique.

DÉFINITION 9.18 Soit (A, \mathcal{D}_A) un opérateur linéaire sur \mathcal{H} , tel que \mathcal{D}_A soit dense dans \mathcal{H} . Soit

$$\mathcal{D}_A^* = \{ \psi \in \mathcal{H}, \text{ t.q. l'application } \varphi \in \mathcal{H} \rightarrow \langle \varphi | A\psi \rangle \text{ est continue} \} . \quad (9.40)$$

1. L'adjoint (A^*, \mathcal{D}_A^*) de (A, \mathcal{D}_A) est défini de la façon suivante :

$$\langle x | A^*y \rangle = \langle Ax | y \rangle , \quad \forall x \in \mathcal{D}_A, y \in \mathcal{D}_A^* . \quad (9.41)$$

2. (A, \mathcal{D}_A) est auto-adjoint si

$$(A^*, \mathcal{D}_A^*) = (A, \mathcal{D}_A) . \quad (9.42)$$

Si A est un opérateur symétrique, il est clair que

$$\mathcal{D}_A \subset \mathcal{D}_A^*$$

et que la restriction de A^* à \mathcal{D}_A est A . Par contre, \mathcal{D}_A^* est souvent strictement plus grand que \mathcal{D}_A .

La plupart des opérateurs linéaires intervenants en physique (bornés ou non-bornés) sont auto-adjoints. On montre qu'il est possible dans ce cas d'en donner une représentation spectrale. Cependant, il en existe une sous-classe pour laquelle cette représentation se simplifie. Cette sous-classe fait appel à la résolvante, que nous avons déjà rencontrée plus haut en (9.31), et dont nous rappelons la définition :

$$R(\cdot, A) : \lambda \in \Omega(A) \rightarrow R(\lambda, A) = (\lambda \mathbf{1} - A)^{-1} .$$

Si il existe λ dans l'ensemble résolvant tel que l'opérateur $R(\lambda, A)$ soit compact, alors la situation se simplifie considérablement.

THÉORÈME 9.11 Soit (A, \mathcal{D}_A) un opérateur linéaire sur un espace de Hilbert séparable \mathcal{H} . Alors les trois propriétés suivantes sont équivalentes.

1. Il existe $\lambda \in \Omega(A)$ tel que $R(\lambda, A)$ soit compact.
2. Pour tout $\lambda \in \Omega(A)$, $R(\lambda, A)$ est compact.
3. a) (A, \mathcal{D}_A) admet une base orthonormée de vecteurs propres $\{u_n, n \in \mathbb{Z}\}$:

$$Au_n = \lambda_n u_n , \quad \text{et } \lambda_n \in \mathcal{D}_A . \quad (9.43)$$

b) Le spectre de (A, \mathcal{D}_A) est un ensemble infini dénombrable $\{\lambda_n, n \in \mathbb{Z}^+\}$ de points isolés.

c) Le domaine de A n'est autre que l'ensemble des éléments $u \in \mathcal{H}$ tels que la suite de terme général $\lambda_n \langle u_n | u \rangle$ soit de module carré sommable :

$$\mathcal{D}_A = \{ u \in \mathcal{H}, \{ \lambda_n \langle u_n | u \rangle, n \in \mathbb{Z}^+ \} \in \ell^2(\mathbb{Z}^+) \} . \quad (9.44)$$

On ne donnera pas la preuve complète de cet important résultat ici. Ceci étant, il est intéressant de comprendre quelle en est l'origine. Pour cela, supposons que le point (1) soit vérifié, et montrons les grandes lignes du point (2). Soit $\lambda \in \Omega(A)$ tel que $R(\lambda, A)$ soit compact. A étant autoadjoint, on vérifie facilement que $R(\lambda, A)$ l'est aussi. Alors on sait que le spectre de $R(\lambda, A)$ est constitué de valeurs propres non-nulles μ_n , avec $\lim_{n \rightarrow \infty} \mu_n = 0$. Notons u_n les vecteurs propres correspondants. On a alors, comme $\mu_n \neq 0$,

$$\frac{1}{\mu_n} u_n = (\lambda \mathbf{1} - A) u_n ,$$

de sorte que

$$Au_n = \left(\lambda - \frac{1}{\mu_n} \right) u_n .$$

Comme A est auto-adjoint, ses valeurs propres sont réelles (notons que $R(\lambda, A)$ n'étant pas nécessairement auto-adjoint, ses valeurs propres sont généralement complexes, tout comme λ).

9.6 Éléments de théorie des perturbations

Il existe des situations dans lesquelles le spectre d'un opérateur n'est pas calculable explicitement, mais peut l'être de façon approchée. C'est en particulier le cas d'opérateurs $A_0 \in \mathcal{L}(\mathcal{H})$ qui peuvent s'écrire comme "perturbations" d'un opérateur A_0 dont la représentation spectrale est connue, c'est à dire comme la somme de celui-ci et d'un "petit" opérateur ϵW (petit au sens où sa norme est petite) :

$$A = A_0 + \epsilon W .$$

Dans ce cas là, il est naturel de s'attendre à ce que le spectre de A soit dans un certain sens proche de celui de A_0 , et de même pour la représentation spectrale. La théorie des perturbations permet de donner une estimation du spectre de A sous forme d'un développement en série suivant les puissances de ϵ , le premier terme (d'ordre zéro) correspondant au spectre de A_0 .

Dans ce qui suit, on va démontrer un résultat de ce type, dans le cas particulier d'opérateurs sur un espace Hermitien (c'est à dire de dimension finie), en nous limitant au premier ordre.

9.6.1 Perturbation d'un vecteur propre non-dégénéré d'un opérateur linéaire auto-adjoint sur un espace Hermitien

Soit \mathcal{H} un espace Hermitien, et soit

$$\epsilon \rightarrow A(\epsilon) \in \mathcal{L}(\mathcal{H})$$

une famille d'opérateurs linéaires auto-adjoints sur \mathcal{H} , que l'on suppose C^∞ par rapport à la variable ϵ .¹ On suppose que la représentation spectrale de A_0 est connue, et on note $\lambda_0, \dots, \lambda_{N-1}$ et x_0, \dots, x_{N-1} les valeurs propres et vecteurs propres correspondants. On s'intéresse dans un premier temps aux valeurs propres non-dégénérées, et on suppose que λ_0 est non-dégénérée.

Il est possible de démontrer (en utilisant la continuité de l'application $\epsilon \rightarrow A(\epsilon)$) que pour ϵ suffisamment petit, il existe un voisinage \mathcal{O} de λ_0 tel que $A(\epsilon)$ admette une unique valeur propre $\lambda(\epsilon)$ dans \mathcal{O} , et on note $x(\epsilon)$ le vecteur propre correspondant. Ce dernier étant défini à une constante multiplicative près, on le normalise, en imposant (par exemple)

$$\langle x_0 | x(\epsilon) \rangle = 1 . \tag{9.45}$$

Le caractère C^∞ de $\epsilon \rightarrow A(\epsilon)$ implique que $\epsilon \rightarrow \lambda(\epsilon)$ et $\epsilon \rightarrow x(\epsilon)$ sont C^∞ également. Ecrivons

$$A(\epsilon)x(\epsilon) = \lambda(\epsilon)x(\epsilon) . \tag{9.46}$$

En différenciant par rapport à ϵ , on obtient

$$(A'(\epsilon) - \lambda'(\epsilon))x(\epsilon) = (A(\epsilon) - \lambda(\epsilon))x'(\epsilon) , \tag{9.47}$$

alors que (9.45) donne

$$\langle x_0 | x'(\epsilon) \rangle = 0 .$$

L'équation (9.47) donne, par produit scalaire avec x_0

$$\lambda'(\epsilon) = \langle x_0 | A'(\epsilon)x(\epsilon) \rangle + \langle x_0 | A(\epsilon)x'(\epsilon) \rangle ,$$

¹Ceci revient à supposer que la matrice de $A(\epsilon)$ est C^∞ .

et dans le cas particulier $\epsilon = 0$,

$$\lambda'(0) = \langle x_0 | A'(0)x_0 \rangle + \langle x_0 | A(0)x'(0) \rangle = \langle x_0 | A'(0)x_0 \rangle . \quad (9.48)$$

Ceci, combiné à un développement de Taylor au premier ordre, donne donc une estimation de la valeur propre perturbée :

$$\lambda(\epsilon) = \lambda_0 + \epsilon \langle x_0 | A'(0)x_0 \rangle + O(\epsilon^2) .$$

Passons maintenant à l'analyse du vecteur propre perturbé. En considérant le cas $\epsilon = 0$ dans (9.47), on a

$$[A'(0) - \langle x_0 | A'(0)x_0 \rangle]x_0 + [A(0) - \lambda_0]x'(0) = 0 ,$$

d'où on déduit

$$(A(0) - \lambda_0)x'(0) = \langle x_0 | A'(0)x_0 \rangle x_0 - A'(0)x_0 = (\Pi_{x_0} - \mathbf{1})A'(0)x_0 , \quad (9.49)$$

où on note Π_{x_0} le projecteur orthogonal sur l'espace (de dimension 1) $\mathbb{C}x_0$ engendré par x_0 .

A ce point, il suffit d'appliquer des deux côtés de cette équation l'inverse de $A(0) - \lambda_0 \mathbf{1}$ pour obtenir une expression de $x'(0)$, d'où on tire une expression de $x(\epsilon)$ par développement de Taylor au premier ordre. Cependant, il faut faire attention, car λ_0 étant valeur propre de A_0 , $A(0) - \lambda_0 \mathbf{1}$ n'est pas inversible.

Commençons par remarquer que le noyau de l'opérateur $A(0) - \lambda_0 \mathbf{1}$ n'est autre que $\mathbb{C}x_0$, et que $x'(0) \perp \mathbb{C}x_0$ (voir plus haut). Par ailleurs, on peut aussi noter que le membre de droite de l'équation (9.49) ci-dessus n'est autre que la projection orthogonale de $A'(0)x_0$ sur le complément orthogonal $(\mathbb{C}x_0)^\perp$ de $\mathbb{C}x_0$ dans \mathcal{H} . Donc cette dernière équation est une équation dans $(\mathbb{C}x_0)^\perp$.

Notons $(A(0) - \lambda_0 \mathbf{1})^\sharp$ l'opérateur défini par

$$(A(0) - \lambda_0 \mathbf{1})^\sharp x = \begin{cases} (A(0) - \lambda_0 \mathbf{1})^{-1}x & \text{si } x \in (\mathbb{C}x_0)^\perp \\ 0 & \text{sinon} \end{cases}$$

Alors on peut écrire

$$x'(0) = (A(0) - \lambda_0 \mathbf{1})^\sharp (\Pi_{x_0} - \mathbf{1})A'(0)x_0 , \quad (9.50)$$

Pour conclure, il nous suffit maintenant d'explicitier le membre de droite de cette dernière équation. Partant de la représentation spectrale de A_0

$$x = \sum_{i=0}^{N-1} \langle x_i | x \rangle x_i , \quad \forall x \in \mathcal{H} ,$$

on peut écrire

$$(\Pi_{x_0} - \mathbf{1})x = - \sum_{i=1}^{N-1} \langle x_i | x \rangle x_i , \quad \forall x \in \mathcal{H} ,$$

et

$$(A(0) - \lambda_0 \mathbf{1})^\sharp (\Pi_{x_0} - \mathbf{1})A'(0)x_0 = \sum_{i=1}^{N-1} \frac{1}{\lambda_i - \lambda_0} \langle x_i | A'(0)x_0 \rangle x_i .$$

Ainsi, nous avons obtenu l'expression suivante pour la dérivée du premier vecteur propre x_0 :

$$x'(0) = \sum_{i=1}^{N-1} \frac{1}{\lambda_i - \lambda_0} \langle x_i | A'(0)x_0 \rangle x_i , \quad (9.51)$$

d'où par développement de Taylor au premier ordre,

$$x(\epsilon) = x_0 + \epsilon \sum_{i=1}^{N-1} \frac{1}{\lambda_i - \lambda_0} \langle x_i | A'(0)x_0 \rangle x_i .$$

Bien entendu, le même traitement s'applique à toutes les valeurs propres non-dégénérées.

REMARQUE 9.12 Notons que dans ce calcul, nous avons uniquement supposé que la valeur propre λ_0 était non-dégénérée, sans aucune hypothèse de la sorte sur les autres λ_i (qui peuvent quant à elles être dégénérées).

9.6.2 Le cas des valeurs propres dégénérées

Relaxons maintenant nos hypothèses, en évitant de supposer que la valeur propre considérée est de multiplicité 1. On se limitera ici à étudier les valeurs propres perturbées, le calcul des vecteurs propres correspondants est laissé en exercice au lecteur intéressé.

Supposons donc que la valeur propre considérée λ_0 a pour multiplicité m . Donc, le sous-espace propre correspondant, noté \mathcal{H}_0 , est de dimension m , et admet une base orthonormée que l'on notera $\{x_0^{(1)}, x_0^{(2)}, \dots, x_0^{(m)}\}$. Pour $\epsilon \neq 0$, il n'y a aucune raison que la valeur propre $\lambda(\epsilon)$ soit elle aussi dégénérée, et on va donc supposer qu'il existe m valeurs propres $\lambda^{(1)}(\epsilon), \dots, \lambda^{(m)}(\epsilon)$, simples ou dégénérées. On note $x^{(1)}(\epsilon), \dots, x^{(m)}(\epsilon)$ les vecteurs propres correspondants.

Le point de départ du calcul est le même que précédemment, partant de

$$[A(\epsilon) - \lambda^{(k)}(\epsilon)]x^{(k)}(\epsilon) = 0,$$

en différentiant par rapport à ϵ puis en prenant le produit scalaire par rapport à un $x_0^{(\ell)}$ quelconque, on aboutit à

$$\langle x_0^{(\ell)} | [A'(\epsilon) - \lambda^{(k)'}(\epsilon)]x^{(k)}(\epsilon) \rangle + \langle x_0^{(\ell)} | [A(\epsilon) - \lambda^{(k)}(\epsilon)]x^{(k)'}(\epsilon) \rangle = 0.$$

En prenant le cas particulier $\epsilon = 0$, et en utilisant le fait que A_0 est auto-adjoint, et que $\lambda^{(k)}(0) = \lambda_0$ pour tout k , on obtient

$$\langle x_0^{(\ell)} | [A'(0) - \lambda^{(k)'}(0)]x_0^{(k)} \rangle = 0, \quad \forall \ell = 1, \dots, m. \quad (9.52)$$

Ceci étant vrai pour tout $\ell = 1, \dots, m$, on en déduit que $\lambda^{(k)'}(0)$ est nécessairement valeur propre de la restriction de $A'(0)$ au sous-espace propre \mathcal{H}_0 .

Sixième partie

Probabilités

10 Introduction aux Probabilités

10.1 Modélisation probabiliste

10.1.1 Généralités

Modèle probabiliste :

Les modèles probabilistes s'introduisent naturellement en physique dans deux contextes différents :

- Il existe une grande quantité de systèmes physiques qui sont trop complexes pour pouvoir être décrits de façon déterministe. L'exemple le plus immédiat est fourni par des systèmes de particules "libres". Supposons simplement que nous ayons à décrire un système de 10^{10} particules, évoluant librement dans un volume fini donné, interagissant par chocs élastiques. Il est clair que les lois fondamentales de la dynamique nous permettent théoriquement de calculer les trajectoires individuelles de chacune des particules. Cependant, un tel calcul s'avère irréalisable en pratique, compte tenu du grand nombre de possibilités dont il faut tenir compte.
- Il existe aussi des systèmes qui ont une nature intrinsèquement probabiliste. C'est en particulier le cas des systèmes quantiques, si on s'en tient aux interprétations classiques. Etant donné un système quantique, on ne peut généralement pas prédire à l'avance ce que sera le résultat d'une mesure. On doit se contenter de connaître les résultats possibles, auxquels on peut affecter une probabilité.

Dans de tels cas de figure, la modélisation est une modélisation probabiliste. Une modélisation probabiliste consiste essentiellement en un ensemble (discret ou continu) de "possibles" (appelés *événements*), auxquels on associe un nombre (la probabilité de l'évènement) qui en caractérise la vraisemblance.

EXEMPLE 10.1 Prenons l'exemple d'une pièce normale, avec laquelle on tire à pile ou face. Il est quasiment impossible de calculer avec une précision suffisante la trajectoire de la pièce pour prédire le résultat. Par contre, il est possible d'associer des probabilités aux résultats possibles. Par exemple, si on note P et F les 2 résultats possibles d'un tirage, et si on effectue un seul tirage, on a donc deux résultats possibles, avec probabilités $1/2$. Si on effectue 2 tirages, on a 4 résultats possibles équiprobables ; et ainsi de suite. On peut faire le tableau suivant :

P	F	PP	PF	FP	FF	...	$PP\dots P$...	$FF\dots F$
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$...	2^{-n}	...	2^{-n}

On associe ainsi une probabilité à chacun des évènements considérés. On vérifie facilement que la somme des probabilités associées à un nombre de tirages n fixé vaut 1.

EXEMPLE 10.2 *Marche au hasard* : Prenons l'exemple d'un ivrogne, à la démarche hésitante. Si il a une jambe plus courte que l'autre, on peut penser qu'il a une probabilité p d'aller vers la gauche (G), et $q = 1 - p$ d'aller vers la droite (D). On peut faire le tableau suivant, représentant les configurations possibles pour un nombre n de pas donné :

G	D	GG	GD	DG	DD	...	GG...G	...	FF...F
p	q	p ²	pq	pq	q ²	...	p ⁻ⁿ	...	q ⁻ⁿ

On vérifie aussi dans ce cas que pour un nombre fixé de pas n , la somme des probabilités vaut 1.

Ces exemples suggèrent de poser la première ébauche de définition suivante.

PRÉ-DÉFINITION Un modèle probabiliste pour une expérience à un nombre fini N de résultats possibles consiste en

1. Un ensemble $\Omega = \{\omega_1, \dots, \omega_N\}$ de résultats possibles.
2. Une application $\mathbb{P} : \Omega \rightarrow [0, 1]$

$$\omega_j \rightarrow \mathbb{P}\{\omega_j\},$$

telle que

$$\sum_{j=1}^N \mathbb{P}\{\omega_j\} = 1. \tag{10.1}$$

Nous allons voir un peu plus loin que cette définition n'est pas suffisante pour couvrir les cas susceptibles de nous intéresser.

EXEMPLE 10.3 On tire à pile ou face avec une pièce ordinaire. Si le résultat est pile (P), on jette un dé à 6 faces. Si le résultat est face (F), on retire à pile ou face. L'ensemble des résultats possibles est

$$\Omega = \{(P, 1), (P, 2), (P, 3), (P, 4), (P, 5), (P, 6), (F, P), (F, F)\}$$

ω	(P,1)	(P,2)	(P,3)	(P,4)	(P,5)	(P,6)	(F,P)	(F,F)
$\mathbb{P}\{\omega\}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$

Probabilités et statistique

La théorie des probabilités et la statistique sont deux sciences complémentaires. Plus précisément, la théorie des probabilités est une *science prédictive* : son objectif est de construire des modèles décrivant la réalité, et prédisant les fréquences d'apparition de phénomènes, ou résultats de mesure.

Comment connaît on les probabilités de tels évènements ? La réponse à cette question est fournie par la théorie statistique, qui permet d'estimer ces probabilités à partir d'expériences. La statistique est donc une *science descriptive*.

Comment fait on le lien entre les deux théories ? la relation est souvent donnée par ce que l'on nomme des *théorèmes limites*, comme par exemple la *loi des grands nombres*, ou encore le *théorème central limite*. L'exemple suivant permet d'illustrer la situation.

On considère un dé usuel à 6 faces. Si le dé est un dé ordinaire, les 6 faces sont équiprobables, et les 6 faces ont donc probabilités $p_1 = p_2 = \dots = p_6 = 1/6$. Supposons maintenant que l'on fasse N expériences, c'est à dire N tirages, et que l'on obtienne n_1 fois la face 1, ... et n_6 fois la face 6. Posons alors pour $i = 1, \dots, 6$: $\hat{p}_i = n_i/N$. En général, on n'a pas $\hat{p}_i = p_i$, mais le "bon sens" suggère que lorsque N est grand, on devrait avoir $\hat{p}_i \approx p_i$. La loi des grands nombres permet de donner un sens plus précis à cette assertion. On montre en fait que pour tout $\delta > 0$,

$$\mathbb{P}\{|\hat{p}_i - p_i| > \delta\} \rightarrow 0 \quad \text{quand } N \rightarrow \infty.$$

Le théorème central limite permet de préciser cette limite.

10.1.2 Calcul des probabilités

Ensembles d'évènements, union, intersection

Etant donné un ensemble Ω et une application \mathbb{P} comme ci-dessus, nous allons maintenant étendre ces notions. \mathbb{P} associe un nombre à tout élément de Ω ; on veut également être capables d'associer une probabilité à des sous-ensembles de Ω : calculer les probabilités pour que le résultat appartienne à un certain ensemble $A \subset \Omega$ de résultats possibles. Les sous-ensembles $A \subset \Omega$ auxquels on peut associer une probabilité seront appelés *évènements*.

Pour cela, examinons cas par cas un certain nombre de situations possibles.

Le cas fini Supposons tout d'abord que Ω soit un ensemble fini, et qu'on dispose d'une application $\mathbb{P} : \Omega \rightarrow \mathbb{R}^+$, définie donc sur les éléments de Ω . Pour tout sous-ensemble $A \subset \Omega$, il est alors facile d'introduire $\mathbb{P}\{A\}$:

DÉFINITION 10.1 Soit Ω un ensemble fini, et soit $A \subset \Omega$. On définit alors $\mathbb{P}\{A\}$ par

$$\mathbb{P}\{A\} = \sum_{\omega \in A} \mathbb{P}\{\omega\}. \quad (10.2)$$

Il est immédiat qu'avec cette définition, on a nécessairement

$$\mathbb{P}\{\Omega\} = 1. \quad (10.3)$$

Ayant posé cette définition, on se retrouve tout naturellement dans un contexte de théorie des ensembles. On a alors :

PROPOSITION 10.1 1. Si $A, B \subset \Omega$ sont deux ensembles disjoints (c'est à dire $A \cap B = \emptyset$), alors

$$\mathbb{P}\{A \cup B\} = \mathbb{P}\{A\} + \mathbb{P}\{B\}. \quad (10.4)$$

2. Plus généralement : si $A, B \subset \Omega$,

$$\mathbb{P}\{A \cup B\} = \mathbb{P}\{A\} + \mathbb{P}\{B\} - \mathbb{P}\{A \cap B\}. \quad (10.5)$$

3. Si A^c est le complémentaire de A dans Ω :

$$\mathbb{P}\{A^c\} = 1 - \mathbb{P}\{A\}. \quad (10.6)$$

Le cas infini dénombrable Supposons maintenant que Ω soit un ensemble infini dénombrable. Supposons aussi qu'on ait introduit une application $\mathbb{P} : \Omega \rightarrow \mathbb{R}^+$. On peut alors étendre \mathbb{P} à des sous-ensembles plus généraux de Ω , comme dans le cas fini, en définissant (comme dans la DÉFINITION 10.1) pour toute famille (finie ou infinie dénombrable) $A = \{a_1, a_2, \dots, a_i \in \Omega\}$ la probabilité de l'évènement A par

$$\mathbb{P}\{A\} = \sum_n \mathbb{P}\{a_n\} = \mathbb{P}\{a_1\} + \mathbb{P}\{a_2\} + \dots. \quad (10.7)$$

La formule (10.3) est bien évidemment toujours valable.

EXEMPLE 10.4 On considère une variante du jeu de pile ou face, dans laquelle les deux faces de la pièce ont probabilités $1/2$, mais la pièce est tirée jusqu'à ce que "face" apparaisse. L'ensemble des résultats possibles est

$$\Omega = \{F, PF, PPF, \dots\},$$

et les probabilités des éléments de Ω sont données par

ω	F	PF	PPF	$PPPF$	\dots	$P \dots PF$ (n fois P)
$\mathbb{P}\{\omega\}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	\dots	$2^{-(n+1)}$

A partir de là, on peut calculer les probabilités des évènements (les sous-ensembles de Ω) : par exemple, la probabilité pour que le nombre de P soit inférieur à 4 est donnée par

$$\mathbb{P}\{\{F, PF, PPF, PPPF\}\} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{15}{16}.$$

On a évidemment

$$\mathbb{P}\{\Omega\} = \sum_{n=1}^{\infty} 2^{-n} = \frac{1}{1 - 1/2} - 1 = 1.$$

Le cas continu Le cas infini continu est un peu plus complexe. Prenons l'exemple d'un nombre tiré au hasard dans l'intervalle $[0, 1]$, toutes les valeurs ayant même probabilité. On a donc $\Omega = [0, 1]$. Mais Ω étant un ensemble continu, les probabilités des valeurs individuelles sont nulles :

$$\mathbb{P}\{t\} = 0 \quad \forall t \in [0, 1].$$

Dans ce cas, il est nécessaire d'utiliser directement des sous-ensembles de Ω pour construire le modèle. Toutes les valeurs étant équiprobables, on définit alors \mathbb{P} par

$$\mathbb{P}\{[a, b[\} = b - a$$

pour tout intervalle $[a, b[\subset \Omega$. Pour définir $\mathbb{P}\{A\}$ pour des sous-ensembles (presque) quelconques $A \subset \Omega$, on peut alors s'inspirer de la PROPOSITION 10.1, et écrire par exemple, pour toute famille d'intervalles disjoints deux à deux $I_1 = [a_1, b_1[$, $I_2 = [a_2, b_2[$, \dots ,

$$\mathbb{P}\{I_1 \cup I_2 \cup \dots\} = \mathbb{P}\{I_1\} + \mathbb{P}\{I_2\} + \dots = (b_1 - a_1) + (b_2 - a_2) + \dots$$

Dans ce cas, nous serons naturellement amenés à considérer des intégrales, comme nous le verrons dans la section 10.1.2.

Définition fondamentale Motivé par les discussions précédentes, on peut maintenant donner la définition suivante d'un modèle probabiliste au sens de Kolmogorov (dans la littérature mathématique, on parle d'espace probabilisé).

DÉFINITION 10.2 Un modèle probabiliste consiste en

1. Un ensemble de résultats possibles Ω (appelé univers, espace des épreuves ou référentiel).
2. Une tribu, c'est à dire une famille \mathcal{F} de parties de Ω , appelées évènements, telle que :
 - $\Omega \in \mathcal{F}$.
 - Si $A \in \mathcal{F}$, alors son complémentaire $A^c = \Omega \setminus A \in \mathcal{F}$.
 - Si $A_1, A_2, \dots \in \mathcal{F}$ est un ensemble (fini ou infini dénombrable) de parties de Ω , alors $A_1 \cup A_2 \cup \dots \in \mathcal{F}$.
 - Si $A_1, A_2, \dots \in \mathcal{F}$ est un ensemble (fini ou infini dénombrable) de parties de Ω , alors $A_1 \cap A_2 \cap \dots \in \mathcal{F}$.
3. Une application $\mathbb{P} : A \in \mathcal{F} \rightarrow \mathbb{P}\{A\} \in [0, 1]$, telle que
 - $\mathbb{P}\{\Omega\} = 1$.
 - Si $A_1, A_2, \dots \in \mathcal{F}$ est un ensemble de parties de Ω deux à deux disjointes,

$$\mathbb{P}\{A_1 \cup A_2 \cup \dots\} = \mathbb{P}\{A_1\} + \mathbb{P}\{A_2\} + \dots \quad (10.8)$$

L'ensemble \mathcal{F} possédant les propriétés ci-dessus est parfois appelé *tribu*. Une conséquence immédiate de la définition est que $\emptyset \in \mathcal{F}$, et que $\mathbb{P}\{\emptyset\} = 0$.

Calcul sur les évènements

La formulation en termes de théorie des ensembles permet de formuler simplement un certain nombre de résultats. En particulier, le résultat suivant est une conséquence immédiate des définitions.

PROPOSITION 10.2 Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un modèle probabiliste.

1. Si $A, B \in \mathcal{F}$ et $A \subset B$, alors $\mathbb{P}\{A\} \leq \mathbb{P}\{B\}$.
2. Si $A \in \mathcal{F}$, $\mathbb{P}\{A^c\} = 1 - \mathbb{P}\{A\}$.
3. Si $A, B \in \mathcal{F}$ et $A \subset B$, alors $\mathbb{P}\{B \setminus A\} = \mathbb{P}\{B\} - \mathbb{P}\{A\}$.
4. Si $A, B \in \mathcal{F}$, $\mathbb{P}\{A \cup B\} = \mathbb{P}\{A\} + \mathbb{P}\{B\} - \mathbb{P}\{A \cap B\}$.

Preuve : 1) On a $B = A \cup (B \setminus A)$, et $A \cap (B \setminus A) = \emptyset$. Donc $\mathbb{P}\{B\} = \mathbb{P}\{A\} + \mathbb{P}\{B \setminus A\} \geq \mathbb{P}\{A\}$.

2) est immédiat : $\mathbb{P}\{A\} + \mathbb{P}\{A^c\} = \mathbb{P}\{A \cup A^c\} = \mathbb{P}\{\Omega\} = 1$.

3) Si $A \subset B$: $B = A \cup (B \setminus A)$ avec $A \cap (B \setminus A) = \emptyset$. Donc $\mathbb{P}\{B\} = \mathbb{P}\{A\} + \mathbb{P}\{B \setminus A\}$.

4) $A \cup B = A \cup (B \cap A^c)$, donc $\mathbb{P}\{A \cup B\} = \mathbb{P}\{A\} + \mathbb{P}\{B \cap A^c\}$. De plus, $\mathbb{P}\{B\} = \mathbb{P}\{B \cap A\} + \mathbb{P}\{B \cap A^c\}$, d'où le résultat. ♠

Indépendance ; probabilités conditionnelles

Rappelons que l'opération d'intersection des ensembles correspond au *et* logique des évènements.

DÉFINITION 10.3 Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un modèle probabiliste

1. On dit que deux évènements $A, B \in \mathcal{F}$ sont indépendants si

$$\mathbb{P}\{A \cap B\} = \mathbb{P}\{A\}\mathbb{P}\{B\}. \quad (10.9)$$

2. Une famille finie (A_1, A_2, \dots, A_n) d'éléments de \mathcal{F} est une famille d'évènements indépendants si

$$\mathbb{P}\{A_1 \cap A_2 \cap \dots \cap A_n\} = \mathbb{P}\{A_1\}\mathbb{P}\{A_2\} \dots \mathbb{P}\{A_n\}. \quad (10.10)$$

REMARQUE 10.1 Il est important de signaler que si (A_1, A_2, \dots, A_n) est une famille d'évènements indépendants, les évènements A_i sont indépendants deux à deux : $\mathbb{P}\{A_i \cap A_j\} = \mathbb{P}\{A_i\}\mathbb{P}\{A_j\}$ pour tous i, j . La réciproque n'est pas vraie, comme le montre l'exemple suivant.

EXEMPLE 10.5 On tire à pile ou faces 2 fois avec une pièce non biaisée. On considère les évènements suivants : $A = \{PP, PF\}$, $B = \{PF, FF\}$, $C = \{FF, PP\}$, qui ont tous probabilité 1/2. Il est facile de voir que A, B, C sont deux à deux indépendants, mais que $\mathbb{P}\{A \cap B \cap C\} = 0$, de sorte que la famille $\{A, B, C\}$ n'est pas une famille d'évènements indépendants.

Lorsque deux évènements A et B sont indépendants, le fait que A soit réalisé n'apporte aucune information sur le fait que B soit réalisé ou pas. Par contre, ça n'est plus le cas si A et B ne sont pas indépendants. Il est dans ce cas utile d'introduire la notion de *probabilité conditionnelle*, qui permet de calculer la probabilité $\mathbb{P}\{B|A\}$ de B , supposant A réalisé.

DÉFINITION 10.4 Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un modèle probabiliste. Si A et B sont deux évènements quelconques, la probabilité de A sachant B (ou conditionnée par B) est le nombre

$$\mathbb{P}\{A|B\} = \frac{\mathbb{P}\{A \cap B\}}{\mathbb{P}\{B\}} \quad (10.11)$$

EXEMPLE 10.6 On considère deux dés non pipés. Calculer la probabilité pour que le résultat obtenu avec le premier soit un 4, sachant que le résultat du "bi-tirage" appartient à l'ensemble

$$B = \{(3, 6), (4, 5), (4, 6), (5, 4), (5, 5), (5, 6), (6, 3), (6, 4), (6, 5), (6, 6)\}.$$

L'espace de configurations est $\Omega = \{(i, j), i, j = 1, \dots, 6\}$ et $\text{card}(\Omega) = 36$. Pour tout $\omega \in \Omega$, $\mathbb{P}\{\omega\} = 1/36$.

$$A = \{(4, 1), (4, 2), \dots, (4, 6)\}.$$

$\mathbb{P}\{A \cap B\} = 2/36$, et $\mathbb{P}\{B\} = 10/36$. On a donc

$$\mathbb{P}\{A|B\} = \frac{1}{5}.$$

Sous forme multiplicative, on peut aussi écrire

$$\mathbb{P}\{A \cap B\} = \mathbb{P}\{A|B\}\mathbb{P}\{B\}, \quad (10.12)$$

d'où on déduit la formule de Bayes

$$\mathbb{P}\{A|B\}\mathbb{P}\{B\} = \mathbb{P}\{B|A\}\mathbb{P}\{A\}. \quad (10.13)$$

On en déduit également le résultat plus général suivant :

PROPOSITION 10.3 Pour toute famille finie d'évènements (A_1, A_2, \dots, A_n) :

$$\mathbb{P}\{A_1 \cap A_2 \cap \dots \cap A_n\} = \mathbb{P}\{A_1\} \mathbb{P}\{A_2|A_1\} \mathbb{P}\{A_3|A_1 \cap A_2\} \dots \times \mathbb{P}\{A_n|A_1 \cap \dots \cap A_{n-1}\}. \quad (10.14)$$

Preuve : Il suffit de procéder de proche en proche :

$$\begin{aligned} \mathbb{P}\{A_1 \cap \dots \cap A_n\} &= \mathbb{P}\{A_n|A_1 \cap \dots \cap A_{n-1}\} \mathbb{P}\{A_1 \cap \dots \cap A_{n-1}\} \\ &= \mathbb{P}\{A_n|A_1 \cap \dots \cap A_{n-1}\} \mathbb{P}\{A_{n-1}|A_1 \cap \dots \cap A_{n-2}\} \\ &\quad \times \mathbb{P}\{A_1 \cap \dots \cap A_{n-2}\} \\ &= \dots \\ &= \mathbb{P}\{A_1\} \mathbb{P}\{A_2|A_1\} \mathbb{P}\{A_3|A_1 \cap A_2\} \dots \\ &\quad \times \mathbb{P}\{A_n|A_1 \cap \dots \cap A_{n-1}\}, \end{aligned}$$

ce qui prouve la proposition. ♠

Une autre propriété utile est la formule de probabilité totale

PROPOSITION 10.4 Pour toute famille dénombrable d'évènements (A_1, A_2, \dots) telle que $A_i \cap A_j = \emptyset$ pour tous i, j et $\cup A_i \subset \Omega$, on a pour tout $B \in \mathcal{F}$

$$\mathbb{P}\{B\} = \sum_n \mathbb{P}\{B|A_i\} \mathbb{P}\{A_i\}. \quad (10.15)$$

Preuve : Ecrivons la propriété de distributivité

$$\mathbb{P}\{B\} = \mathbb{P}\{B \cap (A_1 \cup A_2 \cup \dots)\} = \mathbb{P}\{(B \cap A_1) \cup (B \cap A_2) \cup \dots\}.$$

Les $B \cap A_i$ étant disjoints deux à deux, il suffit d'appliquer (10.8) puis la formule de Bayes. ♠

Probabilités et intégrales

La façon la plus usuelle de construire un modèle probabiliste dans le cas continu est basée sur l'heuristique suivante, que nous décrivons tout d'abord en dimension 1, par exemple dans le cas $\Omega \subset \mathbb{R}$. On a déjà vu que dans ce cas, les nombres $\mathbb{P}\{t\}$, $t \in \Omega$ n'ont pas de signification. On considère plutôt un petit intervalle de la forme $\Delta t = [t, t + \delta t]$, et on écrit

$$\mathbb{P}\{\Delta t\} = \mathbb{P}\{[t, t + \delta t]\} \approx \rho(t) \delta t;$$

ceci revient à écrire que la probabilité se comporte, en première approximation, de façon linéaire par rapport à la taille de l'intervalle, tout du moins lorsque celui-ci est petit, les variations sur des distances plus grandes étant décrites par une fonction ρ .

Etant donné maintenant un intervalle $I = [a, b]$, on commence par le "découper" en petits intervalles de tailles $\delta t_1, \delta t_2 \dots \delta t_N$, centrés sur les valeurs t_1, t_2, \dots, t_N , et on écrit

$$\mathbb{P}\{I\} = \rho(t_1) \delta t_1 + \rho(t_2) \delta t_2 + \dots + \rho(t_N) \delta t_N.$$

Or, ceci n'est autre qu'une approximation de l'intégrale

$$\int_I \rho(t) dt = \int_a^b \rho(t) dt$$

par une somme finie (à la Riemann). Ceci suggère de définir

$$\mathbb{P}\{I\} = \int_I \rho(t) dt .$$

De plus étant donnés K intervalles I_1, \dots, I_K , disjoints deux à deux, et si $A = I_1 \cup I_2 \cup \dots \cup I_K$, on peut écrire

$$\mathbb{P}\{A\} = \int_{I_1} \rho(t) dt + \int_{I_2} \rho(t) dt + \dots + \int_{I_K} \rho(t) dt = \int_A \rho(t) dt .$$

Ainsi, dans les cas où Ω est un ensemble infini continu, on peut (sauf dans certains cas pathologiques) définir l'application \mathbb{P} à partir d'une fonction ρ , par la formule

$$\mathbb{P}\{A\} = \int_A \rho(t) dt . \quad (10.16)$$

DÉFINITION 10.5 La fonction ρ ci-dessus est appelée densité de probabilités.

EXEMPLE 10.7 Si on reprend l'exemple ci-dessus d'un nombre tiré au hasard dans l'intervalle $[0, 1]$, on a bien

$$\mathbb{P}\{[a, b[\} = b - a = \int_a^b dt ,$$

de sorte que cette probabilité est bien construite à partir d'une densité, qui est dans ce cas

$$\rho(t) = \chi_{[0,1]}(t) .$$

REMARQUE 10.2 Nous avons utilisé l'intégrale de Riemann pour introduire la notion de densité de probabilités. Cependant, une formulation plus rigoureuse et générale de la théorie doit être basée sur la notion de mesure et l'intégrale de Lebesgue. Ceci ne change toutefois pas grand chose tant qu'il s'agit de faire des calculs explicites.

Lorsqu'un modèle probabiliste est construit à partir d'une densité de probabilités, celle-ci doit satisfaire un certain nombre de conditions. On en donne les deux les plus importantes ci-dessous.

PROPOSITION 10.5 Soit ρ une densité de probabilités sur $\Omega \subset \mathbb{R}$.

1. $\rho(\omega) \geq 0$ pour presque tout $\omega \in \Omega$.
2. $\int_{\Omega} \rho(t) dt = 1$

Preuve : Tout d'abord, supposons qu'il existe un ensemble $A \subset \Omega$ de mesure non nulle, tel que $\rho(\omega) < 0$ pour tout $\omega \in A$. Alors on aurait $\mathbb{P}\{A\} = \int_A \rho(\omega) d\omega < 0$, ce qui est impossible. Par ailleurs, on a évidemment $\int_{\Omega} \rho(\omega) d\omega = \mathbb{P}\{\Omega\} = 1$ ♠

REMARQUE 10.3 une densité n'est pas à proprement parler une probabilité. Elle est certes à valeurs positive, mais n'a aucune raison d'être toujours inférieure à 1 : il peut exister des valeurs t telles que $\rho(t) \geq 1$.

En dimension supérieure, on peut procéder de façon similaire. Par exemple, si $\Omega \subset \mathbb{R}^n$, on est conduit à introduire une densité de probabilités ρ , qui est cette fois une fonction de plusieurs variables, ou d'une variable vectorielle $\underline{x} = (x_1, \dots, x_n)$:

$$\rho : \Omega \rightarrow \mathbb{R}^+ ,$$

qui est telle que

$$\int_{\Omega} \rho(\underline{x}) d\underline{x} = 1 .$$

et

$$\rho(\underline{x}) \geq 0, \quad \forall \underline{x} \in \Omega .$$

On peut aussi s'intéresser à des Ω plus compliqués (par exemple en considérant des points tirés au hasard sur des sphères ou dans des objets géométriques complexes). On en verra des exemples par la suite.

Variables aléatoires

Nous étudierons les variables aléatoires en détails dans le chapitre suivant. On se contente ici de donner une idée grossière de la définition.

On considère un modèle probabiliste $(\Omega, \mathcal{F}, \mathbb{P})$. On dira qu'une fonction $X : \Omega \rightarrow U \subset \mathbb{R}$ (ou parfois \mathbb{C}) est une variable aléatoire si on peut lui associer une distribution de probabilités, c'est à dire des nombres

$$\mathbb{P}_X\{V\} = \mathbb{P}\{\omega \in \Omega : X(\omega) \in V\}$$

pour certains sous ensembles $V \subset U$, telle que

$$\mathbb{P}_X\{U\} = 1 .$$

On verra de multiples utilisations de cette notion essentielle.

10.1.3 Exemples

Les exemples suivants illustrent la souplesse qu'offre la modélisation probabiliste, qui permet de couvrir des situations très différentes. On va voir ici des exemples qui couvrent les trois cas de figures que nous avons vus, à savoir des espaces Ω finis, infinis dénombrables, et continus.

Modèles finis

Quelques formules utiles Les modèles finis sont souvent construits à partir de dénombrements. On voit ici quelques exemples simples.

1. Nombre d'applications $E \rightarrow F$ (où E, F sont deux ensembles finis) :

$$\text{Card}(F^E) = \text{Card}(F)^{\text{Card}(E)} .$$

2. Nombre d'injections $E \rightarrow F$ (où E, F sont deux ensembles finis, avec $r = \text{Card}(E) \leq n = \text{Card}(F)$) :

$$A_n^r = \frac{n!}{(n-r)!} .$$

Ce nombre représente le nombre de façons différentes de ranger r objets différents dans n boîtes, avec au maximum un objet par cellule.

En particulier, le nombre de bijections $E \rightarrow E$ est donné par $A_n^n = n!$ (aussi appelé nombre d'arrangements de E).

3. Nombre de parties à r éléments de F :

$$C_n^r = \binom{n}{r} = \frac{A_n^r}{r!} = \frac{n!}{r!(n-r)!} .$$

Ce nombre représente le nombre de façons différentes de ranger r objets identiques dans n boîtes (permutations), avec au maximum un objet par boîte.

Le modèle uniforme fini Le modèle uniforme consiste en un ensemble fini $\Omega = \{\omega_1, \dots, \omega_N\}$, auquel on associe des probabilités

$$\mathbb{P}\{\omega_i\} = \frac{1}{N}, \quad i = 1, \dots, N.$$

On prend alors pour \mathcal{F} l'ensemble (fini) de toutes les parties de Ω , et il est facile de voir que pour tout k -uplet de ω_i distincts $\omega_{i_1}, \dots, \omega_{i_k}$, on a $\mathbb{P}\{\omega_{i_1}, \dots, \omega_{i_k}\} = k/N$.

Le modèle binomial On considère tout d'abord une expérience à 2 résultats possibles $\Omega_0 = \{\omega_1, \omega_2\}$, avec probabilités

$$\mathbb{P}\{\omega_1\} = p_1, \quad \mathbb{P}\{\omega_2\} = p_2 = 1 - p_1.$$

Considérons maintenant N réalisations de cette expérience. Les résultats possibles sont des n -uplets $(\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_N})$, qui engendrent l'ensemble

$$\Omega_1 = \{(\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_N}) \in \Omega_0^N\},$$

et on a

$$\mathbb{P}\{(\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_N})\} = p_1^{n_1} p_2^{n_2},$$

où n_i est le nombre d'occurrences de ω_i dans le n -uplet $(\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_N})$. Supposons maintenant que l'on ne s'intéresse qu'aux nombres n_1 et n_2 . Comme $n_2 = N - n_1$, il suffit de s'intéresser à n_1 . Les valeurs possibles de n_1 forment l'ensemble

$$\Omega = \{0, 1, \dots, N\},$$

et la probabilité d'une valeur particulière de n_1 est de la forme

$$\mathbb{P}\{n_1\} = C_N^{n_1} p_1^{n_1} p_2^{N-n_1}. \quad (10.17)$$

On a noté ici C_N^n le nombre de façons différentes d'obtenir n fois le résultat ω_1 en N tirages. Il est bien connu que l'on a

LEMME 10.1

$$C_N^n = \frac{N!}{n!(N-n)!}. \quad (10.18)$$

Preuve : Procédons par récurrence et supposons le résultat valide jusqu'à l'ordre m , pour tout $n = 0, \dots, m$. Pour obtenir n fois ω_1 en m tirages, il faut soit avoir obtenu n fois ω_1 en $m-1$ tirages (C_{m-1}^n chances), soit avoir obtenu $n-1$ fois ω_1 en $m-1$ tirages (C_{m-1}^{n-1} chances), et l'obtenir au m -ième tirage.

$$C_m^n = C_{m-1}^n + C_{m-1}^{n-1} = \frac{(m-1)!}{n!(m-1-n)!} + \frac{(m-1)!}{(n-1)!(m-n)!} = \frac{m!}{n!(m-n)!}.$$

Comme $C_0^0 = 1$, ceci prouve le résultat. ♠

La formule du binôme donne immédiatement

$$\sum_{n=0}^N C_N^n p_1^n p_2^{N-n} = (p_1 + p_2)^N = 1.$$

Comme précédemment, on prend pour \mathcal{F} l'ensemble (fini) de toutes les parties de Ω . Le modèle correspondant est appelé *modèle binomial*, et est donc caractérisé par

$$\mathbb{P}\{n\} = C_N^n p_1^n p_2^{N-n}.$$

Le modèle multinomial On considère tout d'abord une expérience à K résultats possibles $\Omega_0 = \{\omega_1, \omega_2, \dots, \omega_K\}$, avec probabilités

$$\mathbb{P}\{\omega_1\} = p_1, \mathbb{P}\{\omega_2\} = p_2, \dots, \mathbb{P}\{\omega_K\} = p_K,$$

avec

$$p_1 + p_2 + \dots + p_K = 1.$$

Considérons maintenant N réalisations de cette expérience. Les résultats possibles sont des n -uplets $(\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_N})$, qui engendrent l'ensemble

$$\Omega_1 = \{(\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_N}) \in \Omega_0^N\},$$

et on a

$$\mathbb{P}\{(\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_N})\} = p_1^{n_1} p_2^{n_2} \dots p_K^{n_K},$$

où n_i est le nombre d'occurrences de ω_i dans le n -uplet $(\omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_N})$. On a évidemment

$$n_1 + n_2 + \dots + n_K = N.$$

Supposons maintenant que l'on ne s'intéresse qu'aux nombres n_1, \dots, n_K . Leurs valeurs possibles forment un ensemble noté Ω , et la probabilité d'un $\omega = \{n_1, n_2, \dots, n_K\}$ particulier est de la forme

$$\mathbb{P}\{n_1, n_2, \dots, n_K\} = \mathcal{N}(N; n_1, n_2, \dots, n_K) p_1^{n_1} p_2^{n_2} \dots p_K^{n_K}. \quad (10.19)$$

Les nombres $\mathcal{N}(N; n_1, \dots, n_K)$ sont obtenus pas un calcul similaire au précédent :

LEMME 10.2

$$\mathcal{N}(N; n_1, \dots, n_K) = \frac{N!}{n_1! n_2! \dots n_K!}. \quad (10.20)$$

Le modèle correspondant est appelé *modèle multinomial*.

Le modèle hypergéométrique Le modèle hypergéométrique décrit des situations où un choix aléatoire est effectué dans une population d'objets de deux types : par exemple, choisir 8 individus dans une population de 10 garçons et 7 filles. Plus généralement, on considère une population de N objets, dont n sont de type 1 et $N - n$ sont de type 2. On en tire m au hasard, et on s'intéresse à la probabilité d'en avoir tiré k de type 1 (et donc $m - k$ de type 2). On peut montrer que cette probabilité vaut

$$\mathbb{P}\{k\} = \begin{cases} \frac{C_m^k C_{N-m}^{n-k}}{C_N^n} & \text{si } \max(0, n - (N - m)) \leq k \leq \min(m, n) \\ 0 & \text{sinon.} \end{cases} \quad (10.21)$$

On parle alors de *modèle hypergéométrique* $\mathcal{H}(N, n, m)$.

Un autre exemple est donné par les problèmes de type "loterie", dans lesquelles on effectue n tirages (sans remise), parmi N résultats possibles. Un joueur choisit (ou mise sur) m résultats possibles (bien entendu, $m \leq N$), et la probabilité qu'il ait k bons numéros vaut précisément le $\mathbb{P}\{k\}$ donné ci-dessus.

Modèles infinis dénombrables

Le modèle de Pascal Le modèle de Pascal (on parle aussi de modèle géométrique) est donné par

$$\Omega = \mathbb{Z}^+,$$

et \mathcal{F} est l'ensemble (infini dénombrable) des parties de Ω . La distribution de probabilités correspondante est complètement caractérisée par un nombre $p \in]0, 1[$, par

$$\mathbb{P}\{n\} = p^n (1 - p). \quad (10.22)$$

On vérifie facilement que

$$\sum_{n=0}^{\infty} \mathbb{P}\{n\} = (1 - p) \sum_{n=0}^{\infty} p^n = 1,$$

car il s'agit de la somme d'une série géométrique, qui converge puisque $0 < p < 1$.

Le modèle de Poisson Le modèle de Poisson est généralement utilisé pour décrire la fréquence d'apparition d'événements rares. C'est un modèle dans lequel l'ensemble des résultats possibles Ω est infini dénombrable :

$$\Omega = \mathbb{Z}^+,$$

et on a, pour $n \in \mathbb{Z}^+$,

$$\mathbb{P}\{n\} = \frac{\lambda^n}{n!} e^{-\lambda}, \quad (10.23)$$

où λ est un réel positif fixé. On vérifie que

$$\sum_{n=0}^{\infty} \mathbb{P}\{n\} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = 1.$$

REMARQUE 10.4 Le modèle de Poisson peut être obtenu comme une certaine limite d'un modèle binomial. En effet, partons d'un modèle binomial comme plus haut, et supposons que la probabilité p_1 soit très faible et que le nombre d'expériences N soit très grand :

$$p_1 \ll 1, \quad N \gg 1.$$

Alors, pour $n \ll N$, et en posant $\lambda = Np_1$, on peut écrire

$$\log \mathbb{P}\{0\} = \log \left((1 - p_1)^N \right) = N \log \left(1 - \frac{\lambda}{N} \right) = -N \left(\frac{\lambda}{N} + O \left(\left(\frac{\lambda}{N} \right)^2 \right) \right),$$

de sorte que l'on peut écrire

$$\mathbb{P}\{0\} \approx e^{-\lambda},$$

l'approximation étant d'autant plus valide que λ^2/N est petit. De même, on a aussi

$$\frac{\mathbb{P}\{n+1\}}{\mathbb{P}\{n\}} = \frac{(N-k)!n!}{(N-n-1)!(n-1)!} \frac{p_1}{p_2} = \frac{\lambda}{n+1} \left[\left(1 - \frac{n}{N} \right) \frac{1}{p_2} \right] \approx \frac{\lambda}{n+1},$$

si on utilise l'hypothèse $n \ll N$ et $p_1 \ll 1$. Ceci conduit alors à

$$\mathbb{P}\{n\} \approx \frac{\lambda^n}{n!} e^{-\lambda},$$

c'est à dire à un modèle Poissonien.

Modèles continus

Le modèle exponentiel Le modèle exponentiel est généralement utilisé pour décrire le temps d'attente avant un évènement rare. Il s'agit d'un modèle continu :

$$\Omega = \mathbb{R}^+,$$

et

$$\mathbb{P}\{[t, \infty[\} = e^{-\lambda t}, \quad (10.24)$$

où $\lambda \in \mathbb{R}^+$ caractérise le modèle. On voit facilement que

$$\mathbb{P}\{[a, b] \} = \mathbb{P}\{[a, \infty[\} - \mathbb{P}\{[b, \infty[\} = e^{-\lambda a} - e^{-\lambda b} = \int_a^b \rho(t) dt,$$

d'où une densité de probabilités de la forme

$$\rho(t) = \lambda e^{-\lambda t}. \quad (10.25)$$

de nouveau, il est facile de vérifier que

$$\int_0^\infty \rho(t) dt = \lambda \int_0^\infty e^{-\lambda t} dt = 1.$$

Le modèle exponentiel peut aussi être obtenu comme une limite de modèles discrets, comme le montre l'exemple suivant.

EXEMPLE 10.8 On s'intéresse à la désintégration d'un atome (disons, un atome de radium), et on s'intéresse au temps de survie de l'atome. Les valeurs possibles sont des nombres positifs (ou nuls), de sorte que l'on prend $\Omega = \mathbb{R}^+$.

Pour décrire le processus en termes probabilistes, on commence par découper Ω en intervalles de taille δt , et on fait le modèle suivant : si l'atome ne s'est toujours pas désintégré à l'instant t_0 , on suppose qu'il a une probabilité $p_{\delta t}$ de se désintégrer entre t_0 et $t_0 + \delta t$, et une probabilité égale à $q_{\delta t} = 1 - p_{\delta t}$ de ne pas se désintégrer entre t_0 et $t_0 + \delta t$. L'hypothèse essentielle qui est faite ici est que ces deux probabilités ne dépendent pas de t_0 . Par contre, il est normal de supposer que ces deux probabilités dépendent de δt (et que $p_{\delta t}$ soit une fonction croissante de δt).

On peut alors calculer la probabilité pour que l'atome ne se soit toujours pas désintégré après le temps t :

$$\mathbb{P}\{[t, \infty[\} = (q_{\delta t})^{t/\delta t} = \left(q_{\delta t}^{1/\delta t}\right)^t.$$

Supposons que la limite $\lim_{\delta t \rightarrow 0} q_{\delta t}^{1/\delta t}$ existe, elle est nécessairement comprise entre 0 et 1. On pose alors

$$\lim_{\delta t \rightarrow 0} q_{\delta t}^{1/\delta t} = e^{-\lambda},$$

où λ est un nombre positif, et on obtient ainsi un "modèle limite" exponentiel.

Le modèle uniforme continu Le modèle uniforme est un autre exemple de modèle à ensemble continu, et correspond à la notion d'équiprobabilité : dans un modèle uniforme, les résultats possibles ont tous même probabilité.

Commençons par le cas le plus simple : Etant donné un intervalle

$$\Omega = [a, b] \subset \mathbb{R},$$

on définit le modèle uniforme sur $[a, b]$ en associant une probabilité à tout intervalle $[u, v[\subset [a, b]$:

$$\mathbb{P}\{[u, v[\} = \frac{v - u}{b - a}, \quad (10.26)$$

et en utilisant (10.8) pour associer des probabilités à des sous-ensembles plus compliqués de Ω . Il est facile de vérifier que ce modèle uniforme peut être construit à partir d'une densité de probabilités ρ de la forme

$$\rho(x) = \frac{1}{b-a} \chi_{[a,b]}(x). \quad (10.27)$$

Plus généralement, il est possible de construire des modèles uniformes dans des domaines plus complexes que des intervalles. Par exemple, étant donné un domaine borné V de dimension n , le modèle uniforme sur V est donné par les probabilités $\mathbb{P}\{A\}$ des sous-ensembles $A \subset V$

$$\mathbb{P}\{A\} = \frac{\text{vol}(A)}{\text{vol}(V)}, \quad (10.28)$$

où l'application *volume* vol associe à $A \subset V$ sa longueur (en dimension 1), son aire (en dimension 2), son volume (en dimension 3) ou les généralisations en dimension supérieure.

Par exemple, prenons le cas d'un disque D de rayon fixé R . Sa surface vaut πR^2 . Étant donné un sous-ensemble $A \subset D$, sa probabilité vaut

$$\mathbb{P}\{A\} = \frac{1}{\pi R^2} \int_A d\underline{x},$$

de sorte que l'on peut caractériser le modèle par la densité de probabilités

$$\rho(\underline{x}) = \frac{1}{\pi R^2} \chi_D(\underline{x}).$$

Notons qu'il s'agit d'une fonction qui est constante sur son support.

REMARQUE 10.5 Il est important de signaler ici une propriété "désagréable" de la densité de probabilités : elle change de forme lorsque l'on change de variables : en effet, considérons l'écriture de \underline{x} en coordonnées polaires : $\underline{x} = (r, \theta)$. En effectuant le changement de variables $(r, \theta) = \phi(x)$, on a alors

$$\mathbb{P}\{A\} = \frac{1}{\pi R^2} \int_A r \, dr \, d\theta,$$

de sorte qu'une fois écrite en coordonnées polaires, la densité de probabilités correspondante est donnée par

$$\rho_p(r, \theta) = \rho(\phi(\underline{x})) = \frac{r}{\pi R^2} \chi_{[0,R]}(r) \chi_{[0,2\pi]}(\theta),$$

et n'est donc plus une fonction constante de θ et r . On verra cette propriété de façon plus précise en discutant les variables aléatoires.

Le modèle normal (ou Gaussien) Le modèle normal joue un rôle central en théorie des probabilités et en physique, car il est obtenu comme cas "limite" de bon nombre d'autres modèles. Il s'agit d'un modèle à espace Ω infini continu :

$$\Omega = \mathbb{R}, \quad \text{ou } \Omega = \mathbb{R}^n.$$

Dans le cas unidimensionnel, il est donné par la densité de probabilités Gaussienne :

$$\rho(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(t-\mu)^2/2\sigma^2}, \quad (10.29)$$

où $\sigma \in \mathbb{R}^+$ et $\mu \in \mathbb{R}$ sont deux paramètres caractérisant respectivement la largeur de ρ et sa localisation. Pour tout intervalle $[a, b[\subset \mathbb{R}$, on a

$$\mathbb{P}\{[a, b[\} = \frac{1}{\sigma\sqrt{2\pi}} \int_a^b e^{-(t-\mu)^2/2\sigma^2} dt,$$

et de nouveau $\int_{-\infty}^{\infty} \rho(t) dt = 1$.

10.2 Variables aléatoires

10.2.1 Généralités

Les modèles probabilistes que nous avons décrits dans le chapitre précédent ne donnent pas un cadre suffisamment large pour les applications. On introduit alors les *variables aléatoires*, qui permettent d'enrichir le calcul des probabilités.

Pour introduire la notion de variable aléatoire, il est bon d'illustrer tout d'abord par quelques exemples.

EXEMPLE 10.9 Dans une expérience de N tirages de "pile ou face", à chaque $\omega \in \Omega$ on peut associer le nombre de fois où on a obtenu face dans $\omega = (\omega_{i_1}, \dots, \omega_{i_N})$. Ceci définit donc une fonction

$$N : \omega \in \Omega \rightarrow N(\omega) \in \mathbb{Z}^+,$$

qui est un premier exemple de variable aléatoire.

EXEMPLE 10.10 On tire au hasard, selon un modèle uniforme, un point dans un carré $[0, 1] \times [0, 1]$, et on s'intéresse à la distance à l'origine :

$$R : \omega \in [0, 1] \times [0, 1] \rightarrow R(\omega) \in [0, \sqrt{2}].$$

R est encore une fonction $\Omega \rightarrow [0, \sqrt{2}]$, et est un autre exemple de variable aléatoire.

En termes imagés, les variables aléatoires permettent de manipuler des "produits dérivés" des modèles probabilistes.

DÉFINITION 10.6 *Etant donné un modèle probabiliste $(\Omega, \mathcal{F}, \mathbb{P})$, on appelle variable aléatoire toute fonction mesurable*

$$X : \omega \in \Omega \rightarrow X(\omega) \in \mathbb{R} \text{ (ou } \mathbb{Z} \text{)},$$

c'est à dire telle que pour tout intervalle $I \subset \mathbb{R}$, son image inverse $X^{-1}(I)$ par X appartienne à \mathcal{F} .

Soit maintenant un sous ensemble Borélien $A \subset \mathbb{R}$ (c'est à dire une partie de \mathbb{R} à laquelle on sait associer une mesure). On peut alors s'intéresser à la quantité

$$\mathbb{P}_X\{A\} = \mathbb{P}\{\omega \in \Omega : X(\omega) \in A\}, \quad (10.30)$$

qui est bien définie.

DÉFINITION 10.7 *La collection des nombres $\mathbb{P}_X\{A\}$, où $A \subset \mathbb{R}$ est un sous-ensemble Borélien de \mathbb{R} , est appelée distribution de probabilités de la variable aléatoire X .*

Dans la suite, on utilisera indistinctement les notations

$$\mathbb{P}_X\{A\} = \mathbb{P}\{\omega \in \Omega : X(\omega) \in A\} = \mathbb{P}\{X(\omega) \in A\} = \mathbb{P}\{X \in A\}.$$

REMARQUE 10.6 Il est utile de faire ici un certain nombre de remarques.

1. En pratique, on n'a pas à se donner les nombres $\mathbb{P}_X\{A\}$ pour tout $A \subset \mathbb{R}$. On se contente d'une règle permettant de calculer ces nombres. On en verra de nombreux exemples dans ce qui suit.
2. Dans le cas où $\Omega \subset \mathbb{R}$, la variable aléatoire la plus simple correspond à la fonction identité : $X(\omega) = \omega$.
3. Souvent, il n'est pas nécessaire de connaître les détails du modèle probabiliste sous-jacent pour caractériser la distribution d'une variable aléatoire. C'est d'ailleurs préférable, dans la mesure où les ensembles Ω sont souvent gigantesques et difficiles à manipuler (en particulier pour ce qui est des applications en physique).

EXEMPLE 10.11 Si nous reprenons un des exemples que nous avons déjà abordés précédemment, à savoir celui qui nous a conduits au modèle binomial, on considère l'ensemble Ω_1 de toutes les configurations possibles en N tirages. A chaque $a \in \Omega_1$ on peut alors associer le nombre $X(a)$ de fois où ω_1 est apparu. Cette fonction X définit une variable aléatoire dont la distribution est donnée par

$$\mathbb{P}_X\{n\} = \mathbb{P}\{a \in \Omega_1, X(a) = n\} = C_N^n p_1^n p_2^{N-n},$$

et il s'agit d'une distribution binomiale.

De même que nous avons vu des modèles probabilistes discrets et continus, on peut considérer des variables aléatoires prenant leurs valeurs dans un espace discret (par exemple \mathbb{Z}), ou continu. Plus précisément, on distingue trois types de variables aléatoires : les variables aléatoires discrètes (c'est à dire les fonctions X prenant un nombre fini ou infini dénombrable de valeurs), et deux catégories de variables aléatoires continues : des variables aléatoires dites "à densité", qui sont caractérisées par une fonction appelée *densité de probabilités*, et des variables aléatoires dites "mixtes", dont la distribution de probabilités ne peut être caractérisée par une densité.

On discutera ici les deux premiers cas (discret et continu) séparément, évitant les variables aléatoires mixtes, dont le traitement est plus complexe. Le cas discret nous permettra d'introduire des notions dont la plupart se généralisent facilement au cas continu.

10.2.2 Variables aléatoires discrètes

Commençons donc par le cas des variables aléatoires discrètes. Etant donné un modèle probabiliste $(\Omega, \mathcal{F}, \mathbb{P})$, on considère donc la fonction mesurable $X : \Omega \rightarrow \mathbb{R}$, dont on suppose qu'elle ne prend qu'un nombre fini ou infini dénombrable de valeurs $\{x_1, x_2, \dots\}$. Il est clair que la distribution de probabilités de X est complètement caractérisée par la donnée des nombres $\mathbb{P}_X\{x_i\}$: en effet, on peut facilement en déduire la probabilité pour que X prenne une des valeurs $x_{k_1}, x_{k_2}, \dots, x_{k_n}$ grâce à la règle

$$\mathbb{P}_X\{x_{k_1}, x_{k_2}, \dots, x_{k_n}\} = \mathbb{P}_X\{x_{k_1}\} + \mathbb{P}_X\{x_{k_2}\} + \dots + \mathbb{P}_X\{x_{k_n}\}.$$

Fonction de répartition

La fonction de répartition constitue un intermédiaire de calcul souvent utile (dans le cas discret et plus encore dans le cas continu).

DÉFINITION 10.8 Soit X une variable aléatoire sur $(\Omega, \mathcal{F}, \mathbb{P})$. Sa fonction de répartition est la fonction $F_X : \mathbb{R} \rightarrow [0, 1]$ définie par

$$F_X(x) = \mathbb{P}_X\{] - \infty, x[\} = \mathbb{P}\{\omega \in \Omega : X(\omega) < x\}. \quad (10.31)$$

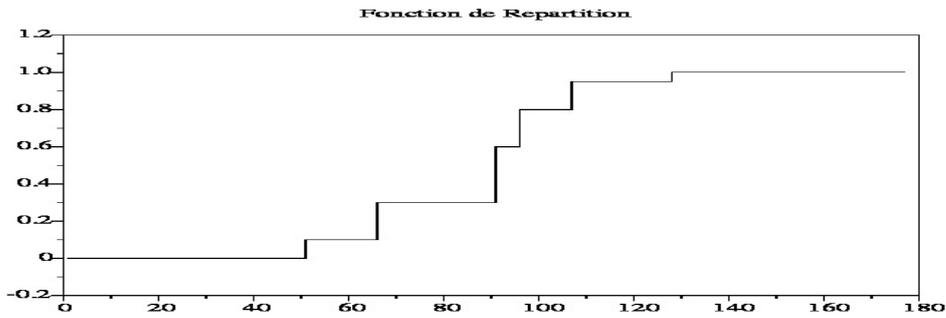


FIG. 10.1: Fonction de répartition d'une variable aléatoire discrète

Si X est une variable aléatoire discrète à valeurs dans $\{x_1, x_2, \dots\}$, on a alors

$$F_X(x) = \sum_{k: x_k < x} \mathbb{P}_X\{x_k\}. \quad (10.32)$$

La fonction de répartition possède un certain nombre de propriétés intéressantes, faciles à démontrer :

- PROPRIÉTÉ 10.1**
1. F_X est monotone : si $x \leq y$, alors $F_X(x) \leq F_X(y)$.
 2. Par conséquent, pour tout x , $0 \leq F_X(x) \leq \lim_{y \rightarrow \infty} F_X(y) = 1$.
 3. La fonction de répartition est continue à gauche : pour tout x ,

$$\lim_{y \rightarrow x, y \leq x} F_X(y) = F_X(x).$$

En fait, quand X est une variable aléatoire discrète, sa fonction de répartition est une fonction constante par morceaux, comme on peut le voir dans l'exemple donné en FIG. 10.1.

Fonction d'une variable aléatoire : transfert de loi

Etant donnée la distribution d'une variable aléatoire X , et une fonction $\varphi : \mathbb{R} \rightarrow \mathbb{R}$, on peut en déduire la distribution de probabilités d'une nouvelle variable aléatoire (on parle alors de transfert de loi)

$$Y = \varphi(X) : \omega \in \Omega \rightarrow Y(\omega) = \varphi(X(\omega)),$$

à condition que φ vérifie des conditions de mesurabilité appropriées. On peut alors calculer

$$\mathbb{P}_Y\{y\} = \mathbb{P}\{\omega : \varphi(X(\omega)) = y\} = \mathbb{P}\{\omega : X(\omega) \in \varphi^{-1}(\{y\})\} = \mathbb{P}_X\{\varphi^{-1}(\{y\})\}, \quad (10.33)$$

où $\varphi^{-1}(\{y\})$ est l'image inverse du singleton $\{y\}$ par la fonction φ . Si φ est bijective, les choses sont plus simples car $\varphi^{-1}(\{y\})$ est le singleton $\varphi^{-1}(y)$, et on peut écrire

$$\mathbb{P}_Y\{y\} = \mathbb{P}\{\omega : X(\omega) = \varphi^{-1}(y)\} = \mathbb{P}_X\{\varphi^{-1}(y)\}.$$

EXEMPLE 10.12 Prenons l'exemple d'une variable aléatoire X à valeurs entières : $X(\omega) \in \mathbb{Z}$ pour tout $\omega \in \Omega$, et la fonction $\varphi(x) = x^2$. Alors pour $y \neq 0$ on a

$$\mathbb{P}_Y\{y\} = \mathbb{P}\{\omega : X(\omega) = \pm\sqrt{y}\} = \mathbb{P}_X\{\sqrt{y}\} + \mathbb{P}_X\{-\sqrt{y}\},$$

et $\mathbb{P}_Y\{0\} = \mathbb{P}_X\{0\}$.

Prenons le cas particulier d'une variable aléatoire X Poissonnienne : $\mathbb{P}_X\{n\} = e^{-\lambda} \lambda^n / n!$. Comme X ne prend que des valeurs positives, φ est une bijection de \mathbb{Z}^+ sur $\varphi(\mathbb{Z}^+)$, et on peut écrire

$$\mathbb{P}_Y\{y\} = \begin{cases} e^{-\lambda} \lambda^{\sqrt{y}} / (\sqrt{y}!) & \text{si } \sqrt{y} \in \mathbb{Z}^+ \\ 0 & \text{sinon.} \end{cases}$$

Espérance, variance

Espérance mathématique L'espérance mathématique est la notion probabiliste qui permet de décrire la valeur moyenne d'une variable aléatoire. La définition est donnée par

DÉFINITION 10.9 Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{F}, \mathbb{P})$. Son espérance mathématique est donnée par

$$\mathbb{E}\{X\} = \sum_k x_k \mathbb{P}_X\{x_k\}, \quad (10.34)$$

où les x_k sont les valeurs prises par X .

L'espérance mathématique est une application linéaire : si X, Y sont deux variables aléatoires sur le même modèle probabiliste, alors $\mathbb{E}\{X + Y\} = \mathbb{E}\{X\} + \mathbb{E}\{Y\}$. De même, si $\lambda \in \mathbb{R}$, on a $\mathbb{E}\{\lambda X\} = \lambda \mathbb{E}\{X\}$.

Une autre propriété importante est la propriété suivante :

$$\mathbb{E}\{|X|\} = 0 \text{ implique } X(\omega) = 0 \text{ presque sûrement.} \quad (10.35)$$

La signification précise de cette dernière assertion est la suivante : si $\mathbb{E}\{|X|\} = 0$, alors $\mathbb{P}\{X(\omega) = 0\} = 1$.

EXEMPLE 10.13 Prenons l'exemple de la roulette : 37 résultats possibles ($\Omega = \{0, d \dots 36\}$). Si on mise une somme x_0 sur l'une des 37 cases, le gain est de $36x_0$ en cas de tirage favorable, et de 0 sinon. On peut donc modéliser le gain comme une variable aléatoire X qui vaut $36x_0$ avec probabilité $p = 1/37$, et 0 avec probabilité $1 - p = 36/37$. L'espérance de X vaut alors

$$\mathbb{E}\{X\} = \frac{36}{37} x_0.$$

On peut également miser sur rouge ou noir, ou sur pair ou impair. Dans ce cas (comme le 0 est considéré comme "ni pair ni impair", et est associé à la couleur verte), la probabilité de gagner est $p = 18/37$. Si on mise x_0 , le gain est une variable aléatoire X qui vaut $2x_0$ en cas de tirage gagnant, et 0 sinon. L'espérance du gain vaut de nouveau $\mathbb{E}\{X\} = 36x_0/37$.

EXEMPLE 10.14 *Espérance d'une variable aléatoire de loi géométrique* : Considérons le cas d'une variable aléatoire X de distribution géométrique. X prend des valeurs entières positives ou nulles, et sa distribution est complètement caractérisée par un nombre $p \in]0, 1[$, par

$$\mathbb{P}_X\{k\} = p^k (1 - p), \quad k = 0, 1, \dots$$

On peut alors calculer

$$\begin{aligned}
 \mathbb{E}\{X\} &= \sum_{k=0}^{\infty} k p^k (1-p) \\
 &= p(1-p) \sum_{k=1}^{\infty} k p^{k-1} \\
 &= p(1-p) \frac{d}{dp} \sum_{k=1}^{\infty} p^k \\
 &= p(1-p) \frac{d}{dp} \left(\frac{1}{1-p} - 1 \right) \\
 &= \frac{p(1-p)}{(1-p)^2} = \frac{p}{1-p},
 \end{aligned}$$

où on a utilisé la somme de la série géométrique, qui converge puisque $p < 1$.

REMARQUE 10.7 Il existe des variables aléatoires pour lesquelles l'espérance mathématique n'est pas définie, ou plutôt telles que la somme dans l'équation (10.34) soit divergente. C'est le cas par exemple de la variable aléatoire X à valeurs dans \mathbb{Z}^+ définie par $\mathbb{P}_X\{k\} = 6/(k\pi)^2$, $k = 1, 2, \dots$. On a bien $\sum_1^{\infty} \mathbb{P}_X\{k\} = 1$, mais $\sum_1^{\infty} k \mathbb{P}_X\{k\}$ est divergente.

Espérance d'une fonction d'une variable aléatoire Soit X une variable aléatoire discrète prenant les valeurs x_1, x_2, \dots avec probabilités p_1, p_2, \dots , et considérons la variable aléatoire $Y = X^2$. Si les x_i sont tous positifs, les nombres x_i^2 sont tous différents, et Y prendra donc les valeurs x_1^2, x_2^2, \dots avec probabilités p_1, p_2, \dots , et on pourra écrire

$$\mathbb{E}\{Y\} = \sum_i x_i^2 p_i.$$

Supposons maintenant que les x_i^2 ne soient pas tous différents, par exemple que $x_1^2 = x_2^2$, et que tous les autres x_i^2 soient deux à deux différents. Alors $\mathbb{P}_X\{x_1^2\} = p_1 + p_2$, et $\mathbb{E}\{Y\} = (p_1 + p_2)x_1^2 + \sum_{i \notin \{1,2\}} x_i^2 p_i$, qui coïncide avec l'équation précédente. On vérifie de même que cette expression reste inchangée dans les cas où plus de deux valeurs de x_i^2 coïncident. Ceci suggère la définition suivante :

DÉFINITION 10.10 Soient X et $Y = \varphi(X)$ deux variables aléatoires discrètes. On note x_1, x_2, \dots les valeurs prises par X . L'espérance de $Y = \varphi(X)$ est donnée par

$$\mathbb{E}\{Y\} = \sum_i \varphi(x_i) \mathbb{P}_X\{x_i\}. \quad (10.36)$$

On définit ainsi (lorsque les séries correspondantes sont convergentes) les moments de X

$$\mathbb{E}\{X^n\} = \sum_i x_i^n \mathbb{P}_X\{x_i\}, \quad (10.37)$$

la variance de X :

$$\text{Var}\{X\} = \mathbb{E}\{(X - \mathbb{E}\{X\})^2\}, \quad (10.38)$$

et l'écart type de X

$$\sigma_X = \sqrt{\text{Var}\{X\}}.$$

On peut en particulier remarquer que

$$\text{Var} \{X\} = \mathbb{E} \{X^2\} - \mathbb{E} \{X\}^2 .$$

La variance est une opération non-linéaire : $\text{Var} \{X + Y\} \neq \text{Var} \{X\} + \text{Var} \{Y\}$ en général (lorsque l'égalité est vérifiée, on dit que X et Y sont décorréliées). De plus, si $\lambda \in \mathbb{R}$, on a $\text{Var} \{\lambda X\} = \lambda^2 \text{Var} \{X\}$.

EXEMPLE 10.14 (SUITE) On peut calculer la variance d'une variable aléatoire de loi géométrique en utilisant les mêmes techniques que pour le calcul de l'espérance :

$$\begin{aligned} \mathbb{E} \{X^2\} &= \sum_{k=0}^{\infty} k^2 p^k (1-p) \\ &= \mathbb{E} \{X\} + \sum_{k=0}^{\infty} k(k-1) p^k (1-p) \\ &= \frac{p}{1-p} + p^2(1-p) \sum_{k=2}^{\infty} k(k-1) p^{k-2} \\ &= \frac{p}{1-p} + p^2(1-p) \frac{d^2}{dp^2} \sum_{k=2}^{\infty} p^k \\ &= \frac{p}{1-p} + p^2(1-p) \frac{d^2}{dp^2} \left(\frac{1}{1-p} - 1 - p \right) \\ &= \frac{p}{1-p} + \frac{2p^2}{(1-p)^2} = \frac{p+p^2}{(1-p)^2} , \end{aligned}$$

d'où on déduit la variance :

$$\text{Var} \{X\} = \frac{p}{(1-p)^2} .$$

EXEMPLE 10.15 Considérons une variable aléatoire Poissonnienne X , de paramètre $\lambda > 0$. Les valeurs prises par X sont donc $0, 1, 2, \dots$, avec des probabilités données par une loi de Poisson :

$$\mathbb{P}_X \{k\} = \frac{\lambda^k}{k!} e^{-\lambda} .$$

Un calcul direct donne

$$\mathbb{E} \{X\} = \sum_0^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_1^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda .$$

De même,

$$\mathbb{E} \{X^2\} = \mathbb{E} \{X\} + \sum_0^{\infty} k(k-1) \frac{\lambda^k}{k!} e^{-\lambda} = \lambda + \lambda^2 e^{-\lambda} \sum_2^{\infty} \frac{\lambda^{k-2}}{(k-2)!} = \lambda + \lambda^2 ,$$

de sorte que

$$\text{Var} \{X\} = \lambda .$$

REMARQUE 10.8 On peut construire de même l'espérance de fonctions $\varphi(X)$ de X dans le cas où φ est à valeurs complexes, comme on le verra en étudiant les fonctions caractéristiques.

La fonction caractéristique La fonction caractéristique (parfois appelée fonction génératrice) est un intermédiaire de calcul souvent très utile lorsqu'il s'agit de calculer les moments d'une variable aléatoire.

DÉFINITION 10.11 Soit X une variable aléatoire discrète. La fonction caractéristique de X est la fonction $G_X : \mathbb{R} \rightarrow \mathbb{C}$ définie par

$$G_X(u) = \mathbb{E} \left\{ e^{-iuX} \right\} = \sum_k e^{-iux_k} \mathbb{P}_X \{ x_k \}. \quad (10.39)$$

Il est facile de voir que cette série est absolument convergente, puisque

$$|G_X(u)| \leq \sum_k \left| e^{-iux_k} \mathbb{P}_X \{ x_k \} \right| = \sum_k \mathbb{P}_X \{ x_k \} = 1.$$

Remarquons que cette définition est réminiscente de la transformation de Fourier. Cette analogie sera plus claire dans le cas des variables aléatoires à densité.

La fonction caractéristique permet de calculer facilement les moments de X (lorsqu'ils existent) :

PROPOSITION 10.6 Soit G_X la fonction caractéristique de la variable aléatoire X . Alors

$$G_X(0) = 1, \quad (10.40)$$

et plus généralement, pour tout $n \in \mathbb{Z}^+$ tel que $\mathbb{E} \{ X^n \}$ existe

$$\left[\left(i \frac{d}{du} \right)^n G_X \right] (0) = \mathbb{E} \{ X^n \}. \quad (10.41)$$

Preuve : Partant de la définition de G_X , on a

$$\sum_k \left(i \frac{d}{du} \right)^n e^{-iux_k} \mathbb{P}_X \{ x_k \} = \sum_k x_k^n e^{-iux_k} \mathbb{P}_X \{ x_k \}.$$

Si cette série est absolument convergente, elle vaut $\mathbb{E} \{ X^n \}$ pour $u = 0$, et on peut aussi intervertir la sommation et la dérivation, ce qui prouve le résultat. ♠

EXEMPLE 10.15 (suite). Reprenons l'exemple de la variable aléatoire Poissonnienne. La fonction caractéristique vaut

$$G_X(u) = \sum_0^{\infty} e^{-iuk} \frac{\lambda^k}{k!} e^{-\lambda} = \sum_0^{\infty} \frac{(\lambda e^{-iu})^k}{k!} e^{-\lambda} = e^{-\lambda} e^{\lambda e^{-iu}}.$$

Comme attendu, $G_X(0) = 1$. En dérivant une première fois,

$$iG_X'(u) = ie^{-\lambda} (-i)\lambda e^{-iu} e^{\lambda e^{-iu}} = \lambda e^{-\lambda} e^{-iu} e^{\lambda e^{-iu}},$$

de sorte que $iG_X'(0) = \lambda$. En dérivant une seconde fois,

$$-G_X''(u) = i\lambda e^{-\lambda} \left(-ie^{-iu} e^{\lambda e^{-iu}} - i\lambda e^{-2iu} e^{\lambda e^{-iu}} \right) = \left(\lambda e^{-iu} + \lambda^2 e^{-2iu} \right) e^{-\lambda} e^{\lambda e^{-iu}},$$

ce qui redonne bien $-G_X''(0) = \lambda + \lambda^2$, et donc $\sigma_X = \lambda$.

Paaires de variables aléatoires

Généralités On s'intéresse maintenant à des familles de variables aléatoires. Considérons pour commencer une paire de variables aléatoires X, Y sur un même modèle probabiliste $(\Omega, \mathcal{F}, \mathbb{P})$. Les distributions de probabilités \mathbb{P}_X et \mathbb{P}_Y donnent des indications sur le comportement de X et Y , mais pas sur les relations existant entre X et Y . C'est pourquoi il est important de s'intéresser au comportement de la paire (X, Y) .

DÉFINITION 10.12 Soient X, Y deux variables aléatoires sur un même modèle probabiliste $(\Omega, \mathcal{F}, \mathbb{P})$. La distribution de probabilités jointe de X, Y est l'ensemble des nombres

$$\mathbb{P}_{X,Y}\{U\} = \mathbb{P}\{\omega \in \Omega : (X(\omega), Y(\omega)) \in U\}, \quad (10.42)$$

où $U \subset \mathbb{R}^2$ est un ensemble mesurable quelconque. Les distributions de probabilités \mathbb{P}_X et \mathbb{P}_Y sont appelées distributions marginales de $\mathbb{P}_{X,Y}$.

Il est facile d'obtenir les distributions marginales à partir de la distribution jointe. Supposons que X et Y soient des variables aléatoires discrètes, à valeurs dans $\{x_1, x_2, \dots\}$ et $\{y_1, y_2, \dots\}$ respectivement. Soit V un sous-ensemble mesurable de \mathbb{R} . On peut écrire

$$\begin{aligned} \mathbb{P}_X\{V\} &= \sum_{x_i \in V} \mathbb{P}_X\{x_i\} \\ &= \sum_{x_i \in V} \mathbb{P}\{X(\omega) = x_i, Y(\omega) \in \{y_1, y_2, \dots\}\} \\ &= \sum_{x_i \in V} \sum_j \mathbb{P}\{X(\omega) = x_i, Y(\omega) = y_j\} \\ &= \sum_j \mathbb{P}_{X,Y}\{V \times \{y_j\}\} \end{aligned}$$

Notons que la réciproque est fautive : la connaissance des distributions marginales ne permet pas de retrouver la distribution jointe en général. Il suffit pour s'en convaincre de considérer le cas extrême $Y = X$. La connaissance des distributions marginales \mathbb{P}_X et \mathbb{P}_Y (au demeurant identiques) ne permet pas de savoir que $Y = X$.

EXEMPLE 10.16 Soient deux variables aléatoires X, Y , prenant des valeurs entières $0, 1, \dots, N$, de distribution jointe multinômiale

$$\mathbb{P}_{X,Y}\{n, m\} = \frac{N!}{n! m! (N - n - m)!} p_1^n p_2^m (1 - p_1 - p_2)^{N - n - m},$$

où $p_1, p_2 \in [0, 1]$ sont tels que $0 \leq p_1 + p_2 \leq 1$.

On obtient de là la distribution de X :

$$\begin{aligned} \mathbb{P}_X\{n\} &= \sum_{m=0}^{N-n} \frac{N!}{n! m! (N - n - m)!} p_1^n p_2^m (1 - p_1 - p_2)^{N - n - m}, \\ &= \frac{N!}{n!} \frac{1}{(N - n)!} p_1^n \sum_{m=0}^{N-n} \frac{(N - n)!}{m! (N - n - m)!} p_2^m (1 - p_1 - p_2)^{(N - n) - m}, \\ &= \frac{N!}{n! (N - n)!} p_1^n (1 - p_1)^{N - n}. \end{aligned}$$

Il s'agit là encore d'une distribution binômiale.

Espérance de fonction de deux variables aléatoires Etant données deux variables aléatoires X, Y , et une fonction mesurable $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$, on considère la variable aléatoire $Z = \varphi(X, Y)$, définie par : pour tout $\omega \in \Omega$,

$$Z(\omega) = \varphi(X(\omega), Y(\omega)).$$

Pour calculer l'espérance de Z , une première démarche consiste à calculer tout d'abord la distribution de probabilités de Z , puis de calculer l'espérance grâce à l'expression classique : dans le cas discret :

$$\mathbb{E}\{Z\} = \sum_i z_i \mathbb{P}_Z\{z_i\},$$

où les z_i sont les valeurs possibles de Z . On peut en fait montrer que l'espérance de Z peut également être obtenue de la façon suivante :

DÉFINITION 10.13 Soient $X, Y, Z = \varphi(X, Y)$ trois variables aléatoires discrètes sur $(\Omega, \mathcal{F}, \mathbb{P})$. On note $\{x_1, x_2, \dots\}$ et $\{y_1, y_2, \dots\}$ les valeurs possibles de X et Y respectivement. L'espérance mathématique de Z est donnée par

$$\mathbb{E}\{Z\} = \sum_{i,j} \varphi(x_i, y_j) \mathbb{P}_{X,Y}\{x_i, y_j\}. \quad (10.43)$$

Ces notions nous sont utiles pour introduire la covariance et la corrélation :

DÉFINITION 10.14 Soient X, Y deux variables aléatoires sur $(\Omega, \mathcal{F}, \mathbb{P})$.

1. La covariance de X et Y est le nombre

$$\Gamma_{X,Y} = \mathbb{E}\{(X - \mathbb{E}\{X\})(Y - \mathbb{E}\{Y\})\}. \quad (10.44)$$

2. X et Y sont décorrélées si $\Gamma_{X,Y} = 0$.

3. Le coefficient de corrélation de X et Y est donné par

$$r(X, Y) = \frac{\Gamma_{X,Y}}{\sigma_X \sigma_Y}. \quad (10.45)$$

Remarquons que

$$\Gamma_{X,X} = \text{Var}\{X\},$$

et que

$$\text{Var}\{X + Y\} = \text{Var}\{X\} + \text{Var}\{Y\} + 2\Gamma_{X,Y}.$$

La covariance donne des indications sur une possible "association" systématique des variables aléatoires X et Y . Le coefficient de corrélation (qui est toujours un nombre sans dimension) fournit une version "normalisée" de la covariance. Il a les propriétés suivantes.

PROPOSITION 10.7 Soient X et Y deux variables aléatoires sur $(\Omega, \mathcal{F}, \mathbb{P})$.

1. $-1 \leq r(X, Y) \leq 1$.

2. Si $r(X, Y) = 1$, alors il existe une relation du type $Y = aX + b$, où $a, b \in \mathbb{R}$ et $a > 0$.

3. Si $r(X, Y) = -1$, alors il existe une relation du type $Y = aX + b$, où $a, b \in \mathbb{R}$ et $a < 0$.

Preuve : 1. Si on note $\tilde{X} = X - \mathbb{E}\{X\}$ et $\tilde{Y} = Y - \mathbb{E}\{Y\}$, on a clairement $\Gamma_{\tilde{X}, \tilde{Y}} = \Gamma_{X, Y}$, et $r(\tilde{X}, \tilde{Y}) = r(X, Y)$. Calculons

$$\mathbb{E} \left\{ \left(\frac{\tilde{X}}{\sigma_X} - \frac{\tilde{Y}}{\sigma_Y} \right)^2 \right\} = \frac{\mathbb{E}\{\tilde{X}^2\}}{\sigma_X^2} + \frac{\mathbb{E}\{\tilde{Y}^2\}}{\sigma_Y^2} - 2 \frac{\mathbb{E}\{\tilde{X}\tilde{Y}\}}{\sigma_X \sigma_Y} = 2 - 2 \frac{\mathbb{E}\{\tilde{X}\tilde{Y}\}}{\sigma_X \sigma_Y} \geq 0.$$

Donc, $\mathbb{E}\{\tilde{X}\tilde{Y}\} \leq \sigma_X \sigma_Y$, ce qui montre que $r(X, Y) \leq 1$. De même, l'inégalité

$$\mathbb{E} \left\{ \left(\frac{\tilde{X}}{\sigma_X} + \frac{\tilde{Y}}{\sigma_Y} \right)^2 \right\} \geq 0$$

permet de montrer que $r(X, Y) \geq -1$.

2. Supposons $r(X, Y) = 1$. Ceci implique $\Gamma_{X, Y} = \sigma_X \sigma_Y$. Par conséquent

$$\mathbb{E} \left\{ \left(\frac{\tilde{X}}{\sigma_X} - \frac{\tilde{Y}}{\sigma_Y} \right)^2 \right\} = 0,$$

ce qui implique, d'après (10.35) que

$$\tilde{Y} = \frac{\sigma_Y}{\sigma_X} \tilde{X},$$

et donc le résultat cherché, avec $a = \sigma_Y / \sigma_X > 0$, et $b = \sigma_Y \mathbb{E}\{X\} / \sigma_X - \mathbb{E}\{Y\}$.

3. La démonstration est identique, et conduit à $a = -\sigma_Y / \sigma_X < 0$. ♠

Vecteurs aléatoires

On considère maintenant des familles de plusieurs variables aléatoires, que l'on nomme vecteurs aléatoires.

DÉFINITION 10.15 Soient X_1, X_2, \dots, X_n n variables aléatoires sur un même modèle probabiliste $(\Omega, \mathcal{F}, \mathbb{P})$. La distribution de probabilités jointe de X_1, \dots, X_n est l'ensemble des nombres

$$\mathbb{P}_{X_1, \dots, X_n}\{U\} = \mathbb{P}\{(X_1(\omega), \dots, X_n(\omega)) \in U\}, \quad (10.46)$$

où $U \subset \mathbb{R}^n$ est un ensemble mesurable quelconque. Les distributions de probabilités \mathbb{P}_{X_i} sont appelées distributions marginales.

Comme dans le cas des paires de variables aléatoires, il est facile d'obtenir les distributions marginales à partir de la distribution jointe. La réciproque est toujours fautive : la connaissance des distributions marginales ne permet pas de retrouver la distribution jointe en général.

La distribution jointe permet également de calculer l'espérance de fonctions du vecteur aléatoire considéré. Si $Z = \varphi(X_1, \dots, X_n)$ est fonction du vecteur aléatoire (X_1, \dots, X_n) , on a alors

$$\mathbb{E}\{Z\} = \sum_{x_1, \dots, x_n} \varphi(x_1, \dots, x_n) \mathbb{P}_{X_1, \dots, X_n}\{x_1, \dots, x_n\}. \quad (10.47)$$

10.2.3 Variables aléatoires à densité

On s'intéresse maintenant aux variables aléatoires prenant leurs valeurs dans un ensemble infini continu, et on s'intéresse aux variables aléatoires dites "à densité".

DÉFINITION 10.16 Une variable aléatoire X sur $(\Omega, \mathcal{F}, \mathbb{P})$ est dite à densité si il existe une fonction $\rho_X \in L^1(\mathbb{R})$ telle que pour tout ensemble mesurable $U \subset \mathbb{R}$,

$$\mathbb{P}_X\{U\} = \int_U \rho_X(x) dx . \quad (10.48)$$

Propriétés de la densité

Une densité de probabilités ρ_X est à valeurs réelles. Cependant, $\rho_X(x) \geq 0$ presque partout (par contre, rien ne contraint la densité de probabilités à prendre des valeurs inférieures à 1). De plus, on a

$$\int_{-\infty}^{\infty} \rho_X(x) dx = \mathbb{P}_X\{\mathbb{R}\} = 1 . \quad (10.49)$$

Fonction de répartition La fonction de répartition d'une variable aléatoire à densité est définie comme dans le cas discret (voir DÉFINITION 10.8), et les propriétés énoncées dans la section PROPRIÉTÉS 10.1 restent valables. La nouveauté est que la fonction de répartition est maintenant donnée par

$$F_X(x) = \mathbb{P}_X\{]-\infty, x]\} = \int_{-\infty}^x \rho_X(x) dx ,$$

et est continue. De plus, on a

$$\rho_X(x) = \frac{d}{dx} F_X(x) . \quad (10.50)$$

EXEMPLE 10.17 Soit X une variable aléatoire exponentielle, de paramètre α : pour tout $x > 0$,

$$\mathbb{P}\{X(\omega) \geq x\} = e^{-\alpha x} .$$

On a donc la fonction de répartition

$$F_X(x) = (1 - e^{-\alpha x})\Theta(x) ,$$

et la densité

$$\rho_X(x) = F_X'(x) = \alpha e^{-\alpha x} \Theta(x) .$$

La fonction de répartition (définie sur \mathbb{R}^+) est montrée en FIGURE 10.2.

Densité d'une fonction de variable aléatoire Etant données deux variables aléatoires à densité X et $Y = \varphi(X)$ sur $(\Omega, \mathcal{F}, \mathbb{P})$, on cherche à relier la densité de Y à la densité ρ_X de X . Dans le cas général, il faut procéder en deux étapes : dans un premier temps, calculer la fonction de répartition de Y en fonction de la densité ρ_X de X ; puis en déduire la densité par dérivation.

Prenons l'exemple d'une variable aléatoire à valeurs réelles (positives ou négatives) X , de densité ρ_X , et de $Y = X^2$. La fonction de répartition de Y est donnée par

$$F_Y(y) = \mathbb{P}\{X(\omega)^2 < y\} = \mathbb{P}\{X(\omega) < \sqrt{y}\} \Theta(y) = \Theta(y) \int_{-\sqrt{y}}^{\sqrt{y}} \rho_X(x) dx .$$

On en déduit la densité de Y par dérivation :

$$\rho_Y(y) = \Theta(y) \left(\frac{1}{2\sqrt{y}} \rho_X(\sqrt{y}) - \frac{-1}{2\sqrt{y}} \rho_X(-\sqrt{y}) \right) = \Theta(y) \frac{1}{2\sqrt{y}} (\rho_X(\sqrt{y}) + \rho_X(-\sqrt{y})) .$$

On voit bien sur l'exemple ci-dessus la difficulté introduite par le fait que le φ choisi ne soit pas bijectif. Lorsque φ est bijectif, les choses sont un peu plus simples, et on peut énoncer le résultat suivant :

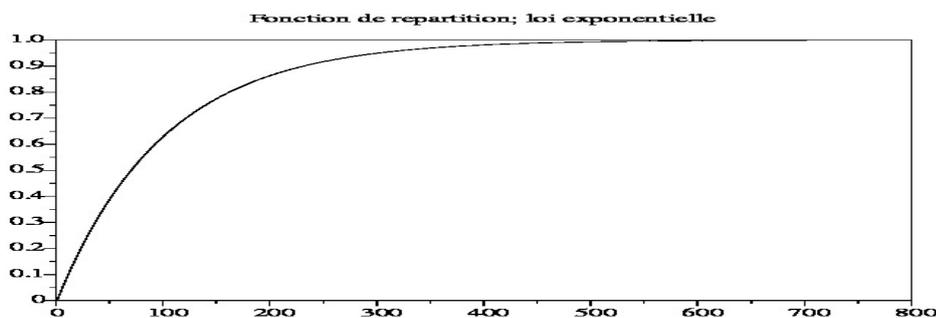


FIG. 10.2: Fonction de répartition d’une variable aléatoire de distribution exponentielle

THÉORÈME 10.1 Si φ est une bijection différentiable, à inverse différentiable, la densité ρ_Y de $Y = \varphi(X)$ est donnée par

$$\rho_Y(y) = (\varphi^{-1})'(y)\rho_X(\varphi^{-1}(y)) . \tag{10.51}$$

Preuve : Les hypothèses faites sont précisément celles qui nous permettent de faire un changement de variables :

$$\rho_Y(y) = \frac{d}{dy} \int_{\varphi^{-1}(-\infty, y]} \rho_X(x) dx = \frac{d}{dy} \int_{-\infty}^y \rho_X(\varphi^{-1}(u)) (\varphi^{-1})'(u) du ,$$

ce qui prouve le résultat. ♠

Une convention de notation

Il est parfois utile d’introduire une notation permettant de traiter simultanément les cas discrets et continus. On introduit pour cela le symbole $d\mathbb{P}_X(x)$ (que l’on note parfois $\mathbb{P}_X(dx)$), auquel on donne la signification suivante :

$$\mathbb{P}_X\{U\} = \int_U d\mathbb{P}_X(x) = \begin{cases} \int_U \rho_X(x) dx & \text{si } X \text{ est à densité } \rho_X \\ \sum_{x_i \in U} \mathbb{P}_X\{x_i\} & \text{si } X \text{ prend les valeurs } x_1, x_2, \dots \end{cases} \tag{10.52}$$

De même, étant donnée une fonction $f : \mathbb{R} \rightarrow \mathbb{C}$, on notera

$$\int_U f(x) d\mathbb{P}_X(x) = \begin{cases} \int_U f(x)\rho_X(x) dx & \text{si } X \text{ est à densité } \rho_X \\ \sum_{x_i \in U} f(x_i)\mathbb{P}_X\{x_i\} & \text{si } X \text{ prend les valeurs } x_1, x_2, \dots \end{cases} \tag{10.53}$$

Bien entendu, des hypothèses sont nécessaires dans les deux cas pour assurer l’existence du membre de gauche. Nous utiliserons cette notation dans la suite, en particulier en manipulant l’espérance mathématique.

Espérance, moments,...

Les définitions des espérances mathématiques des variables aléatoires à densité sont similaires à celles vues dans le cas discret.

DÉFINITION 10.17 1. L'espérance mathématique d'une variable aléatoire continue X de densité ρ_X est définie par

$$\mathbb{E}\{X\} = \int_{-\infty}^{\infty} x \rho_X(x) dx . \quad (10.54)$$

2. Etant donnés deux variables aléatoires continues X et $Y = \varphi(X)$, l'espérance mathématique de Y est définie par

$$\mathbb{E}\{\varphi(X)\} = \int_{-\infty}^{\infty} \varphi(x) \rho_X(x) dx . \quad (10.55)$$

En utilisant la notation introduite plus haut, on peut noter l'espérance d'une fonction de variable aléatoire quelconque (continue ou discrète)

$$\mathbb{E}\{\varphi(X)\} = \int_{-\infty}^{\infty} \varphi(x) d\mathbb{P}_X(x) . \quad (10.56)$$

On introduit par exemple les moments

$$\mathbb{E}\{X^n\} = \int_{-\infty}^{\infty} x^n \rho_X(x) dx ,$$

et la variance

$$\text{Var}\{X\} = \int_{-\infty}^{\infty} (x - \mathbb{E}\{X\})^2 \rho_X(x) dx .$$

La fonction caractéristique est introduite de façon similaire :

DÉFINITION 10.18 Soit X une variable aléatoire continue de densité ρ_X . La fonction caractéristique de X est la fonction $G_X : \mathbb{R} \rightarrow \mathbb{C}$ définie par

$$G_X(u) = \mathbb{E}\{e^{-iuX}\} = \int_{-\infty}^{\infty} e^{-iux} \rho_X(x) dx . \quad (10.57)$$

En utilisant le symbole $d\mathbb{P}_X$, on notera dans le cas général

$$G_X(u) = \int_{-\infty}^{\infty} e^{-iux} d\mathbb{P}_X(x) .$$

Dans le cas d'une variable aléatoire à densité, la fonction caractéristique est proportionnelle à la transformée de Fourier de la densité :

$$G_X(u) = \sqrt{2\pi} \hat{\rho}_X(u) .$$

On peut également en déduire (si la formule d'inversion de Fourier a un sens dans ce contexte) une expression de la densité à partir de la fonction caractéristique :

$$\rho_X(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} G_X(u) e^{iux} du .$$

Il est facile de montrer que la PROPOSITION 10.6 reste valide dans le cas des variables aléatoires à densité (la démonstration est similaire à celle donnant la dérivée d'une transformée de Fourier).

EXEMPLE 10.18 Considérons une variable aléatoire exponentielle X , de paramètre α . La densité est donc

$$\rho_X(x) = \alpha e^{-\alpha x} \Theta(x),$$

et la fonction caractéristique est donnée par

$$G_X(u) = \alpha \int_0^{\infty} e^{-\alpha x} e^{-iux} dx = \frac{\alpha}{\alpha + iu}.$$

On vérifie facilement que $G_X(0) = 1$. En dérivant une première fois, on a

$$iG'_X(u) = \frac{\alpha}{(\alpha + iu)^2},$$

d'où

$$\mathbb{E}\{X\} = \frac{1}{\alpha}.$$

De même,

$$-G''_X(u) = \frac{2\alpha}{(\alpha + iu)^3}$$

donne

$$\mathbb{E}\{X^2\} = \frac{2}{\alpha^2},$$

et donc $\text{Var}\{X\} = \alpha^{-2}$.

Vecteurs aléatoires

Ici encore, la discussion faite dans le cas discret se transpose sans grande difficulté au cas continu.

Densité jointe

DÉFINITION 10.19 Un vecteur aléatoire (X_1, \dots, X_n) sur $(\Omega, \mathcal{F}, \mathbb{P})$ est dit vecteur aléatoire à densité si il existe une fonction de n variables ρ_{X_1, \dots, X_n} telle que pour toute partie mesurable $V \subset \mathbb{R}^n$, on puisse écrire

$$\mathbb{P}\{(X_1(\omega), \dots, X_n(\omega)) \in V\} = \int_V \rho_{X_1, \dots, X_n}(x_1, \dots, x_n) dx_1 \dots dx_n. \quad (10.58)$$

La fonction ρ_{X_1, \dots, X_n} est appelée densité de probabilités jointe de X_1, \dots, X_n .

Au même titre que la densité de probabilités d'une variable aléatoire, la densité jointe d'un vecteur aléatoire est positive presque partout. Elle a également un comportement remarquable vis à vis des changements de variables, comme le montre le résultat suivant, qui généralise le THÉORÈME 10.1

THÉORÈME 10.2 Soit (X_1, \dots, X_n) un vecteur aléatoire sur $(\Omega, \mathcal{F}, \mathbb{P})$, à valeurs dans $U \subset \mathbb{R}^n$, et soit $\Phi : U \rightarrow V = \Phi(U) \subset \mathbb{R}^n$ une application bijective différentiable, à inverse différentiable. On note $(Y_1, \dots, Y_n) = \Phi(X_1, \dots, X_n)$, et $J_{\Phi^{-1}}$ le Jacobien de Φ^{-1} . Alors la densité jointe de (Y_1, \dots, Y_n) est donnée par

$$\rho_{Y_1, \dots, Y_n}(y_1, \dots, y_n) = |J_{\Phi^{-1}}(y_1, \dots, y_n)| \rho_{X_1, \dots, X_n}(\Phi^{-1}(y_1, \dots, y_n)). \quad (10.59)$$

Preuve : La démonstration de ce résultat est relativement simple. Soit f une fonction bornée à support borné définie sur U . Calculons

$$\begin{aligned}\mathbb{E} \{f(Y_1, \dots, Y_n)\} &= \int_V f(y_1, \dots, y_n) \rho_{Y_1, \dots, Y_n}(y_1, \dots, y_n) dy_1 \dots dy_n \\ &= \int_U f(\Phi(x_1, \dots, x_n)) \rho_{X_1, \dots, X_n}(x_1, \dots, x_n) dx_1 \dots dx_n \\ &= \int_V f(y_1, \dots, y_n) \rho_{X_1, \dots, X_n}(\Phi^{-1}(y_1, \dots, y_n)) \\ &\quad \times |J_{\Phi^{-1}}(y_1, \dots, y_n)| dy_1 \dots dy_n ,\end{aligned}$$

par changement de variables. Ceci étant vrai pour toute fonction f continue à support borné, on en déduit le résultat. ♠

Il est utile de considérer quelques exemples pour illustrer ce résultat. Les exemples les plus simples sont fournis par les changements de systèmes de coordonnées (du système Cartésien au système angulaire en particulier).

EXEMPLE 10.19 Supposons que des fléchettes soient tirées au hasard sur une cible ronde de rayon r_0 , avec une distribution uniforme sur le disque. On s'intéresse à la répartition des impacts en fonction de la distance au centre de la cible. La modélisation la plus adaptée consiste à introduire un espace Ω égal au disque de rayon r_0 . La distribution étant uniforme sur Ω , elle est associée à une densité de probabilités ρ , de sorte que pour tout $A \subset \Omega$, on ait

$$\mathbb{P}\{A\} = \frac{\text{Vol}(A)}{\text{Vol}(\Omega)} = \frac{1}{\pi r_0^2} \int_A dx dy = \int_A \rho(x, y) dx dy .$$

Ainsi, la densité ρ est donnée de la façon suivante : si $\omega = (x, y) \in \Omega$,

$$\rho(\omega) = \begin{cases} \frac{1}{\pi r_0^2} & \text{si } \omega \in \Omega \\ 0 & \text{sinon .} \end{cases}$$

Introduisons les variables aléatoires X et Y correspondant aux coordonnées Cartésiennes de ω : $X(\omega) = x$ et $Y(\omega) = y$. La densité jointe de X et Y est évidemment

$$\rho_{X,Y}(x, y) = \begin{cases} \frac{1}{\pi r_0^2} & \text{si le point de coordonnées } (x, y) \text{ appartient à } \Omega \\ 0 & \text{sinon .} \end{cases}$$

Si nous introduisons maintenant les variables aléatoires R et Θ représentant les coordonnées polaires de ω , nous avons d'après le théorème précédent

$$\rho_{R,\Theta}(r, \theta) = \begin{cases} \frac{r}{\pi r_0^2} & \text{si le point de coordonnées polaires } (r, \theta) \text{ appartient à } \Omega \\ 0 & \text{sinon .} \end{cases}$$

Il est important de remarquer que la distribution jointe de X, Y est une distribution uniforme, alors que la distribution jointe de R, Θ n'est absolument pas uniforme.

Ceci nous permet entre autres de calculer $\mathbb{E} \{R\} = \mathbb{E} \left\{ \sqrt{X^2 + Y^2} \right\}$ de deux façons différentes :

$$\mathbb{E} \{R\} = \int_{\Omega} \sqrt{x^2 + y^2} \rho_{X,Y}(x, y) dx dy = \frac{1}{\pi r_0^2} \int_0^{r_0} \int_0^{2\pi} r^2 dt d\theta = \frac{2}{3} r_0 ;$$

ou encore

$$\mathbb{E} \{R\} = \int_0^{r_0} \int_0^{2\pi} r \rho_{R,\Theta}(r, \theta) r dr d\theta = \frac{2}{3} r_0 .$$

Ces deux calculs sont en fait tout à fait similaires.

EXEMPLE 10.20 Prenons un autre exemple, issu de la physique. On suppose que l'état d'une particule est décrit par une fonction d'onde $(x, y, z) \in \mathbb{R}^3 \rightarrow \psi(x, y, z) \in \mathbb{C}$. Conformément à l'interprétation donnée par Sommerfeld, on donne à $\rho = |\psi(x, y, z)|^2$ le sens d'une densité de probabilité de présence au point $\omega = (x, y, z) \in \mathbb{R}^3$. De nouveau, on peut interpréter ρ comme la densité jointe des variables aléatoires X, Y, Z représentant les coordonnées Cartésiennes du point considéré. On peut également introduire les variables aléatoires R, Θ et ϕ représentant les coordonnées sphériques. Il est alors facile de voir que

$$\rho_{R, \Theta, \phi}(r, \theta, \varphi) = r^2 \sin \theta \rho(r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta).$$

Fonction de plusieurs variables aléatoires Etant données n variables aléatoires (X_1, \dots, X_n) , et une fonction $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$, on peut de nouveau introduire une variable aléatoire $Z = \varphi(X_1, \dots, X_n)$. Sa distribution peut dans certains cas être déduite simplement de la distribution jointe de (X_1, \dots, X_n) , mais ce cas se présente rarement. Pour calculer l'espérance de Z , il est généralement préférable d'utiliser la relation

$$\mathbb{E} \{ \varphi(X_1, \dots, X_n) \} = \int_{\mathbb{R}^n} \varphi(x_1, \dots, x_n) \rho_{X_1, \dots, X_n}(x_1, \dots, x_n) dx_1 \dots dx_n. \quad (10.60)$$

Si nous reprenons l'exemple inspiré de la mécanique quantique donné plus haut, nous en déduisons par exemple une expression pour l'espérance de la variable aléatoire R :

$$\mathbb{E} \{ R \} = \int_{\mathbb{R}^3} r^3 \sin \theta \rho(r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta) dr d\theta d\varphi.$$

10.3 Variables aléatoires indépendantes, distribution conditionnelle

Nous avons vu au chapitre précédent que lorsque l'on a à étudier plusieurs variables aléatoires définies sur un même modèle probabiliste, il ne suffit pas de connaître leurs distributions individuelles pour les caractériser. Il faut surtout connaître leur distribution jointe (qui permet également de connaître les distributions individuelles, ou marginales). Cependant, il existe un cas où la connaissance des distributions marginales est suffisante : on dit alors que les variables aléatoires sont indépendantes.

10.3.1 Variables aléatoires indépendantes

Nous avons déjà rencontré la notion d'indépendance dans le premier chapitre : deux événements $A, B \in \mathcal{F}$ sont indépendants si $\mathbb{P}\{A \cap B\} = \mathbb{P}\{A\}\mathbb{P}\{B\}$, et on donne une définition similaire pour des familles finies d'événements indépendants. Nous nous intéressons maintenant à l'utilisation de cette notion dans le cas de variables aléatoires.

Définitions

Commençons par donner la définition d'une famille de variables aléatoires indépendantes.

DÉFINITION 10.20 Etant données N variables aléatoires X_1, \dots, X_N définies sur un même modèle probabiliste $(\Omega, \mathcal{F}, \mathbb{P})$, on dit qu'elles sont indépendantes si pour toute famille (U_1, \dots, U_N) de sous-ensembles de \mathbb{R} , on a

$$\mathbb{P}_{X_1, X_2, \dots, X_N} \{ U_1 \times U_2 \times \dots \times U_N \} = \mathbb{P}_{X_1} \{ U_1 \} \mathbb{P}_{X_2} \{ U_2 \} \dots \mathbb{P}_{X_N} \{ U_N \} \quad (10.61)$$

Des variables aléatoires indépendantes possèdent des propriétés simples, que nous décrivons ci-dessous.

1. Si les variables aléatoires X_1, \dots, X_N sont des variables aléatoires discrètes, alors leur indépendance est équivalente à la propriété :

$$\mathbb{P}\{X_1 = x_1, X_2 = x_2, X_N = x_N\} = \mathbb{P}_{X_1}\{x_1\}\mathbb{P}_{X_2}\{x_2\} \dots \mathbb{P}_{X_N}\{x_N\} \quad (10.62)$$

pour tous x_1, \dots, x_N appartenant aux ensembles de valeurs possibles respectifs de X_1, \dots, X_N .

2. Si les variables aléatoires sont toutes continues, alors l'indépendance se traduit en une propriété de leurs densités de probabilités : on a alors

$$\rho_{X_1, X_2, \dots, X_N}\{x_1, x_2, \dots, x_N\} = \rho_{X_1}\{x_1\}\rho_{X_2}\{x_2\} \dots \rho_{X_N}\{x_N\}. \quad (10.63)$$

Ces dernières propriétés sont faciles à démontrer. Prenons l'exemple de deux variables aléatoires continues X_1, X_2 indépendantes. On a alors, pour tous $U_1, U_2 \subset \mathbb{R}$,

$$\begin{aligned} \mathbb{P}_{X_1, X_2}\{U_1 \times U_2\} &= \int_{U_1 \times U_2} \rho_{X_1, X_2}(x_1, x_2) dx_1 dx_2 \\ &= \int_{U_1} \int_{U_2} \rho_{X_1, X_2}(x_1, x_2) dx_1 dx_2 \\ &= \int_{U_1} \int_{U_2} \rho_{X_1}(x_1)\rho_{X_2}(x_2) dx_1 dx_2. \end{aligned}$$

Ceci devant être vérifié pour tous $U_1, U_2 \subset \mathbb{R}$, on en déduit bien la propriété de factorisation $\rho_{X_1, X_2} = \rho_{X_1}\rho_{X_2}$ souhaitée. Le cas des variables aléatoires discrètes se traite de façon similaire.

Il est parfois utile d'étendre cette définition au cas d'une famille infinie de variables aléatoires.

DÉFINITION 10.21 Une famille infinie $\{X_n, n \in \mathbb{Z}\}$ de variables aléatoires définies sur un même modèle probabiliste $(\Omega, \mathcal{F}, \mathbb{P})$ est une famille indépendante si toute sous-famille finie est une famille indépendante.

EXEMPLE 10.21 Le nombre de particules N entrant dans un compteur obéit à une loi de Poisson de paramètre λ donné. Ceci produit une tension aléatoire

$$V = NU$$

où U est une tension aléatoire, indépendante du nombre de particules, et de distribution caractérisée par la densité de probabilités

$$\rho_U(u) = \frac{1}{(1+u)^2} \Theta(u).$$

L'hypothèse d'indépendance nous permet de calculer

$$\begin{aligned} \mathbb{P}\{V < 1\} &= \sum_{n=0}^{\infty} \mathbb{P}\{N = n, U < 1/n\} \\ &= \sum_{n=0}^{\infty} \mathbb{P}\{N = n\}\mathbb{P}\{U < 1/n\} \\ &= \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} e^{-\lambda} \int_0^{1/n} \frac{du}{(1+u)^2} \\ &= e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \left[\frac{-1}{1+u} \right]_0^{1/n} \\ &= e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} \frac{1}{n+1} = \frac{e^{-\lambda}}{\lambda} \sum_{n=0}^{\infty} \frac{\lambda^{n+1}}{(n+1)!} \\ &= \frac{e^{-\lambda}}{\lambda} (e^\lambda - 1) = \frac{1 - e^{-\lambda}}{\lambda} \end{aligned}$$

Conséquences de l'indépendance

1. L'indépendance est une propriété forte d'une famille de variables aléatoires. Il est possible de trouver des exemples de triplets de variables aléatoires (X_1, X_2, X_3) tels que toutes les sous-paires (X_i, X_j) soient des paires de v.a. indépendantes, mais que le triplet (X_1, X_2, X_3) soit quant à lui dépendant.
2. Par contre, étant donnée une famille de v.a. indépendantes, toute sous-famille est également indépendante. Si nous reprenons l'exemple des triplets de v.a. indépendantes, on voit facilement que

$$\begin{aligned} \mathbb{P}\{X_1 \in U_1, X_2 \in U_2\} &= \mathbb{P}\{X_1 \in U_1, X_2 \in U_2, X_3 \in \mathbb{R}\} \\ &= \mathbb{P}\{X_1 \in U_1\}\mathbb{P}\{X_2 \in U_2\}\mathbb{P}\{X_3 \in \mathbb{R}\} \\ &= \mathbb{P}\{X_1 \in U_1\}\mathbb{P}\{X_2 \in U_2\}. \end{aligned}$$

3. Si X_1, \dots, X_N forment une famille de variables aléatoires indépendantes. Alors étant données des fonctions $f_1, \dots, f_N : \mathbb{R} \rightarrow \mathbb{R}$, les variables aléatoires $Y_1 = f_1(X_1), \dots, Y_N = f_N(X_N)$ forment une famille de variables aléatoires indépendantes. Vérifions le sur l'exemple simple d'une paire de variables aléatoires. Calculons

$$\begin{aligned} \mathbb{P}\{f_1(X_1) \in U_1, f_2(X_2) \in U_2\} &= \mathbb{P}\{X_1 \in f_1^{-1}(U_1), X_2 \in f_2^{-1}(U_2)\} \\ &= \mathbb{P}\{X_1 \in f_1^{-1}(U_1)\}\mathbb{P}\{X_2 \in f_2^{-1}(U_2)\} \\ &= \mathbb{P}\{f_1(X_1) \in U_1\}\mathbb{P}\{f_2(X_2) \in U_2\}, \end{aligned}$$

ce qui prouve l'assertion.

4. Plus généralement, si X_1, \dots, X_N forment une famille de variables aléatoires indépendantes, alors $f(X_1, X_2), g(X_3), h(X_5, X_7), \dots$ sont des variables aléatoires indépendantes.

Nous pouvons résumer les résultats obtenus dans cette section dans la proposition importante suivante

PROPOSITION 10.8 Soit (X_1, \dots, X_N) une famille de variables aléatoires indépendantes. Alors toutes fonctions de groupes disjoints de ces variables aléatoires forment des familles de variables aléatoires indépendantes.

Indépendance, variance et espérance

Les espérances de fonctions de variables aléatoires indépendantes offrent des simplifications.

PROPOSITION 10.9 Soient X et Y deux variables aléatoires indépendantes. Alors, on a

$$\mathbb{E}\{XY\} = \mathbb{E}\{X\}\mathbb{E}\{Y\} \quad (10.64)$$

$$\Gamma_{XY} = 0. \quad (10.65)$$

Preuve : On se limitera aux cas de variables aléatoires discrètes ou continues (le cas mixte se traite de façon similaire). Supposons X et Y discrètes. On a alors

$$\begin{aligned} \mathbb{E}\{XY\} &= \sum_{x_i, y_j} x_i y_j \mathbb{P}\{X = x_i, Y = y_j\} \\ &= \sum_{x_i} \sum_{y_j} x_i y_j \mathbb{P}\{X = x_i\} \mathbb{P}\{Y = y_j\} \\ &= \mathbb{E}\{X\} \mathbb{E}\{Y\}. \end{aligned}$$

Dans le cas continu, on a

$$\begin{aligned} \mathbb{E}\{XY\} &= \int_{\mathbb{R}^2} xy \rho_{X,Y}(x, y) dx dy \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} xy \rho_X(x) \rho_Y(y) dx dy \\ &= \mathbb{E}\{X\} \mathbb{E}\{Y\}. \end{aligned}$$

On a immédiatement le corollaire suivant

COROLLAIRE 10.1 1. Si X et Y sont indépendantes, on a pour toutes fonctions $f, g : \mathbb{R} \rightarrow \mathbb{R}$

$$\mathbb{E}\{f(X)g(Y)\} = \mathbb{E}\{f(X)\} \mathbb{E}\{g(Y)\}. \quad (10.66)$$

2. Si X_1, \dots, X_N sont indépendantes, alors

$$\mathbb{E}\{X_1 X_2 \dots X_N\} = \mathbb{E}\{X_1\} \mathbb{E}\{X_2\} \dots \mathbb{E}\{X_N\} \quad (10.67)$$

Preuve : la première assertion est une conséquence directe du fait que X, Y étant indépendantes, $f(X)$ et $g(Y)$ le sont également. La seconde assertion se montre par récurrence : X_1 et $X_2 X_3 \dots X_N$ sont indépendantes, de même que X_2 et $X_3 X_4 \dots X_N$, et ainsi de suite. Donc

$$\mathbb{E}\{X_1 X_2 \dots X_N\} = \mathbb{E}\{X_1\} \mathbb{E}\{X_2 \dots X_N\} = \mathbb{E}\{X_1\} \mathbb{E}\{X_2\} \dots \mathbb{E}\{X_N\}.$$

Les moments d'ordre deux des variables aléatoires indépendantes possèdent aussi d'intéressantes propriétés.

PROPOSITION 10.10 Si X_1, \dots, X_N sont indépendantes, alors

$$\text{Var}\{X_1 + X_2 + \dots + X_N\} = \text{Var}\{X_1\} + \text{Var}\{X_2\} + \dots + \text{Var}\{X_N\}. \quad (10.68)$$

Preuve : Il suffit de calculer

$$\begin{aligned} \text{Var}\{X_1 + X_2 + \dots + X_N\} &= \mathbb{E}\left\{[(X_1 - \mathbb{E}\{X_1\}) + \dots + (X_N - \mathbb{E}\{X_N\})]^2\right\} \\ &= \text{Var}\{X_1\} + \dots + \text{Var}\{X_N\} + 2 \sum_{i \neq j} \Gamma_{X_i X_j} \end{aligned}$$

ce qui termine la preuve, car $\Gamma_{X_i X_j} = 0$ pour tous $i \neq j$.

10.3.2 Sommes de variables aléatoires indépendantes

Généralités

Considérons tout d'abord une paire de variables indépendantes discrètes X et Y , prenant des valeurs entières, et soit $Z = X + Y$. On a alors de façon évidente

$$\mathbb{P}\{Z = n\} = \sum_k \mathbb{P}\{X = k, Y = n - k\} = \sum_k \mathbb{P}_X\{k\} \mathbb{P}_Y\{n - k\}.$$

On aboutit donc à un produit de convolution discret. On a aussi, en utilisant les fonctions caractéristiques de X, Y, Z :

$$G_Z(u) = \mathbb{E} \left\{ e^{-iuZ} \right\} = \sum_n e^{-iun} \sum_k \mathbb{P}_X\{k\} \mathbb{P}_Y\{n - k\} = G_X(u) G_Y(u).$$

De même, si nous considérons deux variables aléatoires continues X, Y , indépendantes, de densités ρ_X et ρ_Y on a en posant de nouveau $Z = X + Y$

$$\begin{aligned} F_Z(z) = \mathbb{P}\{Z < z\} &= \int_{x+y < z} \rho_X(x) \rho_Y(y) dx dy \\ &= \int_{-\infty}^z \int \rho_X(x) \rho_Y(u - x) dx du, \end{aligned}$$

de sorte que la densité ρ_Z de Z peut s'écrire sous la forme d'un produit de convolution

$$\rho_Z(z) = \int \rho_X(x) \rho_Y(z - x) dx.$$

Au niveau des fonctions caractéristiques, on a aussi

$$G_Z(u) = \mathbb{E} \left\{ e^{-iuZ} \right\} = \int e^{-iuz} \rho_Z(z) dz = G_X(u) G_Y(u).$$

On peut montrer une propriété similaire dans le cas de variables aléatoires mixtes. En fait, on a dans le cas général

$$G_Z(u) = \mathbb{E} \left\{ e^{-iu(X+Y)} \right\} = \mathbb{E} \left\{ e^{-iuX} \right\} \mathbb{E} \left\{ e^{-iuY} \right\} = G_X(u) G_Y(u).$$

En étendant ce résultat à plusieurs variables aléatoires (ce qui se fait par récurrence), on peut donc montrer :

PROPOSITION 10.11 Si X_1, \dots, X_N sont indépendantes, et si on pose

$$Z = X_1 + X_2 + \dots + X_N$$

alors on a

$$G_Z(u) = G_{X_1}(u) G_{X_2}(u) \dots G_{X_N}(u). \quad (10.69)$$

Cette propriété nous sera utile par la suite.

La loi des grands nombres

C'est un résultat d'une grande importance historique et pratique. C'est le plus simple des théorèmes limites, et il est à la base de la théorie des statistiques. Il montre essentiellement qu'une moyenne de variables aléatoires indépendantes, identiques, tend vers leur espérance commune quand le nombre de variables aléatoires tend vers l'infini. Ce résultat est assez intuitif, et la difficulté essentielle réside dans le sens précis qu'on lui donne. En effet, la moyenne des variables aléatoires est toujours aléatoire, alors que l'espérance est quant à elle un nombre déterministe. Ecrire une égalité entre ces deux quantités demande donc un minimum de précautions.

EXEMPLE 10.22 Considérons une expérience de pile ou face, avec une pièce non pipée, effectuée un grand nombre N de fois. On note S_N le nombre de fois où la pièce retombe sur "pile". On s'attend intuitivement à ce que lorsque N tend vers l'infini, S_N/N soit de plus en plus proche de $1/2$, qui est la probabilité de "pile". La loi des grands nombres précise cette intuition, en montrant que pour tout nombre $\delta > 0$,

$$\mathbb{P}\{[S_N/N - 1/2] < \delta\} \rightarrow 1 \quad \text{quand } N \rightarrow \infty .$$

Pour donner un énoncé dans le cas général, on a besoin de la définition suivante :

DÉFINITION 10.22 1. Les variables aléatoires X_1, \dots, X_N sur un même modèle probabiliste $(\Omega, \mathcal{F}, \mathbb{P})$ sont dites *identiquement distribuées* si pour tout $U \subset \mathbb{R}$ on a

$$\mathbb{P}_{X_1}\{U\} = \mathbb{P}_{X_2}\{U\} = \dots = \mathbb{P}_{X_N}\{U\} \quad (10.70)$$

2. Si elles sont de plus indépendantes, on dit qu'elles sont *i.i.d.*

THÉORÈME 10.3 Soient X_1, \dots, X_N des variables aléatoires *i.i.d.* ; soit $\delta > 0$. Alors, en posant

$$S_N = \sum_{n=1}^N X_n \quad (10.71)$$

on a

$$\mathbb{P}\left\{\left(\frac{S_N}{N} - \mathbb{E}\{X_1\}\right) < \delta\right\} \rightarrow 1 \quad \text{quand } N \rightarrow \infty . \quad (10.72)$$

Preuve : La preuve est relativement simple : posons

$$Z = \frac{S_N}{N} - \mathbb{E}\{X_1\} ,$$

et soit

$$A = \{\omega \in \Omega, Z(\omega) \geq \delta\} .$$

On a alors $\mathbb{E}\{Z^2\} \geq \delta^2 \mathbb{P}\{A\}$, de sorte que

$$\mathbb{P}\{|Z| \geq \delta\} \leq \delta^{-2} \mathbb{E}\{Z^2\} .$$

Mais on a aussi, comme $\mathbb{E}\{S_N\} = N\mathbb{E}\{X_1\}$,

$$\mathbb{E}\{Z^2\} = \text{Var}\{Z\} = \text{Var}\left\{\frac{S_N}{N}\right\} = \frac{1}{N} \text{Var}\{X_1\} .$$

Comme cette dernière quantité tend vers 0 quand $N \rightarrow \infty$, le théorème est donc démontré.

Le théorème "Central Limit"

La loi des grands nombres montre que la moyenne d'une famille de variables aléatoires *i.i.d.* tend vers leur espérance commune quand la taille de la famille tend vers l'infini. Cependant, cette moyenne possède quand même un caractère aléatoire, et il est donc utile de s'intéresser au comportement asymptotique de sa distribution. Le théorème "Central Limit" montre que celui-ci obéit (généralement) à une "loi générique".

On ne donnera pas ici de preuve du théorème "Central Limit" en toute généralité. On s'appuiera plutôt sur l'exemple suivant.

EXEMPLE 10.23 Considérons N variables aléatoires *i.i.d.*, uniformes sur l'intervalle $[-1, 1]$, et soit

$$Z = \frac{1}{\sqrt{N}} (X_1 + X_2 + \cdots + X_N) .$$

Intéressons nous tout d'abord à la distribution de X/\sqrt{N} , où X est une v.a. continue donnée. On voit facilement que

$$F_{X/\sqrt{N}}(x) = \mathbb{P}\{X < x\sqrt{N}\} = \int_{-\infty}^{x\sqrt{N}} \rho_X(y) dy = \sqrt{N} \int_{-\infty}^x \rho_X(u\sqrt{N}) du ,$$

de sorte que

$$\rho_{X/\sqrt{N}}(x) = \sqrt{N} \rho_X(x\sqrt{N}) ,$$

et, au niveau de la fonction caractéristique

$$G_{X/\sqrt{N}}(\lambda) = \sqrt{N} \int \rho_X(x\sqrt{N}) e^{-i\lambda x} dx = G_X\left(\frac{\lambda}{\sqrt{N}}\right) .$$

Revenant à notre famille X_1, \dots, X_N :

$$G_{X_1}(\lambda) = G_{X_2}(\lambda) = \cdots = G_{X_N}(\lambda) = \frac{1}{2} \int_{-1}^1 e^{-i\lambda x} dx = \frac{\sin \lambda}{\lambda} .$$

On a donc

$$G_Z(\lambda) = G_{X_1}\left(\frac{\lambda}{\sqrt{N}}\right) G_{X_2}\left(\frac{\lambda}{\sqrt{N}}\right) \cdots G_{X_N}\left(\frac{\lambda}{\sqrt{N}}\right) = \left(\frac{\sin \lambda/\sqrt{N}}{\lambda/\sqrt{N}}\right)^N .$$

Calculons maintenant, dans la limite des grandes valeurs de N ,

$$\begin{aligned} \log(G_Z(\lambda)) &= N \log\left(\frac{\sin \lambda/\sqrt{N}}{\lambda/\sqrt{N}}\right) \\ &\approx N \log\left(1 - \frac{\lambda^2}{6N} + O\left(\frac{\lambda^4}{N^2}\right)\right) \\ &\approx -\frac{\lambda^2}{6} + O\left(\frac{\lambda^4}{N}\right) \end{aligned}$$

On a donc

$$\lim_{N \rightarrow \infty} G_Z(\lambda) = e^{-\lambda^2/6} ,$$

et donc, par transformation de Fourier inverse

$$\lim_{N \rightarrow \infty} \rho_Z(x) = \sqrt{\frac{3}{2\pi}} e^{-3x^2/2} .$$

On a donc montré qu'à la limite $N \rightarrow \infty$, la variable aléatoire Z a un comportement Gaussien.

Le théorème, dans sa version plus générale, est le suivant :

THÉORÈME 10.4 Soit (X_1, X_2, \dots, X_N) une famille de variables aléatoires i.i.d. sur un même modèle probabiliste $(\Omega, \mathcal{F}, \mathbb{P})$, de moyenne μ et d'écart type σ . Soit

$$Z_N = \frac{1}{\sigma\sqrt{N}} (X_1 + X_2 + \dots + X_N - N\mu) . \quad (10.73)$$

Alors, on a

$$\lim_{N \rightarrow \infty} \mathbb{P}\{Z_N < z\} = \frac{1}{2\pi} \int_{-\infty}^z e^{-x^2/2} dx . \quad (10.74)$$

En plus du comportement Gaussien de Z_N , ce théorème permet de préciser le comportement de l'écart type de la variable aléatoire

$$S_N = X_1 + X_2 + \dots + X_N .$$

On a

$$\sigma_{S_N} = \sigma\sqrt{N}$$

(on savait déjà d'après la loi des grands nombres que $\mu_{S_N} = N\mu$).

10.3.3 Variables aléatoires et probabilités conditionnelles

Nous avons déjà rencontré au chapitre 1 la notion de probabilité conditionnelle : soit $(\Omega, \mathcal{F}, \mathbb{P})$ un modèle probabiliste ; si A et B sont deux évènements quelconques, la probabilité de A sachant B (ou conditionnée par B) est le nombre

$$\mathbb{P}\{A|B\} = \frac{\mathbb{P}\{A \cap B\}}{\mathbb{P}\{B\}} \quad (10.75)$$

La notion de probabilité conditionnelle (et de distribution conditionnelle) s'étend aux variables aléatoires.

Distribution conditionnelle d'une variable aléatoire

De là, on déduit les probabilités conditionnelles des variables aléatoires. Par exemple, si X et Y sont des variables aléatoires discrètes :

$$\begin{aligned} \mathbb{P}\{X = x_i | Y = y_j\} &= \mathbb{P}\{\omega \in \Omega, X(\omega) = x_i \text{ et } Y(\omega) = y_j | \{\omega' \in \Omega, Y(\omega') = y_j\}\} \\ &= \frac{\mathbb{P}\{X = x_i, Y = y_j\}}{\mathbb{P}\{Y = y_j\}} \end{aligned}$$

Rappelons que dans ce cas,

$$\mathbb{P}\{Y = y_j\} = \sum_i \mathbb{P}\{X = x_i, Y = y_j\} .$$

Dans le cas général, on a

DÉFINITION 10.23 Etant données deux variables aléatoires X et Y définies sur le même modèle probabiliste, la distribution conditionnelle de X sachant que $Y \in V$ est la collection de nombres

$$\mathbb{P}\{X \in U | Y \in V\} = \frac{\mathbb{P}\{X \in U, Y \in V\}}{\mathbb{P}\{Y \in V\}} \quad (10.76)$$

pour tout ensemble mesurable $U \subset \mathbb{R}$.

On a donc, de façon évidente

PROPOSITION 10.12 Deux variables aléatoires X et Y sont indépendantes si et seulement si pour tous $U, V \subset \mathbb{R}$, mesurables

$$\mathbb{P}\{X \in U | Y \in V\} = \mathbb{P}\{X \in U\}.$$

Exemple : On tire à *pile ou face* avec une pièce tordue, telle que la probabilité d'obtenir "pile" à un tirage vaut p . Soit T le nombre d'essais nécessaires pour obtenir "pile", et soit S le nombre de "pile" en n tirages. Calculons $\mathbb{P}\{T = k | S = 1\}$: pour cela, il suffit de ne considérer que les k premiers tirages, et on calcule

$$\mathbb{P}\{T = k, S = 1\} = p(1-p)^{k-1},$$

et

$$\mathbb{P}\{S = 1\} = \frac{k!}{1!(k-1)!} p(1-p)^{k-1}$$

d'où on déduit

$$\mathbb{P}\{T = k | S = 1\} = \frac{1}{k}$$

La distribution de X sachant $Y = y$

On considère maintenant le cas particulier important où une variable aléatoire X est conditionnée par une valeur particulière y de la variable aléatoire Y . On s'intéresse donc aux nombres

$$\mathbb{P}\{X \in A | Y = y\} = \frac{\mathbb{P}\{X \in A, Y = y\}}{\mathbb{P}\{Y = y\}}.$$

La première question qui se pose est celle du sens que l'on donne à ces nombres. Si Y est une variable aléatoire discrète ou mixte, et si $\mathbb{P}\{Y = y\} \neq 0$, il n'y a pas de problème. Par contre, si $\mathbb{P}\{Y = y\} = 0$, il faut avoir recours à des arguments limite. On procède comme suit. Soit $\Delta y = [y - \delta y, y + \delta y[$ un intervalle de largeur δy centré sur y .

- Si $\mathbb{P}\{Y \in \Delta y\} = 0$, on pose $\mathbb{P}\{X \in A | Y = y\} = 0$.
- sinon, on pose $\mathbb{P}\{X \in A | Y = y\} = \lim_{\delta y \rightarrow 0} \mathbb{P}\{X \in A | Y \in \Delta y\}$.

Si X et Y possèdent une densité de probabilités jointe $\rho_{X,Y}(x, y)$, on sait alors que

$$\rho_Y(y) = \int \rho_{X,Y}(x, y) dx,$$

et on a alors

$$\begin{aligned} \mathbb{P}\{X \in A | Y \in \Delta y\} &= \frac{\int_{A \times \Delta y} \rho_{X,Y}(x, y) dx dy}{\int_{\Delta y} \rho_Y(y) dy} \\ &\rightarrow \frac{\int_A \rho_{X,Y}(x, y) dx \Delta y}{\rho_Y(y) \Delta y}. \end{aligned}$$

On obtient ainsi

PROPOSITION 10.13 Si les variables aléatoires X et Y possèdent une densité de probabilités jointe $\rho_{X,Y}(x, y)$, alors $\mathbb{P}\{X \in A|Y = y\}$ est aussi définie par une densité de probabilités

$$\mathbb{P}\{X \in A|Y = y\} = \int_A \rho(x|y) dx, \quad (10.77)$$

où

$$\rho(x|y) = \frac{\rho_{X,Y}(x, y)}{\rho_Y(y)}. \quad (10.78)$$

DÉFINITION 10.24 La fonction $\rho(x|y)$ est la densité conditionnelle de X sachant Y .

Exemple : On tire au hasard un point sur un disque de rayon 1, avec une densité de probabilités proportionnelle à la distance au centre du disque. Calculer la probabilité pour que la coordonnée x soit supérieure à un x_0 donné sachant que la coordonnée y vaut 0.

On note X, Y les variables aléatoires décrivant les coordonnées x et y du point tiré au hasard. Il faut tout d'abord expliciter la distribution jointe de X et Y : $\rho_{X,Y}(x, y) = K(x^2 + y^2)$. La condition de normalisation $\int \rho_{X,Y}(x, y) dx dy = 1$ donne $K = \pi/2$. La distribution de Y est de la forme

$$\rho_Y(y) = \frac{2}{\pi} \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} \sqrt{1-y^2}(x^2 + y^2) dx = \frac{1}{3\pi} \sqrt{1-y^2}(1 + 2y^2),$$

d'où on déduit

$$\rho(x|y) = \frac{3(x^2 + y^2)}{2(1 + 2y^2)\sqrt{1-y^2}}.$$

Finalement,

$$\mathbb{P}\{X \geq x_0|Y = 0\} = \frac{3}{2} \int_{x_0}^1 x^2 dx = \frac{1}{2}(1 - x_0^2).$$

Notation : Pour unifier les notations entre les cas discret, continu et mixte, on notera

$$\mathbb{P}\{X \in A|Y = y\} = \int_A \mathbb{P}\{dx|Y = y\}.$$

Probabilités conditionnelles d'une fonction de variables aléatoires

Etant donnée deux variables aléatoires X, Y , on se donne une nouvelle variable aléatoire Z , fonction de X, Y : $Z = \varphi(X, Y)$. On s'intéresse à la distribution de Z sachant que $Y = y$. On a alors le résultat suivant :

PROPOSITION 10.14 La distribution conditionnelle de $\varphi(X, Y)$ sachant que $Y = y$ est égale à la distribution conditionnelle de $\varphi(X, y)$ sachant que $Y = y$.

Preuve : si $A \in \mathcal{S}$, on considère

$$\mathbb{P}\{\varphi(X, Y) \in A|Y = y\} = \frac{\mathbb{P}\{\varphi(X, Y) \in A|Y = y\}}{\mathbb{P}\{Y = y\}}.$$

Les ensembles

$$\{\omega : \varphi(X(\omega), Y(\omega)) \in A, Y(\omega) = y\}$$

et

$$\{\omega : \varphi(X(\omega), y) \in A, Y(\omega) = y\}$$

étant identiques, ceci prouve la proposition.

La règle d'addition des probabilités conditionnelles

Exemple : Une urne contient 7 boules noires et 3 boules blanches. Quelle est la probabilité pour que la seconde boule tirée soit blanche ? Le raisonnement le plus simple distingue 2 cas : soit la première boule tirée est blanche (probabilité 3/10) et la seconde est blanche (probabilité 2/9) ; soit la première est noire (probabilité 7/10) et la seconde est blanche (probabilité 3/9). Donc, la probabilité vaut

$$\frac{3}{10} \frac{2}{9} + \frac{7}{10} \frac{3}{9} = \frac{3}{10}.$$

Plus généralement, étant données deux variables aléatoires discrètes X et Y , on a

$$\mathbb{P}\{X = x_i\} = \sum_j \mathbb{P}\{X = x_i, Y = y_j\} = \sum_j \mathbb{P}\{X = x_i | Y = y_j\} \sum_j \mathbb{P}\{Y = y_j\},$$

et dans le cas de variables aléatoires continues

$$\mathbb{P}\{X \in A\} = \int_A \int_{-\infty}^{\infty} \rho_{X,Y}(x, y) dx dy = \int_A \int_{-\infty}^{\infty} \rho(x|y) \rho_Y(y) dx dy,$$

de sorte que

$$\rho_X(x) = \int \rho(x|y) \rho_Y(y) dy.$$

On a donc montré

PROPOSITION 10.15 *Etant données deux variables aléatoires X et Y définies sur le même modèle probabiliste,*

$$\mathbb{P}_X\{A\} = \int \mathbb{P}\{X \in A | Y = y\} d\mathbb{P}_Y\{y\}. \quad (10.79)$$

10.3.4 Espérance conditionnelle et prédiction

Définition et propriétés

La notion d'espérance conditionnelle est celle qui permet de donner un sens à celle de "résultat attendu pour la variable aléatoire X connaissant Y ".

DÉFINITION 10.25 *Etant données deux variables aléatoires X et Y définies sur le même modèle probabiliste, on définit l'espérance de X sachant que $Y = y$ par*

$$\mathbb{E}\{X | Y = y\} = \int x \mathbb{P}\{dx | Y = y\}. \quad (10.80)$$

De même, étant donnée une fonction $\varphi(X)$ de la variable aléatoire X , l'espérance conditionnelle de $\varphi(X)$ sachant Y est définie par

$$\mathbb{E}\{\varphi(X) | Y = y\} = \int \varphi(x) \mathbb{P}\{dx | Y = y\}. \quad (10.81)$$

L'espérance conditionnelle possède un certain nombre de propriétés simples, et faciles à démontrer.

1. Si X et Y sont deux variables aléatoires indépendantes, et si φ est une fonction quelconque,

$$\mathbb{E} \{ \varphi(X) | Y = y \} = \mathbb{E} \{ \varphi(X) \} .$$

En effet, $\mathbb{E} \{ \varphi(X) | Y = y \} = \int \varphi(x) \mathbb{P} \{ dx | Y = y \} = \int \varphi(x) \mathbb{P} \{ dx \} = \mathbb{E} \{ \varphi(X) \} .$

2. Etant données deux variables aléatoires X et Y ,

$$\mathbb{E} \{ X \} = \int \mathbb{E} \{ X | Y = y \} d\mathbb{P}_Y \{ y \} .$$

Par exemple, dans le cas de variables aléatoires à densité, on a $\mathbb{E} \{ X \} = \int x \rho_X(x) dx = \int x \rho(x|y) \rho_Y(y) dx dy = \int \mathbb{E} \{ X | Y = y \} \rho_Y(y) dx dy .$

3. Si X, Y, Z sont des variables aléatoires dépendantes :

$$\mathbb{E} \{ X + Y | Z = z \} = \mathbb{E} \{ X | Z = z \} + \mathbb{E} \{ Y | Z = z \} .$$

4. Pour toute fonction $\varphi(X, Y)$,

$$\mathbb{E} \{ \varphi(X, Y) | Y = y \} = \mathbb{E} \{ \varphi(X, y) | Y = y \} .$$

Prédiction

La prédiction est devenue une thématique scientifique d'une importance considérable dans de multiples domaines (et pas seulement en météorologie, à la bourse ou chez Mme soleil...). On peut en particulier citer le problème du contrôle de la trajectoire des fusées (connaissant leur position et leur vitesse jusqu'à l'instant t , prédire la trajectoire à venir afin de pouvoir éventuellement apporter des corrections).

La formulation probabiliste du problème de prédiction est la suivante :

Etant données deux variables aléatoires X et Y (ou $N + 1$ variables aléatoires X et Y_1, \dots, Y_N), comment construire une fonction de Y (ou Y_1, \dots, Y_N) permettant de prédire au mieux la valeur prise par X , si l'on connaît la valeur prise par Y (ou les valeurs prises par les Y_1, \dots, Y_N).

Pour estimer la qualité de la prédiction, il faut tout d'abord réaliser que tout prédicteur $\varphi(Y)$ est lui aussi une variable aléatoire. Le critère le plus simple consiste à trouver la fonction φ qui minimise l'erreur en moyenne quadratique :

$$\min_{\varphi} \mathbb{E} \{ [X - \varphi(Y)]^2 \} . \tag{10.82}$$

On montre alors le résultat suivant

PROPOSITION 10.16 *La fonction φ^* qui minimise l'erreur de prédiction en moyenne quadratique est donnée par*

$$\varphi^*(y) = \mathbb{E} \{ X | Y = y \} . \tag{10.83}$$

En pratique, cette solution n'est pas toujours facile à expliciter, et il faut se contenter d'approximations.

Septième partie

Annexes

Mesure et convergence

A.1 Notion de mesure

La notion de mesure, essentielle en physique, a également pris un sens précis en mathématiques. On en donne ici très brièvement les éléments essentiels.

DÉFINITION A.1 Soit un ensemble \mathcal{E} , et soit \mathcal{A} une famille de sous-ensembles de E . On dit que \mathcal{A} est une algèbre si

1. \mathcal{A} est non vide.
2. Pour tout $E \in \mathcal{A}$, le complémentaire \bar{E} de E appartient aussi à \mathcal{A} .
3. Pour tous $E_1, E_2 \in \mathcal{A}$, leur réunion $E_1 \cup E_2$ appartient aussi à \mathcal{A} .

On dit que \mathcal{A} est une sigma-algèbre si de plus

4. Pour toute famille dénombrable $\{E_1, E_2, \dots\}$ d'éléments de \mathcal{A} , leur réunion $\bigcup E_i$ appartient aussi à \mathcal{A} .

L'exemple le plus classique est l'exemple des nombres réels : $\mathcal{E} = \mathbb{R}$. On considère la famille des sous-ensembles de la forme $] -\infty, x]$, où $x \in \mathbb{R}$. Il est possible de montrer que cette famille est en fait une sigma-algèbre, que l'on appelle *sigma-algèbre de Borel*.

DÉFINITION A.2 Soit \mathcal{E} un ensemble, et soit \mathcal{A} une sigma algèbre sur \mathcal{E} . On appelle mesure sur \mathcal{A} une application

$$\mu : \mathcal{A} \rightarrow \mathbb{C} \text{ (ou } \mathbb{R})$$

telle que

1. La mesure de l'ensemble vide est nulle.
2. Pour toute famille dénombrable $\{E_1, E_2, \dots\}$ d'éléments de \mathcal{A} , disjoints deux à deux (c'est à dire tels que pour tous i, j , $E_i \cap E_j$ soit vide),

$$\mu\left(\bigcup E_j\right) = \sum_j \mu(E_j). \quad (\text{A.1})$$

De nouveau, l'exemple le plus usuel est l'exemple de la mesure de Lebesgue sur \mathbb{R} . Cette mesure est définie de la façon suivante : à tout intervalle $]a, b]$, on associe sa mesure qui n'est autre que sa longueur :

$$\mu(]a, b]) = b - a. \quad (\text{A.2})$$

Un élément simple t_0 , ou une union dénombrable d'éléments de \mathbb{R} sont alors des ensembles de mesure nulle : pour tous $t_0, t_1, \dots \in \mathbb{R}$,

$$\mu(t_0) = \mu(\{t_0, t_1, \dots\}) = 0.$$

La notion d'ensemble de mesure nulle joue un rôle très important.

A.2 Convergence

La notion de convergence d'une suite de nombres est une notion classique. Par contre, dès lors que l'on s'intéresse à des suites de fonctions, il est possible de donner à la notion de convergence de suites des sens différents, qui traduisent des notions plus ou moins précises (ou fines) de convergence.

On résume ici rapidement les notions de convergence les plus usuelles, dans le cas de fonctions d'une variable réelle. Le cas des fonctions de plusieurs variables s'en déduit sans difficulté.

A.2.1 Convergence simple

On dit que la suite de fonctions $\{f_n, n = 0, 1, \dots\}$ définies sur un domaine A (soit donc $f_n : A \rightarrow \mathbb{C}$) converge simplement vers la fonction f si pour tout t , la suite des nombres $\{f(t), n = 0, 1, \dots\}$ converge vers le nombre $f(t)$. On traduit cela par :

$$f_n \rightarrow f \text{ si } \forall \epsilon > 0, \forall t \in A, \exists N(t, \epsilon) \text{ tel que } n \geq N(t, \epsilon) \implies |f_n(t) - f(t)| \leq \epsilon. \quad (\text{A.3})$$

Cette notion de convergence n'est pas la notion la plus fine, dans la mesure où le nombre $N(t, \epsilon)$ peut ici dépendre du point t considéré. Lorsque cet aspect n'est pas souhaitable, on préfère utiliser la notion de convergence uniforme.

A.2.2 Convergence uniforme

La contrainte de convergence uniforme est donc plus exigeante, et suppose que l'on puisse contrôler l'erreur commise en approximant f par f_n de façon indépendante du point étudié.

$$f_n \rightarrow f \text{ si } \forall \epsilon > 0, \exists N(\epsilon) \text{ tel que } \forall t \in A, n \geq N(\epsilon) \implies |f_n(t) - f(t)| \leq \epsilon. \quad (\text{A.4})$$

Bien évidemment, la convergence uniforme entraîne la convergence simple.

A.2.3 Convergence en moyenne

Les deux notions de convergence ci-dessus sont basées sur les valeurs ponctuelles des fonctions étudiées. Or, il est souvent utile de s'abstraire de l'utilisation des valeurs ponctuelles, et de définir des notions de convergence en moyenne.

Soit p un nombre réel positif. On dit que la suite de fonctions $\{f_n, n = 0, 1, \dots\}$ définies sur un domaine A (soit donc $f_n : A \rightarrow \mathbb{C}$) converge vers f en moyenne d'ordre p si la suite $\int_A |f(t) - f_n(t)|^p dt$ tend vers zéro lorsque n tend vers l'infini :

$$f_n \rightarrow f \text{ si } \forall \epsilon > 0, \exists N(\epsilon) \text{ tel que } n \geq N(\epsilon) \implies \int_A |f_n - f|^p dt \leq \epsilon. \quad (\text{A.5})$$

Pour que ces notions aient un sens, il faut bien entendu que les intégrales $\int_A |f(t) - f_n(t)|^p dt$ existent pour tout n . Ceci est assuré dès que l'on se place dans le cadre des espaces de Lebesgue appropriés, c'est à dire si l'on suppose que $f, f_n \in L^p(A)$. Les espaces L^p sont définis dans l'Annexe B.

La convergence en moyenne d'ordre p n'implique nullement la convergence uniforme ni la convergence simple. On en verra des contre exemples dans l'étude des séries de Fourier. Par contre, la convergence en moyenne d'ordre p entraîne la convergence simple presque partout.

A.2.4 Convergence simple presque partout

On dit que la suite de fonctions $\{f_n, n = 0, 1, \dots\}$ définies sur un domaine A (soit donc $f_n : A \rightarrow \mathbb{C}$) converge presque partout vers f si pour tout $t \in A$, $f_n(t) \rightarrow f(t)$, à l'exception éventuellement d'un ensemble de mesure nulle.

Espaces fonctionnels

La théorie des fonctions (aussi appelée analyse fonctionnelle) repose sur la notion d'espace fonctionnel, ou espace de fonctions. Il s'agit généralement de construire des ensembles de fonctions possédant certaines propriétés génériques (intégrabilité, dérivabilité,...) permettant d'effectuer certaines opérations données. Ces ensembles possèdent une structure d'espace vectoriel (c'est à dire qu'ils sont stables par addition et multiplication scalaire), et sont munis d'une **norme**, qui leur confère une structure d'**espace de Banach**. Dans certains cas, ils sont également munis d'un **produit Hermitien**, qui les munit d'une structure d'**espace de Hilbert**.

On donne ici une description succincte de certains des espaces fonctionnels les plus utiles, ainsi que de leurs propriétés essentielles. Entrer dans plus de détails dépasserait le cadre de ce cours.

B.1 Espaces caractérisant la régularité

Très grossièrement, la notion de régularité décrit la vitesse à laquelle varie une fonction. Plus elle varie lentement, plus elle est régulière. Par exemple, une fonction continue est plus régulière qu'une fonction discontinue, qui varie brusquement à chaque point de discontinuité.

Il existe de multiples façons de décrire (et en fait de définir) la régularité. La plus simple fait appel au degré de différentiabilité.

Dans ce qui suit, r est un nombre entier positif ou nul, appelé *degré de régularité*, ou régularité tout simplement. On note aussi, pour tout entier k et toute fonction f pour laquelle cette expression a un sens

$$f^{(k)}(t) = \frac{d^k f}{dt^k}(t)$$

La première famille d'espaces fonctionnels intéressants est la famille d'espaces de fonctions continûment différentiables :

$$C^r(\mathbb{R}) = \left\{ f : \mathbb{R} \rightarrow \mathbb{C}, f^{(k)} \text{ est continue } \forall k \leq r \right\} \quad (\text{B.1})$$

On vérifie aisément que $C^r(\mathbb{R})$ est un espace vectoriel. On a, de façon évidente

$$C^{r+1}(\mathbb{R}) \subset C^r(\mathbb{R}) . \quad (\text{B.2})$$

De même, si $I \subset \mathbb{R}$ est un intervalle borné, on définit

$$C^r(I) = \left\{ f : I \rightarrow \mathbb{C}, f^{(k)} \text{ est continue } \forall k \leq r \right\} \quad (\text{B.3})$$

$C^r(I)$ est un espace vectoriel, et on a encore

$$C^{r+1}(I) \subset C^r(I) . \quad (\text{B.4})$$

Etant donnée une fonction f , on définit son support $\text{supp}(f)$ par

$$\text{supp}(f) = \overline{\{t \in \mathbb{R}, f(t) \neq 0\}} \quad (\text{B.5})$$

Ceci nous permet d'introduire l'espace des fonctions continûment différentiables à support borné (ou compact)

$$C_0^r(\mathbb{R}) = \left\{ f : \mathbb{R} \rightarrow \mathbb{C}, f^{(k)} \text{ est continue } \forall k \leq r \text{ et } \text{supp}(f) \text{ est borné} \right\} \quad (\text{B.6})$$

Il s'agit de nouveau d'un espace vectoriel, et on a les inclusions

$$C_0^r(\mathbb{R}) \subset C^r(\mathbb{R}) ; \quad C_0^{r+1}(\mathbb{R}) \subset C_0^r(\mathbb{R}) . \quad (\text{B.7})$$

B.2 Les espaces $C^\infty(\mathbb{R})$, $\mathcal{D}(\mathbb{R})$ et $\mathcal{S}(\mathbb{R})$

Les cas limites suivants sont eux aussi très utiles. L'espace des fonctions dont toutes les dérivées sont continues

$$C^\infty(\mathbb{R}) = \left\{ f : \mathbb{R} \rightarrow \mathbb{C}, f^{(k)} \text{ est continue } \forall k \in \mathbb{Z}_+ \right\} , \quad (\text{B.8})$$

et l'espace des fonctions C^∞ à support borné

$$\mathcal{D}(\mathbb{R}) = \{ f \in C^\infty(\mathbb{R}), \text{supp}(f) \text{ est borné} \} , \quad (\text{B.9})$$

Il est maintenant utile d'introduire la notion de fonction à décroissance rapide.

DÉFINITION B.1 (FONCTION À DÉCROISSANCE RAPIDE) Soit f une fonction continue. On dit que f est à décroissance rapide si pour tout entier positif k , il existe une constante C_k telle que l'on ait

$$|f(t)| \leq \frac{C_k}{|t|^k} \quad (\text{B.10})$$

Il s'agit donc de fonctions qui décroissent à l'infini plus vite que toutes les puissances. Ceci nous permet d'introduire l'espace de Schwartz $\mathcal{S}(\mathbb{R})$ (parfois appelé aussi *classe de Schwartz*) :

$$\mathcal{S}(\mathbb{R}) = \left\{ f \in C^\infty(\mathbb{R}), f^{(k)} \text{ est à décroissance rapide } \forall k \in \mathbb{Z}_+ \right\} . \quad (\text{B.11})$$

De nouveau, on a des relations d'inclusion simples

$$\mathcal{D}(\mathbb{R}) \subset \mathcal{S}(\mathbb{R}) \subset C^\infty(\mathbb{R}) \quad (\text{B.12})$$

B.3 Éléments de théorie de l'intégration

Nous avons jusqu'à présent traité l'intégration en évitant d'entrer dans les détails. Cependant, pour pouvoir aborder certains aspects plus fins, il est maintenant nécessaire de préciser quelque peu les choses, sans toutefois entrer dans les détails. On donne ici une description "superficielle" de la différence entre l'intégration "à la Riemann" et l'intégration "à la Lebesgue". Plus de détails peuvent être trouvés dans les références indiquées à la fin de ce chapitre.

La théorie de l'intégration qui nous est familière est la théorie de l'intégration de Riemann. Cette intégrale est construite de la façon suivante. Supposons que nous voulions calculer

$$\int_a^b \varphi(t) dt ,$$

où $a < b$ sont deux nombres finis. On commence par "découper" l'intervalle $[a, b]$ en n sous-intervalles $[a_j, a_{j+1}[$ de taille $a_{j+1} - a_j$ inférieure à un certain δ_n , avec $a_0 = a$ et $a_{n+1} = b$. Puis on forme la quantité

$$\sum_{j=1}^n \varphi(t_j) (a_{j+1} - a_j) .$$

L'intégrale est alors égale à la limite de cette dernière quantité lorsque $n \rightarrow \infty$, quand cette limite existe. C'est en particulier le cas lorsque φ est une fonction continue, mais aussi une fonction bornée admettant une infinité de discontinuités. On montre également que cette intégrale existe pour une certaine classe de fonctions non bornées. Par exemple, si φ est divergente en $t = c \in [a, b]$, on convient de définir $\int_a^b \varphi(t)dt$ la limite quand $\gamma, \gamma' \rightarrow 0$ de la somme $\int_a^{c-\gamma} \varphi(t)dt + \int_{c+\gamma}^b \varphi(t)dt$, quand cette limite existe.

Cependant, cette théorie n'est pas encore assez générale pour permettre d'intégrer certaines fonctions simples. Par exemple, la fonction de Dirichlet $t \rightarrow f(t)$, définie par

$$f(t) = \begin{cases} 1 & \text{si } t \text{ est rationnel} \\ 0 & \text{sinon.} \end{cases}$$

Cette fonction n'est pas intégrable au sens de Riemann. Mais la notion d'intégrale introduite par H. Lebesgue en 1900 permet de donner une définition à l'intégrale de cette fonction, et de montrer qu'elle est en fait nulle. Très grossièrement, la construction de Lebesgue revient à découper l'axe des ordonnées plutôt que l'axe des abscisses. Plus précisément, étant donnée une fonction φ , on commence par partager le domaine $[A, B]$ dans lequel φ prend ses valeurs en sous-intervalles $[\alpha_j, \alpha_{j+1}[$. Étant donné un intervalle $[\alpha_j, \alpha_{j+1}[$, on note E_j l'ensemble des t tels que $\varphi(t) \in [\alpha_j, \alpha_{j+1}[$, et on attribue à l'ensemble E_j sa mesure de Lebesgue μ_j ; par exemple, si E_j est un intervalle, sa mesure est la longueur de l'intervalle. Ou si E_j est une réunion d'intervalles disjoints, sa mesure μ_j est la somme des longueurs des intervalles. Autre exemple, si E_j est réduit à un point ou un ensemble fini ou dénombrable de points, sa mesure est nulle. On considère alors les sommes partielles

$$\sum_j \mu_j \phi_j,$$

où $\phi_j \in [\alpha_j, \alpha_{j+1}[$, et $\int_a^b \varphi(t)dt$ est alors défini comme la limite de ces expressions quand l'on fait tendre tous les $\alpha_{j+1} - \alpha_j$ vers 0, quand cette limite existe.

Très grossièrement, la différence entre l'intégration à la Riemann" et l'intégration "à la Lebesgue" peut être schématisée comme en FIG. B.1. On y voit en particulier que le découpage effectué dans le cadre de l'intégration "à la Lebesgue" est bien plus précis dans les régions où la fonction intégrée varie rapidement.

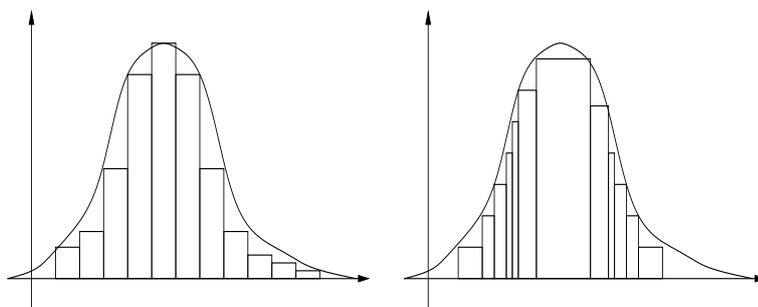


FIG. B.1: L'intégration "à la Riemann" (à gauche) et l'intégration "à la Lebesgue" (à droite)

EXEMPLE B.1 Revenons à la fonction de Dirichlet. Elle ne peut prendre que deux valeurs différentes, 0 et 1. En notant E_1 le domaine des valeurs de t pour lesquelles $\varphi(t) = 1$, on voit que E_1 n'est autre que l'ensemble des nombres rationnels, qui est un ensemble infini dénombrable. Donc sa mesure est égale à 0, et par conséquent, l'intégrale de la fonction de Dirichlet est nulle.

REMARQUE B.1 Nous avons vu intervenir la notion de mesure d'un ensemble. Il découle de la construction même de l'intégrale de Lebesgue que si deux fonctions diffèrent sur un ensemble de mesure nulle (on dit alors qu'elles sont égales presque partout), leurs intégrales seront égales.

On montre que les résultats classiques de la théorie de l'intégration (intégration par parties, lemme de Fubini,...) restent valides dans le cadre de la théorie de Lebesgue. De plus, le résultat suivant, appelé *théorème de convergence dominée de Lebesgue* est très important, car il donne d'une part un moyen de vérifier l'intégrabilité d'une fonction, mais aussi de calculer l'intégrale.

THÉORÈME B.1 (CONVERGENCE DOMINÉE) Soit $\{f_k, k = 0, 1, \dots\}$ une suite de fonctions, telle que pour presque tout t , $f(t) = \lim_{k \rightarrow \infty} f_k(t)$ existe. Supposons de plus qu'il existe une fonction intégrable positive g telle que pour tout k , $|f_k(t)| \leq g(t)$. Alors, f est intégrable, et

$$\int_{-\infty}^{\infty} f(t)dt = \lim_{k \rightarrow \infty} \int_{-\infty}^{\infty} f_k(t)dt . \quad (\text{B.13})$$

B.4 Les espaces de Lebesgue

Ces considérations nous permettent maintenant de nous pencher sur des espaces caractérisant les propriétés d'intégrabilité des fonctions, après avoir examiné des espaces caractérisant leurs propriétés de régularité.

A partir de maintenant, p désigne un nombre réel supérieur à 1 : $1 < p < \infty$. A toute fonction f , on associe le nombre

$$\|f\|_p = \left\{ \int_{-\infty}^{\infty} |f(t)|^p dt \right\}^{1/p} . \quad (\text{B.14})$$

Ceci permet d'introduire les espaces de Lebesgue

$$L^p(\mathbb{R}) = \{f : \mathbb{R} \rightarrow \mathbb{C}, \|f\|_p < \infty\} \quad (\text{B.15})$$

Il est important de remarquer que l'intégrale utilisée ici étant l'intégrale de Lebesgue, les fonctions de $L^p(\mathbb{R})$ doivent être vues plutôt comme des classes d'équivalence de fonctions (deux fonctions étant équivalentes si elles diffèrent sur un ensemble de mesure nulle) que comme des fonctions définies en chaque point. Ceci étant entendu, l'application $f \in L^p(\mathbb{R}) \rightarrow \|f\|_p$ définit une norme sur $L^p(\mathbb{R})$:

1. Inégalité triangulaire : si $f, g \in L^p(\mathbb{R})$, $\|f + g\|_p \leq \|f\|_p + \|g\|_p$.
2. si $f \in L^p(\mathbb{R})$ et $\lambda \in \mathbb{C}$, $\|\lambda f\|_p = |\lambda| \|f\|_p$.
3. $\|f\|_p = 0$ implique $f = 0$.

REMARQUE B.2 Il n'est pas tout à fait exact de dire que l'application $f \in L^p(\mathbb{R}) \rightarrow \|f\|_p$ définit une norme sur les espaces L^p ainsi définis. En effet, si $f = 0$ presque partout, alors $\|f\|_p = 0$, sans que f soit nécessairement uniformément nulle. Pour s'abstraire de ce problème, on convient d'identifier les fonctions qui ne diffèrent que sur un ensemble de mesure nulle. Dans ce cas, la propriété 3 ci dessus est bien vérifiée.

REMARQUE B.3 Les normes introduites plus haut permettent de proposer de nouveaux critères de convergence pour les suites de fonction. On dira par exemple que la suite $\{f_n, n = 0, 1, \dots\}$ de fonctions de $L^p(\mathbb{R})$ converge vers une fonction $f \in L^p(\mathbb{R})$ en moyenne d'ordre p si $\lim_{k \rightarrow \infty} \|f - f_k\|_p = 0$. La vérification de la convergence en moyenne d'ordre p utilise souvent le théorème de convergence dominée que nous avons vu plus haut.

Parmi les espaces $L^p(\mathbb{R})$, l'espace $L^2(\mathbb{R})$ possède un statut particulier : il possède un produit scalaire : si $f, g \in L^2(\mathbb{R})$, on définit

$$\langle f, g \rangle = \int_{-\infty}^{\infty} \overline{f(t)}g(t)dt \quad (\text{B.16})$$

Ce produit scalaire est associé à la norme $\|f\|_2 = \|f\|$, par

$$\|f\| = \sqrt{\langle f, f \rangle}. \quad (\text{B.17})$$

On montre que ce produit est *sesquilinéaire*, c'est à dire qu'il vérifie les conditions suivantes :

1. Hermiticité : si $f, g \in L^2(\mathbb{R})$, on a $\langle f, g \rangle = \overline{\langle g, f \rangle}$.
2. Linéarité à droite : pour tous $f, g_1, g_2 \in L^2(\mathbb{R}), \alpha_1, \alpha_2 \in \mathbb{C}$, $\langle f, \alpha_1 g_1 + \alpha_2 g_2 \rangle = \alpha_1 \langle f, g_1 \rangle + \alpha_2 \langle f, g_2 \rangle$.
3. Positivité : soit $f \in L^2(\mathbb{R})$. Si pour toute $g \in L^2(\mathbb{R})$, on a $\langle f, g \rangle = 0$, alors $f = 0$.

En conséquence des propriétés précédentes, nous avons donc antilinéarité à gauche : pour tous $f_1, f_2, g \in L^2(\mathbb{R}), \alpha_1, \alpha_2 \in \mathbb{C}$, $\langle \alpha_1 f_1 + \alpha_2 f_2, g \rangle = \overline{\alpha_1} \langle f_1, g \rangle + \overline{\alpha_2} \langle f_2, g \rangle$.

Les espaces de Lebesgue possèdent des propriétés intéressantes et utiles, appelées *inégalités de Hölder* et *inégalités d'Young*, qui sont des généralisations de l'inégalité de Cauchy-Schwarz. Nous les donnons sans démonstration ci-dessous.

PROPOSITION B.1 (INÉGALITÉS DE HÖLDER ET YOUNG) 1. Soient $f \in L^p(\mathbb{R})$ et $g \in L^q(\mathbb{R})$. Soit r défini par

$$\frac{1}{r} = \frac{1}{p} + \frac{1}{q}. \quad (\text{B.18})$$

Alors, $fg : t \rightarrow f(t)g(t)$ appartient à $L^r(\mathbb{R})$, et on a (inégalité de Hölder) :

$$\|fg\|_r \leq \|f\|_p \|g\|_q. \quad (\text{B.19})$$

2. Soient $f \in L^p(\mathbb{R})$ et $g \in L^q(\mathbb{R})$. Soit r' défini par

$$1 + \frac{1}{r'} = \frac{1}{p} + \frac{1}{q}. \quad (\text{B.20})$$

Alors, $f \star g \in L^{r'}(\mathbb{R})$, et on a (inégalité d'Young) :

$$\|f \star g\|_{r'} \leq \|f\|_p \|g\|_q. \quad (\text{B.21})$$

De plus, si $p = q = 2$, (donc $r' = \infty$) $f \star g$ est continue.

REMARQUE B.4 On vérifie bien que dans le cas $p = q = 2$, on a $r = 1$, et donc l'inégalité de Hölder correspondante se ramène bien à l'inégalité de Cauchy-Schwarz.

B.5 Relations entre $\mathcal{S}(\mathbb{R})$, $\mathcal{D}(\mathbb{R})$ et $L^2(\mathbb{R})$

Deux remarques sont importantes à faire. Tout d'abord, il est facile de voir que

$$\mathcal{S}(\mathbb{R}) \subset L^2(\mathbb{R}) \quad (\text{B.22})$$

En effet, si $f \in \mathcal{S}(\mathbb{R})$, nous savons qu'il existe une constante A telle que

$$|f(t)| \leq A/(1+t^2).$$

Par conséquent,

$$\|f\|^2 \leq A^2 \int_{-\infty}^{\infty} \frac{dt}{(1+t^2)^2} < \infty,$$

ce qui montre que $f \in L^2(\mathbb{R})$.

Le second point important est contenu dans la proposition suivante, que nous donnons sans démonstration :

PROPOSITION B.2 $\mathcal{D}(\mathbb{R})$ est dense dans $L^2(\mathbb{R})$: pour toute fonction $f \in L^2(\mathbb{R})$ et pour tout nombre $\epsilon > 0$, il existe une fonction $f_\epsilon \in \mathcal{D}(\mathbb{R})$ telle que

$$\|f - f_\epsilon\| \leq \epsilon$$

En d'autres termes, pour toute fonction $f \in L^2(\mathbb{R})$, on peut toujours trouver une suite de fonctions appartenant à $\mathcal{D}(\mathbb{R})$ qui converge vers f en moyenne d'ordre 2. Un corollaire immédiat est que $\mathcal{S}(\mathbb{R})$ est lui aussi dense dans $L^2(\mathbb{R})$. C'est sur cette propriété que l'on s'appuie généralement pour étendre la transformation de Fourier à $L^2(\mathbb{R})$.

C

Elements de géométrie des courbes et surfaces

C.1 Courbes, courbes paramétrées

C.1.1 Courbes planes

Les courbes planes sont des courbes dans le plan, généralement définies en coordonnées Cartésiennes par une équation de la forme $F(x, y) = 0$. Cette relation permet parfois d'exprimer directement y en fonction de x , ou vice-versa. On a aussi souvent recours à des formes paramétrées.

DÉFINITION C.1 On appelle *courbe plane paramétrée*, ou *arc paramétré* une application $\underline{\alpha}$

$$t \in [a, b] \mapsto \underline{\alpha}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \in \mathbb{R}^2,$$

où a, b sont deux réels, et x et y sont deux fonctions continues. Si x et y sont différentiables, on parle de *courbe différentiable*. Si x' et y' ne s'annulent jamais simultanément, la courbe est dite *régulière*.

Des exemples de courbes paramétrées (des cardioïdes), d'équation paramétrique

$$\theta \in [0, 2\pi] \mapsto \underline{\alpha}(\theta) = \begin{cases} x(\theta) = a \cos \theta (1 + \cos \theta) \\ y(\theta) = a \sin \theta (1 + \cos \theta) \end{cases}$$

où a est un paramètre, se trouvent en FIG. C.1.

La longueur d'un arc paramétré $t \in [a, b] \rightarrow \underline{\alpha}(t)$ est donnée par l'expression

$$L(\underline{\alpha}) = \int_a^b \|\underline{\alpha}'(t)\| dt = \int_a^b \sqrt{x'(t)^2 + y'(t)^2} dt. \quad (\text{C.1})$$

Ceci permet de définir l'abscisse curviligne

$$s : t \in [a, b] \rightarrow s(t) = \int_a^t \sqrt{x'(u)^2 + y'(u)^2} du, \quad (\text{C.2})$$

qui fournit une reparamétrisation naturelle de la courbe

$$\underline{\beta}(s) = \underline{\alpha}(t(s)), \quad (\text{C.3})$$

où on a noté $s \rightarrow t(s)$ la réciproque de la fonction $t \rightarrow s(t)$. On montre facilement que le vecteur

$$\underline{N}(s) = \underline{\beta}'(s) \quad (\text{C.4})$$

est un vecteur unitaire, tangent à la courbe en $\underline{\beta}(s)$.

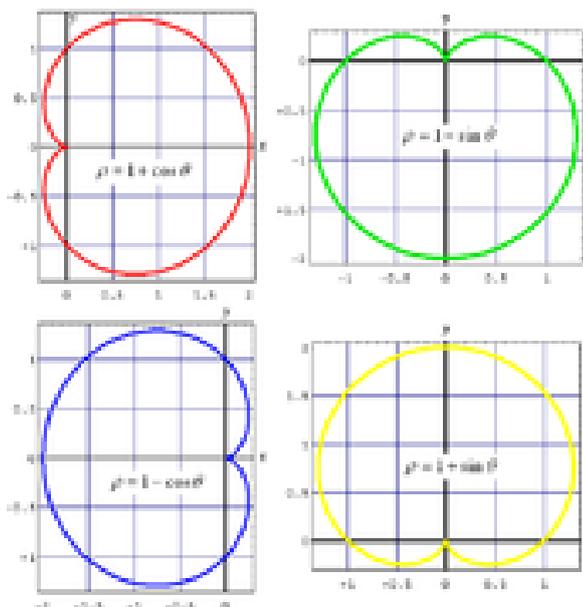


FIG. C.1: Quatre cardioïdes

La courbure est quant à elle définie à partir de l'abscisse curviligne s comme

$$\kappa : s \longmapsto \kappa(s) = \|\beta''(s)\|, \tag{C.5}$$

et on définit le rayon de courbure comme l'inverse de la courbure $\rho = 1/\kappa$. Un calcul explicite donne en outre

$$\kappa(t) = \frac{\alpha''(t) \cdot \alpha'^{\perp}(t)}{\|\alpha'\|^3}. \tag{C.6}$$

Le rayon de courbure est le rayon d'un cercle, appelé **cercle osculateur**, tangent à la courbe au point considéré, et qui en est dans un certain sens le plus proche possible (c'est à dire qui en épouse la forme au mieux). Le centre de ce cercle est appelé **centre de courbure**

Etant donnée une courbe paramétrée, on peut lui associer en tout point son vecteur unitaire tangent

$$\underline{T}(t) = \frac{1}{\sqrt{x'(t)^2 + y'(t)^2}} \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} \tag{C.7}$$

et le vecteur unitaire normal

$$\underline{N}(t) = \frac{1}{\sqrt{x'(t)^2 + y'(t)^2}} \begin{pmatrix} -y'(t) \\ x'(t) \end{pmatrix}. \tag{C.8}$$

On peut alors montrer les **relations de Frénet**

$$\frac{d\underline{T}(s)}{ds} = \frac{\underline{N}}{\rho(s)}, \quad \text{et} \quad \frac{d\underline{N}(s)}{ds} = -\frac{\underline{T}}{\rho(s)}. \tag{C.9}$$

C.1.2 Courbes gauches

Une courbe gauche est une courbe dans l'espace tridimensionnel. Dans un repère orthonormé, une courbe gauche s'écrit naturellement sous forme paramétrique

DÉFINITION C.2 On appelle **courbe gauche paramétrée** une application $\underline{\alpha}$

$$t \in [a, b] \longmapsto \underline{\alpha}(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} \in \mathbb{R}^3,$$

où a, b sont deux réels, et x, y et z sont trois fonctions continues. Si x, y, z sont différentiables, on parle de courbe gauche différentiable. Si x', y' et z' ne s'annulent jamais simultanément, la courbe est dite régulière.

Un exemple de courbe gauche, la courbe de Viviani, d'équation paramétrique

$$\begin{cases} x(t) = R \cos^2(t) \\ y(t) = R \cos(t) \sin(t) \\ z(t) = R \sin(t) \end{cases}$$

où R est un paramètre, se trouve en FIG. C.2.

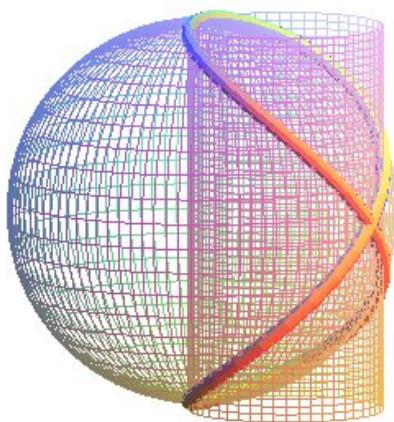


FIG. C.2: La courbe de Viviani est l'intersection d'une sphère de rayon R et d'un cylindre de révolution de diamètre R dont une génératrice passe par le centre de la sphère.

La longueur d'un arc de courbe gauche se définit de façon similaire au cas des courbes planes :

$$L(\underline{\alpha}) = \int_a^b \|\underline{\alpha}'(t)\| dt = \int_a^b \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2} dt, \quad (\text{C.10})$$

ce qui permet cette fois encore de définir l'abscisse curviligne

$$s : t \in [a, b] \rightarrow s(t) = \int_a^t \sqrt{x'(u)^2 + y'(u)^2 + z'(u)^2} du, \quad (\text{C.11})$$

et fournit une reparamétrisation naturelle de la courbe

$$\underline{\beta}(s) = \underline{\alpha}(t(s)), \quad (\text{C.12})$$

où on a noté $s \rightarrow t(s)$ la réciproque de la fonction $t \rightarrow s(t)$.

On définit comme dans le cas des courbes planes le vecteur unitaire tangent

$$\underline{T}(t) = \frac{1}{\sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}} (x'(t), y'(t), z'(t)). \quad (\text{C.13})$$

De même, on montre que la dérivée du vecteur unitaire tangent lui est perpendiculaire, ce qui permet de définir le vecteur normal principal \underline{N} , par

$$\frac{d\underline{T}}{ds} = \frac{1}{\rho} \underline{N}, \quad (\text{C.14})$$

où $\rho = 1/\kappa$ est le **rayon de courbure**, inverse de la **courbure**. En complétant par le vecteur binormal

$$\underline{B} = \underline{T} \times d\underline{N}, \quad (\text{C.15})$$

on obtient ainsi un repère orthonormé direct, appelé **repère de Frénet** $(\underline{T}, \underline{N}, \underline{B})$ de l'espace. On montre facilement que $d\underline{B}/dt$ est proportionnel à \underline{N} , de sorte que l'on écrit

$$\frac{d\underline{B}}{ds} = \frac{1}{r} \underline{N}, \quad (\text{C.16})$$

où $r = 1/\tau$ est le **rayon de torsion**, inverse de la **torsion** τ . Finalement, on montre aussi

$$\frac{d\underline{N}}{ds} = -\frac{1}{\rho} \underline{T} - \frac{1}{r} \underline{B}. \quad (\text{C.17})$$

C.2 Surfaces paramétrées dans l'espace

DÉFINITION C.3 On appelle **surface paramétrée** une application $\underline{\alpha}$

$$(t_1, t_2) \in [a_1, b_1] \times [a_2, b_2] \longmapsto \underline{\alpha}(t_1, t_2) = (x(t_1, t_2), y(t_1, t_2), z(t_1, t_2)) \in \mathbb{R}^3,$$

où a_1, a_2, b_1, b_2 sont quatre réels, et x, y et z sont trois fonctions continues. Si x, y, z sont différentiables, on parle de surface différentiable.

Un exemple de surface paramétrée se trouve en FIG. C.3; il s'agit d'un coquillage, d'équation paramétrique

$$\begin{cases} x(t_1, t_2) &= (a + b \cos(t_2))e^{mt_1} \cos(t_1) \\ y(t_1, t_2) &= (a + b \cos(t_2))e^{mt_1} \sin(t_1) \\ z(t_1, t_2) &= (a + b \sin(t_2))e^{mt_1}, \end{cases}$$

où a, b et m sont des paramètres réels qui contrôlent la forme de la surface.

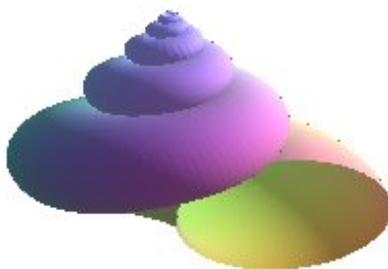


FIG. C.3: Un coquillage

En tout point $\underline{t} = (t_1, t_2)$, on peut calculer les deux vecteurs tangents définis par

$$\frac{\partial \underline{\alpha}}{\partial t_1}(\underline{t}) = \begin{pmatrix} \frac{\partial x(\underline{t})}{\partial t_1} \\ \frac{\partial y(\underline{t})}{\partial t_1} \\ \frac{\partial z(\underline{t})}{\partial t_1} \end{pmatrix}, \quad \frac{\partial \underline{\alpha}}{\partial t_2}(\underline{t}) = \begin{pmatrix} \frac{\partial x(\underline{t})}{\partial t_2} \\ \frac{\partial y(\underline{t})}{\partial t_2} \\ \frac{\partial z(\underline{t})}{\partial t_2} \end{pmatrix}. \quad (\text{C.18})$$

Lorsqu'en un point \underline{t} ces deux vecteurs sont linéairement indépendants (c'est à dire ne sont pas colinéaires), on dit que ce point est régulier, et on définit le plan tangent à la surface par

$$\mathbf{T}_t = \left\{ \underline{\alpha}(\underline{t}) + a \frac{\partial \underline{\alpha}}{\partial t_1}(\underline{t}) + b \frac{\partial \underline{\alpha}}{\partial t_2}(\underline{t}), a, b \in \mathbb{R} \right\}, \quad (\text{C.19})$$

En notant \underline{T}_1 et \underline{T}_2 les vecteurs unitaires correspondant

$$\underline{T}_1(\underline{t}) = \frac{1}{\left\| \frac{\partial \underline{\alpha}}{\partial t_1} \right\|} \frac{\partial \underline{\alpha}}{\partial t_1}(\underline{t}), \quad \underline{T}_2(\underline{t}) = \frac{1}{\left\| \frac{\partial \underline{\alpha}}{\partial t_2} \right\|} \frac{\partial \underline{\alpha}}{\partial t_2}(\underline{t}), \quad (\text{C.20})$$

on en déduit l'expression de la normale à la surface :

$$\underline{N}(\underline{t}) = \underline{T}_1(\underline{t}) \times \underline{T}_2(\underline{t}). \quad (\text{C.21})$$

Il est possible de démontrer que l'aire d'une surface paramétrée est donnée par l'expression

$$\mathcal{A}(\alpha) = \iint d\mathcal{A}(\underline{t}) = \iint \left\| \frac{\partial \underline{\alpha}}{\partial t_1}(\underline{t}) \times \frac{\partial \underline{\alpha}}{\partial t_2}(\underline{t}) \right\| dt_1 dt_2, \quad (\text{C.22})$$

et que cette aire ne dépend pas de la paramétrisation. L'expression

$$d\mathcal{A}(\underline{t}) = \left\| \frac{\partial \underline{\alpha}}{\partial t_1}(\underline{t}) \times \frac{\partial \underline{\alpha}}{\partial t_2}(\underline{t}) \right\| dt_1 dt_2$$

représente ce que l'on nomme parfois **élément d'aire**. Sans entrer dans les détails de la démonstration, remarquons que ceci n'est autre que l'intégrale de la norme du produit vectoriel des deux vecteurs tangents, et que ce produit vectoriel représente l'aire du parallélogramme engendré par les deux vecteurs. Ainsi, calculer cette intégrale revient en quelque sorte à sommer des aires de "parallélogrammes infinitésimaux" utilisés pour paver la surface.

REMARQUE C.1 Il est possible d'associer à une surface paramétrée des notions de courbure locale ; plus précisément, on associe à tout point d'une surface régulière deux courbures dites principales, qui permettent d'en caractériser certaines propriétés géométriques. Nous n'irons pas plus loin ici.

Une surface est dite **orientable** s'il est possible de définir une normale orientée en tout point, et que cette normale dépend continûment des paramètres. Par exemple, un ruban usuel est orientable, alors qu'un ruban de Moëbius (voir FIG. C.4) ne l'est pas.

Étant donnée une surface orientable, la fonction $\underline{t} \rightarrow \underline{N}(\underline{t})$ définit une orientation. A tout champ de vecteur $\underline{x} \in \mathbb{R}^3 \rightarrow \underline{V}$ de l'espace on peut alors associer son **flux** à travers la surface $\underline{\alpha}$:

$$\Phi_{\underline{\alpha}}(\underline{V}) = \int_{\underline{\alpha}} \underline{V}(\underline{\alpha}(\underline{t})) \cdot \underline{N}(\underline{t}) d\mathcal{A}(\underline{t}). \quad (\text{C.23})$$

On montre que le flux est indépendant de la paramétrisation, et ne dépend que de l'orientation.

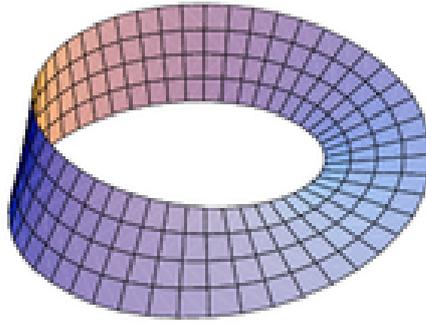


FIG. C.4: Le ruban de Moëbius n'est pas orientable

THÉORÈME C.1 (FORMULE D'OSTROGRADSKY) Soit S une surface fermée qui délimite un domaine U de \mathbb{R}^3 . Soit \underline{V} un champ de vecteurs de classe C^1 défini sur U . On paramétrise $S = \{\underline{t} \rightarrow \underline{\alpha}(\underline{t})\}$ de sorte que le vecteur normal $\underline{N}(\underline{t})$ pointe toujours vers l'extérieur de U . Alors on a

$$\iiint_U \operatorname{div}(\underline{V}(x, y, z)) \, dx dy dz = \iint_S \underline{V}(\underline{\alpha}(\underline{t})) \cdot \underline{N}(\underline{t}) \, d\mathcal{A}(\underline{t}) = \Phi_{\underline{\alpha}}(\underline{V}).$$

Cette formule trouve de nombreuses applications en physique ; elle peut également être utilisée pour calculer le volume du solide délimité par une surface fermée.

EXEMPLE C.1 (AIRE ET VOLUME DE LA SPHÈRE) Considérons une sphère de rayon r , centrée à l'origine. En coordonnées sphériques, elle est paramétrée par

$$\begin{cases} x(\theta, \varphi) = r \sin \varphi \cos \theta \\ y(\theta, \varphi) = r \sin \varphi \sin \theta \\ z(\theta, \varphi) = r \cos \varphi \end{cases}$$

Les deux vecteurs tangents naturellement associés à cette paramétrisation sont donnés par

$$\underline{T}_1(\theta, \varphi) = r \sin \varphi \begin{pmatrix} -\sin \theta \\ \cos \theta \\ 0 \end{pmatrix}, \quad \underline{T}_2(\theta, \varphi) = r \begin{pmatrix} \cos \varphi \cos \theta \\ \cos \varphi \sin \theta \\ -\sin \varphi \end{pmatrix},$$

et

$$\underline{T}_1(\theta, \varphi) \times \underline{T}_2(\theta, \varphi) = r^2 \sin \varphi \begin{pmatrix} -\cos \theta \sin \varphi \\ \sin \theta \sin \varphi \\ \cos \varphi \end{pmatrix}, \quad d\mathcal{A}(\varphi, \theta) = r^2 \sin \varphi \, d\theta d\varphi$$

de sorte que

$$\mathcal{A} = \int_0^\pi \int_0^{2\pi} r^2 \sin \varphi \, d\theta d\varphi = 4\pi r^2.$$

Le volume s'obtient quant à lui grâce à la formule d'Ostrogradsky. Il suffit de prendre le champ de vecteur $\underline{V}(\theta, \varphi) = \underline{r}(\theta, \varphi)$. On a alors $\operatorname{div} \underline{V} = 3$, et

$$\mathcal{V} = \frac{1}{3} \iiint \operatorname{div}(\underline{V}(x, y, z)) \, dx dy dz = \frac{r}{3} \iint d\mathcal{A}(\theta, \varphi) = \frac{4}{3} \pi r^3.$$

D Théorie des perturbations

D.1 Comparaisons : notations de Landau

La comparaison asymptotique consiste à étudier le comportement d'une fonction au voisinage d'un point ou à l'infini, en prenant pour référence le comportement d'une autre fonction fixée. On utilise souvent pour cela des monômes.

Ci dessous, on note $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty; -\infty\}$.

D.1.1 Négligeabilité

Soit $a \in \overline{\mathbb{R}}$. Soient f et g deux fonctions de la variable réelle t , telles que g ne s'annule pas sur un voisinage de a . On dit que f est négligeable devant g en a , et on note

$$f(t) = o_{t \rightarrow a}(g(t))$$

lorsque

$$\frac{f(t)}{g(t)} \xrightarrow{t \rightarrow a} 0.$$

Si le contexte est clair, on ne précise pas le domaine d'étude et on note : $f(t) = o(g(t))$, voire $f = o(g)$ (dire *petit o*). Cependant, la notation est toujours associée à un lieu a et au voisinage de ce lieu : être négligeable est un concept local.

Par abus de langage, on effectue parfois des opérations sur "les petits o ", c'est-à-dire sur les négligeables. En effet, on peut dire que : $o(f) + o(f) = o(f)$ et $o(f) - o(f) = o(f)$.

D.1.2 Equivalence

Soit $a \in \overline{\mathbb{R}}$. Soient f et g deux fonctions de la variable réelle t , telles que g ne s'annule pas sur un voisinage de a . On dit que f équivaut à g en a , ou que g équivaut à f en a , et on note

$$f(t) \underset{t \rightarrow a}{\sim} g(t),$$

lorsque

$$f(t) - g(t) = o_{t \rightarrow a}(g(t)).$$

D.1.3 Domination

La notation *grand O* dénote le caractère dominé d'une fonction par rapport à une autre. Soit $a \in \overline{\mathbb{R}}$. Soient f et g deux fonctions de la variable réelle t . On dit que f est dominée par g en $\pm\infty$, ou que g domine f en $\pm\infty$, et on note

$$f(t) = O_{t \rightarrow \pm\infty}(g(t)),$$

lorsqu'il existe des constantes N et C telles que

$$\forall t \in \mathbb{R} \text{ tel que } |t| > N, \quad |f(t)| \leq C |g(t)|.$$

Intuitivement, cela signifie que f ne croît pas plus vite que g .

Similairement, si a est un réel, on dit que f est dominé par g au voisinage de a , et on écrit

$$f(t) = O(g(t)) \quad \text{quand } t \rightarrow a$$

si et seulement si il existe des constantes $\delta > 0$ et C telles que

$$\forall t \quad |t - a| < \delta \implies |f(t)| \leq C |g(t)|.$$

D.2 Perturbations, équations algébriques

La théorie des perturbations consiste à rechercher des solutions de problèmes complexes, partant de problèmes plus simples que l'on sait résoudre. Par exemple, supposons que x_0 soit solution d'une équation

$$F(x) = 0,$$

pour une certaine fonction F , et que l'on cherche à résoudre une équation légèrement différente

$$F_\epsilon(x) = 0,$$

où la fonction F_ϵ est proche de F , la différence dépendant d'un "petit" paramètre ϵ , par exemple

$$F_\epsilon = F + \epsilon G$$

pour une certaine fonction G . On dit que F_ϵ est une version perturbée de F , et G est la perturbation. On peut alors chercher une solution $x(\epsilon)$ sous forme d'une perturbation de x , sous la forme

$$x(\epsilon) = x_0 + \epsilon x_1 + o(\epsilon),$$

ou plus généralement d'une série

$$x(\epsilon) = \sum_{k=0}^{\infty} x_k \epsilon^k,$$

dont on va chercher à identifier les coefficients x_k .

À SUIVRE...

Index

- Abscisse d'intégrabilité, 92
Abscisse de sommabilité, 92
Action, 32
Application différentielle, 12
Application tangente, 12
- Brachistochrone, 22
- Caténoïde, 25
Chaînette, 25, 34
Circuit RC, 69
Colonne de Cox-Overton, 25
Condition aux bords naturelle, 31
Condition de Neumann, 142
Conditions aux bords, 108, 126
Conditions aux limites, 108
Conditions de Cauchy, 107
Conditions de Dirichlet, 142
Conditions de Robin, 142
Conditions initiales, 107
Convolution des fonctions, 70
Convolution des suites, 47
Convolution des vecteurs, 43
Convolution sur \mathbb{R}^+ , 93
Corde pesante, 33
Courbe intégrale, 107
Cycloïde, 23
- Dérivée fonctionnelle, 20
Dérivées de la transformée de Fourier intégrale, 76
Distribution de Dirac, 160
Distribution impaire, 163
Distribution paire, 163
Distribution régulière, 161, 164
Dual, 159
- Echantillon, 161
Equation aux dérivées partielles, 133
Equation d'Euler-Lagrange, 19, 20
Equation de Beltrami, 21
Equation de Bessel, 120
Equation de la chaleur, 190
Equation de la chaleur dans un milieu semi-infini, 99
Equation de la chaleur finie, 37
Equation de Laplace, 28
Equation de Legendre, 119
Equation de Sturm-Liouville, 127
Equation des ondes, 29
Equation différentielle, 107
Equation différentielle à coefficients constants, 108
Equation différentielle homogène, 108
Equation différentielle non-linéaire, 108
Equation différentielle autonome, 107
Equation indicelle, 117
Equations de Lagrange, 32
Espace affine, 109
Espace de Schwartz, 81
Espace des phases, 115
Espace dual, 160
- Facteur intégrant, 113, 127
Filtrage numérique, 48
Filtre RC, 69
Fonction à croissance au plus exponentielle, 92
Fonction à croissance lente, 92
Fonction de Green, 130
Fonction de Hankel, 122
Fonction de Heaviside, 159
Fonction de Neumann, 121, 122
Fonction génératrice, 125

- Fonction Gamma, 121
 Fonction impaire, 163
 Fonction localement intégrable, 91, 164
 Fonction paire, 163
 Fonctionnelle, 19
 Fonctions de Bessel, 153
 Formalisme Hamiltonien, 32
 Forme résolue d'une équation différentielle, 107
 Formulation variationnelle, 11
 Formule de Rodrigues, 125
 Formule d'échange, 78
 Formule de Plancherel, 81, 82
 Formule de Rodrigues, 119, 125
 Formule intégrale de Poisson, 139

 Hamiltonien, 22, 33
 Hessienne, 16

 Inégalité de Heisenberg, 83
 Intégrale de Fourier, 71

 Jacobien, 13

 Lagrangien, 32
 Lemme d'échange, 78
 Lemme de Carslaw-Jaeger, 99
 Lemme fondamental du calcul variationnel, 20
 Ligne droite, 21
 Loi de conservation, 33
 Lois de Snell-Descartes, 17

 Mécanique Hamiltonienne, 32
 Méthode de Laplace, 110
 Méthode de variation de la constante, 110
 Méthode de Fröbenius, 116
 Matrice Jacobienne, 13
 Mirage, 25
 Modèle probabiliste, 225
 Multiplicateur de Lagrange, 18

 Noyau de Poisson, 140

 Opérateur de convolution, 70
 Opérateur de translation, 163
 Opérateur différentiel, 126
 Opérateur parité, 163
 Opérateur de la chaleur, 190
 Optimisation contrainte, 18, 33

 Partie finie de $1/t^2$, 167
 Peigne de Dirac, 161
 Pendule double, 32
 Point critique, 15

 Point ordinaire, 116
 Point régulier, 116
 point selle, 16
 Point singulier, 116
 Polynômes d'Hermite, 125
 Polynômes de Laguerre, 125
 Principe de Fermat, 17, 25
 Principe de moindre action, 32
 Problème aux bords, 136
 problème de Dirichlet extérieur, 136
 Problème de Dirichlet intérieur, 136
 Problème de Neumann extérieur, 136
 Problème de Neumann intérieur, 136
 Problème de plateau, 27
 Problème de Sturm-Liouville, 125
 Produit tensoriel, 164

 Quantité conservée, 22, 30

 Réponse impulsionnelle, 48, 130

 Signal analogique, 48
 Signal numérique, 48
 Sinus cardinal, 73
 Solution de d'Alembert, 150
 Surface minimale, 27

 Théorème convolution produit : le cas des suites, 47
 Théorème convolution produit : le cas des vecteurs, 44
 Théorème convolution-produit : le cas des fonctions, 80, 83
 Théorème de Riemann-Lebesgue, 75
 Théorème de Liouville, 115
 Traitement du signal, 48
 Transformée de Fourier de la dérivée d'une fonction, 77
 Transformation de Fourier conjuguée, 71
 Transformation de Fourier discrète, 46
 Transformation de Fourier discrète : inversion, 50
 Transformation de Fourier finie, 38
 Transformation de Fourier finie : inversion, 39
 Transformation de Fourier finie : périodicité, 44
 Transformation de Fourier intégrale, 71
 Transformation de Fourier intégrale : inversion, 78
 Transformation de Laplace unilatérale, 91
 Transformation de Laplace : dérivation, 95
 Transformation de Laplace : holomorphie, 94, 95

Valeur principale de $1/t$, 167
Variable dépendante, 19
Variable indépendante, 19
Vibrations d'une corde, 149
Vibrations d'une poutre, 149

Wronskien, 110, 112, 114