



Formation Interuniversitaire de Physique

Année 2013-2014

Mathématiques pour physiciens

Jean-Bernard Zuber

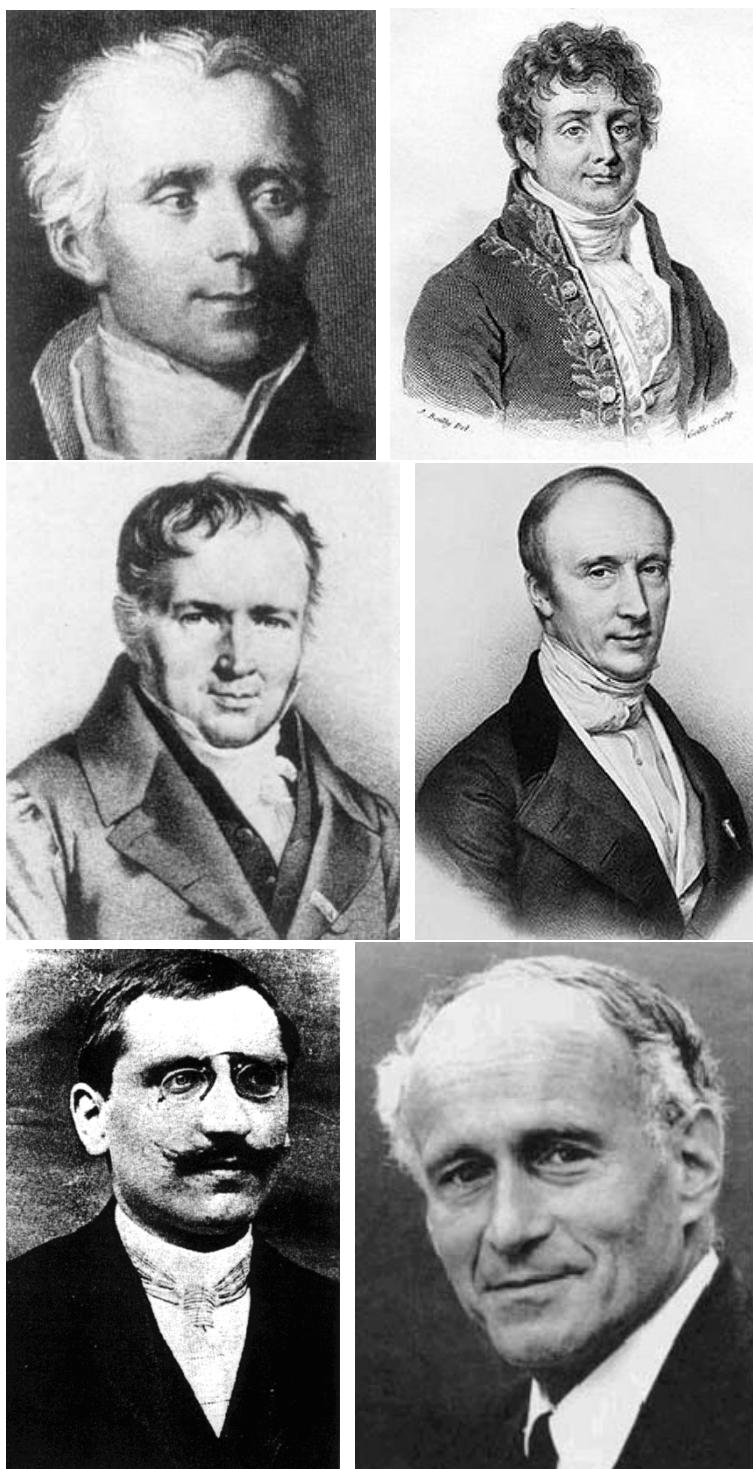


FIGURE 1 –

6 mathématiciens qui ont laissé des contributions fondamentales dans le sujet de ce cours.
 Pierre-Simon Laplace (1749 - 1827) et (Jean-Baptiste) Joseph Fourier (1768 - 1830) ;
 Siméon Denis Poisson (1781 - 1840) et Augustin-Louis Cauchy (1789 - 1857) ;
 Henri Lebesgue (1875 - 1941) et Laurent Schwartz (1915 - 2002)

Bibliographie

- [1] Walter Appel, *Mathématiques pour la physique et les physiciens!*, H& K Éditions
- [2] Claude Aslangul, *Des mathématiques pour les sciences, Cours et Exercices*, de Broeck, 2011
- [3] Claude Aslangul, *Des mathématiques pour les sciences, Exercices corrigés*, de Broeck, 2013
- [4] Henri Cartan, *Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes*, Hermann 1961
- [5] Jacques Gapaillard, *Intégration pour la licence*, Dunod 2002
- [6] John Lamperti, *Probability*, Benjamin 1966
- [7] W. Rudin, *Analyse réelle et complexe*, Masson 1977.
- [8] Laurent Schwartz, *Méthodes mathématiques pour les sciences physiques*, Hermann, 1965.
- [9] Laurent Schwartz, *Théorie des distributions*, Hermann, 1966.
- [10] Laurent Schwartz, *Analyse I. Topologie Générale et Analyse Fonctionnelle*, Hermann, 1991.

Parmi ces ouvrages, certains sont écrits dans un esprit assez proche de celui du présent cours, en particulier [1], dont je me suis beaucoup inspiré. Le lecteur trouvera d'innombrables applications et exercices dans [2] et [3].

Plan du cours

– 1 Rappels, convergence, séries, fonctions	1 cours
– 2 Intégration	2 cours
– 3 Distributions	$1\frac{1}{2}$ cours
– 4 Transformation de Fourier	$1\frac{1}{2}$ cours
– 5 Probabilités	$2\frac{1}{2}$ cours
– 6 Séries entières. Fonctions d'une variable complexe	1 cours
– 7 Fonctions holomorphes. Théorème de Cauchy	$2\frac{1}{2}$ cours
– 8 Fonctions de variables complexes, applications	$1\frac{1}{2}$ cours
– 9 Transformation de Laplace	$1\frac{1}{2}$ cours
– 10 Aspects de la théorie des groupes	1 cours

Table des matières

1 Rappels, convergence, séries, fonctions	1
1.1 Topologie de la droite réelle (rappels)	1
1.1.1 Suites convergentes. Suites de Cauchy	1
1.1.2 Ouverts, fermés. Points d'accumulation	2
1.1.3 Les deux propriétés fondamentales de \mathbb{R}	3
1.1.4 Compléments	3
1.2 Bribes de topologie générale	4
1.2.1 Petit glossaire	4
1.2.2 Espaces vectoriels normés	5
1.2.3 Suites convergentes, suites de Cauchy dans un espace normé	6
1.3 Suites et séries de fonctions. Convergence simple, uniforme, en norme.	7
1.3.1 Convergence d'une suite de fonctions	7
1.3.2 Continuité d'une limite de fonctions continues	8
1.3.3 Dérivabilité d'une limite de fonctions dérivables.	8
1.3.4 Intégrabilité d'une limite de fonctions intégrables.	9
1.3.5 Séries. Séries de fonctions	9
2 Intégration	17
2.1 Intégrale de Riemann	17
2.1.1 Rappels sur l'intégrale de Riemann	17
2.1.2 Intégrales improches	18
2.1.3 Problèmes avec l'intégrale de Riemann	19
2.2 Intégrale de Lebesgue. Mesure	19
2.2.1 Idée intuitive	19
2.2.2 Mesure (Bribes de théorie de la)	21
2.2.3 Retour à l'intégrale de Lebesgue	24
2.2.4 Intégrales de Lebesgue sur \mathbb{R}^2 ou \mathbb{R}^n	26

2.2.5 Espaces \mathcal{L}^p et L^p	27
2.2.6 Comparaison entre intégrales de Riemann et de Lebesgue	28
2.3 Intégrales dépendant d'un paramètre	31
3 Distributions	35
3.1 Introduction	35
3.1.1 Distributions de charges électriques.	35
3.1.2 Diffusion cohérente par un réseau. Peigne de Dirac	36
3.1.3 Choc élastique	37
3.1.4 Autres exemples	38
3.2 Définitions et premières propriétés	38
3.2.1 Espace des fonctions-tests. Définition des distributions.	39
3.3 Opérations sur les distributions	42
3.3.1 Translation, dilatation	42
3.3.2 Dérivation d'une distribution	43
3.4 Distribution delta et distributions reliées	44
3.4.1 Fonction de Heaviside, fonction signe	44
3.4.2 Relations fonctionnelles	44
3.4.3 δ sur une courbe, une surface,	46
3.5 Produit de distributions. Convolution	48
3.6 Exemple : Fonction de Green et potentiel de Coulomb en dimension d	51
4 Transformation de Fourier	55
4.0 Préambule physique. Équation des ondes	55
4.1 Séries de Fourier. Rappels	57
4.2 Transformation de Fourier dans L^1	58
4.2.1 Définition. Conventions	58
4.2.2 Existence et premières propriétés	59
4.2.3 Autres propriétés	61
4.2.4 Transformation de Fourier dans L^2	63
4.2.5 Transformation de Fourier et convolution	65
4.2.6 Diffraction par une fente, par un réseau	66
4.3 Transformées de Fourier des distributions	66
5 Probabilités	71
5.1 Événements. Espace des épreuves.	71
5.2 Probabilités et mesure. Vocabulaire	72

5.2.1	Axiomes de Kolmogorov	72
5.2.2	Probabilité conditionnelle. Événements indépendants	73
5.3	Variables aléatoires. Distributions de v.a.	74
5.3.1	Définition d'une variable aléatoire ; loi de probabilité	74
5.3.2	Les quantités et fonctions importantes attachées à une v.a.	75
5.3.3	Plusieurs variables aléatoires	80
5.3.4	Changement de variable aléatoire	83
5.3.5	Fonction caractéristique	84
5.4	Distributions classiques	85
5.4.1	Distribution uniforme	85
5.4.2	Distribution binomiale	87
5.4.3	Distribution normale	88
5.4.4	Distribution de Poisson	91
5.4.5	Quelques autres lois rencontrées dans les sciences naturelles	93
5.4.6	Limites de la loi binomiale	94
5.4.7	Quelques exemples concrets	96
5.5	Théorèmes limites	100
5.5.1	Convergence presque sûrement, convergence en probabilité, convergence en loi	100
5.5.2	Loi des grands nombres	101
5.5.3	Théorème limite central	103
5.5.4	Marche aléatoire.	108
6	Séries entières. Fonctions analytiques	115
6.1	Séries entières	115
6.1.1	Séries formelles	115
6.1.2	Rayon de convergence	116
6.1.3	Continuité, intégrabilité et dérivabilité de la somme	118
6.1.4	Vitesse de convergence	119
6.2	Les fonctions exp et log	119
6.3	Séries de Taylor. Fonctions analytiques	121
6.3.1	Séries de Taylor	121
6.3.2	Fonctions analytiques	121
6.3.3	Prolongement analytique	123
6.4	Lacunes de ce chapitre.	125
6.4.1	Séries divergentes. Séries asymptotiques	125

6.4.2	Produits infinis	126
7	Fonctions holomorphes. Théorème de Cauchy	127
7.1	Fonctions holomorphes	127
7.1.1	Rappel : fonctions différentiables de \mathbb{R}^n dans \mathbb{R}^p	127
7.1.2	Holomorphie. Conditions de Cauchy–Riemann	128
7.1.3	Dérivations $\partial, \bar{\partial}$	130
7.2	Intégrales sur des chemins	131
7.2.1	Chemins et lacets	131
7.2.2	Intégrales de formes sur des chemins.	133
7.2.3	Intégration sur des chemins dans \mathbb{C}	137
7.3	Théorème de Cauchy et conséquences	139
7.3.1	Théorème de Cauchy	139
7.3.2	Intégrale de Cauchy	140
7.4	Autres propriétés des fonctions holomorphes	141
7.4.1	Fonctions holomorphes et fonctions analytiques	141
7.4.2	Fonctions entières. Principe du maximum	142
7.5	Zéros et singularités des fonctions de variable complexe	143
7.5.1	Zéros d'une fonction holomorphe	143
7.5.2	Pôles, singularités essentielles	144
7.5.3	Séries de Laurent. Fonctions méromorphes. Résidus	145
8	Fonctions de variables complexes, applications	153
8.1	Calcul d'intégrales, de transformées de Fourier, de sommes etc	153
8.1.1	Calcul pratique des résidus	153
8.1.2	Lemmes de Jordan	153
8.1.3	Intégrales sur l'axe réel	154
8.1.4	Intégrales de fractions rationnelles trigonométriques	155
8.1.5	Transformées de Fourier	155
8.1.6	Sommes infinies	156
8.1.7	Intégrales sur un arc. Pôles sur l'axe réel	157
8.2	Fonctions multivaluées	158
8.2.1	Points de branchement, feuillets, surface de Riemann	158
8.2.2	Intégrales de Cauchy de fonctions multivaluées	161
8.3	Fonctions harmoniques	162
8.4	Méthode du col	163

8.4.1	Méthode du col	164
8.4.2	Commentaires et variantes	165
8.4.3	Fonction Γ . Formule de Stirling	166
9	Transformation de Laplace	169
9.1	Définitions et premières propriétés	169
9.1.1	Abscisse de sommabilité et transformée de Laplace	169
9.1.2	Holomorphie de \hat{f} , etc	171
9.1.3	Exemples	172
9.2	Inversion, dérivation, convolution etc	173
9.2.1	Inversion de la transformation de Laplace	173
9.2.2	Translation	174
9.2.3	Convolution	174
9.2.4	Opérations de dérivation et intégration	175
9.2.5	Autres exemples	176
9.3	Transformée de Laplace des distributions	176
9.4	Applications de la transformée de Laplace	177
9.4.1	Équations différentielles, problème de Cauchy	177
9.4.2	Exemple : Circuit LRC	179
9.4.3	Équations linéaires aux dérivées partielles	181
10	Aspects de la théorie des groupes	183
10.1	Représentations linéaires des groupes	184
10.1.1	Définition	184
10.1.2	Représentations réductibles ou irréductibles	184
10.1.3	Lemme de Schur	185
10.2	Groupes et algèbres de Lie	185
10.2.1	Groupes R , $U(1)$ et $SO(3)$. Algèbre de Lie	185
10.2.2	Groupe de rotation à deux dimensions $SO(2)$	186
10.2.3	Générateurs infinitésimaux de $SO(3)$	188
10.3	Algèbre de Lie de $SO(3)$ et ses représentations	190
10.3.1	Algèbre de Lie $so(3)$	190
10.3.2	Représentations de l'algèbre $so(3)$ et représentations du groupe $SO(3)$	192

Chapitre 1

Rappels, convergence, séries, fonctions

1.1 Topologie de la droite réelle (rappels)

1.1.1 Suites convergentes. Suites de Cauchy

On rappelle les définitions d'une suite convergente u_n dans \mathbb{R} :

$$u_n \rightarrow \ell \in \mathbb{R} \iff \forall \epsilon \exists N \text{ tel que } \forall n > N \quad |u_n - \ell| < \epsilon$$

(c'est-à-dire **tous** les u_n sont arbitrairement près de ℓ , dès que n est suffisamment grand) et d'une suite de Cauchy $u_n \in \mathbb{R}$:

$$\forall \epsilon, \exists N \text{ tel que } \forall n, n' > N \quad |u_n - u_{n'}| < \epsilon.$$

Toute suite convergente est de Cauchy (par l'inégalité triangulaire), mais la réciproque n'est pas évidente. Ainsi elle n'est en général pas vraie dans \mathbb{Q} : une suite de Cauchy de rationnels ne converge pas toujours dans \mathbb{Q} . Par exemple, la suite dans \mathbb{Q} définie par la *fraction continue*

$$u_n = 2 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \cdots}}}$$

tronquée au n -ième dénominateur, ou encore définie par la relation de récurrence $u_0 = 2$, $u_n = 2 + \frac{1}{u_{n-1}}$, est une suite de Cauchy (le vérifier), mais ne converge pas dans \mathbb{Q} (puisque elle converge vers $1 + \sqrt{2} \notin \mathbb{Q}$) : on dit que \mathbb{Q} n'est pas *complet*. Au contraire, toute suite de Cauchy de réels converge dans \mathbb{R} par construction, voir plus bas. \mathbb{R} est complet : c'est cette non-convergence de certaines suites de Cauchy dans \mathbb{Q} qu'on corrige par l'introduction des nombres réels.

Théorème 1.1 : (a) Toute suite de Cauchy est bornée : $\exists M \forall n |u_n| < M$.

(b) Toute suite de Cauchy u_n admettant une sous-suite convergeant vers ℓ converge elle-même vers ℓ .

Exercice : le démontrer. (Noter que ces résultats s'appliquent à des suites dans \mathbb{R} ou \mathbb{Q}).

• Rappel : Définition de \mathbb{R} par suites de Cauchy de rationnels équivalentes

Rappelons la construction des nombres réels par les classes d'équivalence de suites de Cauchy de rationnels. On considère l'ensemble des suites u de Cauchy ($u_n \in \mathbb{Q}$) et on dit que $u \sim v$ si $u_n - v_n \rightarrow 0$. On vérifie aisément qu'il s'agit bien d'une relation d'équivalence et on considère alors l'ensemble des classes d'équivalence. On montre qu'on peut le doter des lois d'addition et de multiplication (par addition et multiplication des suites et compatibilité avec la relation \sim), qu'il forme un corps, qu'il contient un sous-corps isomorphe à \mathbb{Q} , etc. Au final l'ensemble \mathbb{R} est défini comme l'ensemble de ces classes d'équivalence. Une propriété utile est que deux suites *adjacentes* (u_n) et (v_n), c'est-à-dire l'une croissante, l'autre décroissante, avec $|u_n - v_n| \rightarrow 0$, convergent vers un même réel x .

1.1.2 Ouverts, fermés. Points d'accumulation

Sur \mathbb{R} on définit les intervalles

- ouverts, $]a, b[: x \in]a, b[\Leftrightarrow a < x < b$, avec éventuellement $a = -\infty$ et/ou $b = +\infty$
- fermés, $[a, b] : x \in [a, b] \Leftrightarrow a \leq x \leq b$, a et b finis.

Un sous-ensemble *borné* E de \mathbb{R} est un ensemble à la fois majoré et minoré :

$$\exists m, M \quad \forall x \in E \quad m \leq x \leq M .$$

La *borne supérieure* B (resp. inférieure b) d'un ensemble borné E est par définition le plus petit majorant (resp. le plus grand minorant) de E . Ce nombre B (resp. b) existe et est unique. (La preuve, classique mais un peu longue, procède par dichotomie ; elle ne sera pas reproduite ici.)

Soit $E \subset \mathbb{R}$ un sous-ensemble (de cardinal¹) infini de \mathbb{R} . Par définition, ξ est un *point d'accumulation* de E s'il existe des points de E arbitrairement proches de ξ mais distincts de ξ , autrement dit

$$\forall \epsilon \exists x \in E, x \neq \xi \text{ tel que } x \in]\xi - \epsilon, \xi + \epsilon[. \quad (1.1)$$

Le point d'accumulation ξ peut appartenir ou non à l'ensemble E .

Théorème 1.2 (Bolzano–Weierstrass) : Tout ensemble $E \subset \mathbb{R}$ infini borné possède au moins un point d'accumulation.

1. la définition de *cardinal* est rappelée à l'Appendice A.1

Ou de façon équivalente, toute suite (u_n) dans E admet une sous-suite convergeant dans E .

Esquisse de preuve : Par dichotomie : si $E \subset [b_1, B_1]$, on coupe $[b_1, B_1]$ en deux moitiés, dont l'une au moins, notons-la $[b_2, B_2]$, a une intersection infinie avec E , et on itère ; cela définit une suite d'intervalles $[b_n, B_n]$ emboîtés, dont les extrémités forment deux suites adjacentes (cf ci-dessus) définissant un point ξ . On montre aisément que ξ est point d'accumulation de E .

On définit alors les sous-ensembles ouverts ou fermés de \mathbb{R} :

Un sous-ensemble $O \subset \mathbb{R}$ est *ouvert* si tout point de O est le centre d'un *intervalle* ouvert entièrement contenu dans O .

Un ensemble fermé $F \subset \mathbb{R}$ est un ensemble qui contient tous ses points d'accumulation. Un intervalle fermé $[a, b]$ est clairement un ensemble fermé.

Le complémentaire dans \mathbb{R} d'un ouvert est un fermé (exercice : le vérifier).

L'union d'un nombre quelconque d'ouverts est un ouvert. L'union d'un nombre fini de fermés est un fermé. En revanche, une union dénombrable² de fermés peut ne pas être fermée, par exemple $\cup_{n \in \mathbb{N}^*} [\frac{1}{n}, 1] =]0, 1]$.

Par passage au complémentaire : l'intersection d'un nombre quelconque de fermés est un fermé. L'intersection d'un nombre fini d'ouverts est un ouvert. En revanche, une intersection dénombrable d'ouverts peut ne pas être ouverte, par exemple $\cap_{n \in \mathbb{N}^*} [-\frac{1}{n}, \frac{1}{n}] = \{0\} = [0, 0]$.

1.1.3 Les deux propriétés fondamentales de \mathbb{R}

Théorème 1.3 : \mathbb{Q} est dense dans \mathbb{R} , c'est-à-dire tout réel est limite de nombre rationnels.

Théorème 1.4 : \mathbb{R} est complet : toute suite de Cauchy y converge.

Le théorème 1.3 découle de la construction de \mathbb{R} que nous avons rappelée, comme ensemble des classes d'équivalence de suites de Cauchy dans \mathbb{Q} : pour tout réel x , il existe une suite de Cauchy de rationnels qui converge vers x , ce qui établit la propriété de densité.

Pour le théorème 1.4, soit une suite u_n de Cauchy de nombres réels. Par le théorème 1.3, pour tout u_p , il existe un rationnel r_p , $|u_p - r_p| < \epsilon$. Une nouvelle fois, l'inégalité triangulaire vient à la rescousse et nous dit que l'on peut rendre $|r_p - r_q| \leq |r_p - u_p| + |u_p - u_q| + |u_q - r_q| < 3\epsilon$ pour p et q assez grands, donc r_n est une suite de Cauchy de rationnels qui définit un nombre réel x , r_n converge vers x et $|u_n - x| \leq |u_n - r_n| + |r_n - x| < 2\epsilon$, donc $\lim u_n = x$.

1.1.4 Compléments

On considère aussi parfois la droite “achevée” $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$, voir App. A.3.

Autres notations : $\mathbb{R}^+ = \{x \in \mathbb{R}; x \geq 0\}$, $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$, etc.

2. Voir la définition 1.8 ci-dessous à l'Appendice A.

1.2 Bribes de topologie générale

L'étude de la topologie de la droite réelle nous amène tout naturellement à des généralisations et à des définitions. Une introduction très complète à la topologie peut être trouvée dans [10].

1.2.1 Petit glossaire

Espace topologique (e.t.) : ensemble E doté d'une collection de sous-ensembles O_i , dits *ouverts*, avec la propriété que toute union d'ouverts ou toute intersection d'un nombre fini d'ouverts est un ouvert, et que E et \emptyset sont des ouverts. Soit \mathcal{O} l'ensemble des ouverts O_i . C'est en fait le couple (E, \mathcal{O}) qui constitue l'e.t. Tout sous-ensemble $F \subset E$ est un e.t. (un sous-e.t. de E) pour la topologie définie par l'ensemble d'ouverts $\{O_i \cap F\}$ (noté abusivement $\mathcal{O} \cap F$).

Exemple $E = \mathbb{R}$, intervalles ouverts $]a, b[$ et leurs unions ("topologie usuelle" de \mathbb{R});

Fermé de E : complémentaire d'un ouvert dans E .

Exemple $E = \mathbb{R}$, fermés $[a, b]$ et leurs unions finies.

Intérieur $\overset{\circ}{A}$ de $A \subset E$: plus grand ouvert contenu dans A .

Adhérence (ou *fermeture*) \bar{A} de A : plus petit fermé contenant A .

Exemple $A = [a, b[, \overset{\circ}{A} =]a, b[, \bar{A} = [a, b]$.

Voisinage d'un point x : sous-ensemble de E contenant un ouvert contenant x . On notera $\mathcal{V}(x)$ l'ensemble des voisinages de x . Exemple $E = \mathbb{R}$, tout $V =]x - \epsilon_1, x + \epsilon_2[$ est un voisinage de x , mais aussi tout $[x - \epsilon_1, x + \epsilon_2]$.

Un espace topologique est *séparé* (ou de Hausdorff) si deux points distincts possèdent deux voisinages disjoints.

Base de voisinages $\mathcal{B}(x)$ d'un point x : sous-ensemble de $\mathcal{V}(x)$ tel que tout $V \in \mathcal{V}(x)$ contient un $W \in \mathcal{B}(x)$. (Intuitivement, une base $\mathcal{B}(x)$ est constituée de "suffisamment" de voisinages de x .)

Exemple $E = \mathbb{R}$, $\mathcal{B}(x) = \{]x - \frac{1}{n}, x + \frac{1}{n}[\}_{n \in \mathbb{N}}$

Continuité : Une fonction f d'un e.t. E dans un e.t. F est continue si pour tout ouvert $O \subset F$, l'ensemble $f^{-1}(O) = \{x \in E \text{ tel que } f(x) \in O\}$ est ouvert dans E .

Exemple $E = F = \mathbb{R}$, on retrouve bien la définition usuelle de la continuité en a :

$\forall \epsilon \exists \eta |x - a| < \eta \text{ (c'est-à-dire } x \in]a - \eta, a + \eta[) \Rightarrow |f(x) - f(a)| < \epsilon \text{ (c'est-à-dire } f(x) \in]f(a) - \epsilon, f(a) + \epsilon[)$.

Voir TD.

Densité : Un sous-ensemble X d'un e.t. E est *dense dans* E si $\forall a \in E$, tout voisinage de a a une intersection non nulle avec X .

Exemple : \mathbb{Q} est dense dans \mathbb{R} : on peut approcher arbitrairement près tout réel par un rationnel.

Connexité : Un e.t. E est connexe s'il ne peut pas être écrit comme union disjointe de deux ouverts non vides. Ou de façon équivalente (vérifier !), si les seuls de ses sous-ensembles à la

fois ouverts et fermés sont E et \emptyset .

Espace compact E : espace topologique (séparé) tel que de tout recouvrement de E par des ouverts, on peut extraire un recouvrement fini.

Conséquences : si E est compact,

- toute suite infinie de points de E admet un point d'accumulation (généralisation du théorème de Bolzano–Weierstrass) ;
- si $f : E \rightarrow F$ est continue, $f(E)$ est compact ;
- toute fonction réelle continue sur E est bornée.

Si E est un sous-espace de \mathbb{R}^n , E compact $\Leftrightarrow E$ borné et fermé (théorème de Heine–Borel).

Espace localement compact : espace topologique (séparé) dont tout point a au moins un voisinage compact.

Exemples : \mathbb{R} n'est pas compact mais localement compact ; \mathbb{Q} n'est ni compact ni localement compact. On peut “compactifier” \mathbb{R} en lui ajoutant un point à l'infini ; et de même, on peut compactifier \mathbb{C} (le plan complexe) en lui ajoutant un point à l'infini, le transformant en la “sphère de Riemann”, voir plus bas au Chap. 7, § 7.5.3.

1.2.2 Espaces vectoriels normés

Un espace vectoriel E sur \mathbb{R} ou \mathbb{C} (de dimension finie ou infinie) est dit normé si on peut y définir une norme.

Définition 1.1 : Norme :

$$\begin{aligned} \|x\| &\geq 0 \\ \|x\| &= 0 \Rightarrow x = 0 \\ \|\lambda x\| &= |\lambda| \|x\| \\ \|x + y\| &\leq \|x\| + \|y\| \quad \text{inégalité triangulaire} \end{aligned} \tag{1.2}$$

Dans la définition d'une *semi-norme*, on relâche le deuxième axiome, $\|x\| = 0$ n'implique pas $x = 0$.

Distance. Dans un e.v. normé E , on peut définir la distance $d(x, y) = \|x - y\|$.

Exemple : les espaces \mathbb{R}^n (ou \mathbb{C} , considéré comme un e.v. de dimension 2 sur les réels) sont des espaces normés par la norme euclidienne $\|x\| = (\sum_{i=1}^n |x_i|^2)^{\frac{1}{2}}$, où les x_i sont les composantes de x dans une base orthonormée. Dans \mathbb{C} , cette norme est le module du nombre complexe.

- **Espaces de fonctions normés**

Sur l'espace $C([a, b])$ des fonctions continues sur $[a, b]$ à valeurs dans \mathbb{R} ou \mathbb{C} , on peut définir des normes

$$\begin{aligned}\|f\|_1 &= \int_a^b |f(x)| dx \\ \|f\|_2 &= \left(\int_a^b |f(x)|^2 dx \right)^{\frac{1}{2}} \\ \|f\|_\infty &= \sup_{x \in [a, b]} |f(x)|\end{aligned}\tag{1.3}$$

Sur l'espace des fonctions continues par morceaux (avec un nombre fini de discontinuités sur $[a, b]$), $\|f\|_1$ et $\|f\|_2$ sont des semi-normes, $\|f\|_\infty$ est une norme, pourquoi ?

1.2.3 Suites convergentes, suites de Cauchy dans un espace normé

Les définitions de suite convergente ou de Cauchy du § 1.1.1 s'étendent à tout espace normé. Soit E un espace vectoriel normé et une suite $u_n \in E$.

Définition 1.2 : La suite (u_n) admet une limite $u \in E$, ou converge vers la limite u , et on écrit $\lim_{n \rightarrow \infty} u_n = u$ ou simplement $u_n \rightarrow u$, si la différence $u_n - u$ tend vers zéro en norme quand n tend vers l'infini

$$\forall \epsilon > 0 \quad \exists N \text{ tel que } \forall n > N \quad \|u_n - u\| < \epsilon.\tag{1.4}$$

Définition 1.3 : Une suite u_n est dite suite de Cauchy (ou elle satisfait le critère de Cauchy) si

$$\forall \epsilon \quad \exists N \text{ tel que } \forall n, n' > N \quad \|u_n - u_{n'}\| < \epsilon.\tag{1.5}$$

Comme précédemment, toute suite convergente est de Cauchy. La réciproque n'est en général pas vraie comme on a vu, une suite de Cauchy peut ne pas converger.

Définition 1.4 : Un espace vectoriel normé E est un espace complet ssi toute suite de Cauchy y converge. Un tel espace est appelé espace de Banach.

Cette notion prend son importance dans l'étude des espaces de fonctions, comme on va le voir. En dimension finie, un corollaire simple du Théorème 1.4 est que

\mathbb{C} ou tout espace vectoriel \mathbb{R}^n de dimension finie sur \mathbb{R} est complet.

Preuve : Il suffit de prendre la norme euclidienne : la condition de Cauchy dans \mathbb{R}^n implique que chaque composante est de Cauchy dans \mathbb{R} donc converge.

1.3 Suites et séries de fonctions. Convergence simple, uniforme, en norme.

1.3.1 Convergence d'une suite de fonctions

On considère l'espace des fonctions à valeurs dans \mathbb{R} ou dans \mathbb{C} , définies sur un sous-ensemble X de \mathbb{R} ou \mathbb{C} , et bornées. Cet espace est muni de la norme $\|f\|_\infty = \sup_{x \in X} |f(x)|$, voir ci-dessus. Soit f_n une suite de telles fonctions (toutes définies sur le même X).

On peut définir deux notions de convergence $f_n \rightarrow f$.

Définition 1.5 : La suite f_n converge simplement vers f si $\forall x \in X$, $f_n(x)$ converge vers $f(x)$.

$$\forall \epsilon > 0 \quad \forall x \in X \quad \exists N \text{ tel que } \forall n > N \quad |f_n(x) - f(x)| < \epsilon \quad (1.6)$$

et on note $f_n \xrightarrow{\text{CV.S.}} f$.

Définition 1.6 : La suite f_n converge uniformément vers f si

$$\forall \epsilon > 0 \quad \exists N \text{ tel que } \forall x \in X \quad \forall n > N \quad |f_n(x) - f(x)| < \epsilon \quad (1.7)$$

ou encore

$$\forall \epsilon > 0 \quad \exists N \text{ tel que } \forall n > N \quad \|f_n - f\|_\infty < \epsilon$$

et on note $f_n \xrightarrow{\text{CV.U.}} f$.

Bien comprendre que dans le cas de la CVU, le nombre N est indépendant de x , alors qu'a priori, il en dépend dans la CVS. La convergence uniforme implique la convergence simple (le vérifier), mais la réciproque n'est pas vraie.

La définition de convergence CVS, CVU s'étend sans difficulté à des suites de fonctions à valeurs dans tout espace normé : il suffit de remplacer dans (1.6,1.7) la valeur absolue par la norme $\|\cdot\|$.

Exemples et contre-exemples.

Exemples : 1. Soit la suite de fonctions $f_n(x) = x^n$ définies sur l'intervalle $[0, 1]$. Pour tout $x < 1$, $x^n \rightarrow 0$, tandis que pour $x = 1$, la valeur limite est 1. La fonction limite $f(x)$ est donc la fonction discontinue qui vaut 0 sur $[0, 1[$ et 1 en 1 (c'est la fonction partie entière $E(x)$, restreinte à l'intervalle $[0, 1]$). La convergence des f_n n'est pas uniforme vers $f(x) = E(x)$, comme on va le montrer plus bas, par l'absurde.

Mais il est instructif de le comprendre directement. Montrons que $\forall a$ fixé, avec $0 \leq a < 1$, la suite $f_n(x) = x^n$ converge uniformément vers 0 sur l'intervalle $[0, a]$. En effet

$$\forall n > N \quad \forall x \in [0, a] \quad |x^n - 0| = x^n \leq a^n < a^N,$$

par des inégalités triviales. Pour assurer que $|x^n - 0| < \epsilon$, il suffit de choisir N tel que $a^N < \epsilon$ ou encore de façon équivalente, $N > \frac{-\ln \epsilon}{-\ln a}$. Ce choix de N assure bien la convergence uniforme des f_n vers 0 dans l'intervalle $[0, a]$. Mais noter que ce N augmente sans limite quand a s'approche de 1. En conséquence, la convergence uniforme ne peut être maintenue sur tout l'intervalle $[0, 1[$, même ouvert à droite.

2. Soit la suite de fonctions $f_n(x) = \begin{cases} (n-1)x & \text{si } x \in [0, \frac{1}{n}] \\ 1-x & \text{si } x \in [\frac{1}{n}, 1] \end{cases}$. Pour $x = 0$, $f_n(0) = 0$ donc

$\lim f_n(0) = 0$. Pour tout $x \neq 0$, il existe un n assez grand à partir duquel tous les $f_n(x) = 1 - x$. La suite converge donc simplement vers la fonction discontinue $f(0) = 0$; $f(x) = 1 - x$, $x \neq 0$, mais la convergence n'est pas uniforme.

3. En revanche les fonctions $f_n(x) = n \sin \frac{x}{n}$ convergent uniformément vers $f(x) = x$ sur l'intervalle $[0, 1]$.

En effet $|f_n(x) - x| = x - n \sin \frac{x}{n} \leq \frac{x^3}{6n^2} < \frac{1}{6n^2}$ quel que soit $x \in [0, 1]$. On peut démontrer l'inégalité $\phi(\alpha) \stackrel{\text{déf}}{=} \sin \alpha - \alpha + \frac{\alpha^3}{6} \geq 0$ utilisée dans cet argument par étude des fonctions dérivées successives $\phi'(\alpha)$ et $\phi''(\alpha)$, ou encore par utilisation de la formule de Taylor-Lagrange.

4. Autres exemples importants, les fonctions $f_n(x) \stackrel{\text{déf}}{=} \sum_{p=0}^n \frac{x^p}{p!}$, et $g_n(x) := (1 + \frac{x}{n})^n$, convergent, pour tout x réel, vers la fonction exponentielle e^x . La convergence est-elle uniforme ?

1.3.2 Continuité d'une limite de fonctions continues

Théorème 1.5 : Une suite uniformément convergente de fonctions continues définies dans un domaine D de \mathbb{R} ou \mathbb{C} et à valeurs dans un espace normé E a pour limite une fonction continue.

Preuve. Prenons $E = \mathbb{R}$ pour simplifier les notations. Soit f la limite uniforme d'une suite de fonctions f_n continues sur $[a, b]$. Montrons la continuité de f en tout $c \in D$. Par la convergence uniforme,

$$\forall \epsilon \exists N \text{ tel que } \forall n > N \quad \forall x \in D \quad |f_n(x) - f(x)| < \frac{\epsilon}{3}.$$

Choisissons un tel $n > N$. La continuité de f_n dit que

$$\forall \epsilon \exists \eta \text{ tel que } \forall x : |c - x| < \eta \quad |f_n(x) - f_n(c)| < \frac{\epsilon}{3}$$

et l'inégalité triangulaire dit alors que

$$|f(x) - f(c)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(c)| + |f_n(c) - f(c)| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon, \quad \text{CQFD.}$$

Voir ci-dessus deux exemples de suites non uniformément convergentes de fonctions continues dont la limite n'est pas continue.

1.3.3 Dérivabilité d'une limite de fonctions dérивables.

Soit une famille de fonctions f_n définies dans un intervalle I de \mathbb{R} à valeurs dans \mathbb{R} ou \mathbb{C} , ce qu'on notera dans la suite $f_n : I \rightarrow \mathbb{R}$ (ou \mathbb{C}). On suppose les f_n dérивables.

Théorème 1.6 : Si les $f_n \xrightarrow{\text{CV.S.}} f$ et les $f'_n \xrightarrow{\text{CV.U.}} g$, alors f est dérivable et $f' = g$.

Contre-exemple (voir [1] p19, 246) : $f_n(x) = \frac{8}{\pi} \sum_{k=1}^n \frac{\sin^2 nx}{4n^2 - 1}$ dont on montre qu'elles convergent uniformément vers $f(x) = |\sin x|$. Les f'_n ne convergent pas uniformément et la fonction f n'est pas dérivable en 0.

Remarque. Ces notions de convergence absolue et de continuité ou dérivabilité associées ne sont pas qu'un problème académique pour le physicien. En Mécanique Statistique, la fonction de partition d'un système à N degrés de liberté $Z = \sum_C \exp -\beta E(C)$ est la somme de poids de Boltzmann sur un nombre \mathcal{N} de configurations, grand mais fini. Cette somme d'exponentielles est une fonction infiniment différentiable (et même "analytique", voir Chap. 6) du paramètre $\beta (= 1/k_B T$, k_B la constante de Boltzmann), donc de la température T . Mais dans la "limite thermodynamique" $\mathcal{N} \rightarrow \infty$, la convergence n'est en général pas uniforme, ce qui peut occasionner l'apparition de discontinuités ou de singularités associées à des transitions de phase. L'étude de ces singularités dans les "phénomènes critiques" est un chapitre important de la physique contemporaine.

1.3.4 Intégrabilité d'une limite de fonctions intégrables.

Théorème 1.7 : Soit une suite de fonctions f_n définies et intégrables (au sens de Riemann) sur un même compact $K \subset \mathbb{R}$, convergeant uniformément vers f sur K . Alors f est intégrable (au sens de Riemann) sur K et

$$\lim_{n \rightarrow \infty} \int_K f_n(x) dx = \int_K f(x) dx \quad (1.8)$$

Autrement dit, sous l'hypothèse de convergence uniforme, "intégration et limite commutent". On donnera une preuve de ce théorème au chapitre 2. Mais on peut amoindrir l'hypothèse de convergence uniforme ("théorème de convergence dominée"), voir Chap. 2.

1.3.5 Séries. Séries de fonctions

Le problème de la convergence d'une série "numérique" $\sum u_k$, où $u_k \in \mathbb{R}$ ou \mathbb{C} , relève de la discussion de la convergence des suites : on s'intéresse à la convergence de la suite des *sommes partielles* $s_n = \sum_{k=1}^n u_k$. On se rappelle que la convergence absolue est une condition suffisante de convergence $\sum |u_k|$ converge $\Rightarrow \sum u_k$ converge. Cela s'étend à des suites dans un e.v. normé **complet** : $\sum \|u_k\|$ converge $\Rightarrow \sum u_k$ converge : on parle de convergence en norme (ou convergence normale) CVN.

Preuve : on utilise l'inégalité triangulaire pour écrire

CVN : $\tilde{s}_n = \sum_1^n \|u_k\|$ converge $\Rightarrow \tilde{s}_n$ est de Cauchy : pour n, n' assez grands $\sum_n^{n'} \|u_k\| < \epsilon \Rightarrow \|\sum_n^{n'} u_k\| \leq \sum_n^{n'} \|u_k\| < \epsilon$ donc $s_n = \sum_1^n u_k$ est de Cauchy donc converge dans l'espace complet.

De même pour des séries de fonctions, on peut définir la CVS ou la CVU comme on l'a fait au § 1.3.1, et aussi une convergence normale CVN dans la norme $\|\cdot\|_\infty$ du sup. Soit une série $\sum f_k$ de fonctions de X ($\subset \mathbb{R}$ ou \mathbb{C}) dans un e.v. normé E .

Définition 1.7 : La série $\sum f_k$ converge

$$\begin{aligned} &\text{simplement vers } F \text{ si } \forall \epsilon \ \forall x \in X \ \exists N \text{ tel que } \forall n > N \quad \left\| \sum_{k=1}^n f_k(x) - F(x) \right\| < \epsilon \\ &\text{uniformément vers } F \text{ si } \forall \epsilon \ \exists N \text{ tel que } \forall x \in X \ \forall n > N \quad \left\| \sum_{k=1}^n f_k(x) - F(x) \right\| < \epsilon \\ &\text{c'est-à-dire} \quad \lim_{n \rightarrow \infty} \left\| \sum_{k=1}^n f_k - F \right\|_\infty = 0 \\ &\text{normalement si} \quad \sum \|f_n\|_\infty \text{ converge .} \end{aligned}$$

N.B. Ne pas confondre la norme $\|\cdot\|$ (sans indice) de l'espace normé E et la norme $\|\cdot\|_\infty$ de l'espace des fonctions dans E .

Théorème 1.8 : Toute série normalement convergente à valeurs dans un e.v. normé complet est uniformément convergente, donc simplement convergente.

Donc CVN \Rightarrow CVU \Rightarrow CVS. (Exercice : le vérifier)

On est alors en mesure d'appliquer à la somme $F(x)$ les théorèmes sur la continuité, la dérivabilité ou l'intégrabilité vus plus haut.

On rencontre souvent en Physique deux types de séries de fonctions : les séries entières $\sum a_k x^k$ et les séries de Fourier $\sum a_k e^{ikx}$. On y reviendra plus bas aux Chapitres 6 et 4, respectivement.

Appendix A. Notations et rappels

• A.1. Notations sur les ensembles

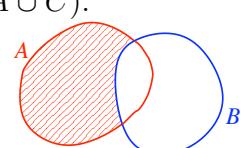
Complémentaire de A dans E : $\complement_E A$ noté aussi \bar{A} quand il n'y a pas d'ambiguïté sur E .

Identités classiques : $\complement_E(A \cap B) = \complement_E A \cup \complement_E B$ et $\complement_E(A \cup B) = \complement_E A \cap \complement_E B$;

distributivité $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ et $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Si $B \subset A$, $A \setminus B \stackrel{\text{déf}}{=} \complement_A B = \{x | x \in A, x \notin B\}$. Plus généralement, si

$A, B \subset E$, $A \setminus B \stackrel{\text{déf}}{=} A \cap \complement_E B = \complement_A(A \cap B)$, voir figure ci-contre.



Exercice : si $A, B, C \subset E$ et la barre désignant le complémentaire dans E , montrer que $A \setminus (B \setminus C) = (A \cap \bar{B}) \cup (A \cap C)$.

Ensemble des parties (= sous-ensembles) de E : $\mathfrak{P}(E) = \{A | A \subset E\}$.

Fonction indicatrice χ_A d'un sous-ensemble $A \subset E$: fonction $E \rightarrow \{0, 1\}$

$$\chi_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \in E \setminus A \end{cases} \quad (\text{A.1})$$

Exercice : démontrer que l'ensemble des points de discontinuité de χ_A pour $A \subset E$ e.t. est la "frontière de A ", soit $\bar{A} \setminus \overset{\circ}{A}$.

Pour un ensemble E fini, on appelle $\text{card}(E)$ (cardinal de E) le nombre d'éléments de E . Deux ensembles finis E et F sont en bijection ssi ils ont même cardinal. Cette notion s'étend à des ensembles infinis, voir sous-§ suivant. Pour E fini, $\text{card}(E) = n$ fini, montrer que $\text{card}(\mathfrak{P}(E)) = 2^n$ (Indication : pour chaque $A \in \mathfrak{P}(E)$ on code chaque élément a de E par 1 ou 0 selon que $a \in A$ ou non, autrement dit, $A \leftrightarrow$ ensemble des $\chi_A(a)|_{a \in E}$, qui est dans $\{0, 1\}^n$. On a donc une bijection $\mathfrak{P}(E) \leftrightarrow \{0, 1\}^n$.)

• A.2. Ensembles dénombrables et non dénombrables. Ensemble de Cantor

Définition 1.8 : Un ensemble est dénombrable s'il est en bijection avec \mathbb{N} .

Autrement dit on peut "numéroter" ses éléments. Ou encore, dans l'énumération de ses éléments, on atteint tout élément au bout d'un nombre fini d'opérations. Exemples : \mathbb{N} , bien sûr, est dénombrable ; \mathbb{Z} l'est aussi, il suffit d'énumérer $0, 1, -1, 2, -2, 3, -3, \dots$. \mathbb{Q} est aussi dénombrable, ce qui est moins évident. On écrit tout rationnel positif sous la forme p/q , avec p et q premiers entre eux. On énumère ces fractions positives selon les valeurs croissantes de $p+q$, (en sautant toutes les paires où p et q ne sont pas premiers) $1/1, 1/2, 2/1, 1/3, 3/1, 4/1, 3/2$, etc. Il est clair que tout rationnel > 0 est atteint au bout d'un nombre fini de pas. On recombine ensuite \mathbb{Q}^+ et \mathbb{Q}^- comme dans \mathbb{Z} . Plus généralement, toute union finie d'ensembles dénombrables est dénombrable, par le même type d'argument.

On étend la notion de cardinal à un ensemble dénombrable. Par définition, le cardinal de \mathbb{N} (ou de tout ensemble dénombrable) est noté \aleph_0 (aleph zéro).

Définition 1.9 : Un nombre algébrique est un nombre (réel ou complexe) solution d'une équation algébrique à coefficients entiers $a_0z^n + a_1z^{n-1} + \dots + a_n = 0$, où $a_i \in \mathbb{Z}$.

Exercice. Montrer que l'ensemble des nombres algébriques est dénombrable.

Le point est que tout ensemble n'est pas dénombrable. Ainsi l'ensemble des réels n'est pas dénombrable. Pour s'en convaincre, on procède par l'absurde (argument de Cantor). On considère les réels de l'intervalle $[0, 1[$, décrits par leur développement décimal $0,abcd\dots$. On suppose qu'on a trouvé une énumération de ces réels sous la forme

$$\begin{aligned} x_1 &= 0, a_1 b_1 c_1 \dots \\ x_2 &= 0, a_2 b_2 c_2 \dots \\ x_3 &= 0, a_3 b_3 c_3 \dots \\ &\vdots \end{aligned} \tag{A.2}$$

et on construit alors un nombre $0, \alpha_1 \alpha_2 \alpha_3 \dots$ où $\alpha_1 \neq a_1$, $\alpha_2 \neq b_2$, $\alpha_3 \neq c_3$, Ce nombre n'est pas dans la liste (A.2) puisqu'il n'est pas à la première ligne car $\alpha_1 \neq a_1$, ni à la deuxième car $\alpha_2 \neq b_2$, ni à la troisième etc³. Il y a contradiction, et l'ensemble des réels de $[0, 1[$ n'est pas dénombrable. On dit qu'il a la *puissance du continu* et on note \mathfrak{c} son cardinal. On montre que tout intervalle de \mathbb{R} non vide est en bijection avec $\mathfrak{P}(\mathbb{N}^*)$ et on écrit $\mathfrak{c} = 2^{\aleph_0}$, voir par exemple [5], Appendice A.2.

Plus généralement, un théorème de Cantor affirme que pour tout ensemble A , $\text{card}(A) < \text{card}(\mathfrak{P}(A))$ avec une inégalité stricte. Ainsi $\mathfrak{P}(\mathbb{R})$ a pour cardinal $2^\mathfrak{c} > \mathfrak{c}$, ce qui signifie qu'il ne peut être en bijection avec \mathbb{R} .

Ensemble de Cantor. Il s'agit d'un ensemble remarquable, qui nous fournira des exemples et contre-exemples dans la suite. On le construit de façon récursive à partir de l'intervalle $[0, 1]$. À l'étape 1, on retranche le tiers médian $\left]\frac{1}{3}, \frac{2}{3}\right[$ ce qui laisse l'union de deux segments fermés $[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$. Puis à chaque étape on retranche le tiers médian ouvert de chaque segment. À la limite on obtient l'ensemble de Cantor \mathcal{K} .

Plus précisément soit γ l'opération qui au segment $I = [a, b]$ associe l'union des deux segments $[a, a + \frac{b-a}{3}] \cup [a + 2\frac{b-a}{3}, b]$. Plus généralement pour une union finie de segments disjoints $\cup_{k=1}^n I_k$, on définit $\gamma(\cup_{k=1}^n I_k) = \cup_{k=1}^n \gamma(I_k)$. Alors si $\mathcal{K}_n = \gamma^n([0, 1])$, $\mathcal{K} = \cap_{n=0}^{\infty} \mathcal{K}_n$.

L'ensemble de Cantor n'est pas vide (par exemple il contient tous les nombres $\frac{1}{3^n}$) et a les propriétés remarquables suivantes (que nous admettrons pour la plupart) :

- il est compact (puisque fermé et borné) et n'a que des points d'accumulation (ensemble “parfait”) ;
- il est non dénombrable, il a la puissance du continu (voir [5] Appendice A) ;

3. Il faut noter que dans la notation décimale, on identifie les paires de nombres $0, ab\dots c9999\dots$ avec $c < 9$ et $0, ab\dots (c+1)$. Vérifier que cela n'affecte pas l'argument de Cantor.

– il est d'intérieur vide et de “mesure nulle” : on définira plus précisément cette notion au Chapitre 2, contentons-nous ici de noter que la somme des longueurs des segments à la n -ième étape de construction de \mathcal{K} vaut $(\frac{2}{3})^n$, donc tend vers zéro quand $n \rightarrow \infty$;

- les points de discontinuité de son indicatrice $\chi_{\mathcal{K}}$ forment l'ensemble \mathcal{K} lui-même ;
- il a des propriétés d'auto-similarité (fractal) : $\mathcal{K}_1 \cap \mathcal{K}_2 \equiv \frac{1}{3}\mathcal{K}_1$, etc.

Ces propriétés semblent contradictoires et cet ensemble est assez déroutant !

• A.3. Bornes. Limites supérieure et inférieure

On rappelle que $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ (“droite réelle achevée”), qui est compact. L'intérêt de cette compactification est clair :

- Tout sous-ensemble de $\overline{\mathbb{R}}$ possède une borne inférieure et une borne supérieure.
- Toute suite croissante (resp. décroissante) dans $\overline{\mathbb{R}}$ converge dans $\overline{\mathbb{R}}$:

$$\lim_{n \rightarrow \infty} u_n = \sup_{n \geq 1} u_n, \quad \text{resp.} \quad \underline{\lim}_{n \rightarrow \infty} u_n = \inf_{n \geq 1} u_n. \quad (\text{A.3})$$

Définition 1.10 : On appelle limite supérieure, resp. limite inférieure, d'une suite (u_n) d'éléments de $\overline{\mathbb{R}}$ l'élément de $\overline{\mathbb{R}}$

$$\overline{\lim}_{n \rightarrow \infty} u_n = \inf_{k \geq 1} (\sup_{n \geq k} u_n) = \lim_{k \rightarrow \infty} (\sup_{n \geq k} u_n) \quad (\text{A.4})$$

$$\text{resp.} \quad \underline{\lim}_{n \rightarrow \infty} u_n = \sup_{k \geq 1} (\inf_{n \geq k} u_n) = \lim_{k \rightarrow \infty} (\inf_{n \geq k} u_n). \quad (\text{A.5})$$

Ces limites supérieure et inférieure existent dans $\overline{\mathbb{R}}$ pour toute suite ! En effet la suite $(\sup_{n \geq k} u_n)$ est une suite en k décroissante, donc convergente dans $\overline{\mathbb{R}}$, on peut donc lui appliquer (A.3), et sa limite, notée $\overline{\lim}_{n \rightarrow \infty} u_n$ est $\inf_{k \geq 1} (\sup_{n \geq k} u_n)$, et vice versa pour la limite inférieure $\underline{\lim}_{n \rightarrow \infty} u_n$.

Une caractérisation utile de la limite supérieure d'une suite $u_n > 0$ (pour l'inférieure, changer les signes d'inégalité) est la suivante :

$$\begin{aligned} \ell &= \overline{\lim}_{n \rightarrow \infty} u_n \text{ finie si } \forall \epsilon \exists N \text{ tel que } \forall n > N \quad u_n < \ell + \epsilon \\ &\quad \text{et pour une infinité de } n > N, \quad u_n > \ell - \epsilon; \\ \ell &= \infty \text{ si } \forall M > 0 \text{ il existe une infinité de } u_n > M. \end{aligned} \quad (\text{A.6})$$

Cette notion nous sera utile en particulier dans l'étude de la convergence des séries entières.

• A.4. Fonctions numériques continues

Fonctions “numériques” = fonctions $E \subset \mathbb{R} \rightarrow \mathbb{R}$.

– toute fonction continue $f : [a, b] \rightarrow \mathbb{R}$ (ou encore de $\mathbb{R} \rightarrow \mathbb{R}$ à support compact) est bornée, $m \leq f(x) \leq M$, avec $M = \sup_{x \in [a, b]} f(x)$ resp. $m = \inf_{x \in [a, b]} f(x)$, et prend toutes les valeurs entre m et M . (On rappelle que support = domaine où la fonction est non nulle).

– toute fonction continue $f : [a, b] \rightarrow \mathbb{R}$ est majorée et minorée par des *fonctions en escalier*.

Rappel : une fonction en escalier est par définition une fonction “constante par morceaux”, c'est-à-dire prenant une valeur constante sur des intervalles $]a_i, b_i[$;

– toute fonction continue $f : [a, b] \rightarrow \mathbb{R}$ est limite uniforme d'une suite de fonctions en escalier ;

– toute fonction continue $f : [a, b] \rightarrow \mathbb{R}$ est limite uniforme d'une suite de polynômes (théorème de Weierstrass). ...

Classes de continuité. Théorèmes de Taylor *et al*

Une fonction numérique $E \subset \mathbb{R} \rightarrow \mathbb{R}$ est dite de classe C^n si toutes les dérivées $f^{(k)}$, $k = 1, \dots, n$ existent sur E et sont continues.

Rappelons alors les différentes formes du théorème de Taylor, qui dit que sous des hypothèses adéquates, une fonction est approximée par son développement de Taylor

$$f(x) = \sum_{k=0}^n \frac{(x - x_0)^k}{k!} f^{(k)}(x_0) + \frac{(x - x_0)^n}{n!} R_n(x)$$

avec un “reste” R_n qui peut prendre plusieurs formes :

$$\text{Taylor--Young } \exists \varphi_n(x), \lim_{x \rightarrow x_0} \varphi_n(x) = 0 : f(x) = \sum_{k=0}^n \frac{(x - x_0)^k}{k!} f^{(k)}(x_0) + \frac{(x - x_0)^n}{n!} \varphi_n(x)$$

$$\text{Taylor--Lagrange } \exists \xi, a < \xi < x : f(x) = \sum_{k=0}^n \frac{(x - x_0)^k}{k!} f^{(k)}(x_0) + \frac{(x - x_0)^{n+1}}{(n+1)!} f^{(n+1)}(\xi)$$

$$\text{Taylor avec reste intégral } f(x) = \sum_{k=0}^n \frac{(x - x_0)^k}{k!} f^{(k)}(x_0) + \int_{x_0}^x \frac{(x - t)^n}{n!} f^{(n+1)}(t) dt$$

Hypothèses : Dans tous les cas, on suppose f n fois dérivable sur un intervalle I , $x_0 \in I$, donc f est (au moins) de classe C^{n-1} . Pour les deux dernières formes, on suppose en outre que f est de classe C^n sur I et que $f^{(n+1)}$ existe sur l'intérieur de I . Pour Taylor-Young, l'existence de la n -ième dérivée suffit à assurer que $f(x) - \sum_{k=0}^n \frac{(x - x_0)^k}{k!} f^{(k)}(x_0) = o((x - x_0)^n)$, ce qui est bien le résultat énoncé. Noter que la forme dite de Taylor-Lagrange généralise la formule dite “des accroissements finis” (qui correspond au cas $n = 0$).

Lectures complémentaires

L'ouvrage de L. Schwartz [10] est une mine d'informations sur les questions d'analyse et de topologie.

Chapitre 2

Intégration

2.1 Intégrale de Riemann

2.1.1 Rappels sur l'intégrale de Riemann

Pour une fonction définie et disons, pour l'instant, continue sur un intervalle $[a, b]$, on introduit d'abord une subdivision de l'intervalle $[a, b]$ en N intervalles $I_k = [x_{k-1}, x_k]$, $x_0 = a$, $x_N = b$, d'union $\cup I_k = [a, b]$, voir figure 2.1, on choisit un point $\xi_k \in I_k$, et on construit la somme de Riemann

$$\Sigma^{(R)} = \sum_{k=1}^N (x_k - x_{k-1}) f(\xi_k). \quad (2.1)$$

Si quand $N \rightarrow \infty$ et quand le découpage devient de plus en plus fin : $\sup_k (x_k - x_{k-1}) \rightarrow 0$, la somme $\Sigma^{(R)}$ a une limite indépendante du choix des intervalles I_k et des points ξ_k , cette limite est appelée intégrale de Riemann et notée $\int_a^b f(x) dx$. Son interprétation est claire : on a coupé l'aire sous la courbe de f en tranches *verticales* de plus en plus fines.

On démontre aisément que toute fonction *en escalier* sur $[a, b]$ est intégrable au sens de Riemann ; mais aussi toute fonction *réglée* (limite uniforme de fonctions en escalier de supports contenus dans un même compact K), ce qui inclut les fonctions continues ou les fonctions monotones sur $[a, b]$, etc ; ou plus généralement toute fonction “pas trop discontinue” en un sens qu'on précisera plus bas au § 2.2.6 où on donnera un critère (condition nécessaire et suffisante) d'intégrabilité au sens de Riemann.

Intégrale et primitive

Toute fonction F de dérivée $F' = f$ est appelée *primitive* de f . Le calcul intégral fournit une méthode de calcul des primitives d'une fonction donnée f . Pour toute fonction continue f

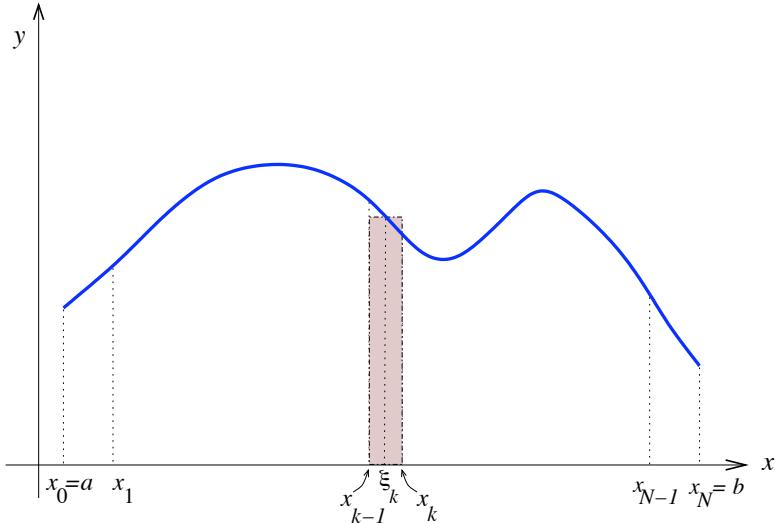


FIGURE 2.1 – Découpage de l'aire dans l'intégration de Riemann

sur un intervalle $[a, b]$, la famille de ses primitives, égales à l'addition d'une constante près, est dénotée par l'*intégrale indéfinie* $\int f(x)dx$, et une primitive quelconque est de la forme

$$F(x) = \int_a^x f(x')dx' + C \quad x \in [a, b], \quad C \text{ constante arbitraire.} \quad (2.2)$$

Il en découle que l'intégrale entre a et b est donnée par la variation d'une primitive quelconque F :

$$\int_a^b f(x)dx = F(b) - F(a). \quad (2.3)$$

2.1.2 Intégrales improches

On étend ensuite cette intégrale de Riemann d'un intervalle compact $[a, b]$ à un intervalle $[a, b[$ avec éventuellement $b = \infty$. L'intégrale “impropre” $\int_a^b f(x)dx$ est par définition la limite, si elle existe, de $\int_a^c f(x)dx$ quand $c \rightarrow b_-$. De même pour la borne inférieure $\int_{-\infty}^b f(x)dx$, etc. Noter que $\int_{-\infty}^{\infty} f(x)dx$ est défini par la limite double $\lim_{\substack{a \rightarrow -\infty \\ b \rightarrow \infty}} \int_a^b f(x)dx$ avec a et b tendant *indépendamment* vers $-\infty$ et ∞ . Rappelons les notions de convergence simple (CVS) et de convergence absolue (CVA) : $\int_a^c f(x)dx$, resp $\int_a^c |f(x)|dx$, converge quand $c \rightarrow b_-$, la CVA impliquant la CVS mais pas l'inverse ; on parle d'intégrale semi-convergente si on a CVS mais pas CVA. Exemple : vérifier (par intégration par parties) que $\int_0^1 \frac{dx}{x} \sin \frac{1}{x}$ est CVS mais pas CVA.

2.1.3 Problèmes avec l'intégrale de Riemann

– Certaines fonctions sont trop irrégulières pour être intégrables au sens de Riemann. Exemple : fonction $1 - \chi_{\mathbb{Q}}$ qui vaut 0 sur les rationnels et 1 sur les irrationnels de $\mathbb{R} \setminus \mathbb{Q}$: son intégrale de Riemann entre 0 et 1 n'est pas définie.

Il suffit pour s'en convaincre de prendre les ξ_k dans (2.1) soit tous rationnels, soit tous irrationnels. On obtient $\Sigma^{(R)} = 0$ ou 1 selon le cas !

– L'intégrale de la limite f d'une suite de fonctions f_n peut être différente de la limite des intégrales. Exemple : $f_n(x) = 2^n$ si $2^{-n} < x < 2^{-(n-1)}$, 0 ailleurs : $\int dx f_n(x) = 1$, mais $f_n \rightarrow f = 0$. En effet *Intégration sur un compact et limite commutent si la convergence est uniforme*, mais pas toujours sinon, cf le Théorème 1.7 du Chap 1.

– Considérons la (semi-)norme sur les fonctions définie par l'intégrale de Riemann et l'espace \mathcal{L}^1 des fonctions de norme $\| \cdot \|_1$ finie : une suite de Cauchy n'y converge pas toujours, l'espace \mathcal{L}^1 n'est pas complet. Exemple : fonctions définies sur \mathbb{R}_+ , $\|f\|_1 = \int_0^\infty |f(x)|dx$; suite $f_n(x) = e^{(i-\frac{1}{n})x}$, $n = 1, 2, \dots$, $\|f_n\|_1 = n$ est finie, mais $\lim_{n \rightarrow \infty} f_n = e^{ix}$ est de norme infinie.

2.2 Intégrale de Lebesgue. Mesure

2.2.1 Idée intuitive

Considérons d'abord une fonction f continue et positive. Au lieu de couper en tranches verticales l'aire sous la courbe de f (Riemann), on effectue une découpe horizontale : soient $\alpha_0 < \alpha_1 < \dots < \alpha_N$ une subdivision des valeurs prises par une fonction positive $f(x)$, et on suppose connue une mesure $\mu(A_k)$ de l'ensemble A_k des x tels que $\alpha_{k-1} \leq f(x) < \alpha_k$. On définit la somme de Lebesgue $\Sigma_N(\alpha) = \sum_{k=0}^{N-1} \mu(A_k) \alpha_k$. Si quand la subdivision devient de plus en plus fine : $N \rightarrow \infty$ avec $\sup_k (\alpha_k - \alpha_{k-1}) \rightarrow 0$, la somme $\Sigma_N(\alpha)$ a une limite indépendante de la subdivision α , cette limite définit l'intégrale de Lebesgue notée $\int f d\mu$ et la fonction f est dite intégrable (au sens de Lebesgue). Voir figure 2.2.

Il est sans doute bon à ce point de citer Lebesgue lui-même¹ :

Les géomètres du XVII^{ème} siècle considéraient l'intégrale de $f(x)$ – le mot intégrale n'était pas encore inventé, mais peu importe – comme la somme d'une infinité d'indivisibles² dont chacun était l'ordonnée, positive ou négative, de $f(x)$. Eh bien ! nous avons tout simplement groupés les indivisibles de grandeur comparable ; nous avons, comme on dit en algèbre, fait la réunion, la réduction des termes semblables. On peut dire encore que, avec le procédé de Riemann, on essayait de sommer les indivisibles en les prenant dans l'ordre où ils étaient fournis par la variation de x ,

1. Henri Lebesgue, *Sur le développement de la notion d'intégrale*, Conférence à la Société Mathématique, Copenhague, 1926. Œuvres Scientifiques, L'Enseignement Mathématique, Genève 1972.

2. nous dirions aujourd'hui ... infinitésimaux

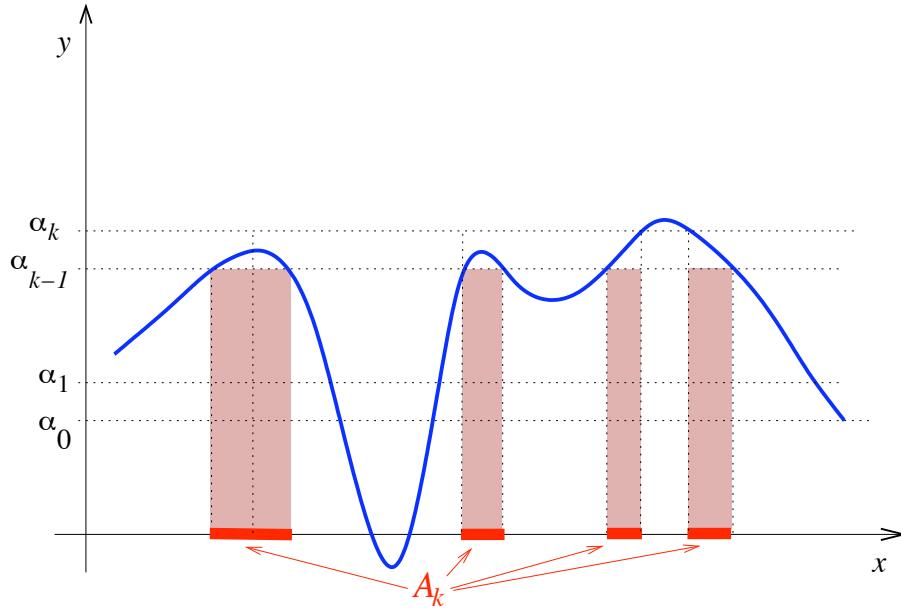


FIGURE 2.2 – Découpage de l'aire dans l'intégration de Lebesgue

on opérait donc comme le ferait un commerçant sans méthode qui compterait pièces et billets au hasard de l'ordre où ils lui tomberaient sous la main ; tandis que nous opérons comme le commerçant méthodique qui dit :

j'ai $\mu(1)$ pièces de 1 couronne valant $1 \cdot \mu(1)$,
 j'ai $\mu(2)$ pièces de 2 couronnes valant $2 \cdot \mu(2)$,
 j'ai $\mu(5)$ pièces de 5 couronnes valant $5 \cdot \mu(5)$,
 etc, j'ai donc en tout

$$S = 1 \cdot \mu(1) + 2 \cdot \mu(2) + 5 \cdot \mu(5) + \dots$$

Les deux procédés conduiront, certes, le commerçant au même résultat parce que, si riche qu'il soit, il n'a qu'un nombre fini de billets à compter ; mais pour nous, qui avons à additionner une infinité d'indivisibles, la différence entre les deux façons de faire est capitale.

Pour une fonction continue (qui peut être approchée de façon uniforme par des fonctions en escalier) et en prenant la mesure sur les intervalles fermés $\mu([\alpha, \beta]) = \beta - \alpha$, on retrouve l'intégrale de Riemann comme on le vérifiera plus bas. Mais l'idée permet de définir une intégration plus générale. Elle repose bien sûr sur une définition plus précise de la mesure μ .

2.2.2 Mesure (Bases de théorie de la)

- **Tribu de Borel**

On se donne un sous-ensemble X de \mathbb{R} , et on cherche à donner une “mesure” à des familles de sous-ensembles de X . On impose à une telle famille \mathcal{T} de sous-ensembles de X , appelée *tribu* (ou σ -algèbre), de satisfaire les axiomes suivants :

- $X \in \mathcal{T}$, $\emptyset \in \mathcal{T}$.
- Si $A \in \mathcal{T}$, son complémentaire dans X : $\bar{A} = \complement_X A \in \mathcal{T}$.
- Si $\{A_n\}$ est un ensemble fini ou dénombrable d’éléments de \mathcal{T} , $\cup_n A_n \in \mathcal{T}$.

Autrement dit, une tribu est une famille de sous-ensembles stable par passage au complémentaire et par union finie ou dénombrable, et contenant \emptyset .

Exercice : montrer que si $A_n \in \mathcal{T}$, (ensemble fini ou dénombrable), $\cap_n A_n \in \mathcal{T}$.

Noter que la tribu est stable par les seules unions (finies ou) dénombrables d’éléments. Cela diffère de la propriété des ouverts d’un e.t., pour lesquels toute union est un ouvert.

Dans la suite, en théorie de l’intégration mais aussi en théorie des probabilités, on fera souvent appel à la *tribu borélienne* $\mathcal{B}(\mathbb{R})$, qui est la plus petite tribu contenant $\mathcal{O}(\mathbb{R})$, l’ensemble des ouverts $]a, b[$ de \mathbb{R} , y compris les cas où $a = -\infty$ ou $b = +\infty$. Par passage au complémentaire, on vérifie que tout $[a, +\infty[,]-\infty, a]$ ou $[a, b]$ appartient à $\mathcal{B}(\mathbb{R})$, qui contient donc tous les ouverts, tous les fermés, et toutes les unions finies ou dénombrables d’ouverts et/ou de fermés (et bien plus encore). On appelle *borélien* un élément de $\mathcal{B}(\mathbb{R})$.

On peut définir de façon analogue la tribu borélienne $\mathcal{B}(E)$ de tout espace topologique E , comme “tribu engendrée par les ouverts” de E , c’est-à-dire la plus petite tribu contenant ces ouverts. Pour un sous-e.t. $F \subset E$, $\mathcal{B}(F) = \mathcal{B}(E) \cap F$ (avec à nouveau une notation un peu abusive).

- **Mesure de Lebesgue.**

On suppose donnée une tribu \mathcal{T} de $X \subset \mathbb{R}$.

Définition 2.1 : Une mesure sur (X, \mathcal{T}) est une fonction μ sur \mathcal{T} à valeurs dans $\overline{\mathbb{R}}^+ \stackrel{\text{déf}}{=} \mathbb{R}^+ \cup \{\infty\}$ telle que

- $\mu(\emptyset) = 0$;
- pour toute famille finie ou dénombrable (A_n) d’éléments de \mathcal{T} deux à deux disjoints :
 $A_i \cap A_j = \emptyset$ si $i \neq j$, on a

$$\mu(\cup_n A_n) = \sum_n \mu(A_n). \quad (2.4)$$

La propriété (2.4) est appelée *additivité dénombrable* (ou σ -additivité).

Exemples de mesures : Tout (X, \mathcal{T}) possède (au moins) une mesure μ , par exemple la mesure triviale nulle ; ou encore la mesure $\mu(A) = \text{card}(A)$, finie ou infinie ; ou bien encore, si $a \in X$, la

mesure de Dirac δ_a telle que pour $A \in \mathcal{T}$, $\delta_a(A) = 1$ ou 0 selon que $a \in A$ ou $a \notin A$. On parle de “masse de Dirac”, une notion que nous retrouverons dans l’étude des distributions. On peut finalement faire des combinaisons linéaires de telles masses de Dirac $\mu = \sum_n \alpha_n \delta_{a_n}$, avec une suite (finie ou infinie) de points $a_n \in X$ et des poids $\alpha_n \in \mathbb{R}^+$.

L’existence de mesure sur tout (X, \mathcal{T}) permet les définitions suivantes

Définition 2.2 : Un couple (X, \mathcal{T}) est dit mesurable ; le triplet (X, \mathcal{T}, μ) est dit mesuré une fois que l’on a choisi une mesure μ .

Exercice : montrer que les axiomes précédents sur la mesure impliquent sa *monotonicité* : si $B \subset C$, $\mu(B) \leq \mu(C)$ (écrire $C = B \cup (C \setminus B)$) ; et sa *sous-additivité* : pour $B, C \in \mathcal{T}$, $\mu(B \cup C) \leq \mu(B) + \mu(C)$, qui se généralise à tout ensemble fini ou dénombrable $\bigcup_n B_n$.

Les notions précédentes de tribu de Borel, de mesure et d’espace mesuré s’étendent à des espaces topologiques plus généraux que \mathbb{R} .

En fait on va s’intéresser surtout à une mesure sur \mathbb{R} généralisant la mesure naturelle pour un intervalle : $\mu([a, b]) = |b - a|$. On peut démontrer le

Théorème 2.1 (Mesure de Lebesgue sur les boréliens de \mathbb{R}) : L’ensemble $(\mathbb{R}, \mathcal{B})$ admet une unique mesure μ telle que $\forall a, b \in \mathbb{R}$, $\mu([a, b]) = |b - a|$.

Il en découle que pour un point $\{a\} = [a, a]$, $\mu(\{a\}) = 0$, puis par union dénombrable, que pour tout sous-ensemble Y fini ou dénombrable de \mathbb{R} , $\mu(Y) = 0$. En particulier

$$\mu(\mathbb{N}) = \mu(\mathbb{Z}) = \mu(\mathbb{Q}) = 0. \quad (2.5)$$

Un exemple d’ensemble de mesure nulle mais non dénombrable est offert par l’ensemble de Cantor \mathcal{K} , voir Appendice A du chap. 1.

Un ensemble A tel que $\mu(A) = 0$ est dit *négligeable* ou de mesure nulle.

Remarques

1. La propriété énoncée dans le Théorème est non triviale. Il s’agit de démontrer que la définition de μ s’étend des intervalles à tout borélien. L’idée est de prendre le $\sup(\mu(K))$ pour tout compact $K \subset A$, ou encore l’ $\inf(\mu(O))$ pour tout ouvert $O \supset A$, [5] et la difficulté est de s’assurer que la σ -additivité est bien satisfaite.
2. Il existe des sous-ensembles de \mathbb{R} non mesurables pour la mesure de Lebesgue, en ce sens qu’aucune extension de la mesure sur les intervalles ne peut être définie sur eux sans incohérence. Leur construction est délicate et pas explicite, faisant appel à l’axiome du choix (Zermelo), voir quelques indications dans [1] (p. 62 et note p. 55).
3. Il peut arriver qu’un ensemble $A \in \mathcal{B}(\mathbb{R})$ soit de mesure nulle, mais qu’il possède des sous-ensembles B non mesurables ! Ainsi si toute partie de l’ensemble de Cantor \mathcal{K} était mesurable, elle serait de mesure nulle puisque contenue dans \mathcal{K} de mesure nulle, donc serait dans la tribu de Borel $\mathcal{B}(\mathbb{R})$. Mais $\text{card}(\mathfrak{P}(\mathcal{K})) = 2^\mathfrak{c} > \mathfrak{c} = \text{card}(\mathcal{B}(\mathbb{R}))$ contredit $\mathfrak{P}(\mathcal{K}) \subset \mathcal{B}(\mathbb{R})$. Noter que cet argument n’est pas “constructif”, il affirme l’existence

de sous-ensembles non mesurables de \mathcal{K} , mais sans les donner explicitement.

On peut donc souhaiter étendre la mesure de Lebesgue à une “tribu de Lebesgue” contenant la tribu de Borel, de telle façon que tout sous-ensemble d’un ensemble de mesure nulle soit lui-même mesurable et de mesure nulle. Le lecteur curieux est renvoyé à la littérature sur cette extension [5].

La mesure de Lebesgue définie sur les boréliens est ensuite étendue à une tribu plus large, que nous ne préciserons pas davantage. Dans la suite la mesure considérée sera cette mesure de Lebesgue, et “ensemble mesurable” signifiera ensemble mesurable pour la mesure de Lebesgue.

• Propriétés vraies presque partout

Définition 2.3 : Une proposition $P(x)$, $x \in X \subset \mathbb{R}$, est vraie presque partout (vraie p.p.) si elle est vraie pour tout x sauf sur un ensemble contenu dans un ensemble de mesure nulle.

On dit encore que $P(x)$ est vraie pour presque tout $x \in X$.

Exemples. 1. La fonction partie entière, notée $E(x)$ (ou $\lfloor x \rfloor$),

$$E(x) = n \quad \text{si } n \leq x < n + 1 \quad (2.6)$$

est dérivable et de dérivée nulle en tout point $x \notin \mathbb{Z}$, elle est donc dérivable p.p.

2. La fonction de Dirichlet $D(x) = \chi_{\mathbb{Q}}(x)$, égale à 1 sur \mathbb{Q} , à 0 sur $\mathbb{R} \setminus \mathbb{Q}$, est nulle p.p.

• Fonctions mesurables

Définition 2.4 : Une fonction $f : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ est mesurable si $\forall B \subset \overline{\mathbb{R}}$ mesurable, $f^{-1}(B) \stackrel{\text{déf}}{=} \{x \in \mathbb{R} : f(x) \in B\}$ est mesurable.

On comparera cette définition à celle donnée plus haut d'une fonction continue (l'image inverse de tout ouvert est un ouvert).

Un exemple utile de fonction mesurable est celui d'une fonction constante, $\forall x \in \mathbb{R}$, $f(x) = a$: pour tout B mesurable, ou bien $a \in B$ et $f^{-1}(B) = \mathbb{R}$, ou bien $a \notin B$ et $f^{-1}(B) = \emptyset$, et dans les deux cas, $f^{-1}(B)$ est mesurable. On montre que toute fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ qui est soit continue, soit monotone, est mesurable ([5], p. 30) ; que si $f : \mathbb{R} \rightarrow \overline{\mathbb{R}}$ est mesurable, $f^\pm = \sup(\pm f, 0)$ le sont, donc aussi $|f| = f^+ + f^-$.

La réciproque n'est pas vraie : $|f|$ peut être mesurable sans que f le soit. Soit $A \subset \mathbb{R}$ mais $A \notin \mathcal{B}(\mathbb{R})$ (un ensemble non mesurable comme ci-dessus un sous-ensemble de l'ensemble de Cantor \mathcal{K}). Considérons la fonction f prenant la valeur ± 1 selon que x appartient ou non à A . Puisque $f^{-1}(1) = A$ n'est pas mesurable, f ne l'est pas. Mais $|f|$ l'est puisque c'est une fonction constante.

Énonçons un fait d'expérience : toutes les fonctions rencontrées en Physique sont mesurables. Cela est dû à la grande difficulté, mentionnée plus haut, de construire des ensembles non mesurables, et donc à leur caractère très artificiel, jamais réalisé en physique. Pour cette raison, nous ne nous étendrons pas davantage sur cette notion.

- Mesures sur \mathbb{R}^n

Sur \mathbb{R}^2 , on définit la tribu notée $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ “engendrée par” (c'est-à-dire la plus petite tribu contenant) tous les produits $A \times B$ de boréliens. Il existe sur cette tribu une mesure héritée de la mesure μ sur les boréliens : $\mu_2(A \times B) = \mu(A)\mu(B)$. Un résultat très utile est l'invariance de la mesure μ_2 non seulement par translation ou réflexion, mais aussi par rotation ou plus généralement par déplacement dans \mathbb{R}^2 . Cela se généralise bien sûr à \mathbb{R}^n .

2.2.3 Retour à l'intégrale de Lebesgue

On va reprendre et préciser la construction du § 2.2.1. Dans ce qui suit, E est un espace mesuré quelconque (donc en fait on se donne (E, \mathcal{B}, μ)). On peut penser de façon concrète à $E = \mathbb{R}$ ou une partie de \mathbb{R} , ou $E = \mathbb{R}^n$, ou $E = \mathbb{C}$, avec la mesure de Lebesgue, comme on l'a vu.

- a. Fonctions étagées

Définition 2.5 : Une fonction de E dans $\overline{\mathbb{R}}$ est dite étagée si il existe un nombre fini d'ensembles mesurables deux à deux disjoints A_j , $j = 1, \dots, n$ et des nombres $\alpha_1, \dots, \alpha_n \in \overline{\mathbb{R}}$ tels que

$$f = \sum_{j=1}^n \alpha_j \chi_{A_j}, \quad (2.7)$$

où χ_A est la fonction indicatrice de l'ensemble A , cf (A.1). Elle est donc mesurable et ne prend qu'un ensemble fini de valeurs. Exemple : la fonction de Dirichlet $D = 0\chi_{\mathbb{R}} + 1\chi_{\mathbb{Q}}$.

Un cas particulier de fonctions étagées est celui des fonctions en escalier. Rappelons la définition déjà donnée à l'App. A.4 :

Définition 2.6 : Une fonction de \mathbb{R} dans $\overline{\mathbb{R}}$ est dite en escalier si elle est constante par morceaux.

C'est donc une fonction étagée dont les ensembles A_j sont des intervalles bornés.

Par exemple, la fonction de Dirichlet déjà rencontrée, $D(x) = \chi_{\mathbb{Q}}(x)$, est étagée mais pas en escalier.

Un théorème important pour la construction de l'intégrale de Lebesgue est le suivant

Théorème 2.2 : Soit E un espace mesuré. Toute fonction mesurable $f : E \rightarrow \overline{\mathbb{R}}^+$ est limite simple d'une suite croissante (f_n) de fonctions étagées positives.

“Suite croissante” signifie que pour tout n , on a $f_n \leq f_{n+1}$. Autrement dit,

Corollaire : L'ensemble des fonctions étagées est dense dans l'ensemble des fonctions mesurables pour la “topologie de la convergence simple”.

La “topologie de la convergence simple” sur l'espace des applications d'un e.t. X dans un e.t. Y est définie comme suit : les ouverts en sont les unions quelconques d'intersections finies de parties de la forme $W(x, V)$ où V est un ouvert quelconque de Y , x un point quelconque de X et $W(x, V)$ l'ensemble des fonctions $f : X \rightarrow Y$ telles que $f(x) \in V$.

Preuve du théorème 2.2 ([7], p. 15) : Considérons d'abord une fonction mesurable f positive ; pour tout entier $n \geq 1$ et tout entier i tel que $1 \leq i \leq n 2^n$ on définit

$$E_{n,i} = f^{-1}\left(\left[\frac{i-1}{2^n}, \frac{i}{2^n}\right]\right) \quad \text{et} \quad F_n = f^{-1}([n, \infty])$$

et la fonction

$$f_n = \sum_{i=1}^{n 2^n} \frac{i-1}{2^n} \chi_{E_{n,i}} + n \chi_{F_n}.$$

f étant mesurable, les ensembles $E_{n,i}$ et F_n sont mesurables, cf. Déf. 2.4. Vérifier que les fonctions f_n sont bien étagées et que la suite f_n est bien croissante. Soit $x \in E$. Si $f(x)$ est fini, $f_n(x) \geq f(x) - 2^{-n}$ pour n assez grand. Si $f(x)$ est infini, $f_n(x) = n$. Donc dans tous les cas, $f_n(x) \rightarrow f(x)$. Si f est bornée sur E , cette même construction fournit une suite uniformément convergente de fonctions étagées positives, puisque $|f - f_n| < 2^{-n}$. Pour une fonction mesurable non positive, on écrit $f = f_+ - f_-$, avec $f_\pm \geq 0$ mesurables, etc. \square

• b. Intégrale de Lebesgue

Pour une fonction étagée positive, $f : E \rightarrow \overline{\mathbb{R}}^+$, donc satisfaisant (2.7) avec $\alpha_i \geq 0$, l'intégrale de Lebesgue est définie comme le nombre positif

$$\int f d\mu \stackrel{\text{déf}}{=} \sum_{i=1}^n \alpha_i \mu(A_i)$$

(avec la convention que si $\alpha_i = \infty$ et $\mu(A_i) = 0$, $\alpha_i \mu(A_i) = 0$ et vice versa). Pour s'assurer de la cohérence de cette définition, il faut vérifier son indépendance par rapport au choix des A_i dans (2.7), ce que nous admettrons. On note de façon équivalente $\int f d\mu = \int_E f d\mu = \int_E f(x) d\mu(x)$ ou même $= \int_E f(x) dx$. Donc pour la fonction de Dirichlet, $\int D d\mu = \mu(\mathbb{Q}) = 0$, $\int_{[0,1]} (1-D) d\mu = 1$, $\int (1-D) d\mu = +\infty$.

Soit une fonction f mesurable à valeurs positives ; on définit son intégrale par

$$\int f d\mu \stackrel{\text{déf}}{=} \sup \left\{ \int g d\mu \mid g \text{ étagée et } 0 \leq g \leq f \right\}. \quad (2.8)$$

C'est un nombre de $\overline{\mathbb{R}}^+$, donc éventuellement infini !

Définition 2.7 : Si ce nombre est fini, f est dite intégrable (ou sommable) au sens de Lebesgue.

Finalement pour une fonction mesurable de signe quelconque, on définit ses parties positive et négative comme ci-dessus, $f^+ = \max(f, 0)$ et $f^- = \max(-f, 0)$, toutes deux positives (ou

nulles) et mesurables. On a $f = f^+ - f^-$, $|f| = f^+ + f^-$. Alors f est intégrable (ou sommable) si séparément f^+ et f^- le sont, autrement dit si $|f|$ est intégrable, et on écrit l'intégrale de Lebesgue de f comme

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu. \quad (2.9)$$

Retenons donc que par définition,

f est intégrable de Lebesgue ssi $|f|$ l'est.

Pour une fonction à valeurs dans \mathbb{C} , on procède de même : on sépare partie réelle et partie imaginaire et f est intégrable si $\Re e f$ et $\Im m f$ le sont, avec $\int f d\mu = \int \Re e f d\mu + i \int \Im m f d\mu$.

2.2.4 Intégrales de Lebesgue sur \mathbb{R}^2 ou \mathbb{R}^n

On a mentionné plus haut la mesure produit μ_2 définie sur les boréliens de \mathbb{R}^2 . Cette mesure est invariante par translation, rotation ou plus généralement par déplacement. On peut alors définir une intégrale de Lebesgue pour les fonctions $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ qu'on notera $\int_{\mathbb{R}^2} f(x, y) d\mu_2(x, y)$. Ici encore, la fonction f est dite *intégrable* (de Lebesgue) ssi $|f|$ l'est. Pour un changement de variable arbitraire $\{x_j\} \mapsto \{y_i\}$ dans \mathbb{R}^n , la formule de changement de variable implique le *jacobien* $J = \det \left(\frac{\partial y_i(x)}{\partial x_j} \right)$, voir Appendice B.

Un résultat très utile est le Théorème de Fubini qui justifie l'interversion des intégrations.

Théorème 2.3 (Fubini) : Pour $A, B \in \mathcal{B}(\mathbb{R})$, et f intégrable sur $A \times B$

$$\int_{A \times B} d\mu_2(x, y) f(x, y) = \int_B d\mu(y) \left(\int_A d\mu(x) f(x, y) \right) = \int_A d\mu(x) \left(\int_B d\mu(y) f(x, y) \right). \quad (2.10)$$

Inversement le fait que les deux intégrales du milieu et de droite ne soient pas égales signale que f n'est pas intégrable. Voir des exemples en TD.

En fait l'hypothèse que f est intégrable découle elle-même de l'hypothèse que l'une ou l'autre des “intégrales itérées” du milieu ou de droite dans (2.10) est **absolument convergente** (théorème de Tonelli).

Autrement dit Fubini nous dit : si f est intégrable, alors $x \mapsto \int dy f(x, y)$ existe p.p. et est intégrable (et *ibid* avec $x \leftrightarrow y$) et on peut intervertir les intégrations ; Tonelli nous dit : si soit $\int_B d\mu(y) (\int_A d\mu(x) |f(x, y)|)$, soit $\int_A d\mu(x) (\int_B d\mu(y) |f(x, y)|)$ est finie, alors f est intégrable et le théorème de Fubini s'applique !

Mais cette condition n'est pas toujours remplie, même en physique, donc prudence ! Voir des exemples en TD.

2.2.5 Espaces \mathcal{L}^p et L^p

- a. Espaces \mathcal{L}^p

Pour f une fonction mesurable $E \rightarrow \mathbb{R}$ (E désignant toujours un espace mesuré), on se rappelle que $\int |f| d\mu$ est défini dans $\overline{\mathbb{R}}^+$. On définit pour p réel ≥ 1

$$\|f\|_p \stackrel{\text{déf}}{=} \left(\int |f|^p d\mu \right)^{\frac{1}{p}}. \quad (2.11)$$

Proposition 2.4 : $\|\cdot\|_p$ est une semi-norme. En particulier

$$\|f\|_p = 0 \Leftrightarrow f \text{ nulle p.p..} \quad (2.12)$$

Preuve : c'est évident dans le sens \Leftarrow ; dans le sens \Rightarrow , on procède par l'absurde. La propriété $\|\lambda f\|_p = |\lambda| \|f\|_p$ est évidente et l'inégalité triangulaire est laissée en exercice (inégalité de Minkowski, voir TD).

On note $\mathcal{L}^p(E)$ l'espace vectoriel des fonctions mesurables f telles que $|f|^p$ soit *intégrable*. (Bien noter qu'on prend la valeur absolue de f , ce qui évite des problèmes quand f a des valeurs négatives et que p est non entier...) Si E est de mesure finie, on montre que $1 \leq p \leq q$ implique que $\|f\|_p \leq \mu(E)^{\frac{1}{p}-\frac{1}{q}} \|f\|_q$ et donc que $\mathcal{L}^q(E) \subset \mathcal{L}^p(E) \subset \mathcal{L}^1(E)$ (cf [5], p.75). Par exemple, pour $E =]0, 1[$, $x \mapsto 1/\sqrt{x} \in \mathcal{L}^1(]0, 1[)$ mais $\notin \mathcal{L}^2(]0, 1[)$. Noter que la condition "E est de mesure finie" exclut $E =]1, \infty[$ ou \mathbb{R} . Ainsi $x \mapsto 1/x \notin \mathcal{L}^1(]1, \infty[)$ mais $\in \mathcal{L}^2(]1, \infty[)$. Donc attention ! il n'existe pas de relation d'inclusion entre les $\mathcal{L}^p(\mathbb{R})$.

Exercice : trouver un exemple de fonction qui $\in \mathcal{L}^1(\mathbb{R})$ mais $\notin \mathcal{L}^2(\mathbb{R})$; qui $\notin \mathcal{L}^1(\mathbb{R})$ mais $\in \mathcal{L}^2(\mathbb{R})$; qui $\in \mathcal{L}^1(\mathbb{R})$ et $\in \mathcal{L}^2(\mathbb{R})$.

Il conviendrait alors d'étudier les questions de convergence : quand une suite $f_n \in \mathcal{L}^p$ convergeant p.p. vers f a-t-elle une limite $f \in \mathcal{L}^p$? (théorème de Fatou) ; quelles sont les relations entre la convergence p.p. et la convergence dans la semi-norme $\|\cdot\|_p$ (ou "convergence en moyenne d'ordre p ")? ; l'espace \mathcal{L}^p est-il complet? (théorème de Riesz-Fischer), voir plus bas. Pour toutes ces questions, nous renvoyons à la littérature mathématique, voir références en fin de chapitre, et nous nous contenterons d'énoncer un résultat remarquable et très utile, dû à Lebesgue :

Théorème 2.5 de Lebesgue de convergence dominée : Soit (f_n) une suite de fonctions mesurables de E , sous-ensemble mesurable de \mathbb{R} , dans \mathbb{R} , qui converge simplement p.p. vers une fonction f . On suppose qu'il existe une fonction g intégrable t.q. $\forall n \geq 1$, $|f_n| \leq g$ p.p. Alors f est intégrable, (f_n) converge pour la semi-norme $\|\cdot\|_1$ vers f et l'on a $\lim \int_E f_n d\mu = \int_E f d\mu$.

La preuve, un peu technique, peut être trouvée par exemple dans [5] pp 94-95, qui en donne aussi une version pour des fonctions f_n et $g \in \mathcal{L}^p$.

Autrement dit, l'existence d'une majoration des $|f_n|$ ("domination") par une *même fonction* g *indépendamment de n* suffit à assurer l'intégrabilité \mathcal{L}^1 de la limite et l'égalité des intégrales à la limite, sans supposer de convergence uniforme comme dans le Théorème 1.8.

Ce théorème a d'importantes conséquences pratiques sur l'intégration terme à terme d'une série, la continuité et la dérivabilité d'une intégrale par rapport à un paramètre, etc, voir ci-dessous § 2.3.

Exemple : ([1] p. 65). Soit $f_n(x) = \frac{e^{-x^2/n^2} \cos \frac{x}{n}}{1+x^2}$ qui converge simplement vers $f(x) = 1/(1+x^2)$ qui est intégrable et est dominée par ce f pour tout x . On en conclut que $\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} f_n(x) dx = \int_{-\infty}^{\infty} f(x) dx = [\text{Arctan } x]_{-\infty}^{\infty} = \pi$. Voir d'autres exemples en TD.

• b. Espaces L^p

On a vu que si deux fonctions intégrables f et g sont égales p.p., leurs intégrales sont égales. On va maintenant identifier deux fonctions égales p.p. et on est amené à la

Définition 2.8 : Espaces L^p . Soit E un espace mesuré. L'espace $L^p(E)$ est l'espace vectoriel des fonctions (à valeurs dans \mathbb{R} ou \mathbb{C}) définies à une égalité p.p. près et telles que f^p soit intégrable.

Exemple. La fonction de Dirichlet $\chi_{\mathbb{Q}}(x)$ s'identifie dans L^1 à la fonction nulle.

On peut munir ces espaces de la norme $\|f\|_p \stackrel{\text{déf}}{=} \int |f|^p d\mu$. Nous aurons à travailler tout particulièrement avec les espaces L^1 et L^2 .

Théorème 2.6 (Riesz–Fischer) : L'espace $L^p(E)$ des fonctions intégrables (à valeurs dans \mathbb{R} ou \mathbb{C}), muni de la norme $\|f\|_p = (\int |f|^p d\mu)^{1/p}$ est un e.v. normé complet. L'espace $C([a, b])$, resp. $C[\mathbb{R}]$ des fonctions continues sur $[a, b]$, resp. \mathbb{R} , est dense dans $L^1([a, b])$, resp. $L^1(\mathbb{R})$.

On a donc résolu une des difficultés de l'intégrale de Riemann mentionnée au § 2.1.3.

2.2.6 Comparaison entre intégrales de Riemann et de Lebesgue

1. Propriétés communes aux intégrales de Riemann et de Lebesgue

L'intégrale de Lebesgue possède les propriétés suivantes :

- (a) linéarité : l'ensemble des fonctions intégrables (sur un certain espace mesuré E) est un espace vectoriel et $f \mapsto \int f d\mu$ est une forme linéaire ;
- (b) croissance : si f et g sont réelles et intégrables et $f \leq g$, alors $\int f d\mu \leq \int g d\mu$;
- (c) si f est intégrable (à valeurs dans \mathbb{R} ou \mathbb{C}), $\left| \int f d\mu \right| \leq \int |f| d\mu$;

- (d) domination : si $g \geq 0$ est intégrable et si f mesurable satisfait p.p. $|f| \leq g$, alors f est intégrable.
- (e) additivité : si A et B sont mesurables et $A \cap B$ est fini ou de mesure nulle, et si f est intégrable sur $A \cup B$, $\int_{A \cup B} f d\mu = \int_A f d\mu + \int_B f d\mu$.
- (f) si f et g sont égales p.p. sur A mesurable, $\int_A f d\mu = \int_A g d\mu$;
- (g) si une fonction mesurable f positive sur A a une intégrale sur A nulle, elle y est nulle p.p.

Les propriétés (a), (b), (c) sont aussi satisfaites par l'intégrale de Riemann ; (e) généralise $\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$ si $c \in]a, b[$. Les propriétés (f) et (g) généralisent au sens du “p.p” des propriétés de l'intégrale de Riemann. La propriété (d) (qui découle du Théorème de Lebesgue de convergence dominée) ne se transpose pas naïvement à l'intégrale de Riemann³. Une fois encore, un contre-exemple nous est fourni par la fonction $f = \chi_{\mathbb{Q} \cap [0,1]}$ qui n'est pas intégrable de Riemann mais satisfait $|f| < \chi_{[0,1]}$ et est intégrable de Lebesgue.

2. Changement de variable

Soit f une fonction intégrable sur $B \subset \mathbb{R}$. Soit φ une fonction de classe C^1 (continûment différentiable) et de dérivée non nulle sur $A = \varphi^{-1}(B)$. Alors $f \circ \varphi$ est intégrable sur A et

$$\int_B f(y) d\mu(y) = \int_A f(\varphi(x)) |\varphi'(x)| d\mu(x). \quad (2.13)$$

Pour l'intégrale de Riemann sur un intervalle image par φ monotone de $[\alpha, \beta]$ on reconnaît la formule familière de changement de variable par $y = \varphi(x)$

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f(y) dy = \int_{\alpha}^{\beta} f(\varphi(x)) \varphi'(x) dx.$$

La formule (2.13) s'étend aux intégrales dans \mathbb{R}^n , voir App. B.

3. Intégrabilité à la Riemann, à la Lebesgue

Un théorème dû à Lebesgue affirme

Théorème 2.7 (Critère de Lebesgue) : Soit f une fonction définie et bornée sur $[a, b]$ et soit Δ l'ensemble des points de discontinuité de f sur $[a, b]$. Alors f est intégrable au sens de Riemann si et seulement si Δ est contenu dans un ensemble de mesure de Lebesgue nulle.

3. Il existe aussi des versions du théorème de convergence dominée pour l'intégrale de Riemann, voir par exemple, http://jf.burnol.free.fr/convergencedominee_v2.pdf

Exemples : la fonction indicatrice de l'ensemble de Cantor $\chi_{\mathcal{K}}$ est intégrable de Riemann bien que non réglée (déf. § 2.1.1), et bien qu'ayant une infinité non dénombrable de points de discontinuité ; celle de $X = \mathbb{Q} \cap [0, 1]$ ne l'est pas.

En effet on se rappelle (App A.1) que l'ensemble des points de discontinuité de χ_X est la “frontière” de X , soit $\Delta = \bar{X} \setminus \overset{\circ}{X}$. Ici $\bar{X} = [0, 1]$, $\overset{\circ}{X} = \emptyset$, $\mu(\Delta) = 1$.

Ce critère permet aussi de démontrer le théorème 1.7 du chapitre 1 sur la commutation de l'intégration et de la limite, sous l'hypothèse de convergence uniforme.

En effet f_n intégrable de \mathbb{R} $\Rightarrow \Delta_n$ est de mesure nulle ; la CVU implique que f est continue à tous les points où les f_n sont continues $\Rightarrow \Delta \subset \cup_n \Delta_n$ est de mesure nulle $\Rightarrow f$ intégrable de \mathbb{R} .

On a vu que toute fonction réglée est intégrable de Riemann ; on montre qu'elle est aussi intégrable de Lebesgue, par application du théorème de Lebesgue de convergence dominée, (voir [5] p 104), et leurs intégrales sont égales. Plus généralement

Théorème 2.8 : *Toute fonction $\mathbb{R} \rightarrow \mathbb{R}$ qui est intégrable au sens de Riemann l'est aussi au sens de Lebesgue et leurs intégrales sont égales*

$$\int_{\mathbb{R}} f(x) dx = \int_{\mathbb{R}} f d\mu.$$

Mais bien sûr, et c'est l'intérêt de l'intégrale de Lebesgue, il existe des fonctions non intégrables au sens de Riemann mais intégrables au sens de Lebesgue. Exemple, la fonction de Dirichlet $D = \chi_{\mathbb{Q} \cap [0, 1]}$.

Par ailleurs, la condition $\mathbb{R} \rightarrow \mathbb{R}$ du Théorème précédent exclut le cas des *intégrales improches* de fonctions sur $[a, b]$, avec b fini ou infini, cf ci-dessus § 2.1.2. Il faut donc reprendre l'analyse dans ce cas-là.

Pour une fonction $f : [a, b] \rightarrow \mathbb{R}$, on note f^* son prolongement à $\mathbb{R} \setminus [a, b]$ par 0. Par définition de l'intégrale de Lebesgue, f^* est intégrable de Lebesgue ssi $\int_a^b f dx$ est absolument convergente. Par conséquent **les intégrales semi-convergentes**, telle $\int_1^\infty \frac{dx}{x} \sin x$, **ne peuvent être interprétées comme intégrales de Lebesgue**.

Remarque. *L'intégrale fonction de sa borne supérieure.*

Comme on l'a rappelé plus haut, pour l'intégrale de Riemann, intégration et dérivation sont des opérations inverses l'une de l'autre : si f est continue donc intégrable au sens de Riemann sur $[a, b]$, pour tout $x \in]a, b[$ $F(x) := \int_a^x f(x') dx'$ est dérivable et $F'(x) = f(x)$. De plus $\int_a^b f(x') dx' = F(b) - F(a)$. Pour l'intégrale de Lebesgue, cela n'est plus toujours le cas : on verra plus bas, au Chap. 5, le cas de la fonction de Cantor F telle que $F(1) - F(0) = 1$ mais dont la dérivée est définie et nulle p.p. Clairement on ne peut avoir $F(1) - F(0) = \int_0^1 F'(x) dx$. D'une manière générale, si f est intégrable de Lebesgue, l'intégrale de Lebesgue $\int_a^x f(x') dx'$ est dérivable avec pour dérivée $f(x)$ sauf sur un ensemble de mesure nulle.

2.3 Intégrales dépendant d'un paramètre

On considère une fonction de deux variables réelles $f : \mathbb{R} \times [a, b] \rightarrow \mathbb{R}$ (si c'est $\rightarrow \mathbb{C}$, il suffit de séparer partie réelle et partie imaginaire), on suppose $x \mapsto f(x, t)$ mesurable pour tout $t \in [a, b]$, et on s'intéresse à la fonction ϕ définie par l'intégrale

$$\phi(t) = \int_{\mathbb{R}} dx f(x, t) \quad t \in [a, b] \quad (2.14)$$

et à ses propriétés de continuité, de dérivabilité et d'intégrabilité etc.

Théorème 2.9 : Pour $t_0 \in [a, b]$, si pour presque tout $x \in \mathbb{R}$, $t \mapsto f(x, t)$ est continue en t_0 , et s'il existe une fonction $g : \mathbb{R} \rightarrow \mathbb{R}^+$, intégrable et dominant $|f|$ dans un voisinage \mathcal{V} de t_0

$$|f(x, t)| \leq g(x) \quad p.p. \text{ en } x \in \mathbb{R}$$

alors $t \mapsto \phi(t)$ est continue en t_0 .

La preuve repose sur le théorème de convergence dominée de Lebesgue, ([1], p 66).

Avec les mêmes notations

Théorème 2.10 : Si dans un voisinage de t_0 la fonction $x \mapsto f(x, t)$ est intégrable et la fonction $t \mapsto f(x, t)$ est dérivable pour presque tout $x \in \mathbb{R}$ et que la dérivée est dominée

$$\left| \frac{\partial f}{\partial t}(x, t) \right| \leq g(x) \quad \forall t \in \mathcal{V} \text{ et p.p. en } x \in \mathbb{R},$$

avec g intégrable, alors la fonction $t \mapsto \phi(t)$ est dérivable en t_0 et

$$\frac{d}{dt} \int_{\mathbb{R}} f(x, t) dx \Big|_{t=t_0} = \int_{\mathbb{R}} \frac{\partial f}{\partial t}(x, t_0) dx, \quad (2.15)$$

autrement dit, on peut “dériver sous le signe somme”.

Il s'agit là d'un résultat important en pratique en physique, où on est souvent amené à dériver sous le signe somme. Il importe donc de connaître les conditions qui justifient cette opération.

En ce qui concerne l'intégrabilité de $\phi(t)$, on a vu plus haut le Théorème de Fubini.

Dans le même esprit, quand peut-on intégrer terme à terme une série de fonctions ? Un théorème donne une condition suffisante

Théorème 2.11 (B. Levi) : Si une suite de fonctions f_n de $\mathcal{L}^1(\mathbb{R})$ est telle que $\sum_{n=1}^{\infty} \int_{\mathbb{R}} |f_n(x)| dx < +\infty$, alors la série $\sum f_n$ converge absolument p.p., sa somme f est intégrable et $\int f dx = \sum \int f_n dx$.

Exemples, contrexemples

a) Considérons la fonction $f : f(x, 0) = 0$ et si $t \neq 0$, $f(x, t) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(x - \frac{1}{t})^2\right]$. La fonction $t \mapsto f(x, t)$ est continue pour tout $x \in \mathbb{R}$. Pour $t \neq 0$, la fonction $x \mapsto f(x, t)$ est une gaussienne d'intégrale $\phi(t) = 1$. Mais pour $t = 0$, $f = 0$ a une intégrale nulle. La fonction ϕ est discontinue en $t = 0$. Dans ce cas, on n'a pas de fonction g dominante.

b) Considérons la fonction $\phi(t) = \int_0^\infty dx \frac{\cos(xt)}{1+x^2}$, dont on montrera (Chap. 8, § 8.1.5) qu'elle vaut $\phi(t) = \frac{\pi}{2} e^{-|t|}$, et donc qu'elle est continue mais non dérivable en $t = 0$. Voir TD2.

Appendice B. Changements de variables dans une intégrale sur un domaine de \mathbb{R}^n . Jacobien

Nous admettrons les résultats suivants, qui constituent une généralisation à \mathbb{R}^n du résultat bien connu (2.13) sur les fonctions $\mathbb{R} \rightarrow \mathbb{R}$.

Soient Ω et Ω' deux ouverts de \mathbb{R}^n et φ une application bijective et continûment différentiable (= de classe C^1) de Ω sur Ω'

$$(x_1, x_2, \dots, x_n) \mapsto (y_1, y_2, \dots, y_n) = \varphi(x_1, \dots, x_n). \quad (\text{B.1})$$

Définition 2.9 : On appelle matrice jacobienne la matrice des dérivées partielles

$$\mathcal{J}_\varphi \equiv \frac{D(y_1, y_2, \dots, y_n)}{D(x_1, x_2, \dots, x_n)} \stackrel{\text{déf}}{=} \left(\frac{\partial y_i}{\partial x_j} \right)_{1 \leq i, j \leq n}, \quad (\text{B.2})$$

et on appelle Jacobien le déterminant de la matrice jacobienne

$$J_\varphi = \det \mathcal{J}_\varphi. \quad (\text{B.3})$$

Si ce jacobien J_φ ne s'annule pas sur Ω , la fonction inverse $\varphi^{-1} : \Omega' \rightarrow \Omega$ est aussi continûment différentiable sur Ω' . On appelle un tel φ un *difféomorphisme* de Ω .

Avec ces notations, on a le

Théorème 2.12 (changement de variables) : Supposons que le jacobien J_φ ne s'annule pas sur Ω . Soit f une fonction $\Omega' \rightarrow \mathbb{R}^p$. Alors

$$f \text{ intégrable sur } \Omega' \Leftrightarrow f \circ \varphi \text{ intégrable sur } \Omega$$

et

$$\int_{\Omega'} f(y) d^n y = \int_{\Omega} (f \circ \varphi)(x) |J_\varphi(x)| d^n x.$$

Cela s'applique en particulier aux changements de coordonnées rectangulaires \leftrightarrow polaires dans \mathbb{R}^2 ou \leftrightarrow cylindriques ou sphériques dans l'espace \mathbb{R}^3 .

Coordonnées polaires dans \mathbb{R}^2

$$(x = r \cos \theta, y = r \sin \theta), \quad r \geq 0, \quad 0 \leq \theta < 2\pi \quad dx dy = r dr d\theta \quad (\text{B.4})$$

Coordonnées cylindriques dans \mathbb{R}^3

$$(x = r \cos \theta, y = r \sin \theta, z), \quad r \geq 0, \quad 0 \leq \theta < 2\pi \quad dx dy dz = r dr d\theta dz \quad (\text{B.5})$$

Coordonnées sphériques dans \mathbb{R}^3

$$(x = r \cos \phi \sin \theta, y = r \sin \phi \sin \theta, z = r \cos \theta) \quad , \quad r \geq 0, \quad 0 \leq \phi < 2\pi, \quad 0 \leq \theta \leq \pi, \\ dx dy dz = r^2 dr \sin \theta d\theta d\phi. \quad (\text{B.6})$$

Exercice : écrire les coordonnées sphériques et leur jacobien dans \mathbb{R}^n , n arbitraire.

Lectures complémentaires

Voir Appel [1], et pour plus de précisions mathématiques, Gapaillard [5] et Rudin, [7].

Chapitre 3

Distributions

3.1 Introduction

Le physicien rencontre fréquemment des situations où les fonctions régulières –continues, une fois, deux fois ... différentiables– de l’analyse classique s’avèrent insuffisantes. Ce sont en général des limites singulières de problèmes bien définis, voir ci-dessous des exemples, et le physicien a donc à sa disposition une “régularisation” naturelle de la singularité, fournie par le problème étudié avant d’en prendre la limite. Cette régularisation n’est pas toujours évidente dans une formulation mathématique générale. Cela va nous amener à introduire des objets mathématiques nouveaux, les *fonctions généralisées* ou *distributions*.

Ces concepts ont été intuités par le physicien P.A.M. Dirac et développés par les mathématiciens Sergei Sobolev et Laurent Schwartz, dans les années 1935-1945.

Commençons par présenter quelques problèmes physiques où se rencontrent ces problèmes. Il s’agit en général d’idéalisations de situations bien définies : limite de charge électrique ponctuelle, de réseau diffuseur infiniment étendu, de choc infiniment bref avec changement instantané de la vitesse, etc.

3.1.1 Distributions de charges électriques.

On considère le potentiel créé en $\mathbf{r} \in \mathbb{R}^3$ par une charge électrique q isolée et localisée en \mathbf{r}_0

$$V_0(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \frac{q}{\|\mathbf{r} - \mathbf{r}_0\|} \quad (3.1)$$

et on le compare au potentiel créé par une distribution continue de charges, de densité $\rho(\mathbf{r})$, i.e. telle que la charge contenue dans un petit volume $d^3\mathbf{r}$ autour de \mathbf{r} est $\rho(\mathbf{r})d^3\mathbf{r}$

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int d^3\mathbf{r}' \frac{\rho(\mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|}. \quad (3.2)$$

Comment passer de (3.2) à (3.1) quand la charge est localisée en le seul point \mathbf{r}_0 ? Il faut imaginer que la fonction $\rho(\mathbf{r})$ a un *support* de plus en plus restreint quand un paramètre ϵ , par exemple le diamètre de ce support, tend vers zéro, et est telle que $\lim_{\epsilon \rightarrow 0} \int d^3\mathbf{r}' \rho(\mathbf{r}') \varphi(\mathbf{r}') = q\varphi(\mathbf{r}_0)$ pour toute fonction φ “pas trop singulière” comme 1 ou comme $1/\|\mathbf{r} - \mathbf{r}'\|$. A la limite, ρ devrait donc être nulle presque partout. On voit qu’au sens de l’intégrale de Lebesgue du chapitre 2, son intégrale devrait alors être nulle, alors qu’on attend que cette intégrale vaille q . Cette limite de ρ ne peut donc pas s’assimiler à une fonction. On écrira $\lim_{\epsilon \rightarrow 0} \rho(\mathbf{r}') = q\delta(\mathbf{r}' - \mathbf{r}_0)$, et δ , la distribution delta de Dirac, nulle p.p. et d’intégrale égale à 1, nous offre un premier exemple de distribution.

3.1.2 Diffusion cohérente par un réseau. Peigne de Dirac

Notre deuxième exemple concerne la diffusion d’une onde électromagnétique, supposée monochromatique de vecteur d’onde \mathbf{k} et de pulsation $\omega = 2\pi\nu$, (avec $|k| = \omega/c$), incidente selon un angle α sur un réseau supposé ici unidimensionnel et constitué de $2L + 1$ centres diffuseurs, espacés d’une longueur a . Un calcul classique montre que l’onde diffusée dans la direction β a une amplitude multipliée par

$$\sum_{\ell=-L}^L e^{i\ell k\Delta} = \frac{e^{i(L+1)k\Delta} - e^{-iLk\Delta}}{e^{ik\Delta} - 1} = \frac{\sin(2L+1)k\frac{\Delta}{2}}{\sin k\frac{\Delta}{2}} \quad (3.3)$$

où $\Delta = a(\sin \alpha - \sin \beta)$ est la différence de chemins optiques entre deux rayons lumineux incidents sur des centres adjacents, voir Fig. 3.1.

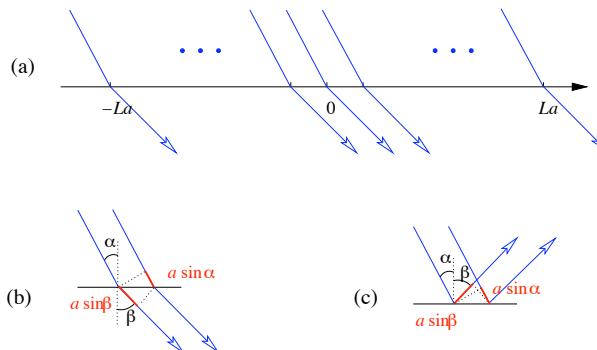


FIGURE 3.1 – (a) Diffusion par un réseau fini fait de $2L + 1$ centres diffuseurs. (b-c) Différence de chemin optique entre deux rayons lumineux adjacents observés en transmission (b) ou en réflexion (c).

Dans la limite où le nombre $2L + 1$ de diffuseurs tend vers l’infini, le facteur (3.3) apparaissant dans l’amplitude devient de plus en plus “piqué” et grand au voisinage de toute valeur $2\pi K$ de x , cf Fig 3.2. La limite (dans un sens qui devra être précisé) est donc infinie en tout point $2\pi K$, nulle ailleurs. On notera $2\pi\delta_P(x)$ cette limite, où l’indice “P” rappelle sa nature périodique, de période 2π ,

$$\delta_P(x) \stackrel{?}{=} \frac{1}{2\pi} \sum_{\ell=-\infty}^{\infty} e^{i\ell x}.$$

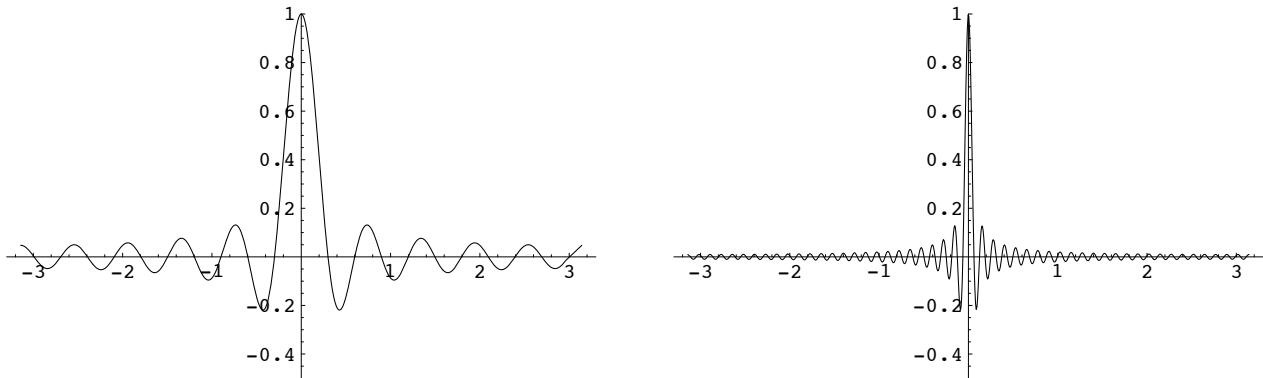


FIGURE 3.2 – La fonction $\sin(2L+1)\frac{x}{2}/[(2L+1)\sin\frac{x}{2}]$ entre $-\pi$ et π pour deux valeurs de $L = 10$ et $L = 50$. On voit que la fonction n'est d'ordre 1 que dans un intervalle $|\Delta x| \approx O(\frac{1}{L})$ autour de chaque multiple de 2π .

Noter que son intégrale vaut 1 sur tout intervalle de longueur 2π (on intervertit hardiment intégration et sommation infinie, il faudrait revenir à la somme de $-L$ à L)

$$\int_{-\pi}^{\pi} \delta_P(x) dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} \sum_{\ell=-\infty}^{\infty} e^{i\ell x} dx = \sum_{\ell=-\infty}^{\infty} \delta_{\ell 0} = 1 , \quad (3.4)$$

et on peut donc écrire –toujours de façon très heuristique– δ_P comme une superposition linéaire de distributions de Dirac localisées aux multiples de 2π

$$\delta_P(x) = \sum_{K=-\infty}^{\infty} \delta(x - 2\pi K) .$$

On donne pour cela à cette fonction généralisée, dont on a tenté de dessiner le graphe à la figure 3.3, le nom de “peigne de Dirac”, et on remplace souvent la notation δ_P par celle, plus suggestive, de III (la lettre cyrillique “cha”). On y reviendra plus bas, équ. (3.7).

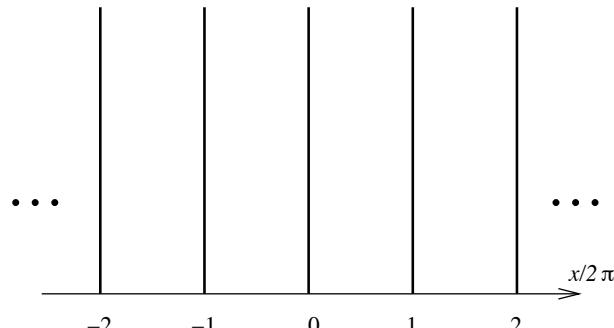


FIGURE 3.3 – Le “peigne de Dirac”

3.1.3 Choc élastique

Une balle rebondit élastiquement sur un obstacle. Sa vitesse avant le choc est \mathbf{v}_0 , elle est $-\mathbf{v}_0$ après (hypothèse de choc *élastique*, sans perte d'énergie cinétique). Supposons pour simplifier que le profil de vitesse

pendant l'intervalle Δt du choc est linéaire, voir figure 3.4. Par la loi de Newton, la force à laquelle la balle est soumise est $\mathbf{F} = m\dot{\mathbf{v}} = m\Delta\mathbf{v}/\Delta t$, donc $\mathbf{F} = -2m\mathbf{v}_0/\Delta t$. La variation de la *quantité de mouvement*, $\Delta_{21}\mathbf{p} = \mathbf{p}_2 - \mathbf{p}_1 = -2m\mathbf{v}_0 = \mathbf{F}\Delta t$ est indépendante de Δt . Dans le cas limite où $\Delta t \rightarrow 0$, la force \mathbf{F} est mal définie, mais son intégrale sur tout intervalle de temps $[t_1, t_2]$ qui mesure cette variation de la *quantité de mouvement* $\Delta_{21}\mathbf{p} = \mathbf{p}_2 - \mathbf{p}_1 = \int_{t_1}^{t_2} dt \mathbf{F}(t)$ reste, elle, bien définie. En particulier, pour tout $\epsilon \neq 0$, $\int_{-\epsilon}^{\epsilon} dt \mathbf{F}(t) = -2m\mathbf{v}_0$. À nouveau l'analyse usuelle faisant appel à des fonctions ordinaires ne peut rendre compte de ce $\mathbf{F}(t)$.

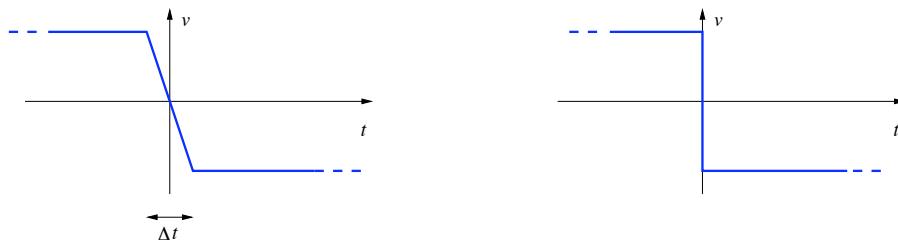


FIGURE 3.4 – (a) profil de vitesse lors d'un choc élastique ; (b) limite d'un choc instantané

3.1.4 Autres exemples

On rencontrera d'autres exemples en mécanique quantique, en physique statistique, puis par la suite, en théorie quantique des champs.

3.2 Définitions et premières propriétés

L'idée est de définir des objets dont l'*intégrale* avec des fonctions suffisamment régulières est bien définie. Plutôt que de définir une fonction f par sa valeur en chaque point, on veut en donner la moyenne locale obtenue par intégration avec une “fonction-test” de support suffisamment concentré. Ainsi dans les exemples précédents, la densité de charge peut être un objet singulier à la limite ponctuelle, mais la charge totale est bien définie ; la force dans un choc instantané est mal définie, mais son intégrale sur le temps de la collision, c'est-à-dire la variation de quantité de mouvement, est bien définie, etc. Pour donner un sens plus précis à cette idée, il faut dire avec quelles fonctions f , quelles “fonctions-tests”, on va travailler.

Remarque : L'intégration d'une fonction est une opération linéaire : si φ est une fonction de \mathbb{R}^n dans \mathbb{C} d'une classe \mathcal{D} à préciser, et f une fonction telle que $f(x)\varphi(x)$ soit intégrable, (fonction “localement intégrable”, voir plus bas), l'application $\varphi \mapsto \int_{\mathbb{R}^n} d^n x f(x)\varphi(x)$ est une *forme linéaire*, c'est-à-dire une application linéaire de \mathcal{D} dans \mathbb{C} . On va exploiter cette remarque pour définir les distributions comme formes linéaires sur des espaces de fonctions bien choisis.

3.2.1 Espace des fonctions-tests. Définition des distributions.

Définition 3.1 : $\mathcal{D}(\mathbb{R}^n)$ (en général on écrira plus simplement \mathcal{D}) est l'espace des fonctions lisses, c'est-à-dire des fonctions de \mathbb{R}^n dans \mathbb{C} de classe C^∞ (infinitement différentiables), et de support borné.

Un exemple standard sur \mathbb{R} est fourni par la fonction de support (a, b) :

$$\phi(x) = \begin{cases} c \exp \frac{1}{(x-a)(x-b)} & \text{si } a < x < b \\ 0 & \text{sinon,} \end{cases} \quad (3.5)$$

voir figure 3.5.

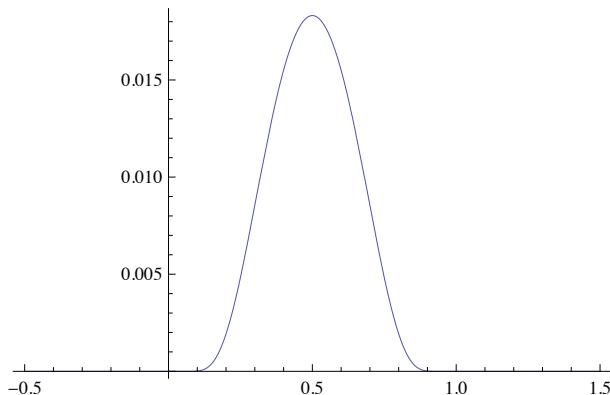


FIGURE 3.5 – La fonction (3.5) pour $a = 0, b = 1, c = 1$

Définition 3.2 : Une distribution est une forme linéaire continue sur \mathcal{D} . L'espace des distributions est noté \mathcal{D}' , c'est le dual de \mathcal{D} .

Si $T \in \mathcal{D}'$ est une distribution, on notera $\langle T, \varphi \rangle$ la forme linéaire $T(\varphi)$.

Mais qu'entend-on par forme linéaire (ou fonctionnelle) continue ?

Définition 3.3 : Une suite de fonctions-tests $\varphi_n \in \mathcal{D}$ converge dans \mathcal{D} vers $\varphi \in \mathcal{D}$ si les supports des φ_n sont contenus dans un même ensemble borné indépendant de n et si toutes les dérivées partielles des φ_n convergent uniformément vers celles de φ , $\partial^{(p)} \varphi_n \xrightarrow{\text{CVU}} \partial^{(p)} \varphi$.

Définition 3.4 : Une forme linéaire (ou “fonctionnelle”) sur \mathcal{D} est continue ssi pour toute suite de fonctions-tests $\varphi_n \in \mathcal{D}$ convergeant dans \mathcal{D} vers $\varphi \in \mathcal{D}$, $\langle T, \varphi_n \rangle \longrightarrow \langle T, \varphi \rangle$.

Ce point complète la définition des distributions. Retenir qu'une distribution doit être une fonctionnelle linéaire **et** continue. Ce dernier point est souvent le plus délicat à vérifier.

Commentaire sur le choix de l'espace \mathcal{D} . Pourquoi ce choix, qui semble assez restrictif ? Noter que plus on restreint la classe des fonctions-tests, plus grand sera l'espace des distributions définies sur ces fonctions¹. Le fait que les fonctions-tests soient de classe C^∞ n'est pas très restrictif, toute fonction continue à support borné étant limite uniforme de telles fonctions² ; le fait qu'elles soient de support borné sera levé plus tard : on considérera au Chap. 4 l'espace \mathcal{S} des fonctions-tests *lisses* (c'est-à-dire de classe C^∞) et à "décroissance rapide" ainsi que toutes leurs dérivées (\mathcal{S} pour Schwartz), et l'espace dual \mathcal{S}' qui en résultera sera plus petit que \mathcal{D}' .

• Distributions régulières

Définition 3.5 : Une fonction mesurable f de \mathbb{R}^n dans \mathbb{C} est localement intégrable si pour tout compact $K \subset \mathbb{R}^n$, la fonction $f \cdot \chi_K$ est intégrable, autrement dit $\int_K f(x)dx$ existe. On note $L_{\text{loc}}^1(\mathbb{R}^n)$ l'ensemble des fonctions localement intégrables sur \mathbb{R}^n .

Si la propriété d'intégrabilité est vraie sur tout \mathbb{R}^n , la fonction est *intégrable*. Par exemple la fonction $1/\sqrt{x}$, ou la fonction constante 1, sont localement intégrables sur \mathbb{R} mais pas intégrables. La fonction $1/x$ n'est pas localement intégrable. Toute fonction de $\mathcal{L}^p(\mathbb{R}^n)$, $p \geq 1$, est localement intégrable.

Toute fonction localement intégrable définit une distribution par intégration :

Théorème 3.1 : Si f est localement intégrable, elle définit une distribution, notée également f , par

$$\forall \varphi \in \mathcal{D} \quad \langle f, \varphi \rangle = \int_{-\infty}^{\infty} f(x)\varphi(x)dx.$$

Cette distribution est appelée la distribution régulière associée à la fonction localement intégrable f .

Preuve : Il est clair que l'intégrale existe, puisque restreinte à un intervalle compact, le support de φ . La fonctionnelle est bien linéaire. Elle est continue car si la suite $\varphi_n \in \mathcal{D}$ converge vers $\varphi \in \mathcal{D}$, avec des supports contenus dans un compact K , $|\int f(\varphi_n - \varphi)| \leq \int |f| |\varphi_n - \varphi| \leq \|\varphi_n - \varphi\|_\infty \int_K |f| = M \|\varphi_n - \varphi\|_\infty$, où $M = \int_K |f|$ existe puisque f est localement intégrable. Mais l'hypothèse CVU : $\|\varphi_n - \varphi\|_\infty \rightarrow 0$ donc $\langle f, \varphi_n \rangle \rightarrow \langle f, \varphi \rangle$. \square

Ces distributions régulières expliquent en quoi les distributions sont des "fonctions généralisées". Elles vont nous guider dans la définition des opérations de dérivation, etc, des distributions.

1. au contraire du cas d'un espace vectoriel E de dimension finie, pour lequel $\dim E' = \dim E$, cf App. B

2. c'est le théorème de Stone–Weierstrass, une généralisation du théorème de Weierstrass cité à l'App. A.4 du Chap. 1.

Noter que deux fonctions localement intégrables égales p.p. définissent la même distribution régulièrre.

• Distributions singulières

Ce sont les distributions que l'on ne peut pas associer à des fonctions localement intégrables. Exemples :

1. Distribution de Dirac sur \mathbb{R} . C'est la distribution notée δ_{x_0} ou encore $\delta(x - x_0)$ définie par

$$\langle \delta_{x_0}, \varphi \rangle = \varphi(x_0) \quad (3.6)$$

comme anticipé plus haut dans les exemples physiques.

2. Distribution de Dirac sur \mathbb{R}^n : *ibidem*, définie par $\langle \delta_{\mathbf{r}_0}, \varphi \rangle = \varphi(\mathbf{r}_0)$.

3. Peigne de Dirac, cf § 3.1.2. Pour $\varphi \in \mathcal{D}(\mathbb{R})$, c'est la distribution définie par

$$\langle \text{III}, \varphi \rangle \stackrel{\text{def}}{=} \sum_{n=-\infty}^{\infty} \varphi(n). \quad (3.7)$$

On peut donc écrire $\text{III} = \sum_{n=-\infty}^{\infty} \delta_n$.

Exercice : calculer la somme $\Sigma(x) = \sum_{k=-\infty}^{\infty} e^{2\pi i k x}$ en la testant sur une fonction φ périodique de période 1 donnée par son développement en série de Fourier : on n'intégrera d'abord que sur une période puis on rétablira la périodicité ; vérifier que $\Sigma(x)$ est égal à $\text{III}(x)$. Comparer avec la discussion du § 3.1.2.

4. Partie principale de Cauchy de $1/x$. Soit $\varphi \in \mathcal{D}(\mathbb{R})$. On définit la distribution $\text{PP} \frac{1}{x}$, *partie principale de Cauchy*, (certains auteurs parlent de “valeur principale”), par

$$\langle \text{PP} \frac{1}{x}, \varphi \rangle = \text{PP} \int \frac{\varphi(x)}{x} dx \stackrel{\text{def}}{=} \lim_{\epsilon \rightarrow 0^+} \int_{|x|>\epsilon} \frac{\varphi(x)}{x} dx. \quad (3.8)$$

(On note aussi l'intégrale de partie principale $\int \frac{\varphi(x)}{x} dx$.) Bien noter que l'intégrale est calculée de façon symétrique en $x = 0$. On définit par extension la $\text{PP} \frac{1}{x-a}$ en tout point a .

• Support d'une distribution

Soit O un ouvert de \mathbb{R}^n , $\mathcal{D}(O)$ l'espace des fonctions lisses de support contenu dans O et $T \in \mathcal{D}'(O)$ une distribution. On dit que T est nulle sur un ouvert $U \subset O$ si $\forall \varphi \in \mathcal{D}(O)$ de support contenu dans U , on a $\langle T, \varphi \rangle = 0$.

Définition 3.6 : On appelle support d'une distribution T sur O le complémentaire du plus grand ouvert sur lequel T est nulle.

Remarque : Le support est bien défini, car si une distribution est nulle sur chacun des ouverts d'une famille, elle est nulle sur leur réunion ; son support est donc le complémentaire de la réunion de tous les ouverts sur lesquels elle est nulle.

Exemples : Si T est une distribution régulière associée à une fonction continue, alors le support qu'on vient de définir s'identifie au support défini précédemment pour les fonctions continues. Si T est une distribution δ ou n'importe laquelle de ses dérivées, son support se réduit au point 0.

3.3 Opérations sur les distributions

On se propose de définir sur les distributions un certain nombre d'opérations : translation, réflexion, dilatation de la variable, dérivation, changement de variable, ... La méthode est d'écrire d'abord ces opérations sur des distributions régulières en les exprimant dans le langage des distributions, puis de les étendre comme définitions dans les cas non réguliers.

3.3.1 Translation, dilatation

Pour une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$, on définit l'action sur f d'une translation de la variable $x \rightarrow x' = x + a$ par $f_a(x') = f(x)$, soit $f_a(x) = f(x - a)$. Pour une fonction f localement intégrable et $\varphi \in \mathcal{D}$, on a donc $\langle f_a, \varphi \rangle = \int_{\mathbb{R}} f(x - a)\varphi(x)dx = \int f(x)\varphi(x + a)dx = \langle f, \varphi_{-a} \rangle$. En général, pour toute distribution $T \in \mathcal{D}'$, on définit sa translatée T_a par

$$\langle T_a, \varphi \rangle \stackrel{\text{déf}}{=} \langle T, \varphi_{-a} \rangle \quad (3.9)$$

qui est bien une fonctionnelle linéaire et continue.

Exemple : la partie principale $\text{PP}_{x-a}^{\frac{1}{x-a}}$ est la translatée de $\text{PP}_x^{\frac{1}{x}}$.

Cette opération de translation s'étend bien évidemment à \mathbb{R}^n .

Pour le changement de variable par dilatation, $x \mapsto x' = x/k$, on procède de même. On définit d'abord la transformation de la fonction $f : \mathbb{R} \rightarrow \mathbb{R}$, $f_{(k)}(x') = f(x)$, donc $f_{(k)}(x) = f(kx)$, puis $\langle f_{(k)}, \varphi \rangle = \int_R f(kx)\varphi(x)dx = \dots = \frac{1}{|k|} \langle f, \varphi_{(k^{-1})} \rangle$, qu'on généralise à une distribution quelconque $T \in \mathcal{D}'$ en posant par définition

$$\langle T_{(k)}, \varphi \rangle = \frac{1}{|k|} \langle T, \varphi_{(k^{-1})} \rangle, \quad (3.10)$$

ou encore

$$\langle T(kx), \varphi(x) \rangle = \frac{1}{|k|} \left\langle T(x), \varphi\left(\frac{x}{k}\right) \right\rangle. \quad (3.11)$$

Ainsi par exemple, pour la distribution δ de Dirac, on écrira

$$\delta(kx) = \frac{1}{|k|} \delta(x). \quad (3.12)$$

À nouveau, l'extension à \mathbb{R}^n est aisée, mais attention, le jacobien est maintenant $\frac{1}{|k|^n}$!

$$\text{dans } \mathbb{R}^n \quad \langle T_{(k)}, \varphi \rangle = \frac{1}{|k|^n} \langle T, \varphi^{(k^{-1})} \rangle. \quad (3.13)$$

3.3.2 Dérivation d'une distribution

On part d'une fonction f localement intégrable ainsi que sa dérivée : $f, f' \in L^1_{\text{loc}}(\mathbb{R})$. On écrit alors pour la distribution régulière associée à f'

$$\langle f', \varphi \rangle = \int_{-\infty}^{\infty} f'(x) \varphi(x) dx = - \int_{-\infty}^{\infty} f(x) \varphi'(x) dx = -\langle f, \varphi' \rangle, \quad (3.14)$$

où on a utilisé une intégration par parties, sans terme de bord puisque φ est à support borné. Cela suggère la

Définition 3.7 : La dérivée d'une distribution T de \mathcal{D}' est définie par

$$\boxed{\langle T', \varphi \rangle \stackrel{\text{déf}}{=} -\langle T, \varphi' \rangle} \quad (3.15)$$

(Elle est appelée d-dérivée dans les TD)

Cette définition, que nous allons beaucoup utiliser, s'étend très naturellement aux dérivées d'ordre supérieur, mais aussi aux dérivées partielles pour des distributions sur \mathbb{R}^n , etc. Nous nous intéresserons plus bas au laplacien de $T \in \mathcal{D}'(\mathbb{R}^n)$

$$\varphi \in \mathcal{D} \quad : \quad \langle \Delta T, \varphi \rangle = \langle T, \Delta \varphi \rangle.$$

On voit que la dérивabilité des distributions ne pose pas de problème. En fait même si une fonction (localement intégrable) n'est pas dérivable comme fonction, elle est toujours dérivable en tant que distribution ! Exemple : la fonction “saut” de Heaviside,

$$H(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{2} & \text{si } x = 0 \\ 1 & \text{si } x > 0 \end{cases} \quad (3.16)$$

est une fonction localement intégrable ; elle est discontinue en 0 et n'est donc pas dérivable en ce point. Sa dérivation en tant que distribution, cependant, ne pose pas de problème :

$$\langle H', \varphi \rangle = -\langle H, \varphi' \rangle = - \int_0^{\infty} \varphi'(x) dx = [-\varphi]_0^{\infty} = \varphi(0) = \langle \delta, \varphi \rangle$$

donc

$$\boxed{H' = \delta} \quad (3.17)$$

La dérivée de H est la distribution de Dirac ! Noter que selon une remarque faite plus haut, on peut modifier la “valeur” de H en $x = 0$ sans modifier la distribution H .

3.4 Distribution delta et distributions reliées

3.4.1 Fonction de Heaviside, fonction signe

On vient de rencontrer la distribution de Heaviside $H(x)$ ou fonction saut, notée aussi souvent $\theta(x)$ par les physiciens. Une fonction ou distribution reliée est la *fonction signe*

$$\epsilon(x) \equiv \text{sgn}(x) = -1 + 2H(x) = \begin{cases} -1 & \text{si } x < 0 \\ 0 & \text{si } x = 0 \\ 1 & \text{si } x > 0 \end{cases} \quad (3.18)$$

C'est au signe près et à une dilatation verticale près la fonction qui décrit la discontinuité de vitesse dans l'exemple du choc du § 3.1.3.

Par le même calcul que plus haut, on a $\epsilon' = 2\delta$.

Les dérivées successives de la distribution δ de Dirac se définissent de même et peuvent être rencontrées en physique.

3.4.2 Relations fonctionnelles

- Anticipant un peu sur la suite, on peut définir le produit de toute distribution T par une fonction ψ de classe C^∞ : $\langle T\psi, \varphi \rangle = \langle T, \psi\varphi \rangle$, voir ci-dessous § 3.5. L'opération est particulièrement simple pour la distribution δ : $\delta(x)\psi(x) = \psi(0)\delta(x)$. En particulier on a $x\delta = 0$; réciproquement, il est naturel de se demander : quelles distributions $T \in \mathcal{D}'$ sont telles que $xT = 0$? ou telles que $x^n T = 0$? Voir TD.

- Composition d'une distribution avec une fonction f

Rappelons d'abord la formule de changement de variable dans une intégrale de fonctions ordinaires (cf Chap. 2). Soient g une fonction localement intégrable et $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction dérivable et monotone, donc inversible au sens des fonctions (f^{-1} existe). On supposera aussi que f' ne s'annule pas. La composée $g \circ f$ est localement intégrable et pour toute $\varphi \in \mathcal{D}$

$$\langle g \circ f, \varphi \rangle = \int_{\mathbb{R}} g(f(x)) \varphi(x) dx = \int_{\mathbb{R}} g(y) \varphi(f^{-1}(y)) \frac{dy}{|f'(f^{-1}(y))|} = \left\langle \frac{g}{|f' \circ f^{-1}|}, \varphi \circ f^{-1} \right\rangle$$

où le dénominateur vient du jacobien dans le changement de variable $x \mapsto y = f(x)$

$$J = \left| \frac{dx}{dy} \right| = |f'(x)|^{-1} = |f'(f^{-1}(y))|^{-1}.$$

De la même façon, pour une distribution $T \in \mathcal{D}'$, on définit

$$\langle T \circ f, \varphi \rangle \stackrel{\text{déf}}{=} \left\langle \frac{T}{|f' \circ f^{-1}|}, \varphi \circ f^{-1} \right\rangle. \quad (3.19)$$

C'est en particulier le cas pour la distribution δ

$$\langle \delta \circ f, \varphi \rangle = \left\langle \frac{\delta}{|f' \circ f^{-1}|}, \varphi \circ f^{-1} \right\rangle = \frac{\varphi(f^{-1}(0))}{|f'(f^{-1}(0))|}, \quad (3.20)$$

ou encore

$$\delta(f(x)) = \frac{1}{|f'(x_0)|} \delta(x - x_0) \quad (3.21)$$

avec x_0 l'*unique* zéro de la fonction (supposée monotone) f . (Si f ne couvre pas tout \mathbb{R} et n'a pas de zéro, l'intégrale est nulle). Noter qu'on a besoin de l'hypothèse que f' ne s'annule pas au zéro x_0 de f : $f'(x_0) \neq 0$. Ainsi $f(x) = x^3$ est exclue.

On peut généraliser au cas d'une fonction f n'ayant que des zéros isolés $x_i \in \mathbb{R}$ où sa dérivée ne s'annule pas. Il suffit de couper le domaine d'intégration en sous-domaines où f est monotone, et la formule (3.21) donne alors

$$\delta(f(x)) \stackrel{\text{déf}}{=} \sum_{\substack{x_i \\ f(x_i)=0}} \frac{1}{|f'(x_i)|} \delta(x - x_i). \quad (3.22)$$

Cette identité est *très* utile au physicien, on la rencontrera chaque fois qu'on imposera une contrainte par l'intermédiaire d'une distribution δ , voir ci-dessous § 3.4.3 des exemples d'application.

On peut retrouver la relation (3.22) par un argument qualitatif ("heuristique") : la distribution $\delta \circ f$ "localise" l'intégration dans $\langle \delta \circ f, \varphi \rangle$ au voisinage des zéros x_i de f . Au voisinage de chacun de ces zéros, on développe au premier ordre non nul $f(x) \approx f(x_i) + (x - x_i)f'(x_i) + \dots = (x - x_i)f'(x_i) + \dots$ puisque $f(x_i) = 0$ et qu'on suppose $f'(x_i) \neq 0$; on conçoit que l'on puisse écrire $\delta(f(x)) = \sum_i \delta(f'(x_i)(x - x_i)) = \sum_i \frac{\delta(x - x_i)}{|f'(x_i)|}$ (par (3.12)), qui n'est autre que (3.22).

• δ comme limite de fonctions variées

Définition 3.8 : (Convergence faible dans \mathcal{D}') : Une suite T_n de distributions de \mathcal{D}' converge faiblement dans \mathcal{D}' si $\forall \varphi \in \mathcal{D}$, la suite $\langle T_n, \varphi \rangle$ converge dans \mathbb{C} .

Cela définit une fonctionnelle linéaire T , on peut montrer que cette fonctionnelle est continue, c'est donc une distribution de \mathcal{D}' et on écrit $T_n \xrightarrow{\mathcal{D}'} T$.

On montre aussi la compatibilité entre limite faible et dérivation *au sens des distributions* : si $T_n \xrightarrow{\mathcal{D}'} T$, il y a aussi convergence faible de toutes les dérivées $T_n^{(p)} \xrightarrow{\mathcal{D}'} T^{(p)}$.

Exemples. Considérons les deux suites de fonctions

- gaussiennes $f_n(x) = \sqrt{\frac{n}{2\pi}} e^{-\frac{nx^2}{2}}$, ce sont des fonctions de classe C^∞ ;
- fonctions "porte" $g_n(x) = \frac{n}{2}(H(x + \frac{1}{n}) - H(x - \frac{1}{n}))$, égales à $\frac{n}{2}$ pour $-\frac{1}{n} \leq x \leq \frac{1}{n}$, nulles ailleurs : ce sont des fonctions en escalier.

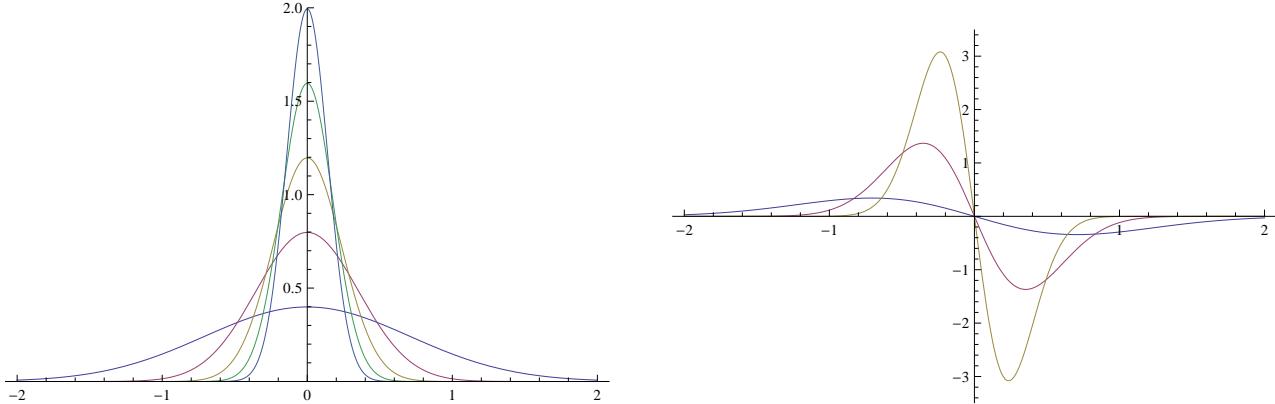


FIGURE 3.6 – Les fonctions gaussiennes f_n de plus en plus piquées, de $n = 1$ à $n = 5$, et leurs dérivées premières de $n = 1$ à $n = 3$.

Ces fonctions f_n et g_n sont positives et normalisées par $\int f_n dx = 1$, $\int g_n dx = 1$. Elles prennent des valeurs $\geq O(1)$ (en un sens qu'il faudrait préciser) dans un intervalle de longueur tendant vers 0 quand $n \rightarrow \infty$. Elles définissent des distributions régulières qui tendent faiblement dans \mathcal{D}' vers δ quand $n \rightarrow \infty$. Inversement elles peuvent constituer des “régularisations” utiles de la distribution de Dirac.

Exercice : étudier analytiquement et avec le logiciel Maple ou Mathematica la convergence des fonctions trigonométrico-rationnelles : $h_n(x) = \sin^2(\pi nx)/n\pi x^2$ et de leurs dérivées. Le cas des fonctions $\sin nx/\sin x$ rencontrées dans l'introduction, ou celui des $\sin nx/x$ (cf [1] p. 210) est plus délicat, mais permet aussi de construire les distributions III ou δ par une limite.

3.4.3 δ sur une courbe, une surface, . . .

- **Courant linéique**

On apprend en électromagnétisme qu'une particule de charge q suivant une trajectoire d'espace $\mathbf{x}(t)$ crée en tout point $y = (\mathbf{y}, t)$ un vecteur courant d'expression

$$\mathbf{j}(\mathbf{y}, t) = q \frac{d\mathbf{x}(t)}{dt} \delta^3(\mathbf{y} - \mathbf{x}(t)).$$

On peut aussi écrire la densité de charge au point y comme

$$\rho(\mathbf{y}, t) = q \delta^3(\mathbf{y} - \mathbf{x}(t)).$$

Cette formule admet une intéressante généralisation relativiste. Une particule de charge q suivant une trajectoire d'espace-temps $x^\mu(\tau)$, avec τ le temps propre, $ds^2 = c^2 d\tau^2$, crée en tout point $y = (\mathbf{y}, t)$ d'espace-temps un quadri-vecteur courant d'expression

$$j^\mu(\mathbf{y}, t) = q \frac{dx^\mu(\tau)}{d\tau} \delta^3(\mathbf{y} - \mathbf{x}(\tau))|_{t=x^0(\tau)} = q \int d\tau \frac{dx^\mu(\tau)}{d\tau} \delta^4(y - x(\tau)).$$

Noter que dans la 3ème expression on a récrit le courant sous forme explicitement covariante, en passant d'une distribution de Dirac à trois dimensions δ^3 à une distribution δ^4 . Exercice : vérifier l'équivalence entre les 2ème et 3ème expressions en utilisant (3.21). Ce courant a une composante temporelle, la densité de charge, $j^0(\mathbf{y}, t) = q\delta^3(\mathbf{y} - \mathbf{x}(\tau))|_{t=x^0(\tau)}$ qui satisfait bien $\int d^3y j^0(\mathbf{y}, t) = q$ et on vérifie qu'il est de (quadri)divergence nulle : $\partial_\mu j^\mu(y) = 0$.

On pourrait discuter de même une distribution "surfique" de charges, etc.

- **Particule relativiste sur sa "couche de masse"**

En mécanique relativiste, on apprend qu'une particule de masse au repos m a une énergie E et une impulsion \mathbf{p} reliées par la relation $E^2 = \mathbf{p}^2 c^2 + m^2 c^4$ qui se réduit, pour une particule au repos, au célèbre $E = mc^2$. On peut redire cela en termes du quadri-vecteur impulsion $p = (p^0 = E/c, \mathbf{p})$ dont la longueur minkowskienne carrée est $p^2 = (p^0)^2 - \mathbf{p}^2 = m^2 c^2$. On choisit généralement des unités où $c = 1$ et finalement

$$p^2 = (p^0)^2 - \mathbf{p}^2 = m^2. \quad (3.23)$$

On a parfois à intégrer sur des quadri-impulsions contraintes à satisfaire la condition (3.23), c'est-à-dire à être sur la "couche de masse" dans le jargon des physiciens, en fait un hyperbolôde à deux nappes dans l'espace de Minkowski. On peut imposer cela en insérant dans l'intégration une distribution $\delta(p^2 - m^2)$. En suivant (3.22) on récrit

$$\delta(p^2 - m^2) = \frac{\delta(p^0 - \sqrt{\mathbf{p}^2 + m^2})}{2\sqrt{\mathbf{p}^2 + m^2}} + \frac{\delta(p^0 + \sqrt{\mathbf{p}^2 + m^2})}{2\sqrt{\mathbf{p}^2 + m^2}}$$

mais on accompagne en général la condition (3.23) de la condition que l'énergie p^0 est positive, en d'autres termes on s'intéresse à

$$\delta(p^2 - m^2) H(p^0) = \frac{\delta(p^0 - \sqrt{\mathbf{p}^2 + m^2})}{2\sqrt{\mathbf{p}^2 + m^2}}.$$

La mesure d'intégration sur la couche de masse s'écrit donc

$$d^4p \delta(p^2 - m^2) H(p^0) = \frac{d^3p}{2\sqrt{\mathbf{p}^2 + m^2}} \quad (3.24)$$

où l'expression du membre de gauche, qui est évidemment invariante relativiste, nous garantit que celle du membre de droite l'est aussi. Il faut comprendre (3.24) au sens des distributions, c'est-à-dire testé dans une intégration avec une fonction $\varphi \in \mathcal{D}(\mathbb{R}^4)$ arbitraire

$$\int d^4p \delta(p^2 - m^2) H(p^0) \varphi(p^0, \mathbf{p}) = \int \frac{d^3p}{2\sqrt{\mathbf{p}^2 + m^2}} \varphi(\sqrt{\mathbf{p}^2 + m^2}, \mathbf{p}).$$

3.5 Produit de distributions. Convolution

Il faut faire attention à ce que l'on entend par produit de distributions. Plusieurs définitions sont en effet possibles

1. Le produit d'une distribution T par une fonction ψ de classe C^∞ ne pose pas de problème : $\langle T\psi, \varphi \rangle \stackrel{\text{déf}}{=} \langle T, \psi\varphi \rangle$, puisque si $\varphi \in \mathcal{D}$, $\psi\varphi \in \mathcal{D}$ et l'expression de droite est bien définie, c'est une fonctionnelle continue et linéaire de φ , elle définit donc la distribution $T\psi$.
Exemple : pour toute fonction $\psi \in C^\infty$, $\delta(x)\psi(x) = \delta(x)\psi(0)$, $\delta_a(x)\psi(x) = \delta_a(x)\psi(a)$.
2. Pour deux fonctions localement intégrables f et g de $\mathbb{R} \rightarrow \mathbb{R}$, (et donc pour les distributions régulières associées) on peut définir aussi un produit “direct” (ou “tensoriel”) $f \otimes g$ qui est une fonction $\mathbb{R}^2 \rightarrow \mathbb{R}$ définie très naturellement par $(f \otimes g)(x, y) = f(x)g(y)$. Plus généralement, pour deux distributions S et T de $\mathcal{D}'(\mathbb{R})$, l'une “agissant” sur la variable x , l'autre sur y , ce qu'on note par $S(x) \otimes T(y)$ est la distribution de $\mathcal{D}'(\mathbb{R}^2)$ agissant sur des fonctions $\varphi(x, y)$ par $\langle S \otimes T, \varphi \rangle = \int dx dy S(x)T(y)\varphi(x, y)$. Exemple, $\delta(x) \otimes \delta(y)$ n'est autre que la distribution notée plus haut $\delta^2(\mathbf{x})$.
3. En revanche, le produit de deux distributions $T_1, T_2 \in \mathcal{D}'(\mathbb{R})$ n'est en général pas défini. Cela apparaît déjà sur le produit de deux fonctions localement intégrables, qui n'a aucune raison d'être lui-même localement intégrable : ainsi $\psi(x) = \frac{1}{\sqrt{x}}$ est localement intégrable, mais $\psi^2(x) = \frac{1}{x}$ ne l'est pas. En général les distributions, rappelons-le, prennent leur sens quand elles sont testées sur des fonctions régulières (appartenant à \mathcal{D}), **pas sur d'autres distributions**. Rien d'étonnant donc à ce que le produit inconsidéré de deux distributions amène à des résultats incontrôlables ... L'exemple le plus simple est donné par le produit de deux distributions de Dirac : $\delta\delta = ?\delta(0)\delta$ qui n'est pas défini.
On rencontrera ce genre de situations en physique, en Mécanique Quantique par exemple dans le calcul de la “règle d'or de Fermi”, ou en théorie quantique des champs où apparaissent des “divergences ultraviolettes”, nécessitant la mise en œuvre de l'opération de renormalisation. Quand il est confronté à cette situation, le physicien doit chercher à revenir au problème de départ et à “régulariser” la quantité responsable de l'apparition des infinis. Ainsi une distribution delta pourra être remplacée par une courbe en cloche très pointue, telle une gaussienne de largeur tendant vers 0, comme on a vu plus haut.
4. Finalement venons-en au produit le plus naturel sur des distributions, celui de convolution.

D'abord pour des fonctions f et g localement intégrables, on définit leur *produit de convolution* $f * g$ par l'intégrale suivante, si elle existe,

$$(f * g)(x) = \int dy f(x - y)g(y). \quad (3.25)$$

Il est aisément de trouver des conditions *suffisantes* pour que cette intégrale existe : par exemple si f et g sont définies p.p. et localement intégrables et de support borné à gauche, (resp. à droite, ou *a fortiori* à gauche et à droite !), l'intégrale en y est restreinte à un ensemble borné et existe presque partout, cf [5], p. 164

Exemples : La fonction de Heaviside H a un support borné à gauche (par 0). Son carré de convolution $H * H$ existe donc. En revanche, $H(x) * H(-x)$ n'existe pas.

Intuitivement, l'effet de la convolution d'une fonction f par une fonction $g \in \mathcal{D}$ est de "lisser" les singularités de f . Voir sur la figure 3.7 la convoluée d'une fonction "porte" χ par une fonction ϕ du type (3.5).

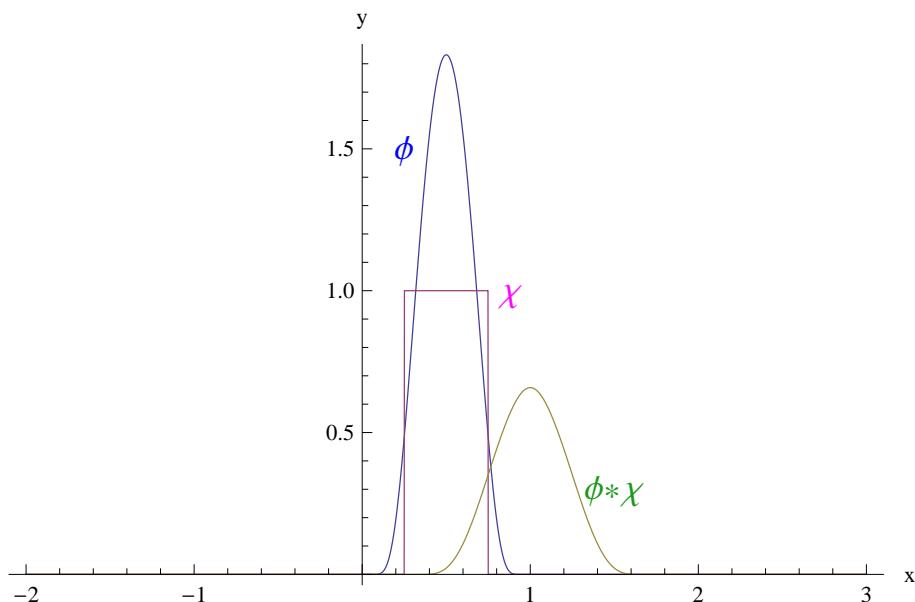


FIGURE 3.7 – Convoluée de la fonction porte $\chi_{[\frac{1}{4}, \frac{3}{4}]}$ avec la fonction (3.5) pour $a = 0$, $b = 1$. La convoluée a un support décalé vers la droite, pourquoi ?

Exercice (facile !) : Vérifier que le produit de convolution, s'il existe, est commutatif : $f * g = g * f$.

Une fois cette définition acquise, on peut envisager de l'étendre à des distributions. Pour deux fonctions localement intégrables f et g et $\varphi \in \mathcal{D}(\mathbb{R})$ considérons

$$\langle f * g, \varphi \rangle = \int dx (f * g)(x) \varphi(x) = \int dx \int dy f(x-y) g(y) \varphi(x) = \int dx \int dy f(x) g(y) \varphi(x+y)$$

par un changement de variable évident $x \rightarrow x+y$ de jacobien 1. Supposons que les supports de f et g sont tels que les domaines d'intégration sur x et sur y sont bornés. Par exemple si f est à support borné, comme φ l'est aussi, les intégrations sur x et sur y ont lieu sur des domaines bornés. Le théorème de Fubini s'applique et on peut écrire

$$\langle f * g, \varphi \rangle = \iint dx dy f(x) g(y) \varphi(x+y)$$

(= $\langle f \otimes g, \varphi(x+y) \rangle$ avec la notation \otimes du point 2 précédent.) Une fois encore nous généralisons cette relation à un produit de deux distributions.

Définition 3.9 : Le produit de convolution de deux distributions T_1 et T_2 est défini par l'intégrale suivante

$$\langle T_1 * T_2, \varphi \rangle \stackrel{\text{def}}{=} \langle T_1(x)T_2(y), \varphi(x+y) \rangle$$

si elle existe ! À nouveau on peut trouver des conditions suffisantes d'existence. Par exemple, (exercice !) T_1 et T_2 sont toutes deux à support borné à gauche (resp. à droite). Ou bien l'une des deux est à support borné. (Voir la définition du support d'une distribution au § 3.2.1.)

Exercice : vérifier que le produit de convolution de distributions est commutatif et associatif : $(T_1 * T_2) * T_3 = T_1 * (T_2 * T_3) = T_1 * T_2 * T_3$, à condition que $T_1 * T_2$, $T_1 * T_3$ et $T_2 * T_3$ soient définis. En revanche, examiner le cas de $1 * \delta' * H$ où 1 est la distribution régulière associée à la fonction constante 1 et montrer que dans ce cas, on n'a pas associativité !

Importance des convolutions en mathématiques et en physique.

L'intérêt des convolutions est d'abord d'ordre mathématique : nous verrons que par transformation de Fourier, produit de convolution et produit ordinaire sont échangés ; la convolution apparaît aussi dans la théorie des probabilités, etc. Mais ce produit est aussi très naturel dans les *systèmes linéaires*, en physique. Considérons un système physique, représenté par une “boîte noire”, qui peut être de nature mécanique, électrique, optique, acoustique, etc. Un signal d'entrée dépendant du temps, noté $X(t)$ (qui peut être multidimensionnel) est transformé par le système en un signal de sortie $Y(t)$ et la transformation $X \mapsto Y$ est supposée linéaire, mais non instantanée, $Y = \mathcal{K}X$. Donc $Y(t)$ à un temps t donné est fonction linéaire des $X(t')$, $t' < t$ pour respecter la causalité. On suppose que l'opération \mathcal{K} n'implique pas de dérivation (nous verrons plus tard comment lever cette restriction). On écrit donc $Y(t) = \int dt' K(t-t')X(t')$. Si on suppose de plus qu'il y a “invariance par translation dans le temps” $K(t+\tau, t'+\tau) = K(t, t')$, le “noyau” K ne dépend en fait que de la différence de ses arguments et, gardant la même notation K , on écrit finalement

$$Y(t) = \int dt' K(t-t')X(t')$$

qui n'est autre qu'un produit de convolution de K et de X . Pour une discussion plus approfondie de cette problématique, voir [8] et [1].

Autre exemple rencontré au début de ce chapitre : le potentiel électrique créé par une distribution de charges de densité ρ est, voir (3.2), le produit de convolution de ρ par le noyau de Coulomb $\frac{1}{4\pi\epsilon_0} \frac{1}{|\mathbf{r}|}$. Cette fois, c'est l'invariance par translation spatiale qui jointe à la linéarité dicte la forme de convolution.

3.6 Exemple : Fonction de Green et potentiel de Coulomb en dimension d .

On apprend dans le cours d'électromagnétisme que le potentiel électrique V , ou le champ électrique \mathbf{E} qui en dérive, $\mathbf{E} = -\vec{\nabla}V$, satisfont à l'équation de Poisson

$$\operatorname{div}\mathbf{E} = -\Delta V = \frac{\rho}{\epsilon_0} \quad (3.26)$$

où ρ est la densité de charge.

Rappelons que cette équation de Poisson est l'expression locale du fait que le flux du vecteur \mathbf{E} à travers une surface fermée S est égale (au facteur $1/\epsilon_0$ près) à la charge totale Q contenue dans le volume englobé par S (théorème de Gauss-Ostrogradsky)

$$\iint d\mathbf{S} \cdot \mathbf{E} = \iiint d^3x \operatorname{div}\mathbf{E} = \iiint d^3x \frac{\rho}{\epsilon_0} = \frac{Q}{\epsilon_0}. \quad (3.27)$$

Pour une densité discrète de charges, c'est-à-dire localisée sur des charges ponctuelles en \mathbf{r}_i , cela nous apprend que ΔV s'annule, donc que V est une *fonction harmonique* (voir ci-dessous au Chap. 8), en tout point *en dehors des charges*.

Mais plus précisément, le potentiel coulombien V créé par les charges ponctuelles q_i en \mathbf{r}_i doit satisfaire

$$\Delta V(\mathbf{x}) = -\frac{1}{\epsilon_0} \sum_i q_i \delta^3(\mathbf{x} - \mathbf{r}_i), \quad (3.28)$$

en accord avec (3.27).

Généralisant le problème à d dimensions, cherchons une solution de l'équation

$$\Delta_x G(x - y) = \delta^d(x - y) \quad (3.29)$$

où $x, y \in \mathbb{R}^d$. (On donne le nom de “fonction de Green” à une telle solution.) Montrons que

$$G(x) = \begin{cases} -\frac{C_d^{-1}}{\|x\|^{d-2}} & \text{si } d \neq 2 \\ -\frac{1}{2\pi} \ln\left(\frac{1}{\|x\|}\right) & \text{si } d = 2 \end{cases} \quad (3.30)$$

est solution, où $C_d = (d-2)\Omega_{d-1}$, Ω_{d-1} l'aire de la sphère unité S^{d-1} dans \mathbb{R}^d :

$$\Omega_{d-1} = \frac{2\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})}. \quad (3.31)$$

Γ désigne une “fonction spéciale” que l'on étudiera en détail plus bas (Chap. 6 et 8). Qu'il suffise de dire ici que $\Gamma(x)$ généralise à x réel quelconque la fonction factorielle, $\Gamma(n) = (n-1)!$ si $n \in \mathbb{N}$ et pour $n = p + \frac{1}{2}$ demi-entier, $\Gamma(p + \frac{1}{2}) = \frac{(2p-1)(2p-3)\cdots 1}{2^p} \sqrt{\pi}$. On retrouve en particulier $\Omega_1 = 2\pi$, $\Omega_2 = 4\pi$ etc.

Remarque. Comme pour des fonctions de la seule variable r ,

$$\Delta = \frac{\partial^2}{\partial r^2} + \frac{d-1}{r} \frac{\partial}{\partial r} = \frac{1}{r^{d-1}} \frac{\partial}{\partial r} r^{d-1} \frac{\partial}{\partial r}, \quad (3.32)$$

(vérifier!), on pourrait penser que $\Delta_{\frac{1}{r^{d-2}}} = -(d-2) \frac{1}{r^{d-1}} \frac{\partial}{\partial r} (r^{d-1} \frac{1}{r^{d-1}}) \equiv 0$. En fait cela n'est vrai que pour $r \neq 0$, et il nous faut faire un calcul de dérivées au sens des distributions pour obtenir le résultat correct.

Preuve de (3.30) : Pour $d \neq 2$ et φ une fonction test de $\mathcal{D}(\mathbb{R}^d)$, on veut calculer $I = \int d^d x \varphi(x) \Delta_{\frac{1}{\|x\|^{d-2}}}$. Comme $\Delta_{\frac{1}{\|x\|^{d-2}}}$ est invariant par rotation et ne dépend que de la variable radiale $r = \|x\|$, seule contribue à cette intégrale la moyenne angulaire $\bar{\varphi}(r)$ de φ sur la sphère de rayon r

$$\Omega_{d-1} r^{d-1} \bar{\varphi}(r) = \int_{\|x\|=r} d^d x \varphi(x),$$

et $I = \int dr \Omega_{d-1} r^{d-1} \bar{\varphi}(r) \Delta_{\frac{1}{r^{d-2}}}$. Noter que $\bar{\varphi}$ est de classe C^∞ sur \mathbb{R}_+ , avec $\bar{\varphi}(0) = \varphi(0)$. On a alors

$$\begin{aligned} I &= \Omega_{d-1} \int_0^\infty dr r^{d-1} \bar{\varphi}(r) \frac{1}{r^{d-1}} \frac{\partial}{\partial r} r^{d-1} \frac{\partial}{\partial r} \frac{1}{r^{d-2}} = -(d-2) \Omega_{d-1} \int_0^\infty dr \bar{\varphi}(r) \frac{\partial}{\partial r} 1 \\ &\stackrel{\text{distrib.}}{=} \Omega_{d-1} (d-2) \int_0^\infty dr \bar{\varphi}'(r) = -\Omega_{d-1} (d-2) \varphi(0) = -C_d \varphi(0), \end{aligned}$$

comme annoncé en (3.30).

On effectue un calcul analogue en $d = 2$ (vérifier!).

On trouvera dans [8] ou [1] une autre démonstration, faisant appel aux dérivées (au sens des distributions) de fonctions discontinues.

On note que la solution (3.30) contient bien la loi de Coulomb en $1/r$ à trois dimensions, mais que le “potentiel coulombien à deux dimensions” est logarithmique !

Lectures complémentaires

La référence de base est bien sûr [9], ou dans une version plus proche de ce cours, [8]. L'exposé de [1] est très clair et complet.

Pour l'histoire du développement des distributions, lire Jean-Michel Kantor, *MATHÉMATIQUES D'EST EN OUEST, théorie et pratique : l'exemple des distributions*, disponible sur <http://www.math.jussieu.fr/~kantor/>

Appendice C. Rappels d'algèbre

Outre les structures algébriques classiques, groupe, anneau, corps, espace vectoriel, ..., on utilise dans les notes celle d'*algèbre* : une algèbre A sur un corps K est un espace vectoriel sur K qui a aussi un produit interne associatif qui en fait un anneau ; les structures d'espace

vectoriel et d'anneau sont compatibles en ce sens que la multiplication par un scalaire λ de K commute avec le produit interne ; et le produit interne est supposé bilinéaire. Donc

$$\begin{aligned} (\lambda x + \mu y)z &= \lambda xz + \mu yz && \text{bilinéarité} \\ x(\lambda y + \mu z) &= \lambda xy + \mu xz && " " " \\ x(yz) &= (xy)z && \text{associativité} \\ (\lambda x)y &= \lambda(xy) = x(\lambda y) \end{aligned} \tag{C.1}$$

Exemples : les polynômes d'une variable à coefficients dans \mathbb{R} forment une algèbre sur \mathbb{R} notée $\mathbb{R}[x]$; l'ensemble $M_n(\mathbb{R})$ des matrices $n \times n$ à éléments dans \mathbb{R} est une algèbre sur \mathbb{R} . On rencontrera au Chap. 6 l'algèbre des séries entières.

Dual d'un espace vectoriel (e.v.)

Soit E un e.v. sur \mathbb{R} . L'espace dual E' est par définition l'espace vectoriel des *formes linéaires* sur E , $T : E \rightarrow \mathbb{R}$, c'est-à-dire des applications linéaires de E dans \mathbb{R} . Si $\dim E = n$, soit e_i , $i = 1, \dots, n$, une base de E : $\forall X \in E$, $X = \sum_{i=1}^n x^i e_i$, et si T est une forme linéaire de E' , on note $T(X)$ ou $\langle T|X \rangle$ son action sur X , c'est une fonction linéaire des x^i : $T(X) = \langle T|X \rangle = \sum_i x^i t_i$ avec des coefficients $t_i \in \mathbb{R}$, ce qu'on peut récrire comme $T(X) = \langle T|X \rangle = \sum_{i,j} x^i t_j \langle f^j | e_i \rangle$ avec f^j tels que $\langle f^j | e_i \rangle = \delta_{ij}$. On montre aisément que ces f^j sont linéairement indépendants et qu'ils forment une base de E' , c'est la *base duale* des e_i dans E' . On note aussi que $\dim E' = \dim E = n$. Cette propriété qui est vraie en dimension finie ne l'est plus forcément en dimension infinie, comme on a vu au § 3.2.1.

Chapitre 4

Transformation de Fourier

On va étudier la transformation intégrale $\tilde{f}(k) = \int dx e^{ikx} f(x)$ qui fait passer d'une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ ou \mathbb{C} à une autre fonction $\tilde{f} : \mathbb{R} \rightarrow \mathbb{C}$. Elle étend la décomposition en série de Fourier $\sum_{n \in \mathbb{Z}} A_n e^{inx}$ des fonctions périodiques, dans laquelle n'apparaissent que des modes multiples d'un fondamental, en une superposition *continue* de modes. Cette transformation de Fourier est extrêmement utile et d'usage constant pour le physicien. Elle permet l'analyse en modes d'un signal, d'un phénomène. Elle joue donc un rôle naturel en optique, en cristallographie, en analyse du signal, ... Elle est aussi très utile dans l'étude des systèmes linéaires (c'est-à-dire régis des équations différentielles ou aux dérivées partielles linéaires) : ce n'est pas pour rien qu'elle porte le nom de Joseph Fourier, qui est un des premiers à avoir étudié l'équation de la chaleur. En physique quantique, elle relie des descriptions alternatives dans des variables x de position et $p = \hbar k$ d'impulsion. Elle intervient aussi en théorie des groupes, en probabilités, etc.

En préambule nous montrons comment l'équation des ondes conduit naturellement aux représentations de Fourier, puis nous rappelons la représentation d'une fonction périodique par sa série de Fourier.

4.0 Préambule physique. Équation des ondes

L'équation des ondes à une dimension s'écrit

$$\left(\frac{\partial^2}{\partial t^2} - v^2 \frac{\partial^2}{\partial x^2} \right) u(x, t) = 0. \quad (4.1)$$

Toute fonction de la forme $F(x+vt) + G(x-vt)$, F et G arbitraires (différentiables) est solution, et réciproquement, ces fonctions constituent les solutions générales de (4.1). Si on cherche une

solution sous la forme $u(x, t) = X(x)T(t)$, on trouve que

$$\frac{1}{X(x)} \frac{\partial^2 X(x)}{\partial x^2} = \frac{1}{v^2 T(t)} \frac{\partial^2 T(t)}{\partial t^2}$$

dont le membre de gauche (resp. de droite) doit être indépendant de t (resp. de x), donc au total, être une constante notée $-k^2$, $k \in \mathbb{C}$. Finalement on a à résoudre

$$\frac{1}{X} \frac{\partial^2 X}{\partial x^2} = \frac{1}{v^2 T} \frac{\partial^2 T}{\partial t^2} = -k^2 \quad (4.2)$$

dont les solutions sont des superpositions linéaires de

$$X(x) = e^{\pm ikx} \quad T(t) = e^{\pm ivkt}.$$

Pour k réel, $A(k) \exp(-ik(x - vt))$ décrit une onde progressive, monochromatique de vecteur d'onde k , se déplaçant vers la droite à la vitesse v . Mais toute superposition

$$\int dk A(k) \exp(-ik(x - vt))$$

est aussi solution de (4.1) : cette solution a une amplitude qui est la transformée de Fourier de l'amplitude $A(k)$ dans l'espace des vecteurs d'onde. Noter que la vitesse v est la vitesse de propagation de l'onde, lumineuse, acoustique, mécanique etc.

Équation d'une corde vibrante.

Cherchons maintenant une solution de (4.1) décrivant les vibrations d'une corde de longueur L et d'extrémités fixées. L'amplitude (transverse) de vibration de la corde est une fonction $u(x, t)$, $0 \leq x \leq L$, $t \geq 0$, sujette aux conditions aux limites

$$u(x, 0) = f(x) \quad \left. \frac{\partial u(x, t)}{\partial t} \right|_{t=0} = g(x) \quad u(0, t) = u(L, t) = 0, \quad (4.3)$$

où les fonctions f et g décrivent comment la corde est excitée à l'instant $t = 0$. Cherchant des solutions réelles et factorisées en $X(x)T(t)$ comme ci-dessus, on prend $k^2 \in \mathbb{R}$ et on écarte les cas où k^2 n'est pas positif car ils ne peuvent décrire une amplitude bornée d'une onde de vibration, $X(x) = A \cos kx + B \sin kx$, $T(t) = a \cos kvt + b \sin kvt$. (On notera $kv = \omega$ dans la suite.) Les conditions aux limites d'espace restreignent à $X(x) = B \sin kx$, $kL = n\pi$, $n \in \mathbb{N}$, donc

$$u(x, t) = \sum_{n=1}^{\infty} (a_n \cos \omega_n t + b_n \sin \omega_n t) \sin k_n x \quad k_n = \frac{n\pi}{L} \quad \omega_n = n \frac{\pi v}{L}$$

et les conditions initiales à $t = 0$ se reflètent dans les coefficients a_n et b_n via les développements en série de Fourier de f et g ¹

$$f(x) = \sum_{n=1}^{\infty} a_n \sin \frac{n\pi}{L} x \quad g(x) = \sum_{n=1}^{\infty} b_n \omega_n \sin \frac{n\pi}{L} x.$$

1. ou plus précisément les développements en série de Fourier des prolongements des fonctions f et g en fonctions impaires sur l'intervalle $[-L, L]$, périodiques de période $2L$.

Les deux types de développements que nous allons discuter maintenant sont donc intimement liés à la physique des ondes.

4.1 Séries de Fourier. Rappels

- Développements en série de Fourier : développement en cos et sin, développement en exponentielles.

On démontre en Analyse que toute fonction $f : \mathbb{R} \rightarrow \mathbb{C}$ “suffisamment régulière” (voir ci-dessous) qui est périodique de période 2π , $f(x + 2\pi) = f(x)$, peut être décomposée sur une base de fonctions périodiques de période 2π , $\{\cos kx, \sin kx\}_{k \in \mathbb{N}}$ ou $\{\exp ikx\}_{k \in \mathbb{Z}}$

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx) \quad (4.4)$$

ou encore

$$f(x) = \sum_{n=-\infty}^{\infty} A_n e^{inx}. \quad (4.5)$$

En utilisant les propriétés d’orthogonalité des fonctions trigonométriques ou exponentielles : si $n, m \in \mathbb{N}$

$$\int_{-\pi}^{\pi} \cos mx \cos nx dx = \pi(\delta_{mn} + \delta_{0n}\delta_{0m}), \quad \int_{-\pi}^{\pi} \sin mx \sin nx dx = \pi\delta_{mn}, \quad \int_{-\pi}^{\pi} \cos mx \sin nx dx = 0 \quad (4.6)$$

ou avec $n, p \in \mathbb{Z}$

$$\int_{-\pi}^{\pi} \frac{dx}{2\pi} e^{i(n-p)x} = \delta_{np}, \quad (4.7)$$

on voit que ces développements, s’ils existent, impliquent les relations suivantes

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx \quad a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx \quad (4.8)$$

et

$$A_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx. \quad (4.9)$$

Mais il s’agit de savoir si ces séries convergent, et si oui, si leur somme reproduit bien la fonction f donnée.

- Théorème de Dirichlet : condition suffisante de convergence

Théorème 4.1 : Si $f(x)$ est définie sur $[-\pi, \pi]$ et n’y a qu’un nombre fini de discontinuités finies et a ailleurs une dérivée continue, (fonction C^1 par morceaux), alors le développement

converge pour tout $x \in [-\pi, \pi]$; sa somme est $\frac{1}{2}(f(x_-) + f(x_+))$ en tout point de $]-\pi, \pi[$, et $\frac{1}{2}(f(-\pi_+) + f(\pi_-))$ en $\pm\pi$. Dans tout sous-intervalle où la fonction f est continue, la convergence de la série de Fourier vers f est uniforme.

- **Relation entre la régularité de la fonction et les propriétés asymptotiques des coefficients.**

Pour une fonction satisfaisant les conditions du théorème de Dirichlet, les coefficients a_n , b_n décroissent en valeur absolue au moins comme $|a_n| \sim 1/n$, $|b_n| \sim 1/n$ pour n grand, ce qui n'assure pas en général la convergence absolue de la série de Fourier, mais n'exclut pas la convergence simple pour certaines valeurs de x . Ces coefficients peuvent décroître plus rapidement, si la fonction est plus régulière. En fait on peut démontrer le

Théorème 4.2 (Stokes) : Si $f^{(k)}$ est la première des dérivées de f à avoir une discontinuité (ou un nombre fini de discontinuités) alors $|a_n|, |b_n|, |A_n| \sim \frac{1}{n^{k+1}}$.

Propriétés de parité, de conjugaison des coefficients

Si la fonction f est réelle, les coefficients a_n , b_n sont réels, $A_{-n} = A_n^*$ et $a_n = 2\Re e A_n$, $b_n = -2\Im m A_n$. Si la fonction f est paire (resp. impaire), les b_n (resp. a_n) s'annulent.

Phénomène de Gibbs. Considérons une fonction périodique, ayant un point de discontinuité x_0 sur sa période. Comment sa série de Fourier, somme de fonctions continues, approxime-t-elle cette discontinuité ? Une expérience numérique avec Mathematica sur une série de Fourier tronquée au 10ème, 20ème, 50ème ordre, montre que la somme présente au voisinage du point x_0 une variation de plus en plus rapide, qui approche le saut de manière excessive ! C'est le très intéressant phénomène de Gibbs, pour lequel nous renvoyons à la littérature, par exemple [1].

4.2 Transformation de Fourier dans L^1

4.2.1 Définition. Conventions

Si f est une fonction $\mathbb{R} \rightarrow \mathbb{R}$ (ou $\rightarrow \mathbb{C}$), on définit

$$\boxed{\tilde{f}(k) = \int_{-\infty}^{\infty} dx e^{ikx} f(x)} \quad (4.10)$$

avec $k \in \mathbb{R}$, si cette intégrale existe (voir plus bas), et on appelle \tilde{f} , notée aussi $\tilde{f} = \mathcal{F}[f]$, la transformée de Fourier de la fonction f . C'est une fonction $\mathbb{R} \rightarrow \mathbb{C}$.

- **Autres conventions. Extension à \mathbb{R}^n**

On rencontre plusieurs conventions dans la définition de cette transformée de Fourier, selon le domaine scientifique concerné.

- signe de l'exponentielle $\int_{-\infty}^{\infty} dx e^{\pm ikx} f(x)$
- normalisation globale : $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx e^{ikx} f(x)$ ou $\frac{1}{2\pi} \int_{-\infty}^{\infty} dx e^{ikx} f(x)$
- normalisation de la variable k : $\int_{-\infty}^{\infty} dx e^{2\pi ikx} f(x)$
- et diverses combinaisons de ces variantes, dont chacune a ses avantages, comme on verra.

Donc attention en utilisant des formules puisées dans la littérature !!

La transformation de Fourier s'étend sans difficulté à des fonctions $\mathbb{R}^n \rightarrow \mathbb{R}$ (ou $\rightarrow \mathbb{C}$) : on écrit pour $\mathbf{k} \in \mathbb{R}^n$,

$$\tilde{f}(\mathbf{k}) = \int_{-\infty}^{\infty} d^n x e^{i\mathbf{k} \cdot \mathbf{x}} f(\mathbf{x}) \quad (4.11)$$

sous réserve que cette intégrale soit bien définie.

- **Premiers exemples.**

Transformée de Fourier de

- une fonction “porte” $\Pi(x) = \chi_{[-1,1]}(x)$, la fonction indicatrice de l'intervalle $[-1, 1]$. On calcule (exercice : vérifier !) que $\tilde{\Pi}(k) = 2 \frac{\sin k}{k}$. Noter que cette fonction $\tilde{\Pi}$ **n'est pas** intégrable (de Lebesgue) sur \mathbb{R} , même si son intégrale est “semi-convergente” ;
- une fonction “lorentzienne” $f(x) = 2a/(x^2 + a^2)$ où $a > 0$, $\tilde{f} = 2\pi e^{-|k|a}$, comme on le verra plus tard par un calcul de résidu (Chap. 8, équ. (8.4)) ;
- une fonction gaussienne $f(x) = \exp -\frac{(x-a)^2}{2b^2}$ a pour transformée de Fourier

$$\tilde{f}(k) = \sqrt{2\pi} |b| e^{iak - \frac{1}{2}k^2 b^2}$$

comme on le trouve par un changement de la variable x par une translation complexe $x' = x - (a + ikb^2)$, ce qu'on justifiera au chapitre 7 par un “changement de contour” ; etc. L'analyse complexe et ses méthodes puissantes nous seront de grande utilité pour les calculs de transformées de Fourier.

Certaines de ces fonctions et de ces formules vont jouer un rôle dans la suite : on rencontre les fonctions lorentziennes en probabilités et en mécanique quantique, les gaussiennes sont omniprésentes en probabilités et statistiques, en mécanique statistique, etc.

4.2.2 Existence et premières propriétés

On se rappelle la définition au Chap. 2 de l'espace $\mathcal{L}^1(\mathbb{R})$ des fonctions intégrables de Lebesgue (c'est-à-dire absolument intégrables sur $(-\infty, \infty)$), et de l'espace $L^1(\mathbb{R})$ de ces mêmes

fonctions définies à égalité p.p. près. On introduit une nouvelle notation commode : $L^\infty(\mathbb{R})$ désigne l'espace (de Banach) des fonctions définies p.p. et bornées sur \mathbb{R} .

Théorème 4.3 : Si $f \in L^1(\mathbb{R})$, \tilde{f} est définie et continue sur \mathbb{R} . Elle est bornée, donc $\tilde{f} \in L^\infty$ et $\|\tilde{f}\|_\infty \leq \|f\|_1$.

Cela découle du théorème 2.9 du chapitre 2, sur la continuité d'une fonction définie par une intégrale : ici, $f(x)e^{ikx}$ est bien dominée par $|f(x)|$ qui est intégrable. En outre on a $|\tilde{f}(k)| = \left| \int_{-\infty}^{\infty} dx e^{ikx} f(x) \right| \leq \int_{-\infty}^{\infty} dx |f(x)| = \|f\|_1$ pour tout $k \in \mathbb{R}$, d'où $\|\tilde{f}\|_\infty = \sup_{k \in \mathbb{R}} |\tilde{f}(k)| \leq \|f\|_1$.

Bien noter que le fait qu'une fonction de L^1 est définie p.p. n'influe pas sur l'intégrale définissant \tilde{f} . Ainsi, pour la fonction porte d'un exemple ci-dessus, qui est non définie en ± 1 , l'intégrale \tilde{f} est bien définie. (En revanche, sa transformée de Fourier $\tilde{f} = \sin k/k$ est bornée, $|\tilde{f}| \leq 1$, $\tilde{f} \in L^\infty$, mais non intégrable de Lebesgue, $\tilde{f} \notin L^1$.)

Théorème 4.4 : La transformation de Fourier est une opération linéaire et continue de $L^1(\mathbb{R})$ dans $L^\infty(\mathbb{R})$.

Qu'elle est continue signifie que pour toute suite de fonctions f_n tendant (au sens de la norme $\|\cdot\|_1$) dans L^1 vers une fonction $f \in L^1$, $\mathcal{F}[f_n] \rightarrow \mathcal{F}[f]$, au sens de la norme $\|\cdot\|_\infty$. Par linéarité, cela découle du Théorème 4.3 : $\|\mathcal{F}[f_n] - \mathcal{F}[f]\|_\infty = \|\mathcal{F}[f_n - f]\|_\infty \leq \|f_n - f\|_1 \rightarrow 0$. On peut réexprimer ce théorème en disant : *la transformée de Fourier transforme la convergence en moyenne (au sens de L^1) en convergence uniforme (au sens de L^∞)*.

Noter encore que la transformation de Fourier fait passer de l'espace L^1 à L^∞ : comme l'a montré l'exemple de la fonction Π plus haut, la transformée de Fourier n'appartient pas toujours à $L^1(\mathbb{R})$!

Théorème 4.5 (Riemann–Lebesgue) : Pour $f \in L^1(\mathbb{R})$, la transformée de Fourier \tilde{f} tend vers zéro à l'infini

$$\lim_{k \rightarrow \pm\infty} |\tilde{f}(k)| = 0.$$

L'idée de la preuve repose sur une astuce, l'insertion dans l'intégrale de $e^{-i\pi} = -1$ sous la forme :

$$\tilde{f}(k) = \frac{1}{2} \int e^{ikx} (1 - e^{-ik \cdot \frac{\pi k}{k^2}}) f(x) dx = \frac{1}{2} \int e^{ikx} f(x) dx - \frac{1}{2} \int e^{ik(x - \frac{\pi k}{k^2})} f(x) dx = \frac{1}{2} \int e^{ikx} \left(f(x) - f\left(x + \frac{\pi k}{k^2}\right) \right) dx$$

Quand $|k| \rightarrow \infty$, l'intégrand de la dernière intégrale tend vers zéro et nous admettrons que cela assure la convergence vers zéro de l'intégrale, voir [5] § 7.3.

- **Inversion de la transformée de Fourier**

Dans les cas où $\tilde{f} \in L^1$, on peut inverser la transformation de façon tout à fait explicite :

Théorème 4.6 d'inversion : Si $f \in L^1$ a une transformée de Fourier \tilde{f} intégrable, $\tilde{f} \in L^1$, alors

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{-ikx} \tilde{f}(k) \quad p.p. \text{ en } x.$$

De plus il y a égalité en tout point x où f est continue.

Autrement dit, au facteur $1/2\pi$ près, l'inversion de la transformation de Fourier est simplement la transformation de Fourier complexe conjuguée : $\mathcal{F}^{-1} = \frac{1}{2\pi} \overline{\mathcal{F}}$, sauf peut-être aux points de discontinuité de f .

Esquisse de la preuve : elle fait appel à l'identité $\int_{\mathbb{R}} f \tilde{g} dx = \int_{\mathbb{R}} \tilde{f} g dx$ pour toute paire de fonctions f et g de \mathcal{L}^1 , conséquence du théorème de Fubini (le vérifier !) ; on applique ce lemme à la fonction f donnée et à la suite de gaussiennes $g_n(y) = \exp -y^2/n^2 e^{-iyx}$ qui tendent vers la fonction e^{-ixy} quand $n \rightarrow \infty$, tandis que leurs transformées de Fourier $\tilde{g}_n(k) = n\sqrt{\pi} e^{-n^2(k-x)^2/4}$ tendent vers la distribution $2\pi\delta(k-x)$, cf Chap. 3, § 3.4.2. Donc $2\pi f(x) = \lim \int_{\mathbb{R}} f(k) \tilde{g}_n(k) dk = \lim \int_{\mathbb{R}} \tilde{f}(k) g_n(k) dk = \int_{\mathbb{R}} dk \tilde{f}(k) e^{-ikx}$. \square

On démontre aussi que la condition d'intégrabilité de \tilde{f} dans le théorème d'inversion est remplie dès que f est de classe C^2 et que f , f' et f'' sont intégrables, cf Th. 4.8 ci-dessous. Exemple : la fonction lorentzienne $f(x) = 2a/(x^2 + a^2)$ vue plus haut. Les fonctions f' et f'' sont continues et intégrables (se comportant respectivement comme x^{-3} , x^{-4} à l'infini), et on tire du théorème d'inversion une transformation de Fourier qu'on vérifie aisément :

$$\frac{2a}{x^2 + a^2} = \int_{\mathbb{R}} dk e^{-ikx} e^{-a|k|}.$$

4.2.3 Autres propriétés

- **Conjugaison, translation, dilatation**

On lit sur l'expression de $\tilde{f}(k) = \mathcal{F}[f](k) = \int_{\mathbb{R}} dk e^{ikx} f(x)$ les propriétés suivantes de conjugaison $\mathcal{F}[f(-x)] = \overline{\mathcal{F}(f(x))} = \tilde{f}(-k)$, $\mathcal{F}[\overline{f(x)}] = \overline{\tilde{f}(-k)}$, $\mathcal{F}[f(x-a)] = e^{ika} \tilde{f}(k)$ et $\mathcal{F}[f(x)e^{-ilx}] = \tilde{f}(k-l)$.

$f(x)$	$\tilde{f}(k)$
$f(x) = f(-x)$ paire	$\tilde{f}(k) = \tilde{f}(-k)$ paire
$f(x) = -f(-x)$ impaire	$\tilde{f}(k) = -\tilde{f}(-k)$ impaire
$f(x) = \overline{f(x)}$ réelle	$\tilde{f}(k) = \overline{\tilde{f}(-k)}$ "hermitienne"
$f(x) = -\overline{f(x)}$ imaginaire	$\tilde{f}(k) = -\overline{\tilde{f}(-k)}$ anti-hermitienne

De même un changement d'échelle sur la variable x se traduit simplement par le changement d'échelle inverse sur la variable k de la t.F., **à un préfacteur près !**

$$a \neq 0 \quad \mathcal{F}[f(ax)] = \frac{1}{|a|} \tilde{f}\left(\frac{k}{a}\right).$$

• Dérivation

On a les relations suivantes entre les transformées de Fourier des dérivées :

$$f(x), xf(x), \dots, x^n f(x) \in L^1 \Rightarrow \tilde{f} \text{ } n \text{ fois dérivable et } \tilde{f}^{(\ell)}(k) = \mathcal{F}[(ix)^\ell f(x)], \ell = 1, \dots, n \quad (4.12)$$

par application de la dérivation sous le signe somme (Théorème 2.10). Inversement, si $f \in L^1$, f de classe C^m et $f^{(\ell)} \in L^1$ pour $\ell = 1, \dots, m$, alors

$$\mathcal{F}[f^{(\ell)}(x)] = (-ik)^\ell \tilde{f}(k) \text{ pour } \ell = 1, \dots, m \quad (4.13)$$

par intégrations par parties répétées de la formule de définition (4.10). (En effet : si $g, g' \in L^1$, $|g(x)| \rightarrow 0$ quand $x \rightarrow \pm\infty$. Exercice : le vérifier.)

En particulier

$$\boxed{\mathcal{F}[f'(x)] = -ik\tilde{f}(k) \quad \text{et} \quad \mathcal{F}[ixf(x)] = \frac{d}{dk}\tilde{f}(k)}$$

Cela implique des relations importantes entre l'existence de dérivées d'une fonction et le comportement asymptotique de sa transformée de Fourier. Ainsi

Proposition 4.7 : Si $f \in L^1$ est à support borné, sa transformée de Fourier est de classe C^∞ .

En effet si $f \in L^1$ est à support borné, tous les $x^n f \in L^1$, donc tous les $\tilde{f}^{(n)}$ existent par (4.12). Pour la réciproque, voir plus bas, Prop. 4.9.

On a vu plus haut (Théorème 4.5 de Riemann–Lebesgue) que si f est continue, $\tilde{f}(k)$ s'annule à l'infini. Peut-on affiner ce résultat et dire *comment* \tilde{f} s'annule ? Comme conséquence de (4.12–4.13), on a le

Théorème 4.8 : (a) Si f et ses dérivées $f^{(\ell)}$, $\ell = 1, \dots, m$ sont intégrables, alors

$$\forall k \in \mathbb{R} \quad |\tilde{f}(k)| \leq |k|^{-m} \int_{\mathbb{R}} |f^{(m)}(x)| dx;$$

(b) si f et les fonctions $x^\ell f$, $\ell = 1, \dots, n$ sont intégrables, alors

$$\forall k \in \mathbb{R} \quad |\tilde{f}^{(n)}(k)| \leq \int_{\mathbb{R}} |x^n f(x)| dx,$$

qui nous dit que des propriétés de régularité (existence d'une dérivée intégrable) d'une fonction se traduisent par des propriétés de décroissance de sa transformée de Fourier et vice versa.

Ces résultats sont l'analogue des résultats concernant le comportement asymptotique des coefficients de la série de Fourier (théorème 4.2 de Stokes).

- **Fonctions à décroissance rapide**

Si dans le (b) du théorème précédent, l'entier ℓ n'est pas borné, on est dans la classe des fonctions à *décroissance rapide*.

Définition 4.1 : Une fonction f est dite à décroissance rapide si $\forall \ell \in \mathbb{N}, \lim_{x \rightarrow \pm\infty} |x^\ell f(x)| = 0$.

Exemples : $f(x) = e^{-x^2}$ est à décroissance rapide et de classe C^∞ ; $f(x) = e^{-|x|}$ est à décroissance rapide, mais pas de classe C^∞ . Toute fonction à support borné est à décroissance rapide.

Par la proposition précédente, si f est intégrable à décroissance rapide, alors toutes les dérivées de \tilde{f} existent.

Proposition 4.9 : La transformée de Fourier \tilde{f} d'une fonction intégrable à décroissance rapide est de classe C^∞ . Réciproquement, si une fonction intégrable est de classe C^∞ , et si toutes ses dérivées sont intégrables, alors \tilde{f} est à décroissance rapide.

4.2.4 Transformation de Fourier dans L^2

- L'espace \mathcal{S}

Définition 4.2 : On désigne par \mathcal{S} (espace de Schwartz) l'espace des fonctions de classe C^∞ qui sont à décroissance rapide ainsi que toutes leurs dérivées.

Exemple : si $f \in \mathcal{D}$, \tilde{f} est de classe C^∞ (Prop. 4.7); toute dérivée $f^{(n)}$ existe et est intégrable, donc par le Théorème 4.8(b), \tilde{f} décroît au moins comme $1/|k|^n$ pour tout n , donc $\tilde{f} \in \mathcal{S}$.

Dans cet espace, on dit qu'une suite de fonctions f_n converge vers 0 si, pour tous $p, q \in \mathbb{N}$,

$$\lim_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} |x^p f_n^{(q)}(x)| = 0$$

ce qui implique la convergence uniforme de f et de toutes ses dérivées vers 0.

L'intérêt d'introduire ce nouvel espace apparaît clairement dans le

Théorème 4.10 : La transformée de Fourier est un opérateur linéaire continu de \mathcal{S} dans \mathcal{S} .

Première preuve : La transformation de Fourier applique \mathcal{S} dans \mathcal{S} : en effet si $f \in \mathcal{S}$, \tilde{f} est de classe C^∞ (Prop. 4.9) ; toute dérivée $f^{(k)}$ est à décroissance rapide et intégrable, donc par le théorème 4.8, \tilde{f} est aussi à décroissance rapide. Il en est de même de toutes ses dérivées $\tilde{f}^{(n)}$, puisqu'elles sont les transformées de Fourier des $x^n f(x)$, à qui l'argument précédent s'applique aussi. La continuité annoncée dans le théorème équivaut au fait que pour toute suite $f_n \in \mathcal{S}$ qui tend vers 0 au sens défini plus haut, les \tilde{f}_n tendent aussi vers 0.

Au lieu de la situation précédente où la transformation de Fourier \mathcal{F} faisait passer de L^1 dans L^∞ , \mathcal{F} est maintenant un *endomorphisme* de \mathcal{S} . De plus, la transformation de Fourier inverse $\mathcal{F}^{-1} = \frac{1}{2\pi} \overline{\mathcal{F}}$ a la même propriété de continuité. \mathcal{F} est donc une application bijective et bicontinue (continue ainsi que son inverse) de \mathcal{S} dans \mathcal{S} .

• Transformation de Fourier dans L^2

Ces résultats sur \mathcal{S} vont nous permettre de définir maintenant la transformée de Fourier sur l'espace $L^2(\mathbb{R})$ des fonctions de carré intégrable (ou “sommable”).

Il convient à ce point de souligner l'importance en physique de cette classe de fonctions. En électromagnétisme et en théorie des ondes, comme en théorie du signal, $\int dt |f(t)|^2$ donne l'énergie de l'onde ou du signal d'amplitude $f(t)$. En mécanique quantique, les états d'un système sont décrits par des fonctions d'onde complexes ψ , fonctions d'une (ou plusieurs) variable(s) de position ou d'impulsion, et ψ est requis à être de module carré intégrable.

On montre, et nous admettrons, (voir [1] p. 269 ou [5] p. 173 et 199 ff. pour plus de détails), que

- l'espace \mathcal{S} est dense dans l'espace $L^2(\mathbb{R})$ pour la norme de L^2 ; ce dernier est complet (cf Chap. 2, Théorème 2.6) ;
- cela permet d'étendre l'opérateur \mathcal{F} de \mathcal{S} à L^2 : c'est un opérateur linéaire continu sur L^2 .

Théorème 4.11 : *La transformée de Fourier est un opérateur linéaire continu de L^2 dans L^2 .*

Enfin on a l'important

Théorème 4.12 (Parseval–Plancherel) : *Si f et g sont dans L^2 , \tilde{f} et \tilde{g} le sont aussi et*

$$\int_{\mathbb{R}} f(x) \overline{g(x)} dx = \frac{1}{2\pi} \int_{\mathbb{R}} \tilde{f}(k) \overline{\tilde{g}(k)} dk \quad (4.14)$$

En particulier pour $g = f$, $\int |f|^2 = (2\pi)^{-1} \int |\tilde{f}|^2$, ou encore, $\|f\|_2 = (2\pi)^{-1} \|\tilde{f}\|_2$.

Première preuve : Pour $f, g \in C^2(\mathbb{R})$ à support borné, $f, g \in L^1 \cap L^2$ ainsi que \tilde{f}, \tilde{g} (cf. Th. 4.8) et on a :

$\int f(x) \overline{g(x)} dx = \frac{1}{2\pi} \int \overline{g(x)} dx \int dk e^{-ikx} \tilde{f}(k) \stackrel{Fubini}{=} \frac{1}{2\pi} \int dk \tilde{f}(k) \int \overline{g(x)} e^{-ikx} dx = \frac{1}{2\pi} \int dk \tilde{f}(k) \overline{\tilde{g}(k)}$. Pour $f, g \in L^2$ quelconques, on les approche par des fonctions de C^2 à support borné, etc, cf [8] p. 216.

On peut redire cette propriété en termes plus géométriques : la transformation de Fourier $f \leftrightarrow (2\pi)^{-\frac{1}{2}}\tilde{f}$ conserve la norme L^2 , c'est une *isométrie*. Sur un plan plus physique, ce théorème nous apprend que l'énergie d'un signal d'amplitude $f(t)$ peut se calculer soit par $\int |f(t)|^2 dt$, soit par la somme des modules carrés des modes de Fourier $\int |\tilde{f}(\omega)|^2 \frac{d\omega}{2\pi}$.

Enfin, et heureusement pour nous, la transformée de Fourier définie dans L^2 coïncide avec celle définie dans L^1 pour toute fonction de $L^1 \cap L^2$.

Il peut arriver que f soit de carré intégrable (c'est-à-dire dans L^2) sans être intégrable (dans L^1). Par exemple la fonction "sinus cardinal" rencontrée plus haut (transformée de Fourier de la fonction porte), $\sin x/x$, n'est pas (absolument) intégrable sur \mathbb{R} , comme on sait, mais son carré l'est (pourquoi?). On peut toutefois le plus souvent calculer la transformée de Fourier de f en prenant la "partie principale à l'infini", c'est-à-dire que ([1], p. 263)

$$\tilde{f}(k) = \lim_{M \rightarrow +\infty} \int_{-M}^M f(x) e^{ikx} dx \quad \text{pour presque tout } k \in \mathbb{R}.$$

4.2.5 Transformation de Fourier et convolution

On a introduit au Chap. 3 la convolution de deux fonctions f et g localement intégrables

$$(f * g)(x) = \int_{\mathbb{R}} dy f(x-y) g(y).$$

Comment cela se marie-t-il avec la transformation de Fourier ?

Théorème 4.13 : Si f et g sont intégrables ainsi que leur convoluée $f * g$, alors

$$(a) \mathcal{F}[f * g] = \mathcal{F}[f]\mathcal{F}[g] \quad \text{ou encore} \quad \widetilde{f * g}(k) = \tilde{f}(k)\tilde{g}(k)$$

et si \tilde{f} et \tilde{g} peuvent être convolées

$$(b) \mathcal{F}[f.g] = 2\pi \mathcal{F}[f] * \mathcal{F}[g] \quad \text{ou encore} \quad \widetilde{f.g}(k) = 2\pi \tilde{f}(k) * \tilde{g}(k)$$

En d'autres termes, produit ordinaire et convolution sont échangés par transformation de Fourier !

Preuve : Calculons la transformée de Fourier de $f * g$:

$$\begin{aligned} \mathcal{F}[f * g](k) &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x-y) g(y) dy \right) e^{ikx} dx \stackrel{\text{Fubini}}{=} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x-y) g(y) e^{ikx} dx \right) dy \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(x-y) e^{ik(x-y)} dx \right) g(y) e^{iky} dy = \int_{-\infty}^{\infty} \tilde{f}(k) g(y) e^{iky} dy \\ &= \tilde{f}(k) \tilde{g}(k) \end{aligned} \quad \text{cqfd.} \quad (4.15)$$

La formule $\mathcal{F}[f.g] = \mathcal{F}[f]*\mathcal{F}[g]$ s'obtient de la même manière si \tilde{f} et \tilde{g} sont intégrables. Nous l'admettrons plus généralement sous les hypothèses ci-dessus.

Ce théorème est de grande importance pratique en physique : c'est lui qui permet le découplage des modes dans l'analyse d'un système linéaire. On a vu au Chap. 3 que dans un système linéaire décrit par un "noyau" $K(t - t')$ et soumis à une excitation $F(t)$, la réponse au temps t est de la forme

$$G(t) = \int dt' K(t - t') F(t'),$$

c'est-à-dire est la convolée $G = K * F$. En passant aux transformées de Fourier, $\tilde{G}(\omega) = \tilde{K}(\omega)\tilde{F}(\omega)$, il y a *découplage* des différentes valeurs de ω , c'est-à-dire des différents modes.

4.2.6 Diffraction par une fente, par un réseau

En Optique, on étudie la diffraction d'une onde lumineuse monochromatique $u(x, t) = Ae^{i(kx - \omega t)}$ par une fente de largeur $d = 2a$. A grande distance, l'amplitude diffractée est donnée par les "formules de Fraunhofer"² qui reposent sur la transformée de Fourier.

L'amplitude de l'onde incidente de longueur d'onde λ et de vecteur d'onde $k = 2\pi/\lambda$ étant décrite par la fonction scalaire $f(x)$, on écrit que l'onde diffractée est la superposition d'ondes sphériques émises par tous les points de l'objet diffractant, ici la fente. Donc l'amplitude diffractée en \vec{x} est, pour $r = |x| \gg a$ et dans la direction θ

$$f_d(\vec{x}) = f_d(r, \theta) = \psi_k(r) \int_{-a}^a dx' f(x') e^{ikx' \sin \theta},$$

avec un préfacteur $\psi_k(r)$ qui ne nous concerne pas ici. On voit que la transformée de Fourier \tilde{f} de l'onde incidente f donne, à un facteur près, l'amplitude à grande distance de la lumière diffractée par la fente, l'intensité diffractée étant proportionnelle à $|\tilde{f}(k \sin \theta)|^2$. Voir TD pour les détails du calcul.

4.3 Transformées de Fourier des distributions

- **Préambule**

On aimeraient maintenant étendre ces transformées de Fourier pour pouvoir les utiliser sur les distributions telles δ , III ou H souvent rencontrées en physique. L'idée première est de partir à nouveau d'une fonction localement intégrable, qui définit une distribution régulière, d'en calculer la transformée de Fourier et de l'appliquer à une fonction test. Mais on se rappelle que l'intégrabilité locale ne suffit pas à assurer l'existence d'une transformée de Fourier. Restreignons-nous donc à des fonctions f intégrables (dans L^1). Selon le Théorème 4.3, sa

2. "The equation was named in honor of Joseph von Fraunhofer although he was not actually involved in the development of the theory." (sic) [Wikipedia]. À ce titre, elle constitue une illustration de la fameuse *loi de Stigler* qui affirme *Une découverte scientifique ne porte jamais le nom de son auteur, ... et cette loi s'applique à elle-même!*

transformée de Fourier \tilde{f} existe et est continue donc localement intégrable. On écrit alors, pour toute fonction-test φ

$$\langle \tilde{f}, \varphi \rangle = \int_{\mathbb{R}} \tilde{f}(k) \varphi(k) dk = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x) e^{ikx} dx \right) \varphi(k) dk,$$

mais on peut utiliser le théorème de Fubini (grâce à l'intégrabilité de f et φ) pour intervertir les deux intégrales et écrire

$$\langle \tilde{f}, \varphi \rangle = \int_{\mathbb{R}} f(x) \left(\int_{\mathbb{R}} e^{ikx} \varphi(k) dk \right) dx = \int_{\mathbb{R}} f(x) \tilde{\varphi}(x) dx.$$

Ceci suggère, pour tout distribution T , la définition suivante de $\mathcal{F}[T]$

$$\langle \mathcal{F}[T], \varphi \rangle \stackrel{\text{déf}}{=} \langle T, \mathcal{F}[\varphi] \rangle. \quad (4.16)$$

Mais une dernière difficulté nous attend là ! La transformée de Fourier $\mathcal{F}[\varphi]$ d'une fonction $\varphi \in \mathcal{D}$ n'est pas toujours dans \mathcal{D} , et le membre de droite n'existe pas toujours ! (le support de $\mathcal{F}[\varphi]$ n'est pas nécessairement borné) : la transformée $\mathcal{F}[T]$ n'est pas toujours définie. Il nous faut donc élargir la classe de fonctions-tests.

• Distributions tempérées

On a introduit plus haut (sect. 4.2.4) l'espace \mathcal{S} des fonctions φ à décroissance rapide, telles que *pour tout* $k \in \mathbb{N}$, φ et toutes ses dérivées $\varphi^{(m)}$ satisfont

$$\lim_{x \rightarrow \pm\infty} |x^k \varphi^{(m)}(x)| = 0.$$

Définition 4.3 : Soit \mathcal{S}' le dual topologique de \mathcal{S} (espace des formes linéaires continues sur \mathcal{S}). On appelle distribution tempérée tout élément de \mathcal{S}' .

Vérifions d'abord que $\mathcal{S}' \subset \mathcal{D}'$: toute distribution tempérée est une distribution “ordinaire”. Cela découle de $\mathcal{D} \subset \mathcal{S}$ et du fait que la continuité en un sens implique celle dans l'autre (à vérifier). Cela implique que les propriétés déjà établies (translation, dilatation, dérivabilité, etc) pour les distributions ordinaires demeurent valables dans \mathcal{S}' .

Qu'est-ce qui remplace dans \mathcal{S}' les distributions “régulières” associées à une fonction localement intégrable ? Définissons les fonctions à croissance lente comme les fonctions croissant au plus à l'infini comme une puissance de x . Le produit d'une telle fonction par une fonction test de \mathcal{S} est encore intégrable. On établit donc

Proposition 4.14 : Toute fonction localement intégrable à croissance lente définit une distribution de \mathcal{S}' .

Mais de même que \mathcal{D}' contenait des distributions non régulières, \mathcal{S}' en contient aussi qui ne sont pas associées à des fonctions intégrables à croissance lente.

Exemples. La distribution δ de Dirac est évidemment dans \mathcal{S}' , son application sur toute $\varphi \in \mathcal{S}$ étant bien définie (et linéaire et continue). Mais elle n'est pas une fonction à croissance lente.

- **Transformées de Fourier des distributions tempérées**

Nous avons maintenant un résultat important qui justifie la démarche adoptée :

Théorème 4.15 : *Toute distribution tempérée admet une transformée de Fourier au sens des distributions (4.16) qui est elle aussi une distribution tempérée.*

Autrement dit (4.16) s'applique pour tout $T \in \mathcal{S}'$ et $\varphi \in \mathcal{S}$ et définit $\mathcal{F}[T] \in \mathcal{S}'$. On vérifie aussi que cette transformée de Fourier coïncide bien avec celle définie plus haut pour les distributions régulières de L^1 ou de L^2 ([1], p 273).

On vérifie enfin que les formules d'inversion du Théorème 4.6 s'appliquent encore aux distributions tempérées : si $T \in \mathcal{S}'$, soit $\tilde{T} = \mathcal{F}[T]$, alors

$$2\pi T = \overline{\mathcal{F}}[\tilde{T}]$$

avec $\overline{\mathcal{F}}$ la transformation conjuguée. Ou encore, de façon plus cavalière, si $\tilde{T}(k) = \mathcal{F}[T](k)$, $2\pi T(x) = \mathcal{F}[\tilde{T}](-x)$.

- **Exemples et applications. Transformée de Fourier de δ , etc**

Considérons la fonction constante $\mathbf{1}$, $x \mapsto 1$: c'est une fonction à croissance lente, donc dans \mathcal{S}' . Calculons sa transformée de Fourier. Soit $\varphi \in \mathcal{S}$ une fonction à décroissance rapide. On a $\langle \mathbf{1}, \varphi \rangle \stackrel{\text{déf}}{=} \langle \mathbf{1}, \widetilde{\varphi} \rangle = \int_{\mathbb{R}} \widetilde{\varphi}(k) dk$, ce qui, par les formules d'inversion (Théorème 4.6), vaut $\int_{\mathbb{R}} \widetilde{\varphi}(k) dk = 2\pi \varphi(0) = 2\pi \langle \delta, \varphi \rangle$. On en conclut que $\mathbf{1} = 2\pi \delta$. Exercice : établir selon la même méthode que $\mathcal{F}[e^{-ika}] = 2\pi \delta_a$.

Proposition 4.16 : *La transformée de Fourier de $\mathbf{1}$ est $2\pi\delta$*

$$\mathcal{F}[\mathbf{1}] = 2\pi\delta, \quad \mathcal{F}[e^{-ika}] = 2\pi\delta_a. \quad (4.17)$$

Réiproquement la transformée de Fourier de δ est la fonction $\mathbf{1}$.

La preuve de la réciproque s'effectue de la même manière :

$$\langle \tilde{\delta}, \varphi \rangle = \langle \delta, \tilde{\varphi} \rangle = \tilde{\varphi}(0) = \int e^{i0} \varphi(x) dx = \langle \mathbf{1}, \varphi \rangle.$$

Des théorèmes de dérivation et de translation découlent alors des formules utiles

$$\mathcal{F}[\delta'] = -ik, \quad \mathcal{F}[\delta^{(m)}] = (-ik)^m \quad \mathcal{F}[\delta_a] = e^{ika}.$$

(Attention : ces formules dépendent des conventions dans la définition de \mathcal{F} !)

Exercice. Établir les formules suivantes pour les distributions fréquemment rencontrées H , III , PP , ...

$$\begin{aligned} \mathcal{F}[H(x)] &= i \text{PP} \frac{1}{k} + \pi\delta & \mathcal{F}[\text{PP} \frac{1}{x}] &= i\pi \operatorname{sgn} k & \mathcal{F}[\operatorname{sgn}(x)] &= 2i \text{PP} \frac{1}{k}, \\ \widetilde{\text{III}}(k) &= \text{III}\left(\frac{k}{2\pi}\right). \end{aligned}$$

On verra d'autres exemples en TD.

Lectures complémentaires

Mon exposé a suivi de très près celui de [1]. Pour plus de détails, preuves, etc, voir [5], [7], [8].

Exercices

1. Calculer la transformée de Fourier des fonctions suivantes :

- (a) : $1/(\mathbf{p}^2 + m^2)$, $\mathbf{p} \in \mathbb{R}^3$, qui représente (à un facteur près) l'amplitude d'une particule de masse m et d'impulsion \mathbf{p} .
- (b) $\frac{e^{-m\|\mathbf{r}\|}}{\|\mathbf{r}\|}$, pour $\mathbf{r} \in \mathbb{R}^3$.

Ces fonctions jouent un rôle considérable en théorie quantique des champs.

2. Changement de normalisation.

Si on définit une transformée de Fourier avec d'autres normalisations $\widehat{\mathcal{F}}[f] = A \int_{\mathbb{R}} e^{iBxk} dx$, que peut-on dire de

- la transformée inverse ?
- le théorème de Parseval–Plancherel ?

Chapitre 5

Probabilités

5.1 Événements. Espace des épreuves.

La théorie des probabilités s'est construite à partir de considérations sur la théorie des jeux. Dans ces cas-là, (jeux de dés, de cartes, paris, etc), on cherche à évaluer la probabilité d'observer un type d'événements donné dans un ensemble *fini* d'événements possibles. Par exemple dans un jeu où on lance deux dés, quelle est la probabilité que la somme des deux dés soit 6 ? Il suffit d'énumérer tous les événements possibles (ici 6^2) et tous les “événements favorables” (ici 5 : 1+5, 2+4, 3+3, 4+2, 5+1) : si tous ces événements se produisent a priori de façon ... équiprobable (!), (dés non pipés), on définit la probabilité comme le rapport

$$P = \frac{\#\text{cas favorables}}{\#\text{cas possibles}}. \quad (5.1)$$

En général, dans ce genre de situations, évaluer une probabilité équivaut à un problème de *dénombrément*, relevant de la combinatoire -élémentaire ou non¹-.

On souhaite généraliser cette définition à une classe d'événements décrits par une ou des variables prenant des valeurs, discrètes ou continues, en nombre infini. Se pose la question de compter, ou plutôt de *mesurer* ces événements. On voit réapparaître la notion de mesure abordée au chapitre 2.

On va donc devoir axiomatiser un peu notre approche.

1. Exemple d'un problème simple à énoncer mais très ardu, le *problème des ménages* : n couples sont placés de façon aléatoire autour d'une table circulaire. Quelle est la probabilité qu'aucune femme ne soit assise à côté de son conjoint ?

- **Espace des épreuves (ou *univers*) Ω**

Cet “espace” (qui n’est en général ni vectoriel, ni topologique !) est la donnée de base d’un problème de probabilités. C’est l’ensemble de tous les événements possibles. Plus précisément, on appellera *événement élémentaire* (ou *atomique*) tout élément de Ω , et *épreuve* le choix d’un élément a dans Ω . En général, un événement (*événement composé*) est associé à une partie (ou sous-ensemble) de Ω . On va chercher à associer une probabilité $\mathbf{P}(A)$ à tout $A \subset \Omega$. En fait cela n’est pas toujours possible et on doit se contenter de le faire pour des sous-ensembles bien particuliers, des sous-ensembles “mesurables”.

Exemples : (a) Ω ensemble fini des résultats de lancers de deux dés, où on tient compte de l’ordre (on distingue les deux dés) ; le cardinal de Ω est $6^2 = 36$; A = sous-ensemble des lancers contenant un 2 ; $|A| = 11$; $\mathbf{P}(A) = 11/36$.
(b) $\Omega = \mathbb{R}$, $A = \mathbb{Q}$, comment définir $\mathbf{P}(\mathbb{Q})$, la probabilité qu’un réel “tiré au hasard” (il faudra préciser) soit un rationnel ?

5.2 Probabilités et mesure. Vocabulaire

5.2.1 Axiomes de Kolmogorov

Dans le cas où l’espace d’épreuves Ω n’est pas un ensemble fini, nous venons de voir que le concept de probabilités sur Ω est lié à celui de mesure sur cet ensemble. En général on dira qu’on a probabilisé l’espace Ω si on a pu définir une famille \mathcal{F} de ses parties (sous-ensembles) A et leur attacher un nombre $0 \leq \mathbf{P}(A) \leq 1$, la probabilité *a priori* de l’événement A , avec les axiomes suivants (Kolmogorov)

$$K(i) : \emptyset \in \mathcal{F}, \text{ et } \mathbf{P}(\emptyset) = 0 ;$$

$$K(ii) : \Omega \in \mathcal{F} \text{ et } \mathbf{P}(\Omega) = 1 ;$$

$$K(iii) : \text{Si } A \in \mathcal{F}, \text{ son complémentaire } \bar{A} \stackrel{\text{déf}}{=} \Omega \setminus A \in \mathcal{F}, \text{ et } \mathbf{P}(\bar{A}) = 1 - \mathbf{P}(A) ;$$

$$K(iv) : \text{Si } A, B \in \mathcal{F}, A \subset B \text{ implique } \mathbf{P}(A) \leq \mathbf{P}(B) ;$$

$$K(v) : \mathcal{F} \text{ est stable par union dénombrable ; si } A, B \in \mathcal{F}, \text{ donc } A \cap B \text{ et } A \cup B \in \mathcal{F}, \text{ on a} \\ \mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B) ;$$

toutes propriétés qui généralisent celles familières pour un ensemble fini.

On voit que \mathcal{F} satisfait les axiomes de définition d’une tribu (cf Chap. 2) et \mathbf{P} ceux d’une mesure, avec la condition supplémentaire de normalisation

$$\mathbf{P}(\Omega) = 1 .$$

On peut alors transcrire dans ce nouveau contexte le vocabulaire introduit au chapitre 2 :

- Une propriété vraie p.p. devient une propriété vraie *presque sûrement* (p.s.) ;
- A tel que $\mathbf{P}(A) = 0$ est un événement presqu'impossible, A tel que $\mathbf{P}(A) = 1$ est presque sûr ;
- une fonction mesurable X sur l'espace (l'univers) Ω à valeurs réelles est appelée *variable aléatoire*, voir § suivant ;
- l'intégrale d'une telle v.a. X est son *espérance* $E(X) = \int_{\Omega} X dP$. En physique, on dit aussi sa *valeur moyenne* et on la note généralement $\langle X \rangle$, ou quelquefois \bar{X} , donc $\langle X \rangle \equiv E(X) \equiv \bar{X}$.

A cela s'ajoutent quelques éléments de terminologie propres aux probabilités :

- on vient de le voir, un élément de \mathcal{F} (donc une partie probabilisée de Ω) est un *événement* ;
- deux événements A et B tels que $A \cap B = \emptyset$ sont dits *incompatibles* ou *mutuellement exclusifs* ;
- $A \cap B$ est l'événement “ A et B ” ; $A \cup B$ est l'événement “ A ou B ”.

Dans le cas où Ω est un ensemble fini, on retrouve les définitions et résultats de la théorie élémentaire des probabilités. La donnée de la “loi” \mathbf{P} se ramène à celle des probabilités des événements élémentaires $p_a = \mathbf{P}(\{a\})$, complétée par $\mathbf{P}(A) = \sum_{a \in A} p_a$, selon l'axiome K(v). La loi de probabilité est bien normalisée $\sum_{a \in \Omega} p_a = \mathbf{P}(\Omega) = 1$, etc.

Remarques. On n'a pas cherché à donner une liste minimale d'axiomes, de laquelle les autres axiomes découlent. L'axiome K(v) se généralise en une formule de Poincaré pour une union finie $P(\cup_i A_i) = \dots$, voir TD.

5.2.2 Probabilité conditionnelle. Événements indépendants

On l'a dit, $\mathbf{P}(A)$ représente la probabilité *a priori* de l'événement A , compte tenu de la seule information dont nous disposons : $A \subset \Omega$. Si maintenant une information supplémentaire est apportée, à savoir qu'un événement B s'est produit, le calcul de la probabilité de A est à reconsidérer, et on peut parler de probabilité *a posteriori*, ou de probabilité de A *conditionné par B*, ce qu'on symbolise par $A|B$. La définition suivante semble naturelle

Définition 5.1 : (Axiome de Bayes) Si $A, B \in \Omega$ sont deux événements, avec $\mathbf{P}(B) \neq 0$, la probabilité conditionnelle de A sachant B (ou “ A si B ”) est

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}.$$

Il découle de cette définition une propriété simple mais utile à laquelle on donne aussi le nom de Bayes : $\mathbf{P}(A \cap B) = \mathbf{P}(A|B)\mathbf{P}(B) = \mathbf{P}(B|A)\mathbf{P}(A)$ donc

Proposition 5.1 (Bayes) : $\mathbf{P}(A|B) = \frac{\mathbf{P}(B|A)\mathbf{P}(A)}{\mathbf{P}(B)}$.

Définition 5.2 : Deux événements A et B de Ω sont indépendants si

$$\mathbf{P}(A|B) = \mathbf{P}(A) \text{ ou de façon équivalente } \mathbf{P}(B|A) = \mathbf{P}(B) \text{ ou encore } \mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B). \quad (5.2)$$

Une famille finie d'événements A_i est dite indépendante si pour toute sous-famille des A_i , $\mathbf{P}(\cap_i A_i) = \prod_i \mathbf{P}(A_i)$.

N.B. Bien noter le “toute sous-famille”, et non pas seulement toutes les intersections deux à deux, ou l'intersection de tous les A_i !

5.3 Variables aléatoires. Distributions de v.a.

Dans de nombreux cas concrets, en physique notamment, l'espace Ω est gigantesque et ne se prête pas bien à une description quantitative directe. (Qu'on pense aux $O(10^{23})$ variables de vitesse des molécules dans un volume macroscopique d'un gaz...) On va alors chercher à décrire les propriétés (statistiques) de ce système à l'aide d'une ou de quelques fonctions à valeurs réelles sur l'espace Ω . Cela nous amène à la définition d'une variable aléatoire.

5.3.1 Définition d'une variable aléatoire ; loi de probabilité

Définition 5.3 : Si Ω est un espace probabilisé, une variable aléatoire (v.a) (unidimensionnelle) X est une fonction mesurable de Ω dans \mathbb{R} . On définit de même une v.a. n -dimensionnelle $X : \Omega \rightarrow \mathbb{R}^n$.

Dans la suite, sauf mention explicite et pour la simplicité des notations, nous discuterons le cas d'une v.a. unidimensionnelle. Par l'hypothèse de mesurabilité de X , pour toute partie borélienne B de \mathbb{R} (c'est-à-dire tout élément de la tribu (borélienne) $\mathcal{B}(\mathbb{R})$), $A = X^{-1}(B) \in \mathcal{F}$, la tribu de Ω . Cela permet alors de définir une probabilité pour toute partie $B \in \mathcal{B}(\mathbb{R})$: $\mathbf{P}_X \stackrel{\text{déf}}{=} \mathbf{P} \circ X^{-1}$, une notation un peu cavalière signifiant $\mathbf{P}_X(B) = \mathbf{P}(A)$. Autrement dit on a remplacé l'espace probabilisé initial $(\Omega, \mathcal{F}, \mathbf{P})$ par l'espace $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbf{P}_X)$, et on dit que X “suit la loi” (de probabilité) \mathbf{P}_X . Ce \mathbf{P}_X n'est pas en général la mesure de Lebesgue. Bien sûr, la nouvelle loi \mathbf{P}_X obéit tous les axiomes de Kolmogorov, en particulier $\mathbf{P}_X(\mathbb{R}) = 1$, condition de normalisation qui exprime que X prend sûrement une valeur réelle. Donc

Définition 5.4 : On appelle loi de probabilité \mathbf{P}_X de la v.a. $X : \Omega \rightarrow \mathbb{R}$ la probabilité image de \mathbf{P} par X , $\mathbf{P}_X \stackrel{\text{déf}}{=} \mathbf{P} \circ X^{-1}$.

(Dans la suite on notera \mathbf{P} au lieu de \mathbf{P}_X chaque fois qu'il n'y aura pas d'ambiguïté.)

5.3.2 Les quantités et fonctions importantes attachées à une v.a.

- Valeur moyenne, moments, variance et écart-type, etc.

Comme on l'a vu plus haut, on définit l'espérance ou valeur moyenne d'une v.a. X par

$$E(X) \equiv \langle X \rangle \stackrel{\text{déf}}{=} \int_{\Omega} X dP = \int_{\mathbb{R}} x dP_X(x) \quad (5.3)$$

ou plus généralement celle de toute fonction $G(X)$, pourvu qu'elle soit intégrable !

$$E(G) = \langle G(X) \rangle \stackrel{\text{déf}}{=} \int_{\Omega} G(X) dP = \int_{\mathbb{R}} G(x) dP_X(x). \quad (5.4)$$

En particulier on définit les *moments*

$$\mu_k \stackrel{\text{déf}}{=} \langle X^k \rangle. \quad (5.5)$$

Attention ! certains de ces moments peuvent ne pas exister : μ_k existe si $X \in L^k$. On verra par la suite des exemples (distribution lorentzienne) de loi pour lesquelles seuls les moments les plus bas existent.

Dans le cas d'une v.a. discrète, pour laquelle la loi \mathbf{P}_X est donnée par les probabilités p_i que X prenne la valeur x_i , $i = 1, 2, \dots$, les expressions de ces espérance/valeur moyenne et moments se réduisent à

$$E(X) = \langle X \rangle = \sum_i p_i x_i, \quad \mu_k = \sum_i p_i x_i^k,$$

sous réserve que si i prend une infinité de valeurs, ces sommes convergent !

Le deuxième moment entre dans la définition de la *variance*, notée $\text{var}(X)$ ou $\sigma^2(X)$:

$$\text{var}(X) \equiv \sigma^2(X) \stackrel{\text{déf}}{=} \langle X^2 \rangle - \langle X \rangle^2 = \mu_2 - \mu_1^2 \quad (5.6)$$

ou encore, en développant le carré

$$\text{var}(X) \equiv \sigma^2(X) = \langle (X - \langle X \rangle)^2 \rangle. \quad (5.7)$$

L'*écart-type* $\sigma(X)$ est la racine carrée positive de cette variance.

L'intérêt de ces deux quantités –variance ou écart-type– est de donner une indication sur la dispersion des valeurs de X autour de sa valeur moyenne $\langle X \rangle$. Une variance (donc un écart-type) faible signale des valeurs de la v.a. resserrées autour de la moyenne $\langle X \rangle$. À l'inverse, une variance élevée signifie que X peut prendre (avec une probabilité non négligeable) des valeurs éloignées de $\langle X \rangle$.

Il est utile d'observer comment se transforment $\langle X \rangle$, μ_k , σ^2 par translation ou dilatation de la v.a. X : La variance est insensible à une translation de la v.a. X par une constante c . En effet, si $X' = X + c$, $\langle X' \rangle = \langle X \rangle + c$ et $X - \langle X \rangle = X' - \langle X' \rangle$, donc les variances de X et X' sont les mêmes. Par une dilatation par un facteur λ , la moyenne et l'écart-type sont dilatées par ce même facteur $X \mapsto \lambda X$, $\langle X \rangle \mapsto \lambda \langle X \rangle$, $\text{var}X \mapsto \lambda^2 \text{var}X$, $\mu_k \mapsto \lambda^k \mu_k$, donc $\langle X \rangle / \sigma$, μ_k / σ^k sont invariants.

On peut aussi construire des combinaisons des moments, homogènes de degré k par dilatation et invariantes par translation : ce sont les *cumulants*. Par exemple, $\kappa_3 = \langle (X - \langle X \rangle)^3 \rangle$ est le cumulant d'ordre 3 (ou skewness) ; $\kappa_4 = \langle (X - \langle X \rangle)^4 \rangle - 3\sigma^4$ est le cumulant d'ordre 4, qui intervient en physique, et κ_4 / σ^4 est le *kurtosis*, qui décrit de façon plus fine la forme de la distribution de probabilité. [Voir TD ?](#)

On peut encore définir d'autres quantités qui donnent telle ou telle information quantitative sur la v.a. X . La *médiane* de la loi d'une v.a. X est par définition la valeur x_m telle que $\mathbf{P}(X \leq x_m) = \mathbf{P}(X > x_m) = \frac{1}{2}$. À ne pas confondre avec la valeur moyenne.

Exemple. Considérons une v.a. discrète prenant les valeurs $\{0, 1, 2, 3, 4\}$ avec les probabilités $\frac{1}{12}\{1, 5, 1, 2, 3\}$. Sa moyenne est $25/12 > 2$, tandis que sa médiane est entre 1 et 2.

• Inégalité de Tchebychev

Supposons que la v.a. est de classe $\mathcal{L}^k(\mathbb{R})$, donc $E(|X|^k)$ existe. Pour $a > 0$,

$$\mathbf{P}(|X| > a) = E(\theta(|X| - a)) \leq E(\theta(|X| - a) \left| \frac{X}{a} \right|^k) \leq E\left(\left| \frac{X}{a} \right|^k\right)$$

(on rappelle que $\theta \equiv H$, la distribution de Heaviside), donc

$$\mathbf{P}(|X| > a) \leq \frac{1}{a^k} E(|X|^k). \quad (5.8)$$

Cette inégalité nous sera très utile dans l'étude des Théorèmes limites, voir § 5.5.

• Covariance, coefficient de corrélation.

Considérons maintenant deux v.a. X et Y attachées au même espace Ω . On définit leur *covariance* et leur coefficient de corrélation par les formules

$$\text{cov}(X, Y) = \langle (X - \langle X \rangle)(Y - \langle Y \rangle) \rangle = \langle XY \rangle - \langle X \rangle \langle Y \rangle \quad (5.9)$$

$$r(X, Y) = \frac{\text{cov}(X, Y)}{\sigma(X)\sigma(Y)} \quad (5.10)$$

Ces quantités servent à évaluer la corrélation, au sens commun du mot, entre les deux variables. Noter que si les deux variables sont indépendantes, le numérateur $\text{cov}(X, Y)$ se factorise en $\langle X - \langle X \rangle \rangle \langle Y - \langle Y \rangle \rangle = 0$ puisque $\langle X - \langle X \rangle \rangle = 0$. Mais la réciproque : “ $\text{cov}(X, Y) = 0 \Rightarrow X, Y$ indépendantes” n'est en général pas vraie, on en verra un contre-exemple plus bas.

• Fonction de répartition

On définit alors la *fonction de répartition*² $F(x)$

$$F(x) = \mathbf{P}(X \leq x) \quad (5.11)$$

correspondant donc à un intervalle (borélien) $(-\infty, x]$ ³. Il en découle par l'axiome K(v) que pour tout intervalle $[a, b]$ (attention au sens des crochets !)

$$\mathbf{P}([a, b]) \equiv \mathbf{P}(a < X \leq b) = F(b) - F(a),$$

et par passage au complémentaire, $\mathbf{P}(X > x) = 1 - F(x)$.

Cette fonction jouit des propriétés suivantes

- elle est positive et prend ses valeurs dans $[0, 1]$ (c'est une probabilité !) ;
- $F(-\infty) = 0$; $F(\infty) = 1$.

Plus précisément, si la v.a. prend ses valeurs dans l'intervalle (a, b) (a, b finis ou infinis), $F(a) = 0$, $F(b) = 1$;

- F est monotone croissante au sens large (“non décroissante”) : si $x < y$, $F(x) \leq F(y)$.

Cela découle de l'axiome K(v) :

$$F(y) = \mathbf{P}(X \leq y) = \mathbf{P}(X \leq x) + \mathbf{P}(x < X \leq y) \geq \mathbf{P}(X \leq x) = F(x);$$

- F est partout continue à droite : $\forall \epsilon \quad F(x + \epsilon) - F(x) = \mathbf{P}(x < X \leq x + \epsilon)$ donc à la limite $\epsilon \rightarrow 0$, la mesure de l'intervalle $[x, x + \epsilon]$ tend vers zéro,

$$\lim_{\epsilon \rightarrow 0} F(x + \epsilon) = F(x),$$

ce qui est la définition de la continuité à droite.

Ces propriétés sont caractéristiques, en ce sens que toute fonction F les possédant peut être considérée comme la fonction de répartition d'une v.a.

F est donc continue à droite. En revanche, elle peut ne pas être continue à gauche. C'est le cas quand la variable X peut prendre des valeurs discrètes x_k avec une probabilité non nulle $p_k = \mathbf{P}(X = x_k)$:

$$\lim_{\epsilon \rightarrow 0} (F(x_k) - F(x_k - \epsilon)) = \lim_{\epsilon \rightarrow 0} (\mathbf{P}(x_k - \epsilon < X \leq x_k)) = \lim_{\epsilon \rightarrow 0} \mathbf{P}(\{x_k\}) = p_k.$$

Exemple. La fonction de répartition d'un dé bien équilibré est représentée sur la figure 5.1. Elle est discontinue en $x = 1, 2, 3, 4, 5, 6$.

2. appelée aussi “fonction de distribution cumulative”, le danger étant que “fonction de distribution” est aussi utilisé dans un autre sens, voir plus bas.

3. Autre définition fréquemment rencontrée dans la littérature : $F(x) = \mathbf{P}(X \in (-\infty, x[)$ avec une inégalité stricte $X < x$!

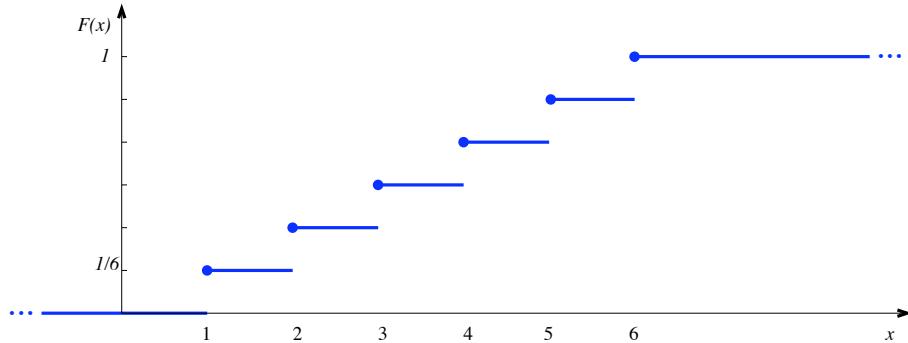


FIGURE 5.1 – Fonction de répartition d'un dé bien équilibré. Elle est partout continue à droite, discontinue à gauche aux 6 points marqués.

En général on peut démontrer (Lebesgue) que la fonction F peut être décomposée de façon unique en la somme de trois composantes

$$F = F_{\text{at.}} + F_{\text{abs. cont}} + F_{\text{sing. cont}} \quad (5.12)$$

- chaque composante est non-décroissante ;
- la composante *atomique* $F_{\text{at.}}$ est une fonction en escalier, somme des contributions d'un nombre fini ou dénombrable de sauts en x_k : elle représente la partie discontinue de F , $F - F_{\text{at.}}$ est continue ;
- la composante *absolument continue* $F_{\text{abs. cont}}$ est une fonction continue et dérivable p.p., $F'_{\text{abs. cont}}$ est intégrable de Lebesgue et on peut écrire

$$F_{\text{abs. cont}}(x) = \int_{-\infty}^x f(x') dx' \quad (5.13)$$

avec une fonction f égale p.p. à $F'_{\text{abs. cont}}$, positive ou nulle (puisque $F_{\text{abs. cont}}$ est non-décroissante) ;

- enfin la composante *singulière continue* $F_{\text{sing. cont}}$ décrit le reste ! Elle est continue mais non dérivable, et elle ne varie que sur un ensemble de mesure nulle. On peut rencontrer à l'occasion cette situation “exotique” en physique...⁴

Compléments.

1. On dit que la fonction F est *absolument continue* sur l'intervalle $[a, b]$ si, pour tout réel $\epsilon > 0$, il existe un $\eta > 0$ tel que, pour toute suite $([a_n, b_n])_{n \in \mathbb{N}}$ de sous-intervalles de $[a, b]$ d'intérieurs disjoints, $\sum_{n \geq 0} (b_n - a_n) < \eta \Rightarrow \sum_{n \geq 0} |F(a_n) - F(b_n)| < \epsilon$.

On montre (théorème dû à Lebesgue) que

F abs. continue sur $[a, b] \Leftrightarrow F$ dérivable p.p., F' intégrable au sens de Lebesgue et $F(x) - F(a) = \int_a^x F'(x') dx'$; alors F absolument continue $\Rightarrow F$ continue, et dérivable p.p.

4. Voir par exemple S. Aubry, J. Physique, 44 (1983) 147-162.

2. Fonction de Cantor ou escalier du Diable : exemple de fonction continue F sur $[0, 1]$ telle que $F(0) = 0$, $F(1) = 1$, qui est dérivable presque partout, la dérivée étant presque partout nulle. On reprend la construction de l'ensemble de Cantor \mathcal{K} (cf App. A2) et on construit à chaque étape une fonction linéaire par morceaux : à l'étape 0, c'est la fonction $f_0(x) = x$; à l'étape 1, f_1 est la fonction continue affine par morceaux qui vaut 0 en 0, 1 en 1, et $\frac{1}{2}$ sur $[\frac{1}{3}, \frac{2}{3}]$; à la $n + 1$ -ième étape, la fonction $f_{n+1} = f_n$ sur les intervalles où f_n est constante, et sur chaque intervalle $[a, b]$ où f_n n'est pas constante, f_{n+1} est la fonction linéaire par morceaux qui vaut $\frac{f_n(a) + f_n(b)}{2}$ sur $[\frac{2a}{3} + \frac{b}{3}, \frac{a}{3} + \frac{2b}{3}]$. Il est aisément de voir que pour tout x , $|f_{n+1}(x) - f_n(x)| \leq 2^{-n}$, ce qui montre que la série de fonctions $\sum_{n \geq 0} (f_{n+1} - f_n)$ converge uniformément, et donc que la suite f_n converge uniformément. La fonction limite F est continue, monotone, et l'on a $F(0) = 0$, $F(1) = 1$. De plus, F a une dérivée nulle sur le complémentaire de l'ensemble de Cantor \mathcal{K} , puisque ce complémentaire est une réunion d'intervalles sur lesquels f , par construction, est constante (d'où le nom d'escalier du Diable!). Elle a donc bien les propriétés annoncées : $F(0) = 0$, $F(1) = 1$, $F' = 0$ p.p.; les points où elle n'est pas dérivable sont les points de l'ensemble de Cantor, non dénombrable mais de mesure nulle. Comme anticipé au § 2.2.6, cette fonction est telle que $1 = F(1) - F(0) \neq \int_0^1 F'(x) dx = 0$. Clairement cette fonction n'est pas absolument continue! Voir son graphe sur la figure 5.2.

3. La fonction de Cantor est la fonction de répartition d'une variable aléatoire réelle X entre 0 et 1 dont les chiffres du développement en base trois sont obtenus par des tirages indépendants équiprobables avec $\mathbf{P}(0) = \mathbf{P}(2) = \frac{1}{2}$, $\mathbf{P}(1) = 0$.

Mais ces cas exotiques mis à part, la situation rencontrée le plus souvent concerne une fonction de répartition de la forme $F = F_{\text{at}} + F_{\text{abs.cont}}$, ce qu'on peut encore récrire sous la forme unifiée (5.13) à condition d'autoriser dans f des contributions de distributions delta $f(x) \rightarrow f(x) + \sum_k p_k \delta(x - x_k)$.

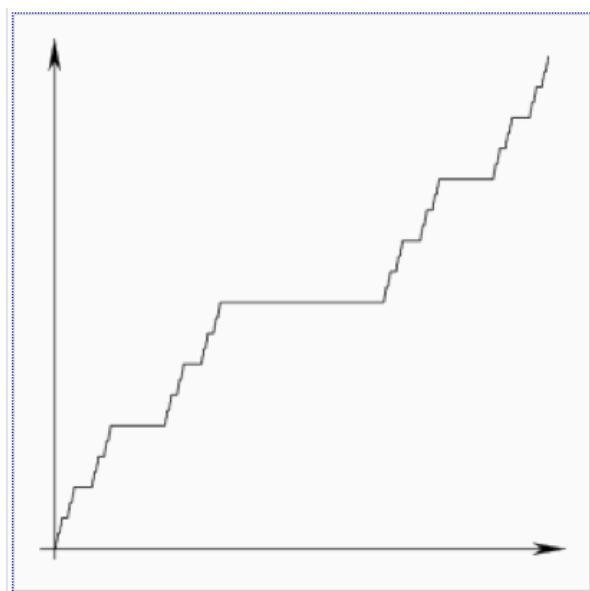


FIGURE 5.2 – Graphe de la fonction de Cantor

• Fonction de répartition continue et dérivable

Supposons maintenant que la fonction $F(x)$ ne comporte que la composante absolument continue : elle est non seulement continue mais aussi dérivable. Soit $f(x) = F'(x)$. Inversement $F(x)$ est la primitive de $f(x)$ s'annulant pour $x \rightarrow -\infty$, donc

$$F(x) = \int_{-\infty}^x f(x')dx' , \quad (5.14)$$

et en général

$$\mathbf{P}(a \leq X \leq b) = \int_a^b f(x)dx . \quad (5.15)$$

La fonction f , appelée *densité de probabilité*⁵, est positive (puisque F est croissante) et telle que $\mathbf{P}(\mathbb{R}) = \int_{-\infty}^{\infty} f(x)dx = 1$, on dit qu'elle est *normalisée*. Noter que $f(x)dx$ représente la probabilité que la v.a. X appartienne à l'intervalle $(x, x + dx)$.

Remarques.

1. Bien comprendre que dans ce cas d'une v.a. continue X à fonction de répartition F absolument continue, on ne parle pas de la probabilité que X prenne une *valeur* x , mais seulement qu'elle soit dans un *intervalle*, qu'il soit fini (a, b) , infini $(-\infty, a)$ ou infinitésimal $(x, x + dx)$.
2. Le *mode* de la loi d'une v.a. X est par définition la (ou les) valeur(s) de x où $f(x)$ atteint son maximum. Par extension, pour une v.a. discrète, c'est la (ou les) valeur(s) de k telles que $\mathbf{P}(X = k) = p_k$ soit maximale. Bien voir que les trois notions de valeur moyenne, mode et médiane sont distinctes. Exercice : dessiner le graphe d'une loi f dont le mode est inférieur à la valeur moyenne.
3. L'espérance et les moments s'expriment aisément en termes de f

$$E(X) \equiv \langle X \rangle = \int xf(x)dx , \quad \mu_k = \int x^k f(x)dx .$$

Graphe d'une loi f . Il est commode de visualiser une loi de probabilité par le graphe de la fonction f . Noter la relation entre ce graphe et les histogrammes que l'on peut construire à partir d'échantillons de N événements : tout histogramme à N fini donne (à un facteur N près) une approximation du graphe de f . Intuitivement, on s'attend à ce que dans la limite où $N \rightarrow \infty$ et où la largeur des intervalles tend vers 0, l'histogramme approche le graphe, puisque $f(x) = \lim_{h \rightarrow 0} \frac{F(x+h)-F(x)}{h}$, donc pour h petit, $f(x)h \approx \mathbf{P}(x \leq X < x + h)$, mais cela mériterait d'être justifié plus précisément ...

5.3.3 Plusieurs variables aléatoires

Ce qui vient d'être dit pour une v.a. continue peut être répété pour deux v.a., qui appliquent l'espace Ω sur \mathbb{R}^2 . Ce \mathbb{R}^2 peut être visualisé comme le plan muni de deux coordonnées x et y , et

5. appelée aussi *probability distribution function*, ou PDF, en anglais...

les deux v.a. X, Y prenant des valeurs x et y avec une certaine loi décrivent un point aléatoire dans le plan.

La fonction de répartition $F(x, y)$ est alors définie comme

$$F(x, y) = \mathbf{P}(X \leq x \cap Y \leq y) ; \quad (5.16)$$

quand F est continue et dérivable en x et en y , on lui associe la *densité de probabilité jointe* $f(x, y)$

$$f(x, y)dx dy = \mathbf{P}(x < X \leq x + dx \cap y < Y \leq y + dy) , \quad (5.17)$$

$f dx dy$ est la probabilité que le point aléatoire soit dans le “pavé” infinitésimal $(x, x + dx) \times (y, y + dy)$, et $F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(x', y')dx'dy'$.

Si f est factorisée, $f(x, y) = f_1(x)f_2(y)$, les deux v.a. X et Y sont indépendantes, puisque

$$\mathbf{P}(a \leq X < b \cap c \leq Y < d) = \int_a^b dx \int_c^d dy f(x, y) = \int_a^b f_1(x)dx \int_c^d f_2(y)dy = \mathbf{P}(a \leq X < b) \mathbf{P}(c \leq Y < d)$$

qui est bien la définition de l’indépendance, voir plus haut, (5.2). Réciproquement, vérifier que si X et Y sont indépendantes, leur loi f est factorisée.

• Distributions marginales

Soit $f(x, y)$ la densité de probabilité d’une paire de v.a. (X, Y) . On étudie la “loi marginale” sur la seule v.a. X quand on ne s’intéresse pas à la v.a. Y , autrement dit quand on somme sur les valeurs possibles de Y . La valeur moyenne/espérance de toute fonction $G(X)$ est donc donnée par

$$\langle G(X) \rangle = \int dx dy f(x, y)G(x) = \int dx G(x) \int dy f(x, y)$$

ou encore

$$\langle G(X) \rangle = \int dx f_1(x)G(x) \quad \text{avec} \quad f_1(x) \stackrel{\text{def}}{=} \int dy f(x, y) .$$

La loi marginale en x a donc pour densité $\int dy f(x, y)$, le mot “marginal” venant des tableaux comptables à deux entrées où les sommes selon une variable étaient reportées dans la ... marge du tableau !

• Fonctions de corrélation de plusieurs v.a.

Quand on dispose de plusieurs v.a., comme X et Y ci-dessus, on s’intéresse à leur fonction de corrélation $\langle XY \rangle$ qui vaut

$$\langle XY \rangle = \int f(x, y) xy dx dy . \quad (5.18)$$

Sous l'hypothèse de factorisation de la densité $f(x, y) = f_1(x)f_2(y)$, donc d'indépendance des v.a. X et Y , on a

$$\langle XY \rangle = \langle X \rangle \langle Y \rangle .$$

Inversement si $\langle XY \rangle \neq \langle X \rangle \langle Y \rangle$, cela implique que X et Y ne peuvent être indépendantes. Le calcul de

$$\langle XY \rangle_c \stackrel{\text{déf}}{=} \langle XY \rangle - \langle X \rangle \langle Y \rangle \quad (5.19)$$

fournit donc une mesure de la non-indépendance, ou de la corrélation, des v.a. X et Y . Plus généralement, on peut bien sûr définir des fonctions de corrélation plus compliquées, telles $\langle P(X)Q(Y) \rangle$ où P et Q sont par exemple des polynômes, pourvu que ces valeurs moyennes soient bien définies.

Remarque importante : $\langle XY \rangle_c = 0$ est une condition *nécessaire* pour que les v.a. X et Y soient indépendantes, mais elle n'est pas *suffisante* : elle n'implique pas que X et Y sont indépendantes. Ainsi soit φ une v.a. angulaire prenant des valeurs sur $[0, 2\pi[$ avec une densité uniforme $f(\varphi) = \frac{1}{2\pi}$. Soient les v.a. $X = \cos \varphi$ et $Y = \sin \varphi$. Montrer que $\langle XY \rangle_c = 0$. Il est pourtant clair que X et Y ne sont pas indépendantes, étant toutes deux liées à la v.a. φ . Exercice : montrer dans cet exemple que $\langle X^2 Y^2 \rangle \neq \langle X^2 \rangle \langle Y^2 \rangle$, ce qui vérifie la non-indépendance de X et Y .

On passe sans difficulté de 2 à n v.a. X_1, \dots, X_n . On rencontrera souvent dans la suite le cas de n v.a. indépendantes et de même loi, ou "également distribuées"⁶, donc dotées d'une densité factorisée $f_n(x_1, \dots, x_n) = \prod_{i=1}^n f(x_i)$.

Exemple physique : dans un barreau de fer, l'aimantation est due aux moments magnétiques \vec{m}_i portés par les N atomes, plus précisément à l'effet coopératif de ces moments microscopiques. Selon les lois de la Mécanique Statistique, on peut décrire un tel système par une approche probabiliste. À chaque configuration de moments magnétiques $\{\vec{m}_i\}$, on associe son énergie $\mathcal{E}(\{\vec{m}_i\})$ qui dépend de tous ces moments et de leurs interactions, et aussi d'éventuels champs magnétiques appliqués, etc. Le principe fondamental de la Mécanique Statistique est qu'à l'équilibre à une température T (température absolue, mesurée en Kelvin), la probabilité de cette configuration est

$$\mathbf{P}(\{\vec{m}_i\}) = \frac{1}{Z} e^{-\beta \mathcal{E}(\{\vec{m}_i\})} \quad (5.20)$$

où $\beta \stackrel{\text{déf}}{=} (k_B T)^{-1}$, avec k_B la constante de Boltzmann, $k_B = 1,380 \cdot 10^{-23} \text{ J K}^{-1}$. Le facteur exponentiel $e^{-\beta \mathcal{E}(\{\vec{m}_i\})}$ est appelé *poids de Boltzmann* de la configuration. Le facteur Z , appelé *fonction de partition* du système, est défini par

$$Z = \sum_{\text{config.}} e^{-\beta \mathcal{E}(\{\vec{m}_i\})} \quad (5.21)$$

où on somme sur toutes les configurations des moments magnétiques, de telle façon que la probabilité (5.20) est bien normalisée : $\sum_{\text{config.}} \mathbf{P}(\{\vec{m}_i\}) = 1$. Le résultat de l'analyse théorique, en accord avec l'observation, est qu'à haute température les moments microscopiques sont désordonnés, "pointant" dans toutes les orientations,

6. Les probabilistes anglo-saxons parlent de variables *i.i.d., independent, identically distributed*.

l'aimantation totale est alors nulle et le corps est dans sa *phase paramagnétique*; à température inférieure à la *température de Curie*, en revanche, les moments magnétiques ont tendance à s'orienter parallèlement les uns aux autres, créant ainsi une aimantation macroscopique \vec{M} : on est dans la *phase ferromagnétique*. La valeur moyenne $\langle \vec{m}_i \rangle$ du moment de l'atome i est nulle dans la phase paramagnétique, et non nulle et égale à $\frac{1}{N} \vec{M}$ (par définition) dans la phase ferromagnétique. La fonction de corrélation $\langle \vec{m}_i \vec{m}_j \rangle$ des moments de deux atomes i et j distants s'annule avec leur séparation dans la phase paramagnétique, elle tend vers le carré de $\langle \vec{m} \rangle$ dans la phase ferromagnétique.

5.3.4 Changement de variable aléatoire

On considère un phénomène aléatoire dont les événements sont décrits par une v.a. X réelle, par exemple une coordonnée d'un point M . Pour une raison ou une autre, on peut être intéressé à changer la description du phénomène en passant de la variable X à une autre, par exemple $Y = aX + b$ ou toute autre fonction (monotone croissante) $Y = \phi(X)$. Comment la loi de X se modifie-t-elle par ce changement de variable (ce changement de coordonnée)? Appelons $f(x)$ et $\hat{f}(y)$ les densités de probabilité qui décrivent le phénomène dans les deux variables. La quantité $f(x)dx$ qui fournit la probabilité que $x < X \leq x + dx$ doit être indépendante du choix de variable, en ce sens que

$$\mathbf{P}(M \in I) = \mathbf{P}(x < X < x + dx) = \mathbf{P}(y < Y < y + dy)$$

donc

$$f(x)dx = \hat{f}(y)dy \quad \text{avec } y = \phi(x) \quad \text{et} \quad dy = \phi'(x)dx \quad (5.22)$$

$$\text{donc } f(x) = \hat{f}(y) \frac{dy}{dx} = \hat{f}(y)\phi'(x) \iff \hat{f}(y) = \frac{f(x)}{\phi'(x)}. \quad (5.23)$$

Noter que l'on a supposé explicitement que y est une fonction croissante de x , sinon que faudrait-il faire?

Exemples 1. Un point aléatoire sur le segment $[-1, 1]$ est décrit par une coordonnée X ou par un angle Θ , $-\frac{1}{2}\pi \leq \Theta \leq \frac{1}{2}\pi$, avec $X = \sin \Theta$. Les densités sont reliées par $\hat{f}(\theta) = f(x) \frac{dx}{d\theta} = f(\sin \theta) \cos \theta$ ou $f(x) = \frac{\hat{f}(\theta)}{\sqrt{1-x^2}}$.

2. Passage en coordonnées radiales. Supposons qu'on se donne une distribution aléatoire de points dans le plan, décrits par une densité (de probabilité) $f(x, y)$. En coordonnées polaires, $(x, y) \mapsto (r, \theta)$, on va écrire

$$f(x, y) dx dy = \hat{f}(r, \theta) dr d\theta$$

mais on sait bien que l'élément de surface infinitésimal $dxdy = r dr d\theta$, donc $\hat{f}(r, \theta) = rf(x, y)$. Si la distribution en θ est uniforme, on peut se réduire à la seule coordonnée radiale r et écrire $\hat{f}(r) = 2\pi r f(x, y)$. Exemple : des points sont distribués dans le plan avec la densité $f(x, y) =$

$A \exp -\alpha(x^2 + y^2)$ (loi gaussienne, voir paragraphe suivant). Cette densité est évidemment invariante par rotation dans le plan (elle ne dépend que de la distance de l'origine au point M de coordonnées (x, y)), et il est naturel d'utiliser la coordonnée radiale $r = \sqrt{x^2 + y^2}$. La densité dans la v.a. R correspondante est $\hat{f}(r) = 2\pi Ar \exp -\alpha r^2$. De la même façon, une distribution dans l'espace \mathbb{R}^3 peut se récrire en coordonnées sphériques (r, θ, ψ) , avec $0 \leq \theta \leq \pi$, $0 \leq \psi \leq 2\pi$. On a alors $\hat{f}(r, \theta, \psi) = r^2 \sin \theta f(x, y, z)$, ou s'il y a invariance par rotation (indépendance en θ, ψ), $\hat{f}(r) = 4\pi r^2 f(x, y, z)$. On verra plus bas (§ 5.4.7) une application de ces considérations à la distribution de Boltzmann des vitesses des molécules d'un gaz.

5.3.5 Fonction caractéristique

Un outil très utilisé en probabilités est la *fonction caractéristique* de la v.a., qui est la transformée de Fourier de la densité f , notée ici φ

$$\varphi(u) = \int_{-\infty}^{\infty} e^{iux} f(x) dx . \quad (5.24)$$

Cette intégrale converge absolument puisque $|\int e^{iux} f(x) dx| < \int |e^{iux}| f(x) dx = \int f(x) dx = 1$. En étendant aux fonctions complexes la notion d'espérance

$$\varphi(u) = E(e^{iux}) .$$

Comme on a vu, la fonction φ (supposée intégrable) détermine complètement la fonction f , donc la loi de probabilité, grâce aux formules d'inversion

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iux} \varphi(u) du . \quad (5.25)$$

L'utilité de cette fonction caractéristique tient au résultat suivant :

Proposition 5.2 : *La somme de N v.a. indépendantes a pour fonction caractéristique le produit des fonctions caractéristiques.*

Preuve : Soient n v.a. indépendantes X_1, \dots, X_n de densités de probabilité f_1, \dots, f_n . La densité de probabilité jointe est, on l'a vu, le produit $\prod_i f_i(x_i)$. La v.a. $X = \sum_n X_i$ a pour densité

$$f(x) = \int dx_1 \cdots \int dx_n \prod_i f_i(x_i) \delta(x - \sum_i x_i)$$

comme on le vérifie en calculant l'espérance de toute fonction G de X seule :

$$\langle G(X) \rangle = \int \prod_i (dx_i f_i(x_i)) G(\sum_i x_i) = \int \prod_i (dx_i f_i(x_i)) \int dx \delta(x - \sum_i x_i) G(x) = \int dx f(x) G(x) .$$

La v.a. X a donc pour fonction caractéristique

$$\begin{aligned}\varphi(u) &= \int_{-\infty}^{\infty} e^{iux} f(x) dx = \int_{-\infty}^{\infty} dx e^{iux} \int dx_1 \cdots \int dx_n \prod_i f_i(x_i) \delta(x - \sum_i x_i) \\ &= \int dx_1 \cdots \int dx_n \prod_j f_j(x_j) e^{iux_j} = \prod_j \varphi_j(u).\end{aligned}\quad (5.26)$$

En particulier

La somme de N v.a. indépendantes et de même loi f a pour fonction caractéristique la puissance N -ième de la fonction caractéristique de f

$$\varphi_N(u) = \varphi(u)^N. \quad (5.27)$$

On peut définir aussi cette fonction caractéristique dans le cas d'une variable discrète entière : soit X une v.a. discrète prenant la valeur entière k avec la probabilité p_k . On a $\varphi(u) = \sum_k p_k e^{iuk}$ ou, de façon équivalente, $\phi(z = e^{iu}) = \sum_k p_k z^k$. Sous cette dernière forme, comme fonction de z , on l'appelle aussi *fonction génératrice* des probabilités p_k . En effet on retrouve les p_k à partir de cette fonction génératrice par différentiation

$$p_k = \frac{1}{k!} \left(\frac{d}{dz} \right)^k \phi(z) \Big|_{z=0}. \quad (5.28)$$

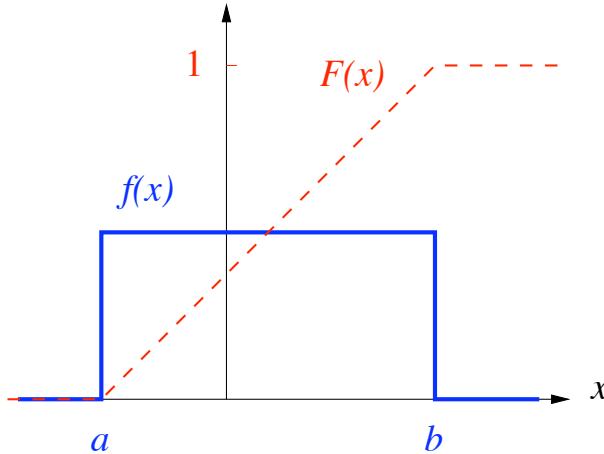
Exemple : loi binomiale. Une v.a. prend la valeur 0 avec la probabilité $1 - p$ et 1 avec la probabilité p , donc $\phi(z) = 1 - p + pz$. La somme de N v.a. indépendantes et de même loi a pour fonction caractéristique $\phi_N(z) = \phi(z)^N = \sum_{k=0}^N C_N^k (1-p)^{N-k} p^k z^k$. Cela donne la réponse à un problème classique : si on tire à pile (0) ou face (1), quelle est la probabilité d'un score total k après N tirages ? Réponse, $p_k = (1-p)^{N-k} p^k C_N^k$. On reviendra sur cette loi binomiale au paragraphe suivant.

5.4 Distributions classiques

Dans ce paragraphe nous définissons et étudions quelques distributions usuelles, ou *lois*, de variables aléatoires discrètes ou continues.

5.4.1 Distribution uniforme

C'est la plus simple des distributions de v.a. : la densité est constante et non nulle dans un intervalle $[a, b]$. La valeur de cette constante p est fixée par la normalisation $\int_a^b f(x) dx =$

FIGURE 5.3 – Densité et fonction de répartition d'une distribution uniforme sur l'intervalle $[a, b]$.

$(b - a)p = 1$ soit $p = \frac{1}{b-a}$, voir Fig. 5.3. La fonction de répartition $F(x)$ est nulle pour $x \leq a$, croît linéairement de a à b et vaut 1 pour $x \geq b$.

Dans sa version discrète, où la v.a. prend un nombre fini de valeurs x_i avec des probabilités p_i , les p_i sont égales, les valeurs x_i sont *équiprobables*. C'est le cas d'une pièce (au jeu de pile ou face) ou d'un dé bien équilibré, avec des probabilités respectives de $p_p = p_f = \frac{1}{2}$ ou de $p_i = \frac{1}{6}$. Dans sa version continue, c'est la distribution des générateurs de nombres aléatoires usuels, qui engendrent des nombres x_i de l'intervalle $[0, 1]$ avec une densité de probabilité constante. À partir de cette densité sur $[0, 1]$, on fabrique une distribution uniforme sur tout autre intervalle $[a, b]$ en translatant et dilatant la variable en $y = x(b - a) + a$ et en prenant $\tilde{f}(y) = \frac{1}{b-a}f(x) = \frac{1}{b-a}f(\frac{y-a}{b-a})$ pour $y \in [a, b]$.

• Moyenne, variance, moments, fonction caractéristique

Le calcul de la moyenne, de la variance et des moments est aisé

$$\langle X \rangle = \frac{1}{b-a} \int_a^b dx x = \frac{a+b}{2} \quad (5.29)$$

$$m_k = \langle X^k \rangle = \frac{1}{b-a} \int_a^b dx x^k = \frac{b^{k+1} - a^{k+1}}{(k+1)(b-a)} = \frac{1}{k+1} (a^k + a^{k-1}b + \cdots + b^k) \quad (5.30)$$

$$\text{var } X = \langle X^2 \rangle - \langle X \rangle^2 = \frac{1}{12} (a-b)^2 \quad (5.31)$$

$$\varphi(u) = \int_a^b dx e^{iux} \frac{1}{b-a} = \frac{e^{iub} - e^{iua}}{iu(b-a)} \quad (5.32)$$

$$(5.33)$$

5.4.2 Distribution binomiale

C'est la distribution rencontrée au paragraphe § 5.3.2 dans la discussion des sommes des valeurs de pile ou face : si on a tiré à pile ou face n fois, quelle est la probabilité de k "faces" ?

Supposons la pièce mal équilibrée, avec une probabilité p de tomber sur face, $0 < p < 1$ et $q = 1 - p$ de tomber sur pile. Dans une série de n lancers, les résultats successifs sont des variables indépendantes (la pièce ne garde pas le souvenir si elle est tombée sur pile ou face les fois précédentes!). Par conséquent la probabilité d'une suite donnée de k faces et donc de $n - k$ piles est indépendante de l'ordre de ces faces ou piles et elle est factorisée de la forme $p^k q^{n-k}$. Comme deux suites de k faces différant par leur ordre sont des événements A et B mutuellement exclusifs, la probabilité de $A \cup B$ est la somme des probabilités de A et de B , chacune égale à $p^k q^{n-k}$, et plus généralement la probabilité totale cherchée $p_k^{(n)}$ est $p^k q^{n-k}$ fois le nombre total de suites avec k faces, soit C_n^k , d'où

$$\mathcal{B}(n, p) : \quad \mathbf{P}(X = k) = p_k^{(n)} \stackrel{\text{déf}}{=} C_n^k p^k q^{n-k} . \quad (5.34)$$

La loi est bien normalisée : $\sum p_k^{(n)} = \sum_{k=0}^n C_n^k p^k q^{n-k} = (p + q)^n = 1$.

La distribution (5.34) est la *loi binomiale* $\mathcal{B}(n, p)$.

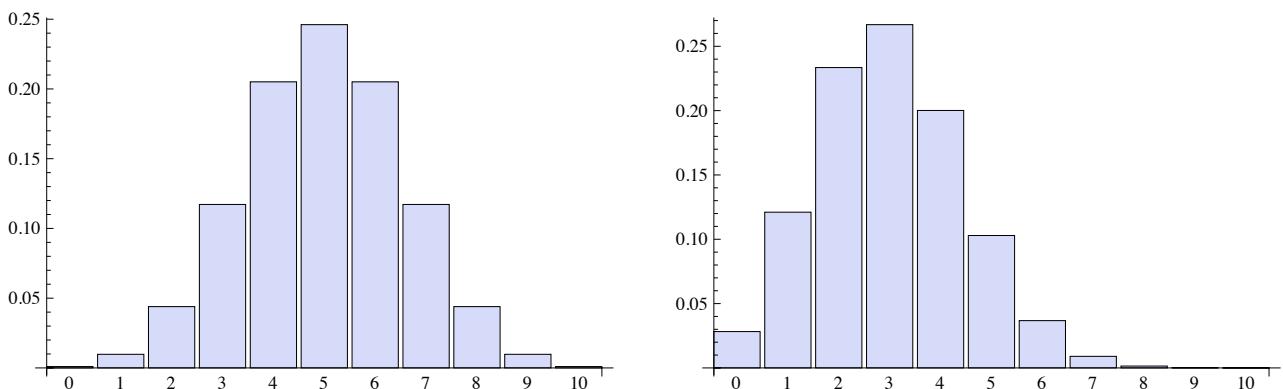


FIGURE 5.4 – Distribution binomiale pour $p = 0,5$ et pour $p = 0,3$ et $n = 10$.

Elle se rencontre chaque fois qu'on s'intéresse à n occurrences d'un événement se produisant ou non avec une probabilité p ou $1 - p$. Par exemple, on considère n noyaux radioactifs dont chacun a la probabilité p de se désintégrer pendant un intervalle Δt . Quelle est la probabilité que k noyaux se soient désintégrés au bout de ce temps ? (on suppose que les désintégrations sont des phénomènes aléatoires indépendants, c'est-à-dire qu'elles ne s'influencent pas mutuellement.) Nous reviendrons en détail sur cet exemple au § 5.4.7.

• **Moyenne, variance, moments. Fonction caractéristique**

À nouveau, le calcul de la moyenne et de la variance est un bon exercice

$$\langle X \rangle = \sum_k kp_k^{(n)} = np \quad (5.35)$$

$$\langle X^2 \rangle = \sum_k k^2 p_k^{(n)} = n(n-1)p^2 + np \quad (5.36)$$

$$\text{var}X = \langle X^2 \rangle - \langle X \rangle^2 = n(p-p^2) = npq \quad (5.37)$$

$$(5.38)$$

dans lequel on pourra utiliser de façon répétée l'identité $kC_n^k = nC_{n-1}^{k-1}$. Pour la fonction caractéristique $\varphi(u) = \sum_k p_k e^{iuk}$, on a pour la distribution binomiale

$$\varphi_{\text{Bin}(n)}(u) = \sum_k p_k^{(n)} e^{iuk} = \sum_{k=0}^n C_n^k e^{iuk} p^k q^{n-k} = (1-p+pe^{iu})^n. \quad (5.39)$$

Nous ferons usage de cette expression dans la suite.

5.4.3 Distribution normale

Rappelons d'abord que

$$I = \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}. \quad (5.40)$$

C'est la très fameuse intégrale gaussienne.

Cette formule très utile se démontre aisément de la façon suivante : calculons le carré de l'intégrale, considérée comme une intégrale dans le plan et appliquons lui le changement de variables en coordonnées polaires. Explicitement $I^2 = \int_{-\infty}^{\infty} dx e^{-x^2} \int_{-\infty}^{\infty} dy e^{-y^2} = \int_{\mathbb{R}^2} dx dy e^{-x^2-y^2} = \int_{\mathbb{R}^2} r dr d\theta e^{-r^2} = 2\pi \int_0^{\infty} \frac{1}{2} d(r^2) e^{-r^2} = \pi \int_0^{\infty} du e^{-u} = \pi$, cqfd.

Par un simple changement de variable, on en déduit que pour tout $\alpha > 0$

$$I(\alpha) = \int_{-\infty}^{\infty} e^{-\alpha x^2} dx = \sqrt{\frac{\pi}{\alpha}} \quad (5.41)$$

et en particulier

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma^2}} dx = 1. \quad (5.42)$$

On définit alors une v.a. gaussienne x par sa densité

$$\mathcal{N}(\mu, \sigma) : \quad f_{\text{Gauss}}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \quad (5.43)$$

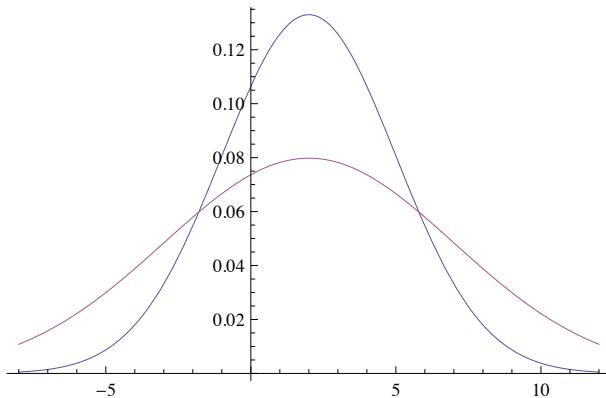


FIGURE 5.5 – Densité d'une distribution gaussienne, $\mu = 2$, $\sigma = 3$ et $\sigma = 5$. Le maximum est atteint en $x = \mu$, la largeur du pic (par exemple à la hauteur e^{-1} fois le maximum) est proportionnelle à σ .

On parle de la *loi normale* $\mathcal{N}(\mu, \sigma)$. Elle dépend de deux paramètres μ et σ^2 que nous allons maintenant interpréter. Calculons pour cela la moyenne et la variance de cette v.a.

$$\langle X \rangle = \int_{-\infty}^{\infty} dx x f_{\text{Gauss}}(x) = \int_{-\infty}^{\infty} dx ((x - \mu) + \mu) f_{\text{Gauss}}(x) = 0 + \mu :$$

le premier terme est nul car dans le calcul de $\langle x - \mu \rangle$ on peut effectuer le changement de variable d'intégration $x' = x - \mu$ et l'intégrale qui en résulte d'une fonction impaire $x' e^{-ax'^2}$ sur $(-\infty, \infty)$ est nulle ; le deuxième terme de la moyenne est égal à μ car l'intégrale est bien normalisée.

Pour le calcul de la variance, remarquons d'abord qu'en différentiant les deux membres de $I(\alpha)$ dans (5.41)⁷ par rapport à α on obtient $-\frac{\partial}{\partial \alpha} I(\alpha) = \int dx x^2 e^{-\alpha x^2} = \frac{1}{2\alpha} I(\alpha)$. Prenant $\alpha = (2\sigma^2)^{-1}$, on est alors en mesure de calculer

$$\text{var}X = \int_{-\infty}^{\infty} dx (x - \mu)^2 f_{\text{Gauss}}(x) = \int \frac{dx}{\sqrt{2\pi\sigma^2}} (x - \mu)^2 e^{\frac{-(x-\mu)^2}{2\sigma^2}} = \int \frac{dx'}{\sqrt{2\pi\sigma^2}} x'^2 e^{\frac{-x'^2}{2\sigma^2}} = \sigma^2 .$$

Par conséquent les deux paramètres μ et σ^2 dont dépend la loi gaussienne (5.43) s'interprètent respectivement comme sa moyenne et sa variance.

Le graphe de $f(x)$ a une allure caractéristique de courbe en cloche, voir Fig. (5.5). Elle est “centrée” en $x = \mu$, et est d'autant plus étroite que l'écart-type σ (ou la variance σ^2) est plus faible. Cette variance peut être mesurée par la “largeur du pic”, définie comme la distance $2\sqrt{2}\sigma$ entre les points $x = \mu \pm \sqrt{2}\sigma$ où f est réduit d'un facteur e par rapport à son maximum :

$$f(\mu \pm \sqrt{2}\sigma) = \frac{f(\mu)}{e} .$$

Définition 5.5 : On dit qu'une v.a. X a une loi normale centrée normalisée si elle est gaussienne avec $\mu = 0$ et $\sigma^2 = 1$ (loi normale $\mathcal{N}(0, 1)$).

7. La “dérivation sous le signe somme” est ici parfaitement justifiée, pourquoi ?

- **Moments.**

Les calculs qui précédent ont préparé le terrain au calcul des moments d'ordre arbitraire. Par changement de variable on ramène tout calcul de moment de X à celui de $X' = X - \mu$. Pour la variable X' qui est centrée, tous les moments impairs sont nuls, par l'argument de parité donné pour la moyenne. Ses moments pairs s'obtiennent à nouveau par différentiation répétée par rapport à α dans (5.41).

Exercice. Selon ce principe, vérifier les formules suivantes

$$\langle X^2 \rangle = \mu^2 + \sigma^2 \quad ; \quad \langle X^3 \rangle = \mu^3 + 3\mu\sigma^2 \quad ; \quad \langle X^4 \rangle = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4 .$$

On verra plus bas une autre manière de calculer ces moments.

La conclusion importante est que *pour la distribution gaussienne, tous les moments existent et sont calculables explicitement*.

Au vu de l'inégalité de Tchebychev (5.8), l'existence de moments d'ordre quelconque implique que la probabilité $\mathbf{P}(|X| > a)$ s'annule très vite quand a croît, plus vite que toute puissance inverse a^{-k} ; ceci est en accord avec notre intuition : dans la distribution gaussienne les événements très éloignés de la moyenne sont extrêmement rares.

- **Fonction caractéristique**

Puisque la fonction caractéristique est une transformée de Fourier, on connaît par le Chap. 4 le résultat pour la loi normale : la transformée de Fourier d'une gaussienne centrée ($\mu = 0$) est une gaussienne centrée, et en général, pour tout μ

$$\varphi_{\text{Gauss}}(u) = e^{i\mu u} e^{-\frac{1}{2}u^2\sigma^2}, \quad (5.44)$$

une expression qui permet aussi de calculer aisément tous les moments de la distribution gaussienne. En effet la fonction caractéristique $\varphi(u)$ peut être considérée comme une *fonction génératrice des moments*, en ce sens que sa dérivée n -ième en 0 vaut

$$\varphi_{\text{Gauss}}^{(n)}(0) = i^n \int dx x^n f(x)$$

qui s'annule pour n impair, comme on a vu; donc $\varphi_{\text{Gauss}}^{(2p)}(0) = (-1)^p m_{2p}$ est au signe près, le $2p$ -ième moment de la distribution $\mathcal{N}(\mu, \sigma)$. Montrer en utilisant l'expression (5.44) que si $\mu = 0$

$$m_{2p} = (2p-1)!! \sigma^{2p}, \quad (5.45)$$

où $(2p-1)!! \stackrel{\text{déf}}{=} 1.3.\dots.(2p-3)(2p-1)$. Ces calculs prendront une grande importance en mécanique statistique et en théorie quantique des champs...

- **Importance des variables gaussiennes**

Pour des raisons que l'on va voir, cette distribution joue un rôle particulièrement important. La distribution gaussienne joue un rôle central aussi bien en probabilités et statistiques qu'en

physique. En probabilités et statistiques, cela tient en particulier au théorème limite central, que nous étudierons au paragraphe suivant, selon lequel cette distribution apparaît comme limite de nombreux phénomènes aléatoires. En physique, elle est au cœur de très nombreuses considérations ; en physique statistique, en relation avec les poids de Boltzmann : pour un système de particules sans interactions, tel un gaz parfait, l'énergie cinétique est $E = \sum \frac{1}{2}mv_i^2$, et les vitesses considérées comme des v.a. obéissent à une loi gaussienne dictée par le poids de Boltzmann $e^{-\frac{E}{k_B T}} = \prod_i e^{-\frac{mv_i^2}{2k_B T}}$, (voir ci-dessous § 5.4.7) ; c'est le cas encore avec l'étude d'ensembles d'oscillateurs harmoniques, que l'on rencontre en physique des solides et en théorie quantique des champs, etc etc.

5.4.4 Distribution de Poisson

Soit λ un paramètre réel positif, et k une v.a. discrète à valeurs entières ≥ 0 . La *distribution de Poisson* est définie par la loi de probabilité

$$\mathcal{P}(\lambda) : \boxed{\mathbf{P}(X = k) = p_k^{\text{Poisson}} \stackrel{\text{def}}{=} \frac{\lambda^k}{k!} e^{-\lambda}} . \quad (5.46)$$

Il convient d'abord de vérifier que cette loi est bien normalisée, ce qui est immédiat : $\sum_{k=0}^{\infty} p_k = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = 1$.

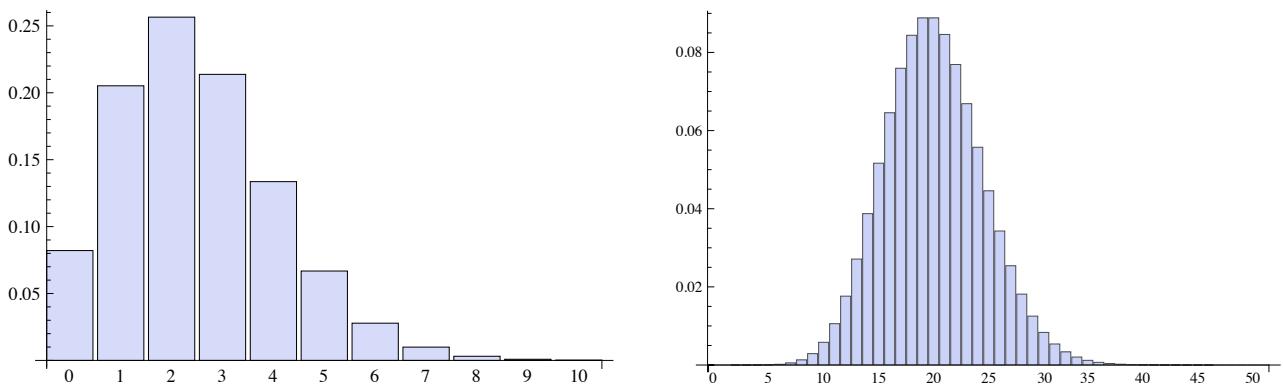


FIGURE 5.6 – Distribution de Poisson pour $\lambda = 2,5$ et $\lambda = 20$.

- **Moyenne, variance, moments, fonction caractéristique**

Une fois encore, on calcule aisément les moyenne et variance de la distribution et sa fonction caractéristique :

$$\langle X \rangle = \sum_{k=0}^{\infty} k \lambda^k \frac{e^{-\lambda}}{k!} = \lambda \quad (5.47)$$

$$\langle X^2 \rangle = \sum_{k=0}^{\infty} k^2 \lambda^k \frac{e^{-\lambda}}{k!} = \lambda^2 + \lambda \quad (5.48)$$

$$\text{var}X = \langle X^2 \rangle - \langle X \rangle^2 = \lambda \quad (5.49)$$

$$\varphi(u) = e^{\lambda(e^{iu}-1)} \quad \text{ou} \quad \phi(z) = e^{\lambda(z-1)}. \quad (5.50)$$

Exercices : vérifier ces formules. Chercher le ou les mode(s) de $\mathcal{P}(\lambda)$, au sens du § 5.3.2. Montrer que si λ est un entier n , les modes sont les deux valeurs $n-1$ et n ; et que si λ n'est pas entier, le mode est $\lfloor \lambda \rfloor$ (partie entière de λ).

- **Importance de la loi de Poisson**

La loi de Poisson apparaît dans de nombreux phénomènes naturels ou de la vie pratique, chaque fois que des événements indépendants se produisent à un taux constant λ (dans l'espace et/ou le temps) : la loi p_k^{Poisson} donne alors la probabilité d'avoir k occurrences de l'événement. Ainsi supposons que N objets ont été distribués au hasard pendant un temps T donné (cela peut s'appliquer à des appels téléphoniques, à des impacts de particules émises par une source, à des gouttes de pluie ou autres projectiles supposés répartis à un rythme constant, etc). On s'attend donc à en recevoir en moyenne $\lambda = N/T$ par unité de temps. La probabilité d'en recevoir k pendant un temps $t \leq T$ est donnée par la loi de Poisson de paramètre λt . Exemple : je reçois en moyenne 50 e-mails par jour, répartis *grossost modo* de façon uniforme dans le temps. Quelle est la probabilité que j'en reçoive k un certain jour ? Réponse : p_k^{Poisson} avec $\lambda = 50$. Que j'en reçoive 9 en une demi-journée ? Réponse : p_9^{Poisson} avec $\lambda' = Nt/T = 25$.

Pour se convaincre que cette interprétation est correcte, montrons que la loi de Poisson jouit d'une propriété d'invariance par "divisibilité" dans le sens suivant. Considérons un processus décrivant l'occurrence d'événements dans un intervalle de temps (ou un intervalle de droite, un élément de surface, etc) suivant une loi de Poisson $\mathbf{P}(X = n) = \frac{\lambda^n}{n!} e^{-\lambda}$, et supposons qu'on subdivise cet intervalle en m sous-intervalles de même taille. On s'attend à ce que les événements se répartissent dans ces sous-intervalles de façon indépendante et avec la même loi. Si le nombre total d'événements X est fixé, le nombre X_i d'événements dans le i -ième sous-intervalle donné suit la loi binomiale : un événement a une probabilité $1/m$ de se produire dans l'intervalle considéré, et la probabilité $(m-1)/m$ de se produire en dehors, donc $\mathbf{P}(X_i = k|X = n) =$

$C_n^k \frac{1}{m^k} \left(\frac{m-1}{m}\right)^{n-k} = \frac{n!(m-1)^{n-k}}{k!(n-k)!m^n}$. Par la formule de Bayes, $\mathbf{P}(X_i = k \cap X = n) = \frac{n!(m-1)^{n-k}}{k!(n-k)!m^n} \frac{\lambda^n}{n!} e^{-\lambda}$ et la probabilité $\mathbf{P}(X_i = k)$ est donc

$$\begin{aligned} \mathbf{P}(X_i = k) &= \sum_{n=k}^{\infty} \mathbf{P}(X_i = k \cap X = n) = \frac{e^{-\lambda} \lambda^k}{k! m^k} \sum_{n=k}^{\infty} \frac{\lambda^{n-k} (m-1)^{n-k}}{(n-k)! m^{n-k}} \\ &= \frac{e^{-\lambda} \lambda^k}{k! m^k} e^{\lambda \frac{m-1}{m}} = \frac{e^{-\lambda}}{k!} \left(\frac{\lambda}{m}\right)^k \end{aligned} \quad (5.51)$$

qui est une loi de Poisson de paramètre λ/m , cqfd. Bien entendu, $\langle X_i \rangle = \text{var}(X_i) = \frac{\lambda}{m}$.

Cette même distribution de Poisson s'applique aussi aux phénomènes quantiques. Ainsi le nombre de particules ou de photons émis par une source avec un taux moyen donné suit une loi de Poisson. On parle de bruit (par exemple de photons) poissonnien.

Cette apparition de la loi de Poisson tient au fait qu'elle décrit une certaine limite de la loi binomiale, comme on le verra au § 5.4.6.

5.4.5 Quelques autres lois rencontrées dans les sciences naturelles

- **Loi exponentielle**

Pour $\alpha \in \mathbb{R}_+$, on définit la loi exponentielle de paramètre α par

$$f_{\text{exp}}(x) = \alpha e^{-\alpha x} \quad \text{si } x > 0, \quad 0 \text{ sinon}. \quad (5.52)$$

Sa fonction de répartition est $F(x) = \int_{-\infty}^x f_{\text{exp}}(x') dx' = 1 - e^{-\alpha x}$. Une propriété remarquable de cette loi est son “absence de mémoire” : $\mathbf{P}(X > x + y | X > y) = \mathbf{P}(X > x) \quad \forall x, y \geq 0$.

- **Loi lorentzienne ou de Cauchy**

Pour $\lambda \in \mathbb{R}_+$, on définit la loi lorentzienne de paramètre λ (aussi appelée loi de Cauchy) par

$$f_{\text{Lorentz}}(x) = \frac{1}{\pi} \frac{\lambda}{x^2 + \lambda^2} \quad x \in \mathbb{R}. \quad (5.53)$$

Sa fonction de répartition est $F(x) = \int_{-\infty}^x f_{\text{Lorentz}}(x') dx' = \frac{1}{2} + \frac{1}{\pi} \text{Arctan} \frac{x}{\lambda}$.

Calcul de l'espérance ? On voit immédiatement que $E(X) \equiv \langle X \rangle$ n'est pas définie. La lorentzienne est un exemple de “loi large” dont l'espérance n'est pas définie.

- **Loi de Pareto**

Une variable aléatoire X suit la loi de Pareto de paramètre $\alpha > 0$ si elle prend des valeurs réelles supérieures à un réel positif x_m et que la probabilité $\mathbf{P}(X > x)$ est de la forme

$$\mathbf{P}(X > x) = \left(\frac{x_m}{x}\right)^\alpha \quad \text{pour } x > x_m, \quad 0 \text{ sinon}. \quad (5.54)$$

avec k un réel positif. On en déduit la fonction $F(x)$ de répartition de X , $F(x) = p(X \leq x) = 1 - \left(\frac{x_m}{x}\right)^\alpha$ et la densité de probabilité $f(x)$ de X , $f(x) = F'(x) = \alpha \frac{x_m^\alpha}{x^{\alpha+1}}$. Cette loi intervient beaucoup en économie

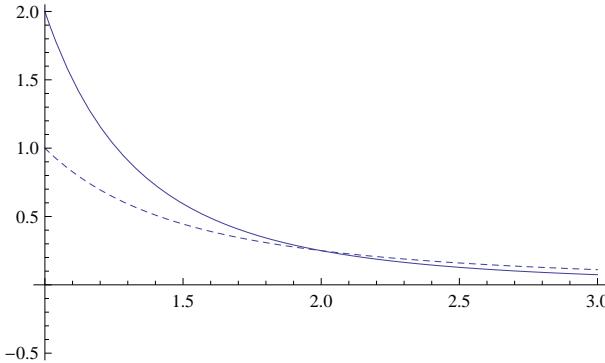


FIGURE 5.7 – Densité de probabilité de la loi de Pareto, $\alpha = 1$ (courbe en tirets), $\alpha = 2$ (continue), avec $x_m = 1$.

(distribution des richesses) mais aussi en géophysique (distributions des ressources en minéraux, des tailles des météorites, ...), etc.

Exercice : déterminer pour quelles valeurs de α les moments d'ordre $\leq k$ sont définis. Calculer la probabilité $\mathbf{P}(X > x + y | X > y)$ pour $x > 0$, $y > x_m$ et sa limite quand $y \rightarrow \infty$.

On verra en TD d'autres de ces lois.

5.4.6 Limites de la loi binomiale

- **Limite gaussienne.**

Comme le suggère le graphe de la distribution binomiale (voir Fig. (5.4)), avec sa forme en cloche très proche de celle de la gaussienne, la distribution gaussienne est une bonne approximation de la loi binomiale $p_k^{(n)}$ dans la limite des n grands, et pour $k \sim np$, avec p et $q = 1 - p$ finis, donc $n \gg 1$, $k \gg 1$, $n - k \gg 1$. C'est donc la partie centrale de la distribution que l'on explore là.

La démonstration repose sur l'étude dans la limite considérée des fonctions caractéristiques calculées plus haut, cf (5.39) et (5.44). Calculons le logarithme de $\varphi_{\text{Bin}(n)}(u)$ pour u petit (qui seul nous intéresse si $k \gg 1$)

$$\begin{aligned} \ln \varphi_{\text{Bin}(n)}(u) &= n \ln(1 + p(e^{iu} - 1)) \simeq n \left(p(e^{iu} - 1) - \frac{1}{2} p^2 (e^{iu} - 1)^2 + \dots \right) \\ &= n \left(p(iu - \frac{u^2}{2} + \dots) + \frac{1}{2} p^2 u^2 + O(u^3) \right) \\ &= n \left(iup - \frac{u^2}{2} p(1-p) + \dots \right). \end{aligned} \quad (5.55)$$

et comparons-le à celui de $\varphi_{\text{Gauss}}(u)$, cf (5.44)

$$\ln \varphi_{\text{Gauss}}(u) = iu\mu - \frac{1}{2}u^2\sigma^2$$

Pour $k \gg 1$, donc $u \ll 1$, on voit que les deux fonctions caractéristiques coïncident pourvu que d'une part les valeurs moyennes $\langle X \rangle = np$ et μ , de l'autre les variances $\text{var } X = np(1-p)$ et σ^2 coïncident.

En pratique, pour comparer la distribution (discrète) binomiale et la distribution (continue) gaussienne, il faut calculer numériquement à l'aide de cette dernière la probabilité que la variable x soit dans un intervalle de longueur 1 centré sur l'entier k

$$p'_k = \mathbf{P}(k - \frac{1}{2} < x < k + \frac{1}{2}) = \int_{k-\frac{1}{2}}^{k+\frac{1}{2}} dx f_{\text{Gauss}}(n) = F(k + \frac{1}{2}) - F(k - \frac{1}{2})$$

et la comparer à $p_k^{(n)}$ calculé avec la loi binomiale. Par exemple pour $n = 50$ et $p = 0,2$, on trouve que p'_k diffère de $p_k^{(n)}$ d'au plus 8% dans l'intervalle $6 \leq k \leq 16$, voir la figure 5.8.

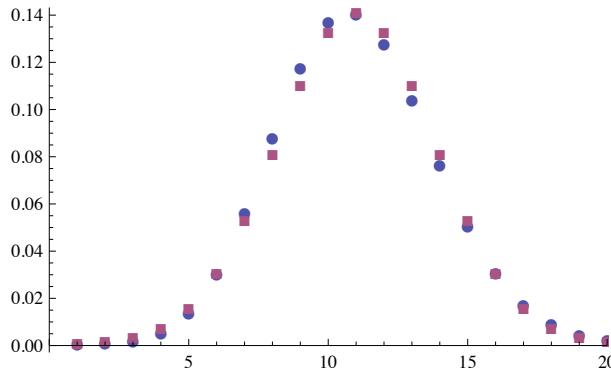


FIGURE 5.8 – Comparaison de la distribution binomiale ($n = 50$, $p = 0,2$: petits disques) et de la gaussienne ($\mu = 10$, $\sigma = \sqrt{np(1-p)}$: petits carrés) intégrée dans chaque intervalle de longueur 1.

Limite poissonnienne

Il peut être aussi intéressant d'étudier la limite de la loi binomiale quand p est très petit et np (la valeur moyenne de k) est finie, tandis que $n \gg 1$. Il faut bien comprendre que l'on explore maintenant le *bord* de la distribution binomiale, loin de son maximum. Poisson a montré que l'on trouve alors la distribution qui porte son nom.

$$p_k^{(n)} \Big|_{n \gg 1} \simeq p_k^{\text{Poisson}} = \frac{\lambda^k}{k!} e^{-\lambda} \quad (5.56)$$

avec $\lambda = np$ supposée finie tandis que $n \rightarrow \infty$.

Démonstration de cette limite :

$$p_k^{(n)} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \frac{n(n-1) \cdots (n-k+1)}{k!} \frac{(np)^k}{n^k} \left(1 - \frac{np}{n}\right)^{n-k} \quad (5.57)$$

$$= \frac{n(n-1) \cdots (n-(k-1))}{k!} \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k} \quad (5.58)$$

$$= \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^n \times \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) \left(1 - \frac{\lambda}{n}\right)^{-k}$$

$$\simeq \frac{\lambda^k}{k!} e^{-\lambda} (1 + O(n^{-1})) \quad (5.59)$$

en se rappelant que $\lim_{n \rightarrow \infty} (1 - \frac{\lambda}{n})^n = e^{-\lambda}$.

A nouveau une petite expérience numérique est utile. Pour $n = 50$, $p = 0,02$, on voit que la distribution de Poisson pour $\lambda = np = 1$ approxime à moins de 2% la binomiale pour $0 \leq k \leq 3$, comme le montre cette table

k	0	1	2	3	4	5
Binomiale $p_k^{(50)}$	0,36417	0,37160	0,18580	0,06067	0,01454	0,002731
Poisson $p_k^{\text{Poisson}} _{\lambda=1}$	0,36788	0,36788	0,18394	0,06131	0,01533	0,003066

TABLE 5.1 – Comparaison des lois binomiale et de Poisson

Récapitulons. Pour n très grand et p finie, la loi binomiale est bien approximée par une gaussienne pour les valeurs de k les plus probables, $k \approx pn \sim n$: les sommets des deux cloches binomiale et gaussienne se superposent bien. Pour p très petit et $\langle k \rangle = pn$ fini, les événements à $k \ll n$ (le bord gauche de la cloche binomiale) sont bien décrits par la loi de Poisson.

5.4.7 Quelques exemples concrets

- **L'aiguille de Buffon**

Voici le problème posé par Buffon : on laisse tomber une aiguille sur un parquet, quelle est la probabilité que l'aiguille chevauche deux lames du parquet ?

De façon plus précise, soit un ensemble de droites parallèles à l'axe des x , espacées d'une distance ℓ . Quelle est la probabilité pour un segment de longueur $b \leq \ell$ uniformément distribué en position et en orientation d'être sécant à une de ces droites ?

Dans la bande (la lame du parquet) où se trouve le milieu du segment, soit y la distance de ce milieu au bord inférieur. Soit θ l'angle qu'il fait avec le bord, $0 \leq \theta < \pi$.

L'hypothèse de distribution uniforme en position et angle signifie que la densité de probabilité autour du point (y, θ) est $f(y, \theta) = \frac{1}{\ell} \frac{1}{\pi}$ qui est bien normalisée puisque $\int dy d\theta f(y, \theta) = \int_0^\ell \frac{dy}{\ell} \int_0^\pi \frac{d\theta}{\pi} = 1$.

La condition que le segment recoupe un bord est $y < \frac{1}{2}b \sin \theta$ (bord inférieur) ou $\ell - y < \frac{1}{2}b \sin \theta$ (bord supérieur). La probabilité pour le segment de *ne pas être sécant* est donc

$$\begin{aligned}
 P_{NC} &= \int_0^\pi \frac{d\theta}{\pi} \int_{\frac{1}{2}b \sin \theta}^{\ell - \frac{1}{2}b \sin \theta} \frac{dy}{\ell} \\
 &= \int_0^\pi \frac{d\theta}{\pi} \left(1 - \frac{b}{\ell} \sin \theta\right) \\
 &= 1 - \frac{2b}{\pi \ell}.
 \end{aligned} \tag{5.60}$$

La probabilité de croisement est donc

$$P_C = \frac{2b}{\pi \ell}.$$

• Distribution de Maxwell–Boltzmann des vitesses dans un gaz

Dans un gaz (parfait) à l'équilibre, soit $f(v_x, v_y, v_z)$ la densité de probabilité des vitesses d'une molécule. Montrons que des considérations de symétrie et d'indépendance conduisent à l'expression de f . Supposons en effet que les v.a. v_x , v_y et v_z sont indépendantes, une hypothèse qu'il faudrait justifier soigneusement... La fonction f est donc factorisée, cf § 5.3.3, et le gaz étant isotrope, les trois directions d'espace jouent le même rôle et les trois fonctions sont identiques

$$f(v_x, v_y, v_z) = g(v_x)g(v_y)g(v_z).$$

Par ailleurs, par invariance par rotation, la fonction $f(v_x, v_y, v_z)$ doit n'être fonction que de $\bar{v}^2 = v_x^2 + v_y^2 + v_z^2$. On cherche donc une fonction

$$f(v_x, v_y, v_z) = h(v_x^2 + v_y^2 + v_z^2) = g(v_x)g(v_y)g(v_z). \quad (5.61)$$

Seule la fonction exponentielle a cette propriété :

$$f(v_x, v_y, v_z) = A e^{-\alpha(v_x^2 + v_y^2 + v_z^2)}$$

et c'est donc une distribution gaussienne !

Preuve. Dérivant (5.61) par rapport à v_x ou v_y , on a $\frac{g'(v_x)}{2v_x g(v_x)} = \frac{g'(v_y)}{2v_y g(v_y)}$, dont le membre de droite ne dépend pas de v_x et est donc une constante. On intègre alors l'équation différentielle ordinaire par rapport à v_x avec le résultat $g(v_x) = a \exp(-\alpha v_x^2)$ etc.

La normalisation de f fixe $A = (\frac{\alpha}{\pi})^{\frac{3}{2}}$ et le coefficient α est identifié avec $\frac{1}{2}m\beta$, $\beta = (k_B T)^{-1}$, de telle sorte qu'apparaît dans l'exponentielle $\beta \times \frac{1}{2}m\bar{v}^2 = \frac{1}{k_B T} \times$ énergie cinétique de la molécule,

$$f(v_x, v_y, v_z) = \left(\frac{m}{2\pi k_B T} \right)^{\frac{3}{2}} e^{-\frac{m(v_x^2 + v_y^2 + v_z^2)}{2k_B T}}. \quad (5.62)$$

Enfin il peut être utile d'exprimer cette loi en termes de la seule variable $v = \sqrt{\bar{v}^2}$. Pour cela on effectue un changement de variable des coordonnées $(v_x, v_y, v_z) \mapsto (v, \theta, \phi)$, cf § 5.3.4, on intègre sur les angles θ et ϕ dont la densité f ne dépend pas et on trouve finalement

$$f(v) = 4\pi \left(\frac{m}{2\pi k_B T} \right)^{\frac{3}{2}} v^2 e^{-\frac{mv^2}{2k_B T}}, \quad (5.63)$$

c'est la distribution de Maxwell-Boltzmann des vitesses d'un gaz, que l'on étudie en Mécanique Statistique.
Exercice : utiliser la dérivation sous le signe somme pour calculer $\langle v^2 \rangle$ et montrer que $\langle v^2 \rangle = 3\langle v_x^2 \rangle$ où la moyenne de v_x^2 est calculée avec la distribution (5.62). Cela n'était-il pas attendu ?

• Désintégrations radioactives

On s'intéresse ici à la désintégration de noyaux d'un isotope $X \rightarrow X'$. On décrit en général le phénomène par une loi "empirique" : le nombre de désintégrations par unité de temps est proportionnel au nombre $N(t)$ de noyaux X présents à l'instant t , avec un taux de désintégration constant κ , caractéristique du noyau étudié X et de son mode de désintégration $X \rightarrow X'$. Pendant le temps dt , le nombre de noyaux se désintégrant selon ce mode est

$$|dN| = \kappa N dt = -dN. \quad (5.64)$$

Autrement dit, la fonction $N(t)$ satisfait l'équation différentielle $\dot{N}(t) = -\kappa N(t)$. En l'intégrant, on trouve

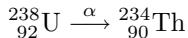
$$N(t) = N(0)e^{-\kappa t}, \quad (5.65)$$

la quantité de l'isotope considéré décroît exponentiellement. On définit la *demi-vie*, notée τ , par le temps au bout duquel le nombre N a décrû par un facteur 2. On a donc $N(\tau) = N(0)/2$, ce qui conduit à

$$\tau = \frac{\ln 2}{\kappa} \quad N(t) = N(0) e^{-t \ln 2 / \tau} = N(0) 2^{-t/\tau}. \quad (5.66)$$

Le nombre de noyaux de l'isotope aura décrû d'un facteur 1000 au bout d'un temps $\frac{\ln 1000}{\ln 2} \tau \approx 10 \tau$ (se rappeler que $2^{10} \approx 10^3$).

Par exemple, pour l'isotope naturel ^{238}U de l'Uranium, qui se désintègre en Thorium selon



en émettant un noyau d'Helium (particule α), la demi-vie est $\tau = 4,5$ Gans, (1 "Gan" = 1 milliard d'années !) donc le nombre N aura décrû d'un facteur 1000 au bout d'un temps $t = 45$ Gans, soit trois fois l'âge actuel de l'Univers... En revanche, pour les isotopes ^{131}I et ^{137}Cs , les demi-vies sont respectivement de 8j et 30 ans...

Le problème de désintégration de noyaux atomiques peut et doit en fait être traité par une approche probabiliste, puisque c'est bien là le fond de la question : un noyau (comme tout système individuel de nature quantique) a une *probabilité de transition* de tel ou tel état vers tel autre. Le traitement qui a précédé s'est appliqué implicitement à une vaste population de noyaux (dont le nombre a été traité comme une variable continue, et non comme un entier discret), et a concerné en fait le nombre moyen (la valeur moyenne ou espérance au sens probabiliste) de la v.a. nombre total de noyaux de l'isotope X. L'équation (5.64) se rédit en termes probabilistes comme suit : la probabilité d'un noyau donné de subir la désintégration étudiée pendant le temps dt est $\mathbf{P}(X \rightarrow X'; dt) = \kappa dt$. Notons $\mathbf{P}_N(t)$ la probabilité d'avoir N noyaux dans l'état initial (radioactif) X au temps t . Les désintégrations des différents noyaux étant supposées des événements indépendants, on a

$$\mathbf{P}_N(t + dt) = (1 - \kappa dt)^N (\mathbf{P}_N(t) + \kappa(N+1)\mathbf{P}_{N+1}(t)dt) + O(dt^2),$$

somme des probabilités des événements exclusifs suivants : aucune désintégration des N noyaux (probabilité $(1 - \kappa dt)$ pour chacun, factorisation pour l'ensemble puisque événements indépendants) + une désintégration d'un des $N+1$ noyaux avec la probabilité κdt fois la probabilité que les N autres restent inchangés + des désintégrations multiples d'ordre plus élevé en dt . Écrivant $\mathbf{P}_N(t + dt) = \mathbf{P}_N(t) + dt \frac{d}{dt} \mathbf{P}_N(t) + O(dt^2)$ et ne gardant que les termes d'ordre dt , on a donc

$$\frac{d}{dt} \mathbf{P}_N(t) = \kappa((N+1)\mathbf{P}_{N+1}(t) - N\mathbf{P}_N(t)) \quad (5.67)$$

On se donne (avec probabilité 1) la valeur initiale N_0 au temps 0, $\mathbf{P}_N(t=0) = \delta_{NN_0}$. On résout l'équation (5.67) en considérant la fonction caractéristique, fonction génératrice des $\mathbf{P}_N(t)$

$$\phi(x, t) = \sum_{N=0}^{N_0} x^N \mathbf{P}_N(t). \quad (5.68)$$

Calculons les dérivées partielles de $\phi(x, t)$

$$\begin{aligned}\frac{\partial}{\partial t} \phi(x, t) &= \sum_{N=0}^{N_0} x^N \frac{d}{dt} \mathbf{P}_N(t), \\ \frac{\partial}{\partial x} \phi(x, t) &= \sum_{N=1}^{N_0} N x^{N-1} \mathbf{P}_N(t) = \sum_{N=0}^{N_0-1} (N+1) x^N \mathbf{P}_{N+1}(t) = \sum_{N=0}^{N_0} (N+1) x^N \mathbf{P}_{N+1}(t) \\ &\quad \uparrow \text{car } p_{N_0+1}(t) = 0! \quad (5.69) \\ x \frac{\partial}{\partial x} \phi(x, t) &= \sum_{N=0}^{N_0} N x^N \mathbf{P}_N(t).\end{aligned}$$

On a donc fait apparaître les trois termes de l'équation (5.67), ce qui conduit donc à une *équation aux dérivées partielles* pour ϕ

$$\frac{\partial}{\partial t} \phi(x, t) = \kappa(1-x) \frac{\partial}{\partial x} \phi(x, t). \quad (5.70)$$

La méthode systématique de résolution fait appel à la “transformation de Laplace”, qui sera étudiée au chapitre 9. Contentons-nous d'observer que toute fonction de la forme $\phi(x, t) = G((1-x)\psi(t))$, avec G une fonction arbitraire, est solution de (5.70) à condition que $\dot{\psi}(t) = -\kappa\psi(t)$, soit $\psi(t) = ae^{-\kappa t}$, a une constante quelconque. La fonction jusque là arbitraire G est fixée par la condition initiale $\phi(x, 0) = x^{N_0} = G(a(1-x))$, soit $G(ax) = (1-x)^{N_0}$, et finalement la solution unique de (5.70) complétée par cette condition initiale est

$$\phi(x, t) = (1 - (1-x)e^{-\kappa t})^{N_0} = ((1 - e^{-\kappa t}) + xe^{-\kappa t})^{N_0} \quad (5.71)$$

$$= \sum_{N=0}^{N_0} C_{N_0}^N x^N e^{-N\kappa t} (1 - e^{-\kappa t})^{N_0-N} \quad (5.72)$$

qui permet d'identifier

$$\mathbf{P}_N(t) = C_{N_0}^N e^{-N\kappa t} (1 - e^{-\kappa t})^{N_0-N}.$$

On voit apparaître la loi binomiale $\mathcal{B}(N_0, p(t))$ pour une probabilité $p(t) = e^{-\kappa t}$! On peut alors calculer l'espérance de la variable aléatoire N

$$\langle N \rangle(t) = \sum_{N=0}^{N_0} N \mathbf{P}_N(t) = \frac{\partial \phi(x, t)}{\partial x} \Big|_{x=1} = N_0 e^{-\kappa t}.$$

On retrouve bien le résultat obtenu plus haut à partir de l'équation différentielle (5.64), qui s'appliquait donc en fait à $\langle N \rangle$.

Si N_0 est très grand et $p = e^{-\kappa t}$ petit, on peut remplacer la loi binomiale $\mathcal{B}(N_0, p(t))$ par sa limite poissonnienne $\mathbf{P}(\lambda = N_0 e^{-\kappa t})$, dont l'espérance est encore $\langle N \rangle(t) = \lambda = N_0 e^{-\kappa t}$.

On pourrait aussi étudier maintenant ce que sont les fluctuations de la v.a. N autour de cette valeur moyenne.

5.5 Théorèmes limites

5.5.1 Convergence presque sûrement, convergence en probabilité, convergence en loi

Considérons d'abord une suite infinie X_n de v.a. réelles, $n = 1, 2, \dots$, définies sur le même espace d'épreuves Ω . On peut définir plusieurs types de convergence de ces v.a.

- On dira que la suite X_n converge presque sûrement vers la v.a. X définie sur Ω s'il existe un $\Omega_0 \subset \Omega$ de mesure nulle : $\mathbf{P}(\Omega_0) = 0$ tel que $\forall \omega \notin \Omega_0$, $X_n(\omega)$ converge vers $X(\omega)$.

$$\forall \omega \notin \Omega_0 \quad \forall \epsilon > 0 \quad \exists N : \forall n > N \quad |X_n(\omega) - X(\omega)| < \epsilon. \quad (5.73)$$

- On dira que la suite X_n converge en probabilité vers la v.a. réelle X si pour tout h , $\mathbf{P}(|X_n - X| > h) \rightarrow 0$:

$$\forall h > 0, \quad \forall \epsilon > 0, \quad \exists N : \forall n > N \quad \forall \omega \in \Omega \quad \mathbf{P}(|X_n(\omega) - X(\omega)| > h) < \epsilon. \quad (5.74)$$

- Enfin on dira que la suite X_n converge en loi vers X si la suite des fonctions de répartition F_n converge (simplement) vers celle, notée F , de X (en tout point de continuité de F)

$$\text{pour presque tout } x, \quad \forall \epsilon > 0, \quad \exists N : \forall n > N \quad |\mathbf{P}(X_n \leq x) - \mathbf{P}(X \leq x)| < \epsilon. \quad (5.75)$$

On démontre la

Proposition 5.3 : $CV\ ps \Rightarrow CV\ p \Rightarrow CV\ l.$

- **Moyenne arithmétique de N v.a.**

Soient X_1, X_2, \dots, X_N N v.a. réelles indépendantes et distribuées selon la même loi de probabilité. (Elles ont donc en particulier même valeur moyenne $\langle X_i \rangle = m$ et même variance $\text{var} X_i = \sigma^2$, si cette moyenne et cette variance existent). Ce peut être par exemple le résultat de N lancers successifs d'un même dé (qu'il soit "honnête" ou pipé), la taille de N individus choisis au hasard dans une population homogène, N mesures répétées d'une quantité physique entachées d'erreurs, etc. On définit alors une nouvelle v.a., leur *moyenne arithmétique*,

$$Y = \frac{1}{N}(X_1 + X_2 + \dots + X_N). \quad (5.76)$$

Bien distinguer cette moyenne arithmétique Y des X_i , qui est une v.a., de leur valeur moyenne $\langle X_i \rangle$, qui est un nombre réel.

On se propose d'étudier la loi que suit Y dans la limite où N est grand.

De façon qualitative (nous allons préciser les hypothèses et démontrer précisément les choses ci-dessous)

- la loi des grands nombres nous dit que la v.a. Y converge en probabilité quand $N \rightarrow \infty$ vers la valeur commune des $\langle X_i \rangle$;
- le théorème limite central affirme que la v.a. Y converge en loi vers une v.a. normale (gaussienne) de moyenne $\langle X \rangle$ et de variance $\frac{1}{N} \text{var}(X_i)$ dans la limite $N \rightarrow \infty$.

5.5.2 Loi des grands nombres

Le théorème et sa preuve

Théorème 5.4 : Soient X_1, X_2, \dots, X_N une suite de v.a. indépendantes et de même loi. On suppose que la valeur moyenne m est bien définie. Alors la moyenne arithmétique Y de (5.76) converge en probabilité vers sa valeur moyenne m quand $N \rightarrow \infty$.

Preuve : Pour simplifier la preuve, on va faire l'hypothèse supplémentaire que la variance σ^2 commune des X_i existe aussi. Par l'hypothèse que les X_i sont indépendantes et ont même loi, donc même valeur moyenne (ou espérance) $\langle X_i \rangle = m$, on a d'abord

$$\langle Y \rangle = \frac{1}{N} \sum_{i=1}^N \langle X_i \rangle = m .$$

Calculons alors la variance de Y

$$\begin{aligned} \text{var}Y &= \langle (Y - m)^2 \rangle = \left\langle \left(\frac{1}{N} \sum_{i=1}^N (X_i - m) \right)^2 \right\rangle \\ &= \frac{1}{N^2} \left\langle \sum_i (X_i - m)^2 \right\rangle + \frac{2}{N^2} \sum_{i < j} \langle (X_i - m)(X_j - m) \rangle \\ &= \frac{1}{N^2} \left\langle \sum_i (X_i - m)^2 \right\rangle + \frac{2}{N^2} \sum_{i < j} \langle (X_i - m) \rangle \langle (X_j - m) \rangle \quad (\text{indépendance}) \\ &= \frac{1}{N^2} \left\langle \sum_i (X_i - m)^2 \right\rangle = \frac{1}{N^2} N \sigma^2 \quad \text{puisque } \langle (X_i - m) \rangle = 0 \end{aligned} \quad (5.77)$$

$$\text{donc} \quad \text{var } Y = \frac{1}{N} \sigma^2 = \frac{1}{N} \text{var } X \quad (5.78)$$

On est maintenant en mesure d'appliquer l'inégalité de Tchebychev ((5.8), avec $k = 2$) à la v.a. $Y - m$

$$\mathbf{P}(|Y - m| > h) \leq \frac{\langle (Y - m)^2 \rangle}{h^2} = \frac{\sigma^2}{Nh^2} . \quad (5.79)$$

Quel que soit $h > 0$, la probabilité $\mathbf{P}(|Y - m| > h)$ tend vers zéro quand $N \rightarrow \infty$, CQFD.

Sous des hypothèses plus fortes, (par exemple existence du 4ème moment), on peut démontrer (Kolmogorov) la CV ps de Y vers m : c'est la “*loi forte des grands nombres*”.

Deux illustrations

1. Évaluation empirique de la probabilité ou de la moyenne $\langle X \rangle$

La loi des grands nombres nous permet de justifier *a posteriori* la définition “intuitive” d’une probabilité d’un événement comme une fréquence d’occurrence de cet événement. Soit A un événement aléatoire se réalisant avec la probabilité p , associons lui une v.a. X prenant la valeur $X = 1$ si A se réalise, $X = 0$ sinon. Donc $\mathbf{P}(X = 1) = p$, $\mathbf{P}(X = 0) = 1 - p$, $\langle X \rangle = 1 \cdot p + 0 \cdot (1 - p) = p$. Effectuons un grand nombre d’épreuves X_i , $i = 1, \dots, N$ et considérons leur moyenne arithmétique $Y = \frac{1}{N}(X_1 + X_2 + \dots + X_N)$. Cette v.a. est la fréquence de réalisation de A dans nos N épreuves. La loi des grands nombres nous dit que Y tend en probabilité vers $\langle X \rangle = p$ et elle fournit donc une estimation de cette probabilité.

C’est bien de cette façon que nous nous assurons qu’un dé est “honnête” : chacun des 6 nombres vient avec une fréquence $\approx \frac{1}{6}$ dans une série d’un grand nombre de lancers. La même méthode permet d’estimer la (valeur) moyenne d’une v.a. X , en “mesurant” la moyenne (arithmétique) d’un échantillon de taille N suffisamment grande.

Les sondages d’opinion procèdent de manière analogue. Pour savoir quel pourcentage de la population aime les savonnettes Truc, ou a une opinion favorable de M. X, etc, c’est-à-dire quelle est la probabilité qu’un individu tiré au hasard ait telle ou telle opinion, on prend un échantillon de taille N et on mesure l’opinion moyenne de cet échantillon. Noter que (5.77) nous enseigne que l’écart-type de Y (c’est-à-dire la largeur de la distribution des Y) décroît comme $N^{-\frac{1}{2}}$: la précision d’un sondage est donc d’ordre $N^{-\frac{1}{2}}$, c’est-à-dire que pour gagner un facteur 10 sur l’incertitude, il faut multiplier la taille de l’échantillon par 100 !

Grâce à l’inégalité de Tchebychev (5.79), on peut même estimer la taille de l’échantillon nécessaire pour obtenir une précision donnée. Quelle doit être cette taille N (ou le nombre d’épreuves ou de mesures dans les exemples précédents) pour que la probabilité que Y fournisse la valeur de $m = \langle X \rangle$ avec une erreur *relative* de moins de 1% soit de 95%, autrement dit pour que $\mathbf{P}(|Y - m| > 0,01m) < 0,05$? Il suffit que $\frac{\sigma^2}{Nh^2} = \frac{10^4}{N} \left(\frac{\sigma}{m}\right)^2 < 0,05$, soit $N > 200\,000 \left(\frac{\sigma}{m}\right)^2$. On a déjà noté que la quantité “sans dimension” $\frac{\sigma}{m}$ est la bonne façon de mesurer l’étalement de la distribution d’une v.a. On voit que c’est elle encore qui donne la taille de l’échantillon requise pour atteindre une précision donnée.

2. Évaluation Monte-Carlo d’une intégrale.

Supposons les v.a. X_1, \dots, X_N uniformément distribuées entre 0 et 1, (donc avec une densité de probabilité $f(x) = 1$). Soit $g(x)$ une fonction (de carré intégrable), $g(X_i)$ est une nouvelle v.a. dont la valeur moyenne et la variance se calculent aisément

$$m_g = \langle g(X_i) \rangle = \int_0^1 g(x) dx$$

$$\sigma_g^2 = \int_0^1 (g(x) - m_g)^2 dx . \quad (5.80)$$

La loi (forte) des grands nombres nous dit que la moyenne arithmétique des $g(X_i)$

$$\frac{1}{N}(g(X_1) + \cdots + g(X_N))$$

converge (p.s) vers la valeur moyenne $m_g = \langle g(X) \rangle = \int_0^1 g(x)dx$ et que l'écart-type de la différence est $\sigma_g = \sigma/N^{1/2}$. Cela signifie qu'on a une approximation

$$\frac{1}{N}(g(x_1) + \cdots + g(x_N)) \approx \int_0^1 g(x)dx \quad (5.81)$$

dans laquelle les x_i sont des valeurs tirées de tables de nombres aléatoires uniformément distribués entre 0 et 1 ; la précision de cette approximation se comporte en $N^{-1/2}$.

Cette méthode de calcul est appelée *méthode de Monte-Carlo*. La convergence en $N^{-1/2}$ n'est pas très rapide et nécessite beaucoup de points. La méthode prend toute son importance pour des intégrales multiples et est très utilisée. On peut lui apporter toutes sortes d'améliorations et de raffinements...

5.5.3 Théorème limite central

On a vu au chapitre précédent que la limite quand $N \gg 1$ d'une loi binomiale au voisinage de son maximum est une loi normale. Or cette loi binomiale décrit la distribution des événements quand on répète N fois une épreuve dont la probabilité de succès est p . Le théorème limite central, appelé encore (incorrectement) "théorème de la limite centrale", généralise ce résultat.

- **Énoncé et remarques**

Théorème 5.5 : Soient N v.a. X_1, \dots, X_N indépendantes et de même loi, dont la valeur moyenne m et la variance σ^2 existent (c'est-à-dire sont finies). Leur moyenne arithmétique Y , définie en (5.76), est une v.a. dont la loi \tilde{f} tend quand $N \rightarrow \infty$ vers une loi normale de moyenne m et de variance σ^2/N (convergence en loi)

$$\tilde{f}(y) \approx f_G^{(N)}(y) = \frac{1}{\sqrt{2\pi} \frac{\sigma}{\sqrt{N}}} \exp -\frac{(y-m)^2}{2\left(\frac{\sigma}{\sqrt{N}}\right)^2} . \quad (5.82)$$

Les hypothèses excluent donc une loi lorentzienne $f(x) = \frac{\lambda}{\pi(x^2+\lambda^2)}$ dont la valeur moyenne et la variance n'existent pas, cf supra.

Il faut d'abord noter que la fonction gaussienne $f_G^{(N)}$ qui apparaît au second membre de (5.82) est de plus en plus "piquée" quand $N \rightarrow \infty$: en effet elle est bien normalisée,

$\int_{-\infty}^{\infty} dy f_G^{(N)}(y) = 1$, sa variance est de plus en plus petite $\sigma_N^2 = \frac{\sigma^2}{N}$ et son maximum de plus en plus grand $f_G^{(N)}(m) \propto \sqrt{N}$. Le graphe de $f_G^{(N)}$ est donc un pic de plus en plus haut et de plus en plus étroit au dessus de la valeur $y = m$, mais toujours d'aire 1. On a vu au chap. 3 que la suite des fonctions $f_G^{(N)}$ tend vers la distribution $\delta_m : f_G^{(N)}(y) \rightarrow \delta(y - m)$.

On voit alors que le théorème central limite implique la loi des grands nombres : la loi de la moyenne Y est de plus en plus concentrée en $Y = m$, autrement dit la v.a. Y converge en probabilité vers la valeur moyenne m des v.a. X_i .

• Éléments de preuve du théorème

La méthode de preuve est analogue à celle suivie pour la limite de la loi binomiale dans (5.55). Elle repose sur la comparaison des fonctions caractéristiques de la v.a. Y et de la loi normale, quand $N \rightarrow \infty$. Puisque Y est la somme de N v.a. indépendantes $\frac{1}{N}X_i$ de même loi, sa fonction caractéristique vaut

$$\varphi_Y(u) = (\varphi_{X/N}(u))^N.$$

On sait que

$$\begin{aligned}\varphi_{X/N}(u) &= \langle e^{iuX/N} \rangle = 1 + \frac{iu}{N} \langle X \rangle - \frac{u^2}{2N^2} \langle X^2 \rangle + O\left(\frac{u^3}{N^3}\right) \\ &= 1 + im\frac{u}{N} - \frac{u^2}{2N^2}(\sigma^2 + m^2) + O\left(\frac{u^3}{N^3}\right)\end{aligned}\quad (5.83)$$

et donc pour le logarithme de la fonction caractéristique de Y

$$\begin{aligned}\ln \varphi_Y(u) &= N \ln \varphi_{X/N}(u) = N \ln \left(1 + im\frac{u}{N} - \frac{u^2}{2N^2}(\sigma^2 + m^2) + O\left(\frac{u^3}{N^3}\right) \right) \\ &= N \left(im\frac{u}{N} - \frac{u^2}{2N^2}\sigma^2 \right) + O\left(\frac{u^3}{N^3}\right) \\ &= imu - \frac{u^2}{2N}\sigma^2 + O\left(\frac{u^3}{N^2}\right) \\ &= \ln \varphi_{\text{Gauss}}^{(N)}(u) + O\left(\frac{u^3}{N^2}\right)\end{aligned}\quad (5.84)$$

où $\varphi_{\text{Gauss}}^{(N)}(u) = e^{ium}e^{-\frac{1}{2}u^2\sigma^2/N}$ est la fonction caractéristique de la loi normale de moyenne m et de variance σ^2/N , cf équ. (5.44). Ce calcul montre que dans la limite $N \rightarrow \infty$, la fonction caractéristique de la moyenne arithmétique Y des N v.a. X_i tend vers la fonction caractéristique de la loi normale attendue. Nous admettrons que cela suffit à démontrer la convergence de la loi de Y vers la loi normale, c'est-à-dire le théorème limite central (P. Lévy).

• Illustrations

Le résultat du théorème central limite est remarquable : quelle que soit la loi des v.a. X_1, \dots, X_N , la loi limite de leur moyenne (5.76) est une loi normale ! La loi normale a ainsi un caractère d'*universalité*. Testons numériquement ce résultat sur des lois simples : les figures qui suivent présentent la densité de probabilité de la v.a. Y pour N v.a. X obéissant à une loi simple f ,

soit la loi uniforme, soit la loi “triangle”

$$f(x) = \begin{cases} 4x & \text{si } 0 \leq x \leq \frac{1}{2} \\ 4(1-x) & \text{si } \frac{1}{2} \leq x \leq 1 \\ 0 & \text{sinon} \end{cases} \quad (5.85)$$

dont la moyenne est $\frac{1}{2}$ et la variance $\frac{1}{24}$. Pour chacune de ces lois, nous calculons numériquement la densité de Y donnée par l’expression que nous admettrons

$$\tilde{f}(y) = N \int dx_1 \cdots dx_{N-1} f(x_1) f(x_2) \cdots f(x_{N-1}) f(Ny - x_1 - \cdots - x_{N-1})$$

À titre d’exercice, effectuons ce calcul de \tilde{f} pour $N = 2$ v.a. obéissant à une loi uniforme entre 0 et 1

$$\tilde{f}(y) = 2 \int_0^1 dx_1 f(x_1) f(2y - x_1).$$

La première fonction $f(x_1)$ est non nulle (et égale à 1) ssi $0 \leq x_1 \leq 1$, la deuxième $f(2y - x_1)$ ssi $0 \leq 2y - x_1 \leq 1$. Il faut prendre l’intersection des deux domaines $0 \leq x_1 \leq 1$ et $2y - 1 \leq x_1 \leq 2y$, ce qui diffère selon que $y < \frac{1}{2}$ ou $y > \frac{1}{2}$. Si $0 \leq y \leq \frac{1}{2}$, $2y - 1 \leq 0$ mais $2y \leq 1$, donc $\tilde{f}(y) = 2 \int_0^{2y} dx = 4y$, tandis que si $\frac{1}{2} \leq y \leq 1$, $2y \geq 1$ mais $2y - 1 \geq 0$ donc $\tilde{f}(y) = 2 \int_{2y-1}^1 = 4(1-y)$. On constate que la loi \tilde{f} obtenue n’est autre que la “loi triangle” de (5.85), cf figure de gauche de la Fig. 5.9.

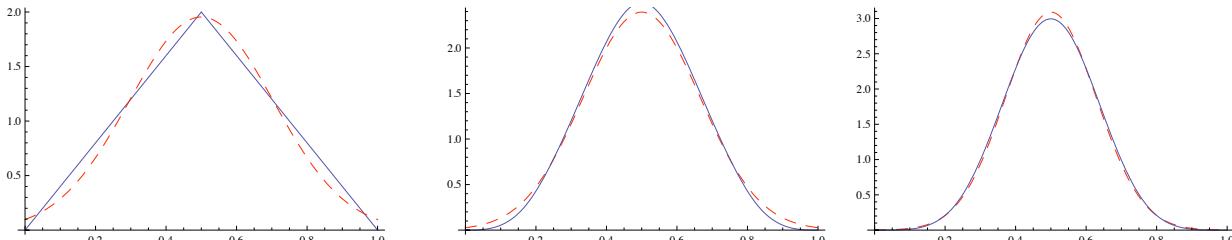


FIGURE 5.9 – Distribution $\tilde{f}(y)$ de la moyenne de N v.a. de loi uniforme, et comparaison avec la loi normale de même moyenne $m = \frac{1}{2}$ et de variance $\frac{1}{12N}$ (en ligne brisée). Successivement $N = 2, 3, 5$.

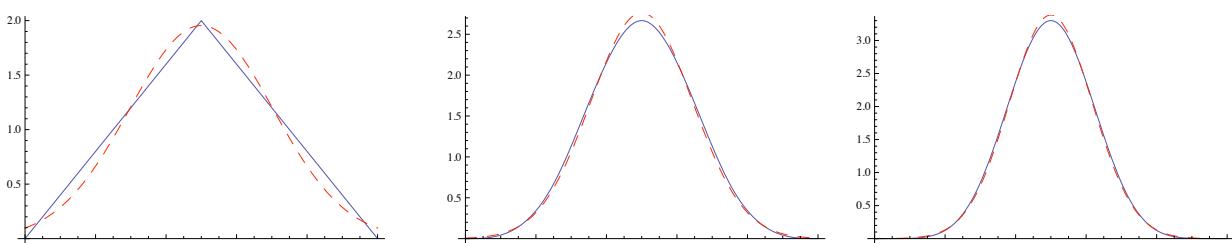


FIGURE 5.10 – Distribution $\tilde{f}(y)$ de la moyenne de N v.a. de loi “triangle” (5.85), et comparaison avec la loi normale de même moyenne $m = \frac{1}{2}$ et de variance $\frac{1}{24N}$ (en ligne brisée). Successivement $N = 1, 2, 3$.

- **Évaluations de l'erreur**

Réexprimons le théorème en termes de la v.a.

$$Z = \frac{Y - m}{\frac{\sigma}{N^{\frac{1}{2}}}} = \frac{\sum X_i - Nm}{\sigma N^{\frac{1}{2}}}$$

Elle est centrée $\langle Z \rangle = 0$ et le théorème nous dit qu'elle obéit asymptotiquement (quand $N \rightarrow \infty$) à une "loi normale centrée normalisée" (de variance 1). On en déduit que

$$\mathbf{P}(|Z| > \zeta) \rightarrow \frac{2}{\sqrt{2\pi}} \int_{\zeta}^{\infty} e^{-\frac{s^2}{2}} ds \quad (5.86)$$

ou de façon équivalente, dans les variables de départ,

$$\mathbf{P}(|Y - m| > \xi) = \mathbf{P}\left(\left|\frac{1}{N} \sum (X_i - m)\right| > \xi\right) \rightarrow \frac{2}{\sqrt{2\pi}} \int_{\xi N^{\frac{1}{2}}/\sigma}^{\infty} e^{-\frac{s^2}{2}} ds \quad (5.87)$$

En raison de leur importance en probabilités et en statistiques, on a donné un nom aux intégrales gaussiennes incomplètes

$$\begin{aligned} \text{erf}(x) &= \frac{2}{\sqrt{\pi}} \int_0^x dt e^{-t^2} \\ \text{erfc}(x) &= \frac{2}{\sqrt{\pi}} \int_x^{\infty} dt e^{-t^2} = 1 - \text{erf}(x) . \end{aligned} \quad (5.88)$$

Ce sont respectivement la *fonction d'erreur* erf et la *fonction d'erreur complémentaire* erfc. Le graphe de la seconde est représenté sur la figure (5.11) et leurs valeurs sont tabulées dans la Table 1. On peut donc récrire (5.86) et (5.87)

$$\mathbf{P}(|Z| > \zeta) \rightarrow \text{erfc}\left(\frac{\zeta}{\sqrt{2}}\right) \quad \text{ou} \quad \mathbf{P}(|Y - m| > \xi) \rightarrow \text{erfc}\left(\frac{\xi N^{\frac{1}{2}}/\sigma}{\sqrt{2}}\right) . \quad (5.89)$$

Avec (5.86) et (5.87) on a retrouvé la loi des grands nombres : la somme $\sum(X_i - m)$ est approximativement normale et centrée avec un écart-type $N^{-\frac{1}{2}}\sigma$. Mais les expressions (5.86) et (5.87) permettent d'affiner l'estimation faite au paragraphe précédent de la taille d'un échantillon nécessaire pour atteindre une précision donnée. Reprenons l'exemple du § 5.5.2 : quel N permet d'obtenir $\mathbf{P}(|Y - m| > 0,01m) < 0,05$? D'après (5.89) il faut chercher N tel que $\text{erfc}\left(\frac{N^{\frac{1}{2}}m}{100\sqrt{2}\sigma}\right) < 0,05$. On lit sur la Table 1 que $\text{erfc}(x) = 0,05$ pour $x \approx \sqrt{2}$, (en fait $x = 1,386$), donc $N^{\frac{1}{2}} > N_0^{\frac{1}{2}} \approx 200\frac{\sigma}{m}$ soit $N_0 \approx 40\,000\left(\frac{\sigma}{m}\right)^2$. On a gagné un facteur 5 par rapport à l'estimation plus grossière venant de l'inégalité de Tchebychev!

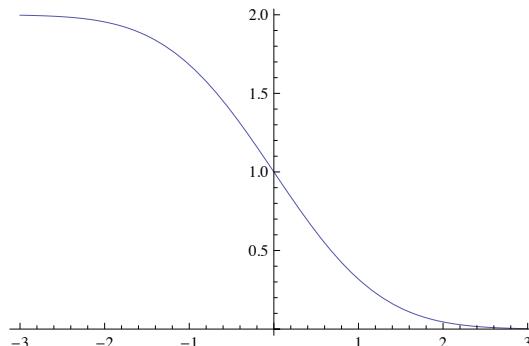


FIGURE 5.11 – Graphe de la fonction $\text{erfc}(x/\sqrt{2})$. On y voit que $\text{erfc}(1,96/\sqrt{2}) \simeq 0,05$ (plus exactement=0,0499958!)

x	$\text{erf}(x)$	$\text{erfc}(x)$	x	$\text{erf}(x)$	$\text{erfc}(x)$
0.	0	1.000000	1.3	0.9340079	0.06599206
0.05	0.05637198	0.9436280	1.4	0.9522851	0.04771488
0.1	0.1124629	0.8875371	1.5	0.9661051	0.03389485
0.15	0.1679960	0.8320040	1.6	0.9763484	0.02365162
0.2	0.2227026	0.7772974	1.7	0.9837905	0.01620954
0.25	0.2763264	0.7236736	1.8	0.9890905	0.01090950
0.3	0.3286268	0.6713732	1.9	0.9927904	0.00720957
0.35	0.3793821	0.6206179	2.	0.9953223	0.00467773
0.4	0.4283924	0.5716076	2.1	0.9970205	0.00297947
0.45	0.4754817	0.5245183	2.2	0.9981372	0.00186285
0.5	0.5204999	0.4795001	2.3	0.9988568	0.00114318
0.55	0.5633234	0.4366766	2.4	0.9993115	0.00068851
0.6	0.6038561	0.3961439	2.5	0.9995930	0.00040695
0.65	0.6420293	0.3579707	2.6	0.9997640	0.00023603
0.7	0.6778012	0.3221988	2.7	0.9998657	0.00013433
0.75	0.7111556	0.2888444	2.8	0.9999250	0.00007501
0.8	0.7421010	0.2578990	2.9	0.9999589	0.00004110
0.85	0.7706681	0.2293319	3.	0.9999779	0.00002209
0.9	0.7969082	0.2030918	3.1	0.9999884	0.00001165
0.95	0.8208908	0.1791092	3.2	0.9999940	$6.025 \cdot 10^{-6}$
1.	0.8427008	0.1572992	3.3	0.9999969	$3.057 \cdot 10^{-6}$
1.1	0.8802051	0.1197949	3.4	0.9999985	$1.522 \cdot 10^{-6}$
1.2	0.9103140	0.08968602	3.5	0.9999993	$7.43 \cdot 10^{-7}$

Table 1. Table des valeurs des fonctions $\text{erf}(x)$ et $\text{erfc}(x)$

Remarques finales

1. Comme noté plus haut, le théorème central limite ne s'applique pas à des “lois larges” comme la loi lorentzienne, dont la moyenne ou la variance n'existent pas.

2. Le fait que des lois normales apparaissent empiriquement (et approximativement) dans de très nombreux phénomènes naturels, taille ou poids d'individus dans une certaine population, erreurs dans les mesures successives d'une même grandeur, etc, pourrait trouver son origine dans le fait que le phénomène en question résulte d'un grand nombre de variables aléatoires indépendantes *et de même loi*. Mais ceci explique-t-il vraiment cela? ...

5.5.4 Marche aléatoire.

Une large classe de phénomènes physiques est modélisée par le problème de la marche au hasard, où les pas successifs d'un marcheur sont des v.a. indépendantes, tirées selon une certaine loi. On s'intéresse au comportement asymptotique de telles marches, dans la limite d'un grand nombre de pas élémentaires. C'est par exemple le cas du célèbre *mouvement brownien*⁸ qu'accomplissent des particules en suspension dans un fluide, sous l'effet des chocs avec les molécules de ce fluide : voir le site

<http://www.phy.ntnu.edu.tw/ntnужava/index.php?topic=24> ou le site de Wikipedia “Brownian motion” pour une très jolie simulation du phénomène ou un fac-simile de son observation par le physicien Jean Perrin.

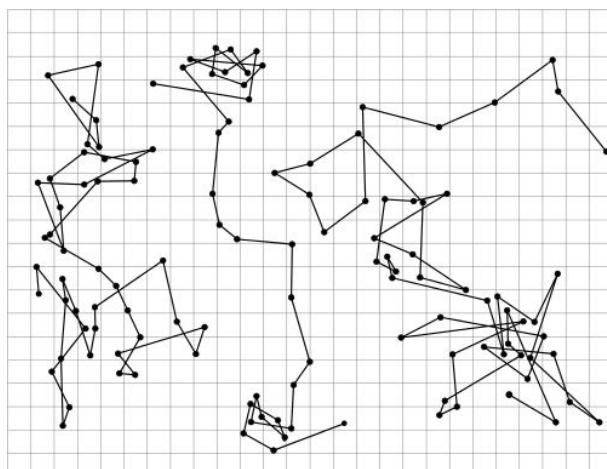


FIGURE 5.12 – Observation du mouvement brownien par le physicien Jean Perrin (1908) : la position de trois particules colloïdales (de diamètre $0,53 \mu\text{m}$), est observée sous le microscope et notée toutes les 30 secondes. Le pas de la grille est de $3,2 \mu\text{m}$. (Source : Wikipedia)

On va d'abord s'intéresser à une version simple du problème, où la marche se fait sur une grille, un *réseau* régulier.

• Marche au hasard sur un réseau

On considère un réseau carré de maille a , de vecteurs de base \vec{e}_1 et \vec{e}_2 (de norme a), voir figure 5.13⁹. Chaque site du réseau a 4 voisins. On considère un mobile (“marcheur”) qui à chaque instant $t_n = n\tau$, $n = 1, 2, 3, \dots$, saute

8. du nom du botaniste Robert Brown (1773–1858) qui a le premier observé le phénomène sous le microscope

9. Toute la discussion qui suit pourrait être menée en dimension arbitraire, sur un réseau “hypercubique” de dimension d . Nous nous restreignons à $d = 2$ dimensions pour la commodité de l'exposition et des dessins.

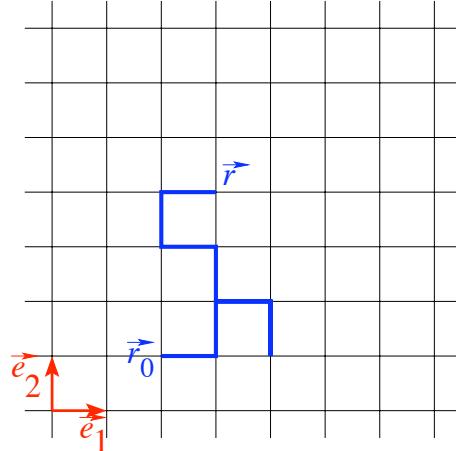


FIGURE 5.13 – Marche aléatoire sur un réseau carré

du site où il se trouve vers un des 4 sites voisins. On suppose ce saut aléatoire et équidistribué : la probabilité de sauter sur n'importe lequel des sites adjacents est égale à 1/4. Les sauts successifs sont supposés indépendants.

Soit $\mathbf{P}(\mathbf{r}, t_n | \mathbf{r}_0, t_0)$ la probabilité conditionnelle que le marcheur se trouve au point \mathbf{r} du réseau à l'instant t_n , sachant qu'il était en \mathbf{r}_0 à l'instant t_0 . On va s'intéresser à la détermination de cette probabilité et à ses propriétés, en particulier dans la limite $n \gg 1$.

On observe d'abord que cette probabilité satisfait une condition initiale

$$\mathbf{P}(\mathbf{r}, t_0 | \mathbf{r}_0, t_0) = \delta_{\mathbf{r}, \mathbf{r}_0} \quad (4.90)$$

où $\delta_{\mathbf{r}, \mathbf{r}_0} = 1$ si $\mathbf{r} = \mathbf{r}_0$, = 0 sinon, c'est un "delta de Kronecker" à deux dimensions : $\delta_{\mathbf{r}, \mathbf{r}_0} = \delta_{x, x_0} \delta_{y, y_0}$ en termes des composantes x et y des vecteurs \mathbf{r} et \mathbf{r}_0 . La probabilité \mathbf{P} satisfait aussi une condition de normalisation

$$\sum_{\mathbf{r}} \mathbf{P}(\mathbf{r}, t_n | \mathbf{r}_0, t_0) = 1 \quad (4.91)$$

où l'on somme sur tous les sites du réseau, condition qui exprime qu'à l'instant t_n , le marcheur se trouve bien quelque part !

Moins trivial est le fait que \mathbf{P} satisfait aussi une relation de récurrence entre les temps t_n et t_{n+1} , qui exprime que le marcheur ne peut être en \mathbf{r} au temps t_{n+1} que s'il était au temps t_n en un des points voisins sur le réseau, $\mathbf{r} \pm \vec{e}_j$, $j = 1, 2$. Comme les événements correspondants (être en $\mathbf{r} \pm \vec{e}_j$ à l'instant t_n) sont exclusifs, leurs probabilités s'ajoutent et on a donc

$$\mathbf{P}(\mathbf{r}, t_{n+1} | \mathbf{r}_0, t_0) = \frac{1}{4} \sum_{\mathbf{r}' = \mathbf{r} \pm \vec{e}_j} \mathbf{P}(\mathbf{r}', t_n | \mathbf{r}_0, t_0) \quad (4.92)$$

(Cette relation rappelle et généralise la relation de récurrence du triangle de Pascal, qui correspondrait à un réseau à une dimension.) Cette relation de récurrence, complétée par la condition initiale, suffit en principe à déterminer complètement la probabilité \mathbf{P} .

Notons encore que le problème étant invariant par translation d'espace et de temps, la fonction $\mathbf{P}(\mathbf{r}, t_n | \mathbf{r}_0, t_0)$ ne dépend que des différences $\mathbf{r} - \mathbf{r}_0$ et $t_n - t_0 = n\tau$, et que pour des raisons dimensionnelles, elle ne peut dépendre que des rapports $\frac{1}{a}(\mathbf{r} - \mathbf{r}_0)$ et $\frac{t_n - t_0}{\tau} = n$

$$\mathbf{P}(\mathbf{r}, t_n | \mathbf{r}_0, t_0) = F\left(\frac{\mathbf{r} - \mathbf{r}_0}{a}, \frac{t_n - t_0}{\tau}\right). \quad (4.93)$$

• Processus stochastiques, propriété de Markov

La marche au hasard sur un réseau que nous venons de définir est un exemple de *processus aléatoire* ou *processus stochastique*, c'est-à-dire d'un système dont l'évolution temporelle est probabiliste, au contraire d'un système *déterministe*. La marche au hasard a de plus la *propriété de Markov*, ce qui signifie que l'évolution au temps t ne dépend pas de l'histoire passée du système. (Ce ne serait pas le cas si on imposait que le marcheur ne recroise jamais sa trajectoire passée.)

Dans notre problème de marche au hasard sur le réseau, on a aussi la propriété que si $t' > t > t_0$,

$$\mathbf{P}((\mathbf{r}', t' | \mathbf{r}_0, t_0) = \sum_{\mathbf{r}} \mathbf{P}((\mathbf{r}', t' | \mathbf{r}, t) \mathbf{P}((\mathbf{r}, t | \mathbf{r}_0, t_0)) \quad (4.94)$$

qui exprime qu'entre les instants t_0 et t' le marcheur passe en un temps intermédiaire t par un point \mathbf{r} , que les événements correspondant à des \mathbf{r} différents sont exclusifs et ont des probabilités qui s'ajoutent, et que les évolutions entre les temps t_0 et t et entre t et t' étant indépendantes (processus de Markov), leurs probabilités se multiplient.

• Limite continue et équation de la chaleur

Si on soustrait aux deux membres de l'équation (4.92) la quantité $\mathbf{P}(\mathbf{r}, t_n | \mathbf{r}_0, t_0)$, elle peut être réécrite sous la forme suivante :

$$\mathbf{P}(\mathbf{r}, t_{n+1} | \mathbf{r}_0, t_0) - \mathbf{P}(\mathbf{r}, t_n | \mathbf{r}_0, t_0) = \frac{1}{4} \sum_{\mathbf{r}' = \mathbf{r} \pm \vec{e}_j} (\mathbf{P}(\mathbf{r}', t_n | \mathbf{r}_0, t_0) - \mathbf{P}(\mathbf{r}, t_n | \mathbf{r}_0, t_0)) \quad (4.95)$$

$$= \frac{1}{4} \sum_{j=1,2} (\mathbf{P}(\mathbf{r} + \vec{e}_j, t_n | \mathbf{r}_0, t_0) + \mathbf{P}(\mathbf{r} - \vec{e}_j, t_n | \mathbf{r}_0, t_0) - 2\mathbf{P}(\mathbf{r}, t_n | \mathbf{r}_0, t_0)). \quad (4.96)$$

Supposons maintenant qu'on laisse les incrément de temps et d'espace τ et a tendre vers zéro. Dans cette limite, les coordonnées de temps et d'espace, qui étaient discrétisées, tendent vers des variables continues. On regarde donc la *limite continue* du problème initial. (De façon équivalente au vu de (4.93), cette limite décrit la situation où les séparations d'espace $\mathbf{r} - \mathbf{r}_0$ ou de temps $t_n - t_0 = n\tau$ sont très grandes par rapport à a et τ .) Dans la limite $\tau \rightarrow 0$ et $a \rightarrow 0$, le membre de gauche de (4.95) est approximé par $\tau \frac{\partial}{\partial t} \mathbf{P}(\mathbf{r}, t_n | \mathbf{r}_0, t_0)$, celui de droite par $\frac{1}{4}a^2 \Delta \mathbf{P}(\mathbf{r}, t_n | \mathbf{r}_0, t_0)$ où Δ est le laplacien à 2 dimensions

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \quad (4.97)$$

agissant sur les coordonnées x et y de \mathbf{r} . Dans cette limite l'équation (4.92) devient

$$\tau \frac{\partial}{\partial t} \mathbf{P} = \frac{1}{4} a^2 \Delta \mathbf{P} \quad (4.98)$$

qui est une équation fondamentale de la physique, décrivant les phénomènes de diffusion, (diffusion de particules dans un milieu, de chaleur etc). Selon le contexte, cette équation est appelée équation de Fick, *équation de la chaleur*, ...

La solution de l'équation (4.98) complétée par une condition initiale qui est la version continue de (4.90), voir plus bas, est bien connue :

$$F\left(\frac{\mathbf{r} - \mathbf{r}_0}{a}, \frac{t - t_0}{\tau}\right) = \mathbf{P}(\mathbf{r}, t | \mathbf{r}_0, t_0) = \frac{A\tau}{(t - t_0)} e^{-\frac{(\mathbf{r} - \mathbf{r}_0)^2}{a^2(t - t_0)/\tau}} \quad (4.99)$$

où on suppose $t - t_0 > 0$ et où la constante A va être fixée plus bas. *Exercice* : vérifier que cette fonction est bien solution de (4.98).

Dans la limite qui nous occupe où \mathbf{r} est une variable (aléatoire) continue, il faut remplacer la probabilité du marcheur d'être en un point donné \mathbf{r} par une densité de probabilité,

$$p(\mathbf{r}, t_n | \mathbf{r}_0, t_0) = \frac{1}{a^2} \mathbf{P}(\mathbf{r}, t_n | \mathbf{r}_0, t_0) \quad (4.100)$$

(qui a bien la dimension d'une *densité* à deux dimensions). La condition de normalisation (4.91), $\sum_{\mathbf{r}} a^2 \frac{\mathbf{P}(\mathbf{r}, t_n | \mathbf{r}_0, t_0)}{a^2} = \sum_{\mathbf{r}} a^2 p(\mathbf{r}, t_n | \mathbf{r}_0, t_0) = 1$ peut être considérée comme une somme de Riemann de l'intégrale de p et donne donc dans la limite $a \rightarrow 0$

$$\int d^2 \mathbf{r} p(\mathbf{r}, t_n | \mathbf{r}_0, t_0) = 1, \quad (4.101)$$

qui est bien la condition de normalisation d'une densité de probabilité, comme nous l'avons vu, mais ici pour une v.a. à deux dimensions \mathbf{r} . Cette condition fixe la constante $A = \frac{1}{\pi}$, (cf les intégrales gaussiennes du § 5.4.3), donc

$$p(\mathbf{r}, t_n | \mathbf{r}_0, t_0) = \frac{\tau}{\pi a^2(t - t_0)} e^{-\frac{(\mathbf{r}-\mathbf{r}_0)^2}{a^2(t-t_0)/\tau}}.$$

Le calcul jusqu'ici a supposé que les limites $a \rightarrow 0$ et $\tau \rightarrow 0$ étaient indépendantes. Afin de se débarrasser complètement de ces échelles "microscopiques", on voit qu'il convient de prendre $\tau \propto a^2$. On choisira l'échelle de temps de telle sorte que

$$\tau = \frac{1}{4} a^2 \quad (4.102)$$

(et plus généralement, en dimension d'espace d , on prendrait $\tau = \frac{1}{2d} a^2$). L'expression finale de la densité de probabilité p est donc

$$p(\mathbf{r}, t_n | \mathbf{r}_0, t_0) = \frac{1}{4\pi(t - t_0)} e^{-\frac{(\mathbf{r}-\mathbf{r}_0)^2}{4(t-t_0)}}, \quad (\text{pour } t > t_0) \quad (4.103)$$

(et plus généralement, en dimension d'espace d , le préfacteur serait $(4\pi(t - t_0))^{d/2}$).

La condition initiale satisfaite par cette solution est obtenue en prenant la limite $t_n - t_0 \rightarrow 0$. Cette limite est une distribution delta de Dirac, cf § 3.4.2. On écrit donc

$$\lim_{t_n - t_0 \rightarrow 0} p(\mathbf{r}, t_n | \mathbf{r}_0, t_0) = \delta(\mathbf{r} - \mathbf{r}_0) = \delta(x - x_0)\delta(y - y_0) \quad (4.104)$$

qui est la version continue de la condition initiale (4.90).

Notons enfin que si $t' > t > t_0$, la propriété (4.94) se récrit pour p selon

$$\int d^2 \mathbf{r} p(\mathbf{r}', t' | \mathbf{r}, t) p(\mathbf{r}, t | \mathbf{r}_0, t_0) = p(\mathbf{r}', t' | \mathbf{r}_0, t_0)$$

comme on le vérifiera en calculant cette *convolution* d'intégrales gaussiennes, un calcul qui après un changement de variables adéquat se ramène à nouveau à une intégrale d'une gaussienne.

L'apparition d'une gaussienne, c'est-à-dire d'une loi normale pour p , comme solution de notre marche au hasard ne doit pas surprendre : elle découle du théorème limite central, la trajectoire de \mathbf{r}_0 à \mathbf{r} résultant de l'addition d'un grand nombre $n = (t - t_0)/\tau$ de pas élémentaires (qui sont des v.a. indépendantes et de même loi).

• Lois d'échelle dans la marche au hasard

Une propriété très remarquable de la solution (4.103) est son invariance par le changement simultané

$$\begin{aligned} \mathbf{r} - \mathbf{r}_0 &\rightarrow \lambda(\mathbf{r} - \mathbf{r}_0); t - t_0 \rightarrow \lambda^2(t - t_0) \\ p(\mathbf{r}, t_n | \mathbf{r}_0, t_0) &\rightarrow \lambda^{-2}p(\mathbf{r}, t_n | \mathbf{r}_0, t_0) \\ p(\mathbf{r}, t_n | \mathbf{r}_0, t_0)d^2\mathbf{r} &\rightarrow p(\mathbf{r}, t_n | \mathbf{r}_0, t_0)d^2\mathbf{r} \end{aligned} \quad (4.105)$$

où nous avons souligné que par cette dilatation, la *densité* de probabilité p est multipliée par un facteur λ^{-2} , tandis que la *probabilité* $p d^2\mathbf{r}$ que \mathbf{r} soit dans un petit domaine au voisinage de \mathbf{r} est invariante. Cette propriété qui relie les dilatations de l'espace et du temps apparaissait déjà dans la relation (4.102) entre les échelles microscopiques a et τ .

Une conséquence de cette invariance est que l'écart-type de la loi (4.103), $\langle (\mathbf{r} - \mathbf{r}_0)^2 \rangle^{\frac{1}{2}}$, qui décrit l'étalement de la distribution, c'est-à-dire le rayon typique R de la région couverte par le marcheur dans sa marche au hasard, croît comme

$$R = \langle (\mathbf{r} - \mathbf{r}_0)^2 \rangle^{\frac{1}{2}} \sim |t - t_0|^{\frac{1}{2}}. \quad (4.106)$$

Inversement la relation $|t - t_0| \sim R^2$ est interprétée en disant que la marche au hasard a une *dimension fractale*, (plus précisément une *dimension de Hausdorff*), égale à 2 : si le marcheur déroule une bobine de fil (fil d'Ariane ?) le long de sa trajectoire, à un taux constant par unité de temps, il aura consommé une quantité $\propto R^2$ de fil quand il se sera éloigné de R de son point de départ...

Cette invariance de dilatation (4.105) est un exemple de ce que l'on appelle une *loi d'échelle* et le nombre $\nu = \frac{1}{2}$ qui apparaît dans la puissance de $(t - t_0)$ dans (4.106) est un exemple d'*exposant critique*. Il est tout à fait remarquable que cet exposant soit indépendant du détail de la marche au hasard : pas discrets égaux à \pm les vecteurs de base du réseau, ou mouvement brownien avec des pas de longueur arbitraire, etc. Il ne dépend pas non plus de la dimension de l'espace où s'effectue cette marche. La raison étant une fois encore le théorème limite central et l'universalité de la loi gaussienne qui en découle. En revanche, cet exposant serait modifié si on imposait que le marcheur ne recoupe jamais sa trajectoire passée...

L'étude de systèmes ayant un comportement d'échelle et des exposants critiques est un thème majeur de la physique contemporaine, qui s'applique par exemple au comportement d'un ferromagnétique au point de Curie, mais cela est une autre histoire !

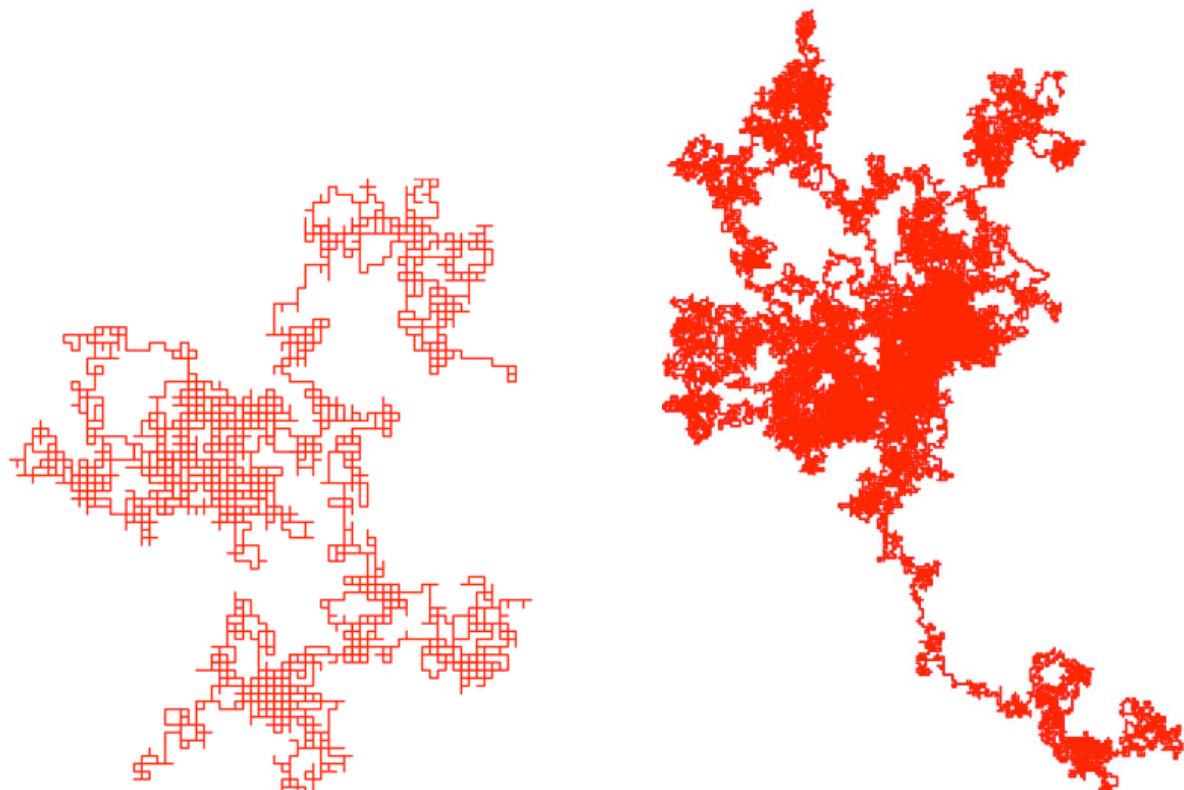


FIGURE 5.14 – Deux marches aléatoires sur réseau, vues à des échelles différentes. Sur la première, où on distingue bien les pas sur le réseau discret, on constate que le mouvement ne s'étale pas beaucoup, le marcheur revenant souvent sur ses pas. Cela est encore plus visible sur la seconde, où apparaît aussi une propriété remarquable du mouvement brownien continu, son invariance d'échelle (propriétés fractales) : le grossissement d'une petite partie de la “courbe” est de même nature que la courbe originale.

Chapitre 6

Séries entières. Fonctions analytiques

Ce chapitre est le premier d'une série qui va porter sur différents aspects des fonctions d'une variable complexe. On va s'intéresser d'abord à des séries $\sum_k a_k z^k$ d'une variable complexe z . Sauf mention du contraire, ces sommes seront supposées commencer à $k = 0$.

6.1 Séries entières

6.1.1 Séries formelles

En préambule, mentionnons d'abord que l'on peut s'intéresser aux objets algébriques $S(X) = \sum_n a_n X^n$ et aux opérations que l'on peut effectuer sur eux, sans se soucier de la nature de la variable X ni de la convergence de la somme. C'est ce qu'on appelle une *série formelle*. Ces séries (avec les coefficients a_n appartenant à un corps K de nombres, typiquement \mathbb{R} ou \mathbb{C}) forment une algèbre notée $K[[X]]$: on peut les ajouter, les multiplier par un nombre de K , les multiplier entre elles selon les règles familières. On peut aussi prendre l'inverse de toute série ayant $a_0 \neq 0$: l'inverse de $S(X) = \sum_n a_n X^n$ ($a_0 \neq 0$) est la série formelle $T(X) = \sum_n b_n X^n$ t.q. $S(X).T(X) = 1$. On peut *composer* deux séries $S(X) = \sum_n a_n X^n$ et $T(Y) = \sum_n b_n Y^n$ en considérant que $Y = S(X)$, à condition que $a_0 = 0$: on obtient une nouvelle série formelle $T \circ S(X)$. Sous les hypothèses $a_0 = 0$, $a_1 \neq 0$ sur S , on peut aussi définir la *série réciproque* T de S (c'est-à-dire inverse pour la composition) : $T \circ S = I$ (la fonction identité $I : x \mapsto x$).

L'intérêt de ces séries formelles est qu'elles dissocient les questions de convergence de ces manipulations algébriques. C'est par exemple intéressant dans des applications à des problèmes de combinatoire, où on définit des "fonctions génératrices". Si on a une collection de nombres a_n indexés par un entier n , il est souvent utile de considérer la fonction génératrice des a_n , définie comme la série formelle $\sum_n a_n X^n$.

Exemple. Le nombre de partitions P_n d'un entier n est le nombre de façons de l'écrire comme somme non ordonnée $n = p_1 \cdot 1 + p_2 \cdot 2 + \cdots + p_n \cdot n$, $p_i \geq 0$. Par exemple $P_5 = 7$ (vérifier). Montrer que la fonction génératrice des P_n est donnée par la formule d'Euler

$$\sum_{n=0}^{\infty} P_n X^n = \frac{1}{\prod_{m=1}^{\infty} (1 - X^m)}. \quad (6.1)$$

Or on peut montrer (Hardy et Ramanujan, voir aussi plus bas la méthode du col) qu'asymptotiquement $P_n \sim \frac{1}{4n\sqrt{3}} e^{\pi\sqrt{\frac{2n}{3}}}$, une croissance beaucoup trop rapide pour que la série converge, pour quelque valeur de X que ce soit. On dira dans la suite que la série a un rayon de convergence nul. La série est et reste une série formelle. Mais telles quelles, ces fonctions génératrices séries formelles sont très utiles.

Exemple : En factorisant le produit $\prod_{m=1}^{\infty} (1 - X^m)$ en deux produits sur les m pairs et impairs, démontrer l'identité entre séries formelles (en fait produits formels ...) $\prod_{m=1}^{\infty} (1 - X^{2m-1}) \prod_{m=1}^{\infty} (1 + X^m) = 1$. En déduire, en écrivant leurs fonctions génératrices, que les nombres de partitions d'un entier en nombres impairs ou en nombres distincts sont égaux. Par exemple pour $n = 5$, $P_{\text{distincts}}(5) = 3$: $\{5\}, \{4, 1\}, \{3, 2, 1\}$ et $P_{\text{impairs}}(5) = 3$: $\{1, 1, 1, 1, 1\}, \{3, 1, 1\}, \{5\}$. Pour le physicien, les nombres de partitions en nombres distincts sont reliées au comptage d'états fermioniques d'énergie donnée, dans un potentiel harmonique...

6.1.2 Rayon de convergence

On considère la série entière $\sum_k a_k z^k$ où les $a_k \in \mathbb{C}$.

Définition 6.1 : Le rayon de convergence de la série $\sum_k a_k z^k$ est le nombre R réel positif ou nul, (éventuellement infini), défini par

$$R \stackrel{\text{déf}}{=} \sup\{t \in \mathbb{R}^+ \text{ tel que } \sum_n |a_n| t^n \text{ est borné}\}. \quad (6.2)$$

(L'ensemble de ces t n'est pas vide puisqu'il contient au moins 0, donc son sup est bien défini dans $\overline{\mathbb{R}}$.) Par exemple, si la suite $|a_n|^{\frac{1}{n}}$ a une limite ℓ , alors $R = \ell^{-1}$.

Mais plus généralement, on va montrer que (formule d'Hadamard)

$$\frac{1}{R} = \overline{\lim}_{n \rightarrow \infty} |a_n|^{1/n} \quad (6.3)$$

où la limite supérieure d'une suite a été définie à l'Appendice 1.3.5,

$$\overline{\lim}_{n \rightarrow \infty} u_n = \lim_{k \rightarrow \infty} (\sup_{n \geq k} u_n).$$

La preuve de (6.3) repose sur la “règle de Cauchy” (comparaison avec une série géométrique) : soit une suite $b_n > 0$; si $\overline{\lim}_{n \rightarrow \infty} b_n^{1/n} < 1$, $\sum_n b_n < +\infty$, et si $\overline{\lim}_{n \rightarrow \infty} b_n^{1/n} > 1$, $\sum_n b_n = +\infty$ (vérifier en utilisant (A.6)). On applique cela à $b_n = |a_n|r^n$ et on en conclut que $\sum |a_n|r^n$ converge pour $r^{-1} > \overline{\lim}_{n \rightarrow \infty} |a_n|^{1/n}$, et diverge si $r^{-1} < \overline{\lim}_{n \rightarrow \infty} |a_n|^{1/n}$ ce qui, compte tenu de la définition (6.2) établit (6.3).

On appellera *disque de convergence* le domaine ouvert $|z| < R$, appellation justifiée par le théorème fondamental

Théorème 6.1 : La série entière $\sum_k a_k z^k$ converge normalement et uniformément dans tout disque $|z| \leq r$ intérieur au disque de convergence ($0 < r < R$) ; elle converge absolument pour tout z dans le disque de convergence $|z| < R$ et diverge pour tout z hors de ce disque, $|z| > R$; sur le bord du disque, $|z| = R$, la série peut être convergente ou divergente.

(Pour les définitions des différents types de convergence, se reporter à la Définition 1.7.) Si $R = 0$, la série ne converge que pour $z = 0$. Si $R = \infty$, elle converge dans tout le plan.

La convergence normale est claire : pour la norme du sup, si $|z| \leq r < R$, $\sum ||a_k z^k|| = \sum |a_k| r^k < \infty$ par définition du rayon de convergence ; et les CVU et CVA découlent du Théorème 1.8. Bien noter que la convergence normale et la convergence uniforme ne sont établies que dans tout disque plus petit que le disque de convergence.

Exemples.

- 1. La série $\frac{z^n}{n!}$ est convergente dans tout le plan, puisque $\ell = 0$, donc $R = \infty$.
Rappelons l'argument : pour tout entier m supérieur à $|z|$, on a $|z^n/n!| = \frac{1}{m!} \frac{|z|^n}{(m+1)\cdots n} < \frac{m^m}{m!} \frac{|z|^n}{m^n}$ dès que $n > m$ d'où $|z^n/n!|^{\frac{1}{n}} < \frac{|z|}{m} \left(\frac{m^m}{m!}\right)^{\frac{1}{n}}$ qui tend vers $|z|/m < 1$. cqfd
- 2. A l'inverse, $z^n n!$ est telle que $\ell = \infty$, $R = 0$, donc la série ne converge qu'en 0.
- 3. $\left(\frac{z}{r}\right)^n$, $r \neq 0$. On a $\ell = 1/r$, donc un rayon de convergence $R = r$. Sur le cercle de convergence, pour $|z| = r$, la série ne converge en aucun point (puisque le terme d'ordre n ne s'annule pas).
- 4. $\frac{z^n}{n}$, $R = 1$, la série converge pour tout z , $|z| < 1$, elle diverge pour $|z| > 1$; pour $z = 1$ elle est divergente ; pour $z = -1$, elle est semi-convergente (convergente mais pas absolument convergente) en tant que série alternée, et plus généralement elle converge (mais pas absolument) sur tout le bord du disque sauf en $z = 1$. On reconnaît la série de $-\log(1 - z)$.
- 5. $\frac{z^n}{n^2}$, $R = 1$. Sur tout le bord du disque de convergence, y compris $z = \pm 1$, convergence absolue.
- 6. nz^n , $R = 1$, la série diverge sur tout le bord du disque.

Il ressort de cette série d'exemples que le comportement sur le bord du disque est une question délicate à étudier avec soin.

On peut additionner, multiplier ou inverser des séries entières de rayon de convergence non nul et on obtient une autre série entière de rayon de convergence non nul. (Pour l'inverse de $\sum a_n z^n$ il faut supposer que $a_0 \neq 0$.) On peut aussi composer la série $T(\zeta) = \sum_{n=0} b_n \zeta^n$ par $S(z) = \sum_{k=1} a_k z^k$,¹ c'est-à-dire construire $T(S(z))$ qui est de rayon de convergence non nul si S et T le sont. En particulier si $a_1 \neq 0$ la série réciproque U de la série $S(z)$, telle que $U(S(z)) = z$, existe et a un rayon de convergence non nul.

1. bien noter que la sommation sur k démarre à 1 !

6.1.3 Continuité, intégrabilité et dérivabilité de la somme

Une série entière convergeant uniformément dans tout disque fermé contenu dans son disque de convergence, on peut lui appliquer les théorèmes de continuité, de dérivabilité et d'intégrabilité énoncés au chapitre 1, en se restreignant d'abord à une variable réelle : pour x réel et $r < R$, (inégalité stricte!), sur le segment $-r \leq x \leq r$, la somme $S(x)$ de la série $\sum a_n x^n$ est continue. En intégrant terme à terme la série $\sum a_n x^n$, la série entière $\sum \frac{a_n}{n+1} x^{n+1}$ converge pour $|x| < R$ vers l'intégrale $\int_0^x S(x') dx'$. Enfin la série dérivée terme à terme, $\sum n a_n x^{n-1}$, est convergente pour $|x| < R$ et a pour somme $S'(x)$. En fait on démontre que ces séries dérivée et intégrée ont le *même* rayon de convergence que la série initiale. (Exercice : le vérifier.)

Exemple. En intégrant terme à terme entre 0 et x la série géométrique $\sum_{n=0}^{\infty} (-1)^n x^n$ qui converge vers $1/(1+x)$ pour $x \in]-1, 1[$, on obtient

$$\ln(1+x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n} \quad (6.4)$$

qui converge pour $|x| < 1$. Pour $x = 1$, la série converge encore, et on a $1 - \frac{1}{2} + \frac{1}{3} - \dots = \ln 2$.

On peut en fait étendre cette dérivation à la série de la variable complexe z . On y reviendra au chapitre suivant. Si $S(z) = \sum_{n=0}^{\infty} a_n z^n$ est la somme d'une série de rayon R , la dérivée de S par rapport à la variable z

$$S'(z) = \lim_{\substack{h \in \mathbb{C}, h \rightarrow 0 \\ h \neq 0}} \frac{S(z+h) - S(z)}{h}$$

existe² et est donnée par la série dérivée

$$S'(z) = \sum_{n=1}^{\infty} n a_n z^{n-1} \quad (6.5)$$

qui converge dans le même disque de rayon R . (Preuve, voir [4], chap. I.)

Coefficients de la série entière

On peut donc dériver terme à terme une série entière à l'intérieur de son disque de convergence, puis dériver à nouveau, etc. La somme de la série $S(z) = \sum a_n z^n$ est donc indéfiniment dérivable et ses dérivées successives en $z = 0$ sont données par $S^{(n)}(0) = n! a_n$, soit $a_n = \frac{S^{(n)}(0)}{n!}$. Pour $|z| < R$, la série entière $S(z) = \sum a_n z^n$ s'identifie au développement de Taylor (– Maclaurin) de $S(z)$

$$S(z) = S(0) + z S'(0) + \frac{z^2}{2} S''(0) + \dots + \frac{z^n}{n!} S^{(n)}(0) + \dots \quad (6.6)$$

2. Bien noter que h est un complexe quelconque tendant vers 0.

6.1.4 Vitesse de convergence

Étant donnée une série entière de rayon de convergence $R \neq 0$ connu, un problème pratique est celui de la *vitesse de convergence* de cette série. Intuitivement, cette vitesse peut s'évaluer à nouveau par comparaison de séries, et nous nous attendons à ce qu'une série dont le terme général a_n se comporte par exemple comme $a_n \sim n^{-\alpha}$ converge d'autant plus vite que α est plus grand. Peut-on aller au delà de cette simple observation ? Combien de termes de la série faut-il garder pour avoir une approximation d'ordre donné de la somme ? Y a-t-il des moyens d'accélérer cette convergence par des procédés de resommation ? Toutes ces questions sont d'un grand intérêt pour le physicien . . . mais nous ne pourrons pas nous y arrêter.

6.2 Les fonctions exp et log

Comme on l'a vu plus haut, la série

$$e^z = \sum_{n=0}^{\infty} \frac{1}{n!} z^n$$

a un rayon de convergence infini. D'après (6.5)

$$S'(z) = S(z)$$

et on retrouve l'équation différentielle satisfaite par la fonction exponentielle.

On vérifie aussi aisément que

$$e^z e^{z'} = e^{z+z'}, \quad e^z e^{-z} = e^0 = 1.$$

ce qui implique que

$$\forall z \in \mathbb{C} \quad e^z \neq 0. \tag{6.7}$$

On peut restreindre la variable z à des valeurs réelles, retrouver les propriétés et inégalités familières (positivité, convexité, . . .) de la fonction exponentielle réelle, définir la fonction réciproque \log sur \mathbb{R}^+ etc etc.

On peut aussi restreindre z à des valeurs purement imaginaires, $\mathbb{R} \ni y \mapsto e^{iy}$ ce qui fournit un *homomorphisme*³ du groupe additif \mathbb{R} dans le groupe multiplicatif $U(1)$ des nombres complexes de module 1 (que l'on peut voir aussi comme le cercle unité) :

$$e^{iy} e^{iy'} = e^{i(y+y')}.$$

3. Rappelons qu'un homomorphisme d'un groupe G dans un groupe H est une application ρ de G dans H qui est compatible avec les deux structures de groupe : $\forall g, g' \quad \rho(g * g') = \rho(g) \cdot \rho(g')$.

Le noyau de cet homomorphisme, c'est-à-dire les y qui sont appliqués sur l'élément 1 sont les multiples de 2π : $y = 2\pi k$, $k \in \mathbb{Z}$. Autrement dit à tout nombre u de module 1 (à tout point du cercle unité) correspond une famille de y définis modulo 2π tels que $u = e^{iy}$: ce sont les différentes déterminations de l'*argument* de u .

$$y = \arg u \mod 2\pi.$$

On étend alors la fonction argument à tout le plan “pointé” à l'origine $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$ par

$$w \neq 0 \quad \arg w = \arg(w/|w|).$$

Logarithme complexe

Finalement intéressons-nous à la fonction réciproque de l'exponentielle (complexe). Étant donné $w \in \mathbb{C}$, on cherche z tel que $e^z = w$. Il est clair par (6.7) que z n'existe que si $w \neq 0$ et que si z est solution, tout $z + 2\pi ik$ l'est aussi. Si on note $z = \log w$, z n'est défini que modulo $2\pi i$. En identifiant les modules et arguments on a

$$e^z = e^{x+iy} = w = |w|e^{i\arg w} \Rightarrow x = \log |w|, \quad y = \arg w \mod 2\pi$$

donc

$$z = \log w = |\log w| + i \arg w \tag{6.8}$$

Attention : on a utilisé la même notation \log pour la fonction déjà bien définie sur les réels et pour le logarithme complexe. De plus, il faut garder à l'esprit que la partie imaginaire de ce $\log w$ est défini à l'addition près d'un multiple entier de $2\pi i$.

Ainsi il faut prendre garde que les identités familières du logarithme peuvent être affectées par ces déterminations multiples, telle

$$\log ww' = \log w + \log w' \mod 2\pi i.$$

Le logarithme complexe ainsi défini est, comme la fonction \arg ci-dessus, un exemple de *fonction multivaluée*. Dans un domaine ouvert connexe du plan complexe, ne contenant pas 0 et “simplement connexe” (heuristiquement, domaine “sans trou”, on verra plus précisément au chapitre suivant ce qu'il faut entendre par là), le logarithme admet une *détermination* qui est une fonction continue de la variable w . Et toute autre détermination en diffère par un multiple entier de $2\pi i$.

Par exemple, dans le domaine ouvert $D : |w| > r > 0$, $|\arg w| < \pi - \alpha$ (r et α petits : plan “coupé” le long de l'axe réel négatif), (6.8) fournit la *détermination principale* du logarithme $\log w$.

Développement en série du logarithme

La série entière

$$T(u) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{u^n}{n}$$

a pour rayon de convergence 1 (cf supra). Elle donne la détermination principale de $\log(1+u)$.

En effet, comme on le sait bien, cette série est la série réciproque de celle de l'exponentielle ; donc $e^{T(u)} = 1 + u$; c'est bien la détermination principale : si $|u| < 1$, $w = 1 + u$ a une partie réelle > 0 donc appartient au domaine D ci-dessus, et en $u = 0$, $T(u)$ s'annule, tout comme la détermination principale.

6.3 Séries de Taylor. Fonctions analytiques

6.3.1 Séries de Taylor

Soit Ω un ouvert de \mathbb{R} ou \mathbb{C} . Soit f une fonction définie sur Ω , et $z_0 \in \Omega$.

Définition 6.2 : La fonction f est développable en série entière centrée en z_0 s'il existe un ouvert $\Omega' \subset \Omega$ contenant z_0 et une suite $a_n \in \mathbb{C}$ tels que

$$\forall z \in \Omega' \quad f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n. \quad (6.9)$$

Restreignons-nous (pour le moment) au cas réel. Soient I un ouvert de \mathbb{R} , $x_0 \in I$, f une fonction de I dans \mathbb{C} développable en série entière centrée en x_0 , $f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n$, de rayon de convergence R . Comme on l'a vu plus haut en (6.6), la fonction est de classe C^∞ (infinitiment dérivable) dans l'ouvert $]x_0 - R, x_0 + R[$, avec

$$\forall k \quad f^{(k)}(x) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} a_n (x - x_0)^{n-k} \quad (6.10)$$

et donc $a_n = \frac{f^{(n)}(x_0)}{n!}$. On appelle la série $\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n$ série de Taylor(-MacLaurin) de f en x_0 .

6.3.2 Fonctions analytiques

Soit une fonction f d'un ouvert Ω de \mathbb{R} ou \mathbb{C} dans \mathbb{C} .

Définition 6.3 : La fonction f est analytique sur Ω si pour tout $z_0 \in \Omega$, elle admet un développement en série entière centré en z_0 .

D'après ce qui précède, si f est analytique dans Ω , elle y est indéfiniment dérivable et ses dérivées successives y sont analytiques ; $1/f$ est analytique dans Ω privé des points z_0 où $f(z_0) = 0$; si f admet une primitive dans Ω connexe (définie à une constante additive près), cette primitive est aussi analytique.

Exemples : les polynômes de z sont analytiques dans \mathbb{C} ; les fractions rationnelles de z sont analytiques dans le plan privé des zéros du dénominateur. On va voir que la classe des fonctions analytiques est bien plus vaste :

Théorème 6.2 : *La somme d'une série entière $\sum a_n z^n$ de rayon de convergence $R > 0$ est analytique en tout point z_0 du disque ouvert de convergence.*

Autrement dit, si $\sum_n a_n z^n$ est convergente dans le disque ouvert $|z| < R$, la fonction $f(z) = \sum_n a_n z^n$ est analytique en tout point z_0 de ce disque, donc tel que $|z_0| < R$. Cela est non trivial : l'existence du développement en série entière centré en 0 n'implique pas évidemment la même propriété en tout point du disque !

Exemple. Soit la série $\sum_{n=0}^{\infty} x^n$. Elle converge dans le disque de rayon 1 vers la fonction $f(x) = \frac{1}{1-x}$. Pour tout x_0 , $-1 < x_0 < 1$, on peut écrire $f(x) = \frac{1}{1-x_0-(x-x_0)}$ et développer $f(x) = \sum_{n=0}^{\infty} \frac{(x-x_0)^n}{(1-x_0)^{n+1}}$. Quel est le domaine de convergence de ce nouveau développement ?

Preuve du Théorème. Montrons que la série $\sum_n \frac{1}{n!} f^{(n)}(z_0) u^n$ a un rayon de convergence $R' \geq R - |z_0|$ et que

$$f(z) = \sum_n \frac{1}{n!} f^{(n)}(z_0) (z - z_0)^n \quad \text{pour } |z - z_0| < R - |z_0|. \quad (6.11)$$

En effet si $r_0 = |z_0|$, $\alpha_n = |a_n|$, on a

$$\begin{aligned} f^{(m)}(z_0) &= \sum_p \frac{(m+p)!}{p!} a_{m+p} z_0^p \\ |f^{(m)}(z_0)| &\leq \sum_p \frac{(m+p)!}{p!} \alpha_{m+p} r_0^p \end{aligned}$$

Si $r_0 \leq r < R$,

$$\begin{aligned} \sum_m \frac{1}{m!} |f^{(m)}(z_0)| (r - r_0)^m &\leq \sum_{m,p} \frac{(m+p)!}{m!p!} \alpha_{m+p} r_0^p (r - r_0)^m \\ &\leq \sum_n \alpha_n \sum_{0 \leq m \leq n} \frac{n!}{m!(n-m)!} r_0^{n-m} (r - r_0)^m = \sum_n \alpha_n r^n < \infty, \end{aligned}$$

La série à coefficients positifs $\sum_n \alpha_n \sum_{0 \leq m \leq n} \frac{n!}{m!(n-m)!} r_0^{n-m} (r - r_0)^m$ converge donc, ce qui justifie *a posteriori* la modification de l'ordre des termes qu'on a effectuée. Cela prouve que la série (6.11) a un rayon de convergence $R' \geq r - r_0$, et puisque r peut être arbitrairement près de R , $R' \geq R - r_0$.

On considère maintenant z tel que $|z - z_0| < R - r_0$. La série double $\sum_{m,p} \frac{(m+p)!}{m!p!} a_{m+p} z_0^p (z - z_0)^m$ étant

absolument convergente d'après ce qui précède, on peut la resommer de deux façons différentes :

$$\begin{aligned} \sum_{m,p} \frac{(m+p)!}{m!p!} a_{m+p} z_0^p (z - z_0)^m &= \sum_n a_n \sum_{0 \leq m \leq n} \frac{n!}{m!(n-m)!} z_0^{n-m} (z - z_0)^m = \sum_n a_n z^n = f(z) \\ &= \sum_m \frac{(z - z_0)^m}{m!} \sum_p \frac{(m+p)!}{p!} a_{m+p} z_0^p = \sum_m \frac{(z - z_0)^m}{m!} f^{(m)}(z_0) \end{aligned}$$

d'où $f(z) = \sum_m \frac{(z - z_0)^m}{m!} f^{(m)}(z_0)$ q.e.d.

Remarques.

1. Le rayon R' de convergence de la série (6.11) peut être plus grand que $R - r_0$. Exemple, considérer la série $f(z) = \sum_n (iz)^n = \frac{1}{1-iz}$ de rayon $R = 1$ et en étudier le développement en un point réel $x_0 < 1$. Montrer que le nouveau rayon de convergence est $R' = \sqrt{1+x_0^2} > 1 - |x_0|$.
2. Pour une fonction de variable réelle, l'hypothèse d'analyticité est plus forte que celle de différentiabilité C^∞ . Ainsi, soit la fonction $f(x) = e^{-1/x^2}$ pour $x \neq 0$, $f(0) = 0$. La fonction est C^∞ , toutes ses dérivées s'annulent en 0, et la série de Taylor est donc convergente, mais sa somme est nulle donc ne converge pas vers f : la fonction n'est pas analytique ! L'origine du problème est que, au voisinage de 0, (dans les directions imaginaires) la fonction et ses dérivées croissent très vite (comme $e^{1/|x|^2}$)... Il existe des conditions (nécessaires et/ou suffisantes) d'analyticité d'une fonction f de classe C^∞ , typiquement des conditions majorant la croissance des dérivées $f^{(n)}$ au voisinage du point considéré ([4], chap. 1).

6.3.3 Prolongement analytique

Problème : Étant donnée une fonction f analytique dans Ω , ouvert connexe, et $\Omega' \supset \Omega$ un ouvert connexe contenant Ω , existe-t-il une fonction g analytique dans Ω' telle que $g = f$ sur Ω ?

Si c'est le cas, on dit que g est un *prolongement analytique* de f à Ω' . Exemple : la fonction $f(z) = \sum_n z^n = (1-z)^{-1}$ est analytique dans $|z| < 1$. Elle admet donc un développement en série autour de tout point du disque Ω : $|z| < 1$, par exemple $z_0 = \frac{1}{2} + \frac{3i}{4}$ (voir figure 6.1) et ce développement a un rayon donné par la distance de z_0 au point 1, soit $r_0 = \sqrt{\frac{1}{4} + \frac{9}{16}} = \frac{\sqrt{13}}{4}$: soit D le disque de centre z_0 et de rayon r_0 . La fonction est analytique en tout point de $\Omega' = \Omega \cup D$. Ce nouveau domaine Ω' déborde largement les limites de Ω : on a donc bien prolongé f au-delà du disque initial. On pourrait itérer cette opération et de proche en proche, prolonger analytiquement f au plan privé du point 1.

Un théorème nous garantit l'unicité de ce prolongement, s'il existe :

Théorème 6.3 : Soit f une fonction analytique dans un ouvert connexe Ω et soit $z_0 \in \Omega$. Les trois conditions suivantes sont équivalentes

1. $\forall n \geq 0 \quad f^{(n)}(z_0) = 0$;
2. $f \equiv 0$ dans un voisinage de z_0 ;
3. $f \equiv 0$ dans Ω .

Rappelons la définition d'un ensemble connexe E : tout sous-ensemble X de E à la fois ouvert et fermé est E lui-même ou \emptyset .

Preuve. Il est clair que 3. \Rightarrow 1. Par ailleurs 1. \Rightarrow 2. selon (6.11) et 2. \Rightarrow 1. est évident. Enfin pour établir 2. \Rightarrow 3., considérons l'ensemble Ω' des points z' de Ω au voisinage desquels $f \equiv 0$; Ω' est non vide (par 2.), et ouvert (par définition). Pour démontrer qu'il est aussi fermé, considérons un point $z_1 \in \overline{\Omega'}$ de son adhérence (déf. § 1.2.1); il existe des points z' de Ω' arbitrairement proches de z_1 et pour lesquels donc $f^{(n)}(z') = 0$ pour tout n (en appliquant 2. \Rightarrow 1. à z'), et par continuité des $f^{(n)}$, il en découle que $f^{(n)}(z_1) = 0$, et (par 1. \Rightarrow 2. appliqué à z_1) $f \equiv 0$ au voisinage de z_1 , donc $z_1 \in \Omega'$ qui est bien fermé. Ω étant connexe, on a $\Omega' = \Omega$ q.e.d.

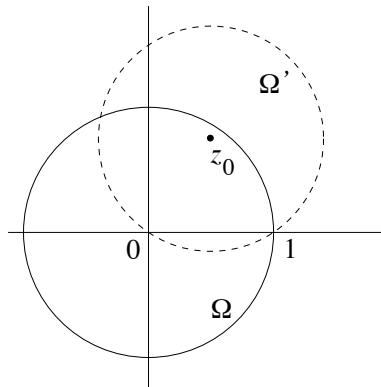


FIGURE 6.1 – Prolongement analytique de la somme de la série $\sum_n z^n$ au delà de son disque de convergence.

Corollaire : Si deux fonctions f et g analytiques dans un ouvert connexe Ω coïncident au voisinage d'un point de Ω , elles sont identiques sur Ω .

Exemples et contre-exemple de prolongement analytique

Considérons la série $f(z) = \sum_{n=0}^{\infty} z^n$, convergente et analytique pour $|z| < 1$ et la fonction $g(z) = \frac{1}{1-z}$ définie et analytique dans le plan pointé $\mathbb{C} \setminus \{1\}$. Elles coïncident dans le disque $|z| < 1$ et g définit donc le prolongement analytique de f au plan pointé.

Considérons ensuite le développement en série de la fonction logarithme. On peut effectuer ce développement autour du point 1, $f_0(z) = \log z = \sum_{n=1}^{\infty} (-1)^{n-1}(z-1)^n/n$, qui converge dans le disque $D_0 = |z-1| < 1$, ou autour du point $e^{i\pi/4}$, $f_1(z) = \frac{i\pi}{4} + \sum_{n=1}^{\infty} (-1)^{n-1}e^{-in\pi/4}(z-e^{i\pi/4})^n/n$ qui converge dans le disque $D_1 = |z - e^{i\pi/4}| < 1$, etc. Ces deux fonctions coïncident au voisinage de $z_0 = \frac{i\pi}{8}$ (vérifier !) donc dans tout $D_0 \cap D_1$ selon la Proposition précédente. f_1 est donc l'unique prolongement de $\log z$ au domaine D_1 . On pourrait itérer et définir de proche en proche les prolongements f_k dans les disques D_k : $|z - e^{ik\pi/4}| < 1$. Mais **attention !** rien ne nous dit que

le prolongement au domaine D_8 autour de $e^{i2\pi} = 1$ coïncide avec la fonction $f_0 = \log$ de départ. En fait comme on sait f_8 définit une autre détermination de la fonction \log , différent de f_0 par $2\pi i$.

À l'inverse, le prolongement analytique d'une fonction somme d'une série entière n'est pas toujours possible. Considérons par exemple la fonction $f(z) \stackrel{\text{def}}{=} \sum_{n=1}^{\infty} z^{2^n}$. Elle converge pour $|z| < 1$, diverge pour tout z , $|z| > 1$ et aussi pour $z = 1$, point où elle a une “singularité”⁴. En fait on démontre aisément (cf Exercice) qu'elle a une singularité pour toute racine 2^p -ième de l'unité. Ces racines étant denses sur le cercle unité $|z| = 1$, ce cercle constitue une *frontière essentielle* de la fonction, au delà de laquelle un prolongement analytique est impossible. On conçoit qu'en général, un prolongement analytique n'est possible que s'il est possible de “se glisser entre les singularités”...

Applications physiques

A côté des applications mathématiques, le prolongement analytique d'une fonction peut aussi être utile en physique.

Exemple : les notions d'impédance Z et de son inverse Y , l'admittance, qui décrivent la réponse d'un circuit électrique à une excitation, sont familières : il s'agit de fonctions (complexes) de la variable réelle de fréquence (“angulaire”) ω . Il se trouve que le prolongement de $Z(\omega)$ et de $Y(\omega)$ à des valeurs complexes de ω est possible, en faisant appel à la *causalité*, et qu'il contient une information utile. On verra au chapitre 8 d'autres exemples de ce cas (relations de dispersion).

D'autres exemples concernent le prolongement de quantités physiques définies pour une variable entière à des valeurs réelles ou complexes arbitraires : c'est le cas du moment cinétique J qui en mécanique quantique prend des valeurs quantifiées, multiples entières ou demi-entières de \hbar , mais qui sous des hypothèses adéquates sur le potentiel de diffusion, peut être prolongé dans le plan complexe (Regge) ; là encore, les propriétés d'analyticité dans la variable J peuvent être riches d'informations ... C'est encore le cas avec d'autres nombres au départ entiers –nombre Q d'états dans le modèle de Potts de la mécanique statistique ; dimension d'espace-temps prolongée à partir de $d = 4$ en mécanique statistique et théorie des champs, etc.

6.4 Lacunes de ce chapitre...

6.4.1 Séries divergentes. Séries asymptotiques

Il conviendrait de discuter aussi le cas des “séries asymptotiques”, qui sont divergentes et qui apparaissent par exemple en mécanique céleste, ou dans les développements “perturbatifs” de mécanique quantique ou de théorie quantique des champs. On se bornera à dire ici qu'une série de rayon de convergence nul peut néanmoins contenir une information utile ... et utilisable avec des outils de “resommation” adéquats et mathématiquement justifiés.

4. On donnera au chapitre 7 un sens précis à ce terme. Contentons-nous pour le moment de la définition suivante : une singularité d'une fonction définie par une série $\sum a_k z^k$ est un point où la série diverge.

6.4.2 Produits infinis

Il conviendrait aussi d'étudier comment on peut donner un sens aux produits infinis $\prod_{n \geq 1} (1 - \frac{z}{a_n})$. Parmi toutes les belles formules que l'on peut démontrer, n'en citons qu'une

$$\sin \pi z = \pi z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2}\right). \quad (6.12)$$

Exercices

1. Considérer la série

$$S = \sum_{k=0}^{\infty} \left(\frac{1}{16}\right)^k \left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6}\right) \quad (6.13)$$

Quel est le comportement asymptotique du terme général ? La convergence est-elle lente ou rapide ?

Pour déterminer sa somme, construire la série entière

$$S(x) = \sum_{k=0}^{\infty} \left(\frac{x^8}{16}\right)^k \left(\frac{4x}{8k+1} - \frac{2x^4}{8k+4} - \frac{x^5}{8k+5} - \frac{x^6}{8k+6}\right). \quad (6.14)$$

- a) Quel est son rayon de convergence ? ;
- b) Montrer que la fonction $S(x)$ satisfait $S'(x) = (4 - 2x^3 - x^4 - x^5)/(1 - x^8/16)$ et $S(0) = 0$;
- c) Vérifier que la fonction $\varphi(x) \stackrel{\text{déf}}{=} -4\text{Arctan}(1-x) + 2\ln(2-x^2) - 2\ln(x^2-2x+2)$ satisfait la même équation différentielle et $\varphi(0) = -\pi$. On pourra utiliser un logiciel de calcul formel comme Mathematica ou Maple pour vérifier les points b) et c), un peu fastidieux à la main...

En conclure que $S(x) = \varphi(x) + \pi$, et en déduire la valeur de la somme S de (6.13). Tester numériquement la convergence de la série (6.13) vers sa limite.

2. Soit I un intervalle de \mathbb{R} et f une fonction $f : I \rightarrow \mathbb{C}$ analytique dans I . Montrer qu'on peut la prolonger en une fonction analytique dans un ouvert $\Omega \subset \mathbb{C}$ contenant I .

3. Soit $f(z) \stackrel{\text{déf}}{=} \sum_{n=1}^{\infty} z^{2^n}$.

Montrer qu'elle converge dans le disque $|z| < 1$ et qu'elle a une singularité en $z = 1$.

Montrer que f satisfait les relations fonctionnelles $f(z) = z^2 + f(z^2)$, et plus généralement $f(z) = z^2 + z^4 + z^8 + \dots + z^{2^p} + f(z^{2^p})$ pour tout p entier. En déduire que f a aussi des singularités en toute racine 2^p -ième de l'unité.

Chapitre 7

Fonctions holomorphes. Théorème de Cauchy

On va voir dans ce chapitre que la condition qu'une fonction de variable complexe est *dérivable* est très contraignante. Les fonctions dérivables ou *holomorphes* possèdent des propriétés remarquables que nous allons explorer.

7.1 Fonctions holomorphes

7.1.1 Rappel : fonctions différentiables de \mathbb{R}^n dans \mathbb{R}^p

Soit Ω un ouvert de \mathbb{R}^n , $a \in \Omega$ et soit une fonction $f : \Omega \rightarrow \mathbb{R}^p$. On dit que f est *différentiable* en a s'il existe une application linéaire notée df_a de \mathbb{R}^n dans \mathbb{R}^p telle que

$$\forall h \in \mathbb{R}^n \quad f(a + h) - f(a) - df_a(h) = o(\|h\|), \quad (7.1)$$

où $o(\|h\|)/\|h\| \rightarrow 0$ quand $h \rightarrow 0$. df_a est appelée la *différentielle* (ou l'application linéaire tangente) de f en a . Si on note x_i , $i = 1, \dots, n$, des coordonnées dans (une base de) \mathbb{R}^n , et si $h = (h_1, \dots, h_n)$ dans cette base, l'application linéaire df_a s'écrit

$$df_a(h) = \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i} \Big|_a \quad (7.2)$$

et les $\frac{\partial f}{\partial x_i} \Big|_a$ s'identifient aux *dérivées partielles* en a comme définies usuellement. La différentiabilité en a implique donc l'existence des dérivées partielles en a ; l'inverse n'est pas vrai, voir ci-dessous. (Remarque : $h = (h_i)$ est ce que l'on appellera plus tard un *vecteur tangent*.)

On peut aussi introduire les (formes) différentielles dx_i , telles que $dx_i(h) = h_i$. (Ce sont des “formes” (linéaires) car elles sont à valeurs dans \mathbb{R} , et elles agissent sur les vecteurs tangents, ici h). On récrit alors (7.2) sous la forme

$$df_a = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \Big|_a dx_i. \quad (7.3)$$

Plus généralement on aura à considérer plus bas des formes différentielles $\omega = \sum_i F_i(x)dx_i$.

Si f est différentiable en tout point de Ω et que l’application $a \mapsto df_a$ est continue, on dit que f est continûment différentiable (ou de classe C^1). Si $n = 1$ (fonction d’une variable réelle), la notion de différentiabilité s’identifie à celle de dérivabilité, définie par l’existence de la dérivée. Si $n > 1$, cependant, cela n’est plus toujours vrai : f peut avoir des dérivées partielles en a sans être différentiable. Considérons par exemple la fonction de \mathbb{R}^2 dans \mathbb{R} :

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

Elle est continue en $(0, 0)$, y admet des dérivées partielles ... mais n’est pas différentiable ! En effet $\partial_x f(0, 0) = 0$, $\partial_y f(0, 0) = 1$, mais $(f(h, k) - 0 - k)/\|(h, k)\| = -kh^2/(h^2 + k^2)^{3/2}$ ne tend pas vers 0 indépendamment de la direction suivie.

Dans la suite de ce chapitre, on va s’intéresser à $n = 2$, $p = 1$ ou 2 (fonctions réelles ou complexes d’une variable complexe).

7.1.2 Holomorphie. Conditions de Cauchy–Riemann

Dans ce qui suit, Ω désigne un ouvert du plan complexe \mathbb{C} , et $f : \Omega \rightarrow \mathbb{C}$ une fonction (à valeurs complexes) de $z \in \Omega$.

Définition 7.1 : La fonction $f : \Omega \rightarrow \mathbb{C}$ est dérivable ou holomorphe en $z_0 \in \Omega$ si

$$f'(z_0) \stackrel{\text{déf}}{=} \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

existe, c’est-à-dire s’il existe un nombre complexe noté $f'(z_0)$ tel que

$$f(z) - f(z_0) - f'(z_0)(z - z_0) = o(|z - z_0|). \quad (7.4)$$

f est holomorphe sur Ω si elle est holomorphe en tout point de Ω .

Bien voir que l’hypothèse est que la limite existe et est indépendante de la façon dont z approche z_0 . Si on écrit $z = x + iy$ et que l’on considère f comme une fonction \tilde{f} des deux variables réelles x et y , $\tilde{f}(x, y) = f(x + iy)$, on va voir que cette condition de dérivabilité de f est plus forte (plus contraignante) que celle de différentiabilité de \tilde{f} .

Théorème 7.1 (conditions de Cauchy–Riemann 1) : Si $f : \Omega \rightarrow \mathbb{C}$ et si $\tilde{f} : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{C}$ est la fonction de x et y correspondante, alors f est dérivable en $z_0 = x_0 + iy_0 \in \Omega$ ssi

$$\tilde{f} \text{ est différentiable dans } \mathbb{R}^2 \text{ et } \frac{\partial \tilde{f}}{\partial x}(x_0, y_0) = -i \frac{\partial \tilde{f}}{\partial y}(x_0, y_0), \quad (7.5)$$

et alors, $f'(z_0) = \frac{\partial \tilde{f}}{\partial x}(x_0, y_0)$.

Preuve. Supposons f dérivable et notons $z - z_0 = h + ik$ et $c = f'(z_0)$: (7.4) se réexprime en termes de \tilde{f} selon

$$\tilde{f}(x_0 + h, y_0 + k) - \tilde{f}(x_0, y_0) - c(h + ik) = o(\sqrt{h^2 + k^2}) \quad (7.6)$$

ce qui prouve que \tilde{f} est différentiable, qui identifie $\frac{\partial \tilde{f}}{\partial x} = c$, $\frac{\partial \tilde{f}}{\partial y}(x_0, y_0) = ic$, et qui établit (7.5). Réciproquement si \tilde{f} satisfait (7.5), alors on a bien (7.6) avec $c = \frac{\partial \tilde{f}}{\partial x}(x_0, y_0)$ et la fonction $f(z)$ est donc dérivable en z_0 avec $f'(z_0) = c$.

Contre-exemple : Soit $f(z) = z\bar{z}$, $\tilde{f}(x, y) = x^2 + y^2$. La fonction \tilde{f} est bien différentiable : $\partial_x \tilde{f}(1, 0) = 2$, $\partial_y \tilde{f}(1, 0) = 0$ et $\tilde{f}(x, y) - \tilde{f}(1, 0) = (1+h)^2 + k^2 - 1 - 2h = h^2 + k^2$. Examinons la dérivabilité en $z_0 = 1$. $\frac{f(z) - f(1)}{z-1} = \frac{z\bar{z}-1}{z-1} = \bar{z} + \frac{\bar{z}-1}{z-1}$. Quand $z \rightarrow 1$, le premier terme $\rightarrow 1$ mais le second est une phase qui dépend de la direction d'approche.

Dans la suite on ne distinguerà plus f et \tilde{f} . Il est souvent commode d'écrire

$$f = P + iQ \text{ avec } P = \Re e f, Q = \Im m f.$$

Les conditions de Cauchy–Riemann se récrivent alors sous la forme

Théorème 7.2 (conditions de Cauchy–Riemann 2) : Si $f : \Omega \rightarrow \mathbb{C}$ est dérivable en $z_0 \in \Omega$, et $f = P + iQ$ avec $P = \Re e f$, $Q = \Im m f$, alors

$$\boxed{\frac{\partial P}{\partial x}(x_0, y_0) = \frac{\partial Q}{\partial y}(x_0, y_0) \quad \frac{\partial Q}{\partial x}(x_0, y_0) = -\frac{\partial P}{\partial y}(x_0, y_0)} \quad (7.7)$$

Interprétation géométrique : les deux courbes $\Re e f(z) = \text{const.}$ et $\Im m f(z) = \text{const.}$ sont orthogonales au point z_0 .

Preuve (de cette interprétation) : voir TD.

Notons finalement que par les formules de Cauchy–Riemann, nous avons plusieurs expressions équivalentes pour la dérivée d'une fonction holomorphe :

$$f'(z) = \frac{\partial P}{\partial x} + i \frac{\partial Q}{\partial x} = \frac{\partial Q}{\partial y} - i \frac{\partial P}{\partial y} = \frac{\partial P}{\partial x} - i \frac{\partial P}{\partial y} = \frac{\partial Q}{\partial y} + i \frac{\partial Q}{\partial x}. \quad (7.8)$$

Exemples, contre-exemples. Tout polynôme $R(z)$ (à coefficients dans \mathbb{C}) est dérivable donc holomorphe en tout point z de \mathbb{C} . Les fonctions e^z , $\sin z$, $\cos z$ etc sont holomorphes sur \mathbb{C} . La fonction $f : z \mapsto 1/z$ est holomorphe sur $\mathbb{C}^* \stackrel{\text{déf}}{=} \mathbb{C} \setminus \{0\}$. En revanche la fonction $f : z \mapsto \bar{z}$

n'est pas dérivable/holomorphe. En effet pour tout z_0 , examinons la limite $\zeta = \frac{\bar{z} - \bar{z}_0}{z - z_0}$ quand $|z - z_0| \rightarrow 0$. Posant $z - z_0 = \rho(z)e^{i\theta(z)}$, on voit que $\zeta = e^{-2i\theta(z)}$ qui peut prendre toute valeur selon la façon dont z approche z_0 . Le rapport ζ n'a pas de limite bien définie et la fonction n'est pas dérivable. Pour la même raison, les fonctions $z \mapsto |z|$, $z \mapsto \Re z$, $z \mapsto \Im z$, ne sont pas non plus holomorphes.

De façon heuristique, sont holomorphes les fonctions (suffisamment régulières) "de z seulement". On voit apparaître une idée intéressante : les variables z et \bar{z} peuvent être considérées comme indépendantes, une idée que l'on retrouvera dans la suite.

7.1.3 Dérivations $\partial, \bar{\partial}$

Comme on l'a rappelé plus haut en (7.3), pour une fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ différentiable au point (x_0, y_0) , on peut écrire la différentielle en (x_0, y_0) comme

$$df = \left. \frac{\partial f}{\partial x} \right|_{(x_0, y_0)} dx + \left. \frac{\partial f}{\partial y} \right|_{(x_0, y_0)} dy. \quad (7.9)$$

Appliquons cela à des fonctions $f : \Omega \rightarrow \mathbb{C}$ d'une variable $z = x + iy$. Pour les fonctions $z \mapsto z$ et $z \mapsto \bar{z}$ on a $dz = dx + idy$ et $d\bar{z} = dx - idy$, soit encore $dx = \frac{1}{2}(dz + d\bar{z})$ et $dy = \frac{1}{2i}(dz - d\bar{z})$. L'expression (7.9) conduit alors à

$$\begin{aligned} df &= \frac{1}{2} \frac{\partial f}{\partial x} (dz + d\bar{z}) + \frac{1}{2i} \frac{\partial f}{\partial y} (dz - d\bar{z}) \\ &= \frac{1}{2} \left(\frac{\partial f}{\partial x} - i \frac{\partial f}{\partial y} \right) dz + \frac{1}{2} \left(\frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} \right) d\bar{z}. \end{aligned} \quad (7.10)$$

Il est donc naturel de noter

$$\frac{\partial}{\partial z} \stackrel{\text{déf}}{=} \frac{1}{2} \left(\frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right) \quad \frac{\partial}{\partial \bar{z}} \stackrel{\text{déf}}{=} \frac{1}{2} \left(\frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right) \quad (7.11)$$

et ces opérateurs différentiels seront notés aussi ∂_z et $\partial_{\bar{z}}$, ou plus simplement ∂ et $\bar{\partial}$ chaque fois qu'il n'y aura pas d'ambiguïté.

On a donc finalement pour la différentielle de f l'expression compacte

$$df = \frac{\partial f}{\partial z} dz + \frac{\partial f}{\partial \bar{z}} d\bar{z}$$

à nouveau en accord avec la remarque ci-dessus : les variables z et \bar{z} peuvent être considérées comme indépendantes.

Ce qui a précédé n'a fait usage que de la différentiabilité dans \mathbb{R}^2 de la fonction f . Supposons maintenant la fonction f dérivable dans $\Omega \subset \mathbb{C}$ (holomorphe) :

Théorème 7.3 (conditions de Cauchy–Riemann 3) : Si $f : \Omega \rightarrow \mathbb{C}$ est dérivable en $z_0 \in \Omega$, alors

$$\boxed{\frac{\partial f}{\partial z}(z_0) = f'(z_0) \quad \text{et} \quad \frac{\partial f}{\partial \bar{z}}(z_0) = 0}, \quad (7.12)$$

puisque compte tenu de (7.10) et (7.5), $\frac{\partial f}{\partial z}(z_0) = \frac{1}{2} \left(\frac{\partial f}{\partial x}(x_0, y_0) - i \frac{\partial f}{\partial y}(x_0, y_0) \right) = \frac{\partial f}{\partial x}(x_0, y_0) = f'(z_0)$, (Théorème 7.1), tandis que $\frac{\partial f}{\partial \bar{z}}(z_0) = \frac{1}{2} \left(\frac{\partial f}{\partial x}(x_0, y_0) + i \frac{\partial f}{\partial y}(x_0, y_0) \right) = 0$, cette dernière relation exprimant à nouveau l’“indépendance de f par rapport à \bar{z} ”.

Exercice. Vérifier à l'aide de cette nouvelle forme des conditions de Cauchy–Riemann que la fonction $z \mapsto |z|$ n'est pas holomorphe.

Application. Utilisons ce qui précède pour établir la

Proposition 7.4 : Soit f holomorphe dans un ouvert connexe Ω ; si $\Re e f$ est constante, alors f est constante.

Preuve. Comme $\Re e f = \frac{1}{2}(f + \bar{f})$, on a $d(f + \bar{f}) = 0$, ce qui compte tenu de l'holomorphie $\partial_z f = 0$ et (par conjugaison) $\partial_z \bar{f} = 0$, conduit à $\partial_z f dz + \partial_{\bar{z}} \bar{f} d\bar{z} = 0$. Mais cela n'est possible que si séparément $\partial_z f = 0$ et $\partial_{\bar{z}} \bar{f} = 0$. Donc $df = 0$ et f est constante.

On démontre de même (voir TD) que sous la même hypothèse f holomorphe dans Ω connexe, $\Im m f = \text{const.} \implies f = \text{const.}, |f| = \text{const.} \implies f = \text{const.};$ si $f \neq 0$ dans Ω , $\log |f| = \text{const.} \implies f = \text{const.}, \arg f = \text{const.} \implies f = \text{const.}$

7.2 Intégrales sur des chemins

7.2.1 Chemins et lacets

On appelle *chemin différentiable* γ dans un ouvert Ω de \mathbb{C} une application continûment différentiable d'un intervalle $[a, b]$ de \mathbb{R} (avec $a < b$) dans Ω . (On peut penser au paramètre $t \in [a, b]$ comme à une variable de temps quand on parcourt le chemin...) Par abus de notation, γ désigne à la fois la fonction et son image dans le plan \mathbb{C} . Les points $\gamma(a)$ et $\gamma(b)$ sont appelés l'*origine* et l'*extrémité* de γ . Le chemin est *fermé* si $\gamma(a) = \gamma(b)$ (on parle aussi de *lacet*). Il est *simple* s'il ne se recoupe pas (l'application γ est injective), sauf en a et b s'il est fermé.

On est souvent amené à effectuer un changement de paramétrisation : si $t \in [a, b]$ et $t = \varphi(u)$, $u \in [a_1, b_1]$ avec $\varphi : [a_1, b_1] \rightarrow [a, b]$, fonction continue, différentiable et strictement croissante, les deux paramétrisations $t \mapsto \gamma(t)$ et $u \mapsto (\gamma \circ \varphi)(u)$ décrivent des *chemins équivalents*. Par exemple les fonctions $t \mapsto e^{2i\pi t}$ de $[0, 1] \rightarrow \mathbb{C}$ et $u \mapsto e^{-2i\pi \cos u}$ de $[0, \frac{1}{2}\pi] \rightarrow \mathbb{C}$ définissent deux chemins équivalents, c'est-à-dire de même image γ (ici le cercle unité) parcourue dans le même sens.

Pour décrire des chemins ayant un nombre fini de points anguleux, tel le bord d'un rectangle, on doit un peu généraliser les définitions qui précèdent et considérer des chemins *différentiables par morceaux* : l'intervalle $[a, b]$ peut être subdivisé en un nombre fini p d'intervalles adjacents $[a_i, a_{i+1}]$, $a_0 = a$, $a_p = b$, tels que γ soit continue sur $[a, b]$ et continûment différentiable sur chaque $[a_i, a_{i+1}]$. Cela sera implicite dans la suite quand on parlera de chemin.

Chemins homotopes dans Ω

Définition 7.2 : On dit que deux chemins $\gamma_0 : [0, 1] \rightarrow \Omega$ et $\gamma_1 : [0, 1] \rightarrow \Omega$ de mêmes extrémités $\gamma_0(0) = \gamma_1(0)$ et $\gamma_0(1) = \gamma_1(1)$ sont homotopes dans Ω s'il existe une déformation continue de γ_0 en γ_1 , c'est-à-dire une fonction $\zeta : [0, 1] \times [0, 1] \rightarrow \Omega$ telle que

$$\forall t \in [0, 1] \quad \zeta(t, 0) = \gamma_0(t) \quad \zeta(t, 1) = \gamma_1(t) \quad (7.13)$$

$$\forall u \in [0, 1] \quad \zeta(0, u) = \gamma_0(0) = \gamma_1(0) \quad \zeta(1, u) = \gamma_0(1) = \gamma_1(1) \quad (7.14)$$

Autrement dit dans $\zeta(t, u)$, u est le paramètre de déformation, et tous les chemins $\gamma(\cdot, u)$ interpolant entre γ_0 et γ_1 ont mêmes extrémités. La relation d'homotopie est une relation d'équivalence entre chemins, qu'on notera simplement $\gamma_0 \sim \gamma_1$.

Définition 7.3 : Un ouvert Ω de \mathbb{C} est dit simplement connexe si tout chemin fermé est homotope à un point.

Heuristiquement, dire que Ω est simplement connexe signifie qu'il n'“a pas de trou”, contrairement au cas représenté sur la Fig. 7.1. Noter que le “trou” peut être réduit à un point : le plan complexe pointé \mathbb{C}^* n'est pas simplement connexe. Remarque : les notions de chemins, d'homotopie et de simple connexité s'étendent à tout espace topologique E .

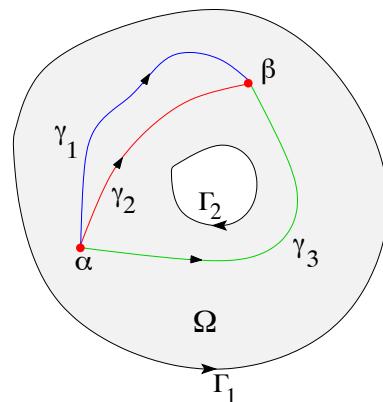


FIGURE 7.1 – Le chemin γ_1 est homotope à γ_2 dans Ω , mais pas à γ_3 . Le domaine Ω n'est pas simplement connexe ; il a pour bord orienté $\Gamma_1 \cup \Gamma_2$.

Bord d'un domaine compact

Soit B un domaine compact du plan. On dit qu'il a pour *bord orienté* $\Gamma = \cup_i \Gamma_i$, où chaque chemin Γ_i est simple et différentiable par morceaux, et où les images des différents Γ_i sont disjointes, si Γ est la frontière de B (c'est-à-dire $\Gamma = \bar{B} - \overset{\circ}{B}$, cf Chap. 1) et (orientation) si quand on parcourt chaque Γ_i dans le sens des t croissants, on a localement à sa gauche des points de B et à sa droite des points de son complémentaire. Cette définition qui peut être formalisée davantage (cf [4] p. 65-66) recouvre un fait assez intuitif, illustré sur la figure 7.1 où le bord de Ω est constitué de $\Gamma = \Gamma_1 \cup \Gamma_2$. En général, il est commode de noter ∂B le bord orienté de B .

7.2.2 Intégrales de formes sur des chemins.

Considérons maintenant une *forme différentielle* dans l'ouvert Ω , c'est-à-dire une expression

$$\omega = P dx + Q dy$$

où P et Q sont à valeurs réelles ou complexes. Par définition l'intégrale de la forme ω sur un chemin différentiable $\gamma : t \in [a, b] \mapsto (x(t), y(t))$ est

$$\int_{\gamma} \omega \stackrel{\text{déf}}{=} \int_a^b \gamma^*(\omega) dt \quad (7.15)$$

où la forme $\gamma^*(\omega)$ est donnée par

$$\gamma^*(\omega) = (P(x(t), y(t))x'(t) + Q(x(t), y(t))y'(t)) dt. \quad (7.16)$$

On note qu'un changement de paramétrisation de la courbe ne modifie pas cette intégrale : si $t = \varphi(u)$, $\gamma^*(\omega) = (P(x(\varphi(u)), y(\varphi(u)))x'(\varphi(u))\varphi'(u) + Q(x(\varphi(u)), y(\varphi(u)))y'(\varphi(u))\varphi'(u)) du = (P(\tilde{x}(u), \tilde{y}(u))\tilde{x}'(u) + Q(\tilde{x}(u), \tilde{y}(u))\tilde{y}'(u)) du$. Ces considérations s'étendent à des chemins différentiables par morceaux : l'intégrale de ω est la somme des intégrales sur les morceaux différentiables successifs de la courbe, $\int_{\gamma} \omega = \sum_{i=1}^p \int_{\gamma_i} \omega$.

Si $\gamma : [0, 1] \rightarrow \Omega$ est un chemin d'origine α et d'extrémité β , on note $-\gamma$ le chemin parcouru en sens inverse $-\gamma(t) = \gamma(1-t)$; il a pour origine β et pour extrémité α . On a bien sûr $\int_{-\gamma} \omega = -\int_{\gamma} \omega$. Dans le même ordre d'idées, si le chemin γ_2 a pour origine l'extrémité du chemin γ_1 , on peut *composer* les deux chemins en un chemin $\gamma_1 + \gamma_2$ ¹ d'origine $\gamma_1(0)$ et d'extrémité $\gamma_2(1)$. Enfin, si γ_1 et γ_2 ont même origine α et même extrémité β , le chemin $\gamma_1 - \gamma_2$ est un chemin fermé de α à α .

Bien comprendre qu'une **intégrale sur un chemin d'extrémités α et β données dépend en général de ce chemin**. C'est une question qui va nous occuper maintenant

1. une notation pas très heureuse, puisqu'on ne peut pas en général définir $\gamma_2 + \gamma_1 \dots$

de trouver quelles conditions doivent être satisfaites pour que l'intégrale ne dépende pas du chemin, et qu'on puisse *déformer le chemin (ou contour) d'intégration*.

Formes exactes, formes fermées, primitives

Supposons que ω est une *forme exacte* dans l'ouvert Ω , c'est-à-dire qu'elle y est la différentielle d'une fonction $f : \omega = df$, continûment différentiable, appelée la *primitive de la forme ω* ; alors pour un chemin γ (différentiable par morceaux)

$$\int_{\gamma} df = f(\gamma(b)) - f(\gamma(a)). \quad (7.17)$$

En particulier, pour un chemin fermé (complètement contenu dans Ω), on voit que $\int_{\gamma} df = 0$.

Attention que la définition de la primitive f de la forme exacte ω impose que f est définie dans tout Ω . Une condition plus faible serait que *localement*, c'est-à-dire dans un voisinage de tout point, il existe une fonction f telle que $\omega = df$. Une telle forme ω peut être appelée *fermée*². Évidemment toute forme exacte est fermée, la réciproque n'étant en général pas vraie. Exemple, dans $\Omega = \mathbb{C}^*$ (plan “pointé” à l'origine), $\omega = dz/z$ est fermée, puisque localement $\omega = d \log z$, mais pas exacte. En effet comme on vérifie aisément, $\int_{\gamma} \omega = 2\pi i n$ où n est le nombre algébrique de fois où le chemin γ tourne autour de l'origine, voir ci-dessous (7.23).

Proposition 7.5 : Si $\omega = df$ est exacte dans Ω , son intégrale le long du chemin γ est la même pour tous les chemins homotopes à γ .

La preuve est simple : si $\gamma_0 \sim \gamma_1$, $\int_{\gamma_0} \omega - \int_{\gamma_1} \omega$ peut être considérée comme l'intégrale le long du chemin fermé $\gamma_0 - \gamma_1$ obtenu en composant γ_0 et $-\gamma_1$, c'est-à-dire le chemin γ_1 parcouru en sens inverse. Il découle de la remarque suivant (7.17) que $\int_{\gamma_0 - \gamma_1} df = 0$. cqfd.

Proposition 7.6 : Pour qu'une forme différentielle ω admette une primitive dans Ω , il faut et il suffit que $\int_{\gamma} \omega = 0$ pour tout chemin fermé γ différentiable par morceaux et contenu dans Ω .

Preuve. On vient de voir que la condition est nécessaire. Qu'elle est suffisante résulte de la construction suivante : soit $\omega = Pdx + Qdy$ une forme satisfaisant la condition du Théorème ; on se donne un point $(x_0, y_0) \in \Omega$ quelconque dans Ω et on définit $f(x, y) = \int_{\gamma} \omega$ où γ est un chemin différentiable dans Ω d'origine (x_0, y_0) et d'extrémité (x, y) . Sous l'hypothèse de la Proposition, $f(x, y)$ ne dépend pas du choix de γ , puisque pour un autre chemin γ_1 de (x_0, y_0) à (x, y) , $\int_{\gamma} \omega - \int_{\gamma_1} \omega = \int_{\gamma - \gamma_1} \omega$ et que $\gamma - \gamma_1$ est un chemin fermé. On vérifie alors aisément que f est continûment différentiable dans Ω et que $\frac{\partial f}{\partial x} = P(x, y)$, $\frac{\partial f}{\partial y} = Q(x, y)$. (Idée : calculer ces dérivées partielles en intégrant ω le long de *segments de droite* entre x et $x + dx$ etc.) f est bien la primitive de ω , cqfd.

Remarque. La condition précédente, qui porte sur *tout* chemin fermé γ , semble d'application difficile. En fait il suffit d'assurer une condition moins forte :

2. Par la suite on donnera une définition différente : ω est fermée si sa différentielle que nous ne définirons pas ici s'annule : $d\omega = 0$, une condition qu'on montre être équivalente à “ ω est localement exacte”.

Proposition 7.7 : Soit D un disque ouvert. Si $\int_{\gamma} \omega = 0$ pour tout chemin γ qui est le bord d'un rectangle complètement contenu dans D , alors ω a une primitive dans D (elle est exacte).

Par "rectangle complètement contenu ...", on veut dire que l'intérieur **et** la frontière du rectangle sont contenus dans D , voir Fig. 7.2(a).

La preuve est instructive : Soit (x_0, y_0) le centre du disque. Tout point (x, y) de D définit avec (x_0, y_0) un rectangle complètement contenu dans D , voir Fig. 7.2(a). Les deux intégrales de (x_0, y_0) à tout $(x, y) \in \Omega$ le long des deux contours $(x_0, y_0) \rightarrow (x, y) \rightarrow (x, y)$ ou $(x_0, y_0) \rightarrow (x_0, y) \rightarrow (x, y)$ sont égales puisque $\int_{\gamma} \omega = 0$ et définissent une primitive $F(x, y)$ de ω (même calcul que dans la Prop. 7.6).

Si au lieu d'un disque on prend un ouvert Ω quelconque, on peut appliquer un raisonnement du même genre, en prenant un point de base $(x_0, y_0) \in \Omega$ tel que le rectangle de sommets (x_0, y_0) et (x, y) soit complètement contenu dans Ω , ce qui est toujours possible pour tout point (x, y) de l'ouvert Ω . Mais la construction sera locale, et la fonction f ainsi construite ne sera pas nécessairement une primitive définie dans tout Ω . Donc

Proposition 7.8 : Soit Ω un ouvert. La forme ω est fermée si et seulement si $\int_{\gamma} \omega = 0$ chaque fois que γ est le bord d'un petit rectangle contenu dans l'ouvert Ω .

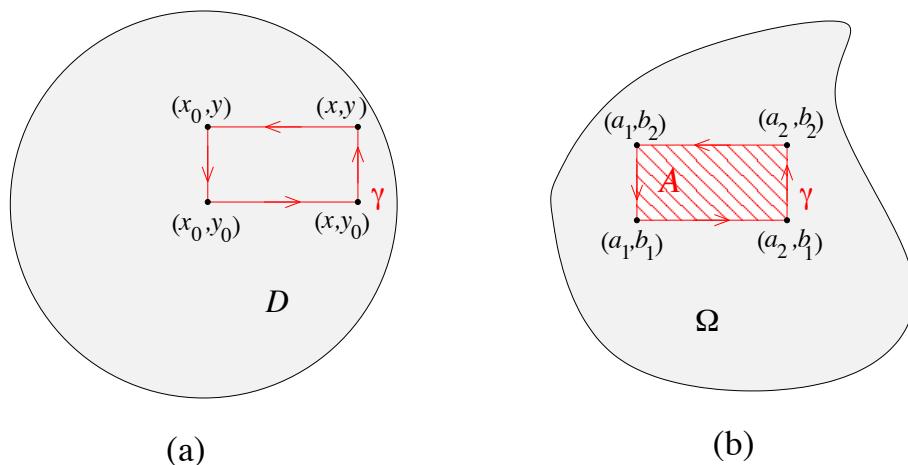


FIGURE 7.2 – (a) Contour rectangulaire γ partant du centre (x_0, y_0) du disque D ; (b) Rectangle A contenu dans l'ouvert Ω .

Formule de Green–Riemann

Théorème 7.9 (Green–Riemann) : Soit $\omega = Pdx + Qdy$ une forme différentielle définie dans un ouvert Ω , avec $\frac{\partial P}{\partial y}$ et $\frac{\partial Q}{\partial x}$ continues dans Ω . Soit K un domaine compact contenu dans Ω , $\gamma = \partial K$ son bord orienté supposé simple (ne se recoupant pas). Alors

$$\int_{\gamma} (Pdx + Qdy) = \iint_K \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dx dy. \quad (7.18)$$

Ce théorème qu'on a déjà rencontré à travers ses applications physiques (la circulation d'un vecteur le long d'un chemin fermé γ égale le flux de son rotationnel à travers la surface dont γ est le bord orienté...) est un cas particulier d'une classe de résultats auxquels on donne le nom de théorème de Stokes, de la forme générale $\int_{\Omega} d\omega = \int_{\partial\Omega} \omega$ pour des domaines Ω et des formes différentielles ω de dimension plus élevée, cf le théorème de Gauss-Ostrogradsky rencontré en électromagnétisme : le flux d'un vecteur à travers une surface fermée égale l'intégrale de sa divergence dans le volume limité par la surface, etc. Nous n'aurons malheureusement pas le temps d'aborder ces questions.

Preuve de la formule de Green–Riemann. On ne donnera la preuve que dans le cas d'un rectangle de côtés parallèles aux axes des x et des y , voir Fig. 7.2(b). Soient $a_1 < a_2$ les abscisses des sommets du rectangle, $b_1 < b_2$ leurs ordonnées. Le chemin fermé γ est fait de 4 segments orientés : $[(a_1, b_1) \rightarrow (a_2, b_1)]$, $[(a_2, b_1) \rightarrow (a_2, b_2)]$, $[(a_2, b_2) \rightarrow (a_1, b_2)]$, $[(a_1, b_2) \rightarrow (a_1, b_1)]$. On a

$$\iint_A \frac{\partial P}{\partial y} dx dy = \int_{a_1}^{a_2} dx \int_{b_1}^{b_2} dy \frac{\partial P}{\partial y} = \int_{a_1}^{a_2} P(x, b_2) dx - \int_{a_1}^{a_2} P(x, b_1) dx = - \int_{\gamma} P dx$$

et une expression analogue pour $\iint_A \frac{\partial Q}{\partial x} dx dy$. En remettant tout ensemble, on a bien (7.18).

Nous ferons par la suite usage de la Proposition suivante :

Proposition 7.10 : Soit une forme différentielle $\omega = Pdx + Qdy$ définie dans un ouvert connexe Ω , telle que $\frac{\partial P}{\partial y}$ et $\frac{\partial Q}{\partial x}$ sont continues dans Ω . Alors

$$\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x} \quad (7.19)$$

est une condition nécessaire pour que $\omega = df$ dans Ω . Elle est suffisante localement (forme fermée) ; elle est suffisante globalement dans tout Ω (forme exacte) si Ω est un disque ouvert.

Preuve : Dire que $\omega = df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$ équivaut à dire que $P = \frac{\partial f}{\partial x}$ et $Q = \frac{\partial f}{\partial y}$, un système de deux équations aux dérivées partielles dont la condition nécessaire et suffisante d'intégrabilité (locale) est $\frac{\partial Q}{\partial x} = \frac{\partial P}{\partial y} = \frac{\partial^2 f}{\partial x \partial y}$, cf Appendice D. Ou encore, autre méthode, si (7.19) est satisfaite, alors le théorème de Green–Riemann nous dit que $\int_{\gamma} \omega = 0$ pour tout chemin rectangulaire γ et la Proposition 7.8 nous dit que ω admet localement une primitive (elle est fermée). La condition (7.19) est nécessaire : si ω est fermée, la Proposition 7.8 et le théorème de G-R nous disent que $\iint_A (\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}) = 0$ pour tout petit rectangle $A \subset \Omega$, ce qui implique (par l'absurde) que $(\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}) = 0$ en tout point (x, y) de Ω . Si le domaine est un disque, (7.19) est une condition suffisante pour que ω soit exacte, car on applique la Proposition 7.7. Comme on va le voir ci-dessous, c'est aussi une condition suffisante pour un domaine Ω simplement connexe.

Exemple : la forme $\omega = dz/z$ satisfait bien (7.19) dans \mathbb{C}^* , elle est fermée (localement exacte) $\omega = d \log z$, mais pas exacte dans tout \mathbb{C}^* , en raison une fois encore des déterminations multiples du log, ou du fait que \mathbb{C}^* n'est pas simplement connexe.

Nous avons finalement le théorème suivant

Théorème 7.11 : Soit ω une forme différentielle fermée dans un ouvert connexe Ω . Alors

- (i) pour deux chemins homotopes γ_1 et γ_2 (de mêmes extrémités), $\int_{\gamma_1} \omega = \int_{\gamma_2} \omega$; ou de façon équivalente, pour tout chemin fermé γ homotope à un point, $\int_{\gamma} \omega = 0$;
- (ii) si Ω est simplement connexe, ω y est exacte (elle admet une primitive dans Ω).

Éléments de preuve : (i) La forme est fermée, donc (Proposition 7.10 et Théorème de Green–Riemann), $\int_{\gamma} \omega = 0$ le long de tout chemin fermé simple complètement contenu dans Ω . Cela s'applique au chemin fermé $\gamma_1 - \gamma_2$ qu'on supposera simple (quitte à le décomposer en un nombre fini de lacets simples s'il a des points multiples). Donc $\int_{\gamma_1} \omega = \int_{\gamma_2} \omega$, cqfd. Pour le point (ii), on observe que si Ω est simplement connexe, tout chemin fermé γ y est homotope à un point, donc $\int_{\gamma} \omega = 0$, et on applique alors la Proposition 7.6. Pour une démonstration plus soigneuse, se reporter à [4], p. 60-61.

7.2.3 Intégration sur des chemins dans \mathbb{C}

Examinons maintenant comment ces formules se transcrivent en variables complexes.

Définition 7.4 : Soit $\omega = \omega_1 dz + \omega_2 d\bar{z}$ une forme définie sur Ω , et γ un chemin dans Ω . L'intégrale de ω sur le chemin γ est définie par

$$\int_{\gamma} \omega \stackrel{\text{déf}}{=} \int_a^b (\omega_1(\gamma(t)) \gamma'(t) + \omega_2(\gamma(t)) \bar{\gamma}'(t)) dt. \quad (7.20)$$

Le chemin γ est appelé contour d'intégration dans le plan complexe.

En particulier, si $\omega = f dz$, (f pas nécessairement holomorphe), $\int_{\gamma} f(z) dz = \int_a^b f(\gamma(t)) \gamma'(t) dt$. Si $\omega = df$, $\int_{\gamma} df = f(\gamma(b)) - f(\gamma(a))$ comme en (7.17).

Cette définition est l'équivalent des formules (7.15, 7.16) et en découle via les formules exprimant dx et dy en termes de dz et $d\bar{z}$. Elle est également insensible à un reparamétrage du chemin γ , grâce aux formules de changements de variable dans une intégrale (Vérifier!).

Avec cette notation, on a le résultat important

Théorème 7.12 : Soit $f : \Omega \rightarrow \mathbb{C}$ une fonction holomorphe et $\gamma : [a, b] \rightarrow \mathbb{C}$ un chemin dans Ω . Alors

$$f(\gamma(b)) - f(\gamma(a)) = \int_{\gamma} \frac{\partial f}{\partial z}(z) dz. \quad (7.21)$$

Autrement dit, sous les hypothèses du théorème, l'intégrale de contour ne dépend pas du contour γ mais seulement de ses extrémités.

Preuve : La relation (7.17) s'applique à df , mais f étant holomorphe, $df = \frac{\partial f}{\partial z}(z) dz$.

Attention : cette formule n'est valable que pour f holomorphe.

Exemple d'intégrale de contour : soit γ le cercle de centre $a \in \mathbb{C}$ et de rayon R , qu'on paramétrise par exemple par $\theta \in [0, 2\pi] \mapsto \gamma(\theta) = a + Re^{i\theta}$. Si f est une fonction (holomorphe ou non) définie dans un ouvert Ω contenant γ , alors (7.20) appliqué à $\omega = f(z) dz$ donne

$$\int_{\gamma} f(z) dz = \int_0^{2\pi} f(a + Re^{i\theta}) iRe^{i\theta} d\theta$$

Cas particulier : intégrons $f(z) = 1/z$ sur le cercle γ de centre O et de rayon R , (ici $\Omega = \mathbb{C}^*$) :

$$\frac{1}{2\pi i} \int_{\gamma} \frac{dz}{z} = \frac{1}{2\pi i} \int_0^{2\pi} \frac{ie^{i\theta}}{e^{i\theta}} d\theta = 1, \quad (7.22)$$

puisque $z = Re^{i\theta}$. On note que le résultat ne dépend pas de R .

Proposition 7.13 : Pour tout chemin fermé γ ne passant pas par l'origine, $\frac{1}{2\pi i} \int_{\gamma} \frac{dz}{z}$ est un entier.

Preuve : la forme dz/z n'est pas exacte, mais localement elle admet comme primitive $\log z$. On se rappelle (chap. 6) que la fonction \log admet plusieurs déterminations. L'intégrale de contour $\frac{1}{2\pi i} \int_{\gamma} \frac{dz}{z}$ est donc égale à la différence de deux déterminations du logarithme divisée par $2\pi i$, soit un entier.

Corollaire : Pour tout chemin γ ne passant pas par l'origine, $\frac{1}{2\pi} \int_{\gamma} \frac{x dy - y dx}{x^2 + y^2}$ est un entier.

Preuve : calculer la différentielle de $\Im m \log z = \arg z = \text{Arctan} \frac{y}{x}$.

Indice d'un chemin fermé par rapport à un point

Théorème 7.14 : Soit γ un chemin fermé et soit U le complémentaire de γ dans le plan complexe. Soit z un point n'appartenant pas à γ (donc $z \in U$). L'indice de γ par rapport à z , défini par

$$\text{Ind}_{\gamma}(z) \stackrel{\text{def}}{=} \frac{1}{2\pi i} \int \frac{d\zeta}{\zeta - z} \quad (7.23)$$

est un nombre entier, constant dans chaque composante connexe de U , nulle sur la composante connexe non bornée de U et variant de ± 1 à chaque traversée du chemin γ en dehors des points multiples.

Heuristiquement l'indice mesure le nombre algébrique de tours qu'effectue le chemin γ autour du point z .

Preuve du Théorème. Que ce soit un entier découlé de la proposition 7.13. Par ailleurs l'indice est une fonction continue de z tant que $z \notin U$ et sa valeur est un entier, donc c'est une constante dans chaque composante connexe de U . Ou inversement, c'est une fonction qui ne varie pas quand γ est déformé sans passer par z . Si z appartient à la composante connexe non bornée de U , γ peut être déformé en un point sans croiser z , donc l'indice est nul. Enfin la traversée par z de γ en dehors d'un point multiple équivaut à modifier γ en $\gamma + \gamma_1$, avec γ_1 déformable en un cercle parcouru une fois dans le sens positif ou négatif, voir Figure 7.3 droite. On a fait ici appel à des déformations de contour d'intégration, une méthode très utilisée dans la suite. Pour une autre démonstration, voir [1], Annexe D.

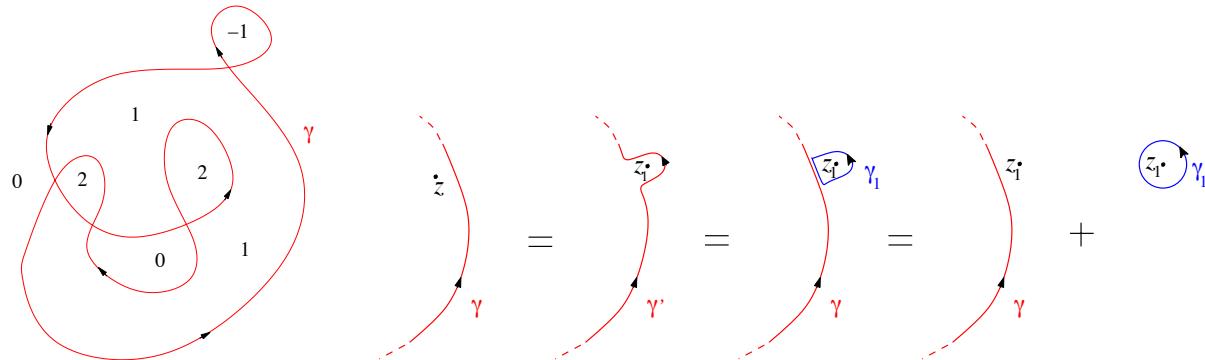


FIGURE 7.3 – À gauche : L’indice du chemin γ par rapport à un point est constant dans chaque composante connexe du complémentaire U de γ . À droite : La traversée par z du contour γ équivaut à ajouter la contribution du lacet γ_1 : $\text{Ind}_\gamma(z) = \text{Ind}_{\gamma'}(z_1) = \text{Ind}_\gamma(z_1) + \text{Ind}_{\gamma_1}(z_1) = \text{Ind}_\gamma(z_1) + 1$.

7.3 Théorème de Cauchy et conséquences

7.3.1 Théorème de Cauchy

Proposition 7.15 : Soit f une fonction holomorphe sur un ouvert Ω , γ un chemin fermé dans Ω , alors $\int_\gamma f'(z)dz = 0$.

En effet cela est une conséquence du Théorème 7.12, $\int_\gamma f'(z)dz = f(\gamma(b)) - f(\gamma(a)) = 0$ puisque le chemin est fermé. En revanche, si f n’est pas holomorphe, $\oint \frac{\partial f}{\partial z} dz$ n’a aucune raison de s’annuler. (Remarquer la notation \oint qui signale que l’on intègre sur un chemin fermé.) Exemple : $f = z\bar{z}$, l’intégrale de $\frac{\partial f}{\partial z} = \bar{z}$ le long du cercle de centre O et de rayon R est (on paramétrise $z = Re^{i\theta}$ comme ci-dessus) $\int_0^{2\pi} iR^2 d\theta = 2\pi iR^2$. Noter que sur le cercle, $\bar{z} = R^2/z$, et on vient juste de répéter le calcul de (7.22) !

Corollaire : Pour tout chemin fermé γ et tout entier $n \neq -1$, on a $\int_\gamma z^n dz = 0$. (Pour $n < 0$ on suppose que γ ne passe pas par 0.)

La preuve est claire : pour $n \neq -1$, z^n a pour primitive $z^{n+1}/n + 1$, qui est holomorphe dans \mathbb{C} (resp. \mathbb{C}^* si $n < -1$) et la proposition précédente s’applique.

Théorème 7.16 (Cauchy) : Si f est une fonction holomorphe sur Ω , la forme $f(z)dz$ est fermée dans Ω : localement elle admet une primitive $F(z)$ telle que $dF(z) = f(z)dz$.

Il s’agit là d’un théorème fondamental dont vont découler beaucoup de conséquences et d’applications. Démontrons-le en faisant l’hypothèse supplémentaire que $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont continues dans Ω . (On verra plus bas que cette propriété est en fait toujours satisfaite par une fonction holomorphe. Mais on peut aussi établir directement, avec un peu plus d’effort, le théorème de Cauchy sous la seule hypothèse d’holomorphie, voir [4], p. 70–71.) Il suffit alors d’écrire la forme

$f(z)dz = f(z)dx + if(z)dy$, soit dans les notations de la Prop. 7.10, $P = f(z)$, $Q = if(z)$. Les conditions d'holomorphie de Cauchy–Riemann $\frac{\partial f}{\partial y} = i \frac{\partial f}{\partial x}$ et la Proposition 7.10 nous assurent alors que $f dz$ est fermée.

Corollaire : Si f est holomorphe dans Ω , on a $\int_{\gamma} f(z)dz = 0$ pour tout chemin fermé homotope à un point dans Ω . De façon équivalente, pour deux chemins γ_1 et γ_2 de mêmes extrémités et homotopes, $\int_{\gamma_1} f(z)dz = \int_{\gamma_2} f(z)dz$.

C'est la transcription en variables complexes du Théorème 7.11, et c'est sous cette forme, ou sous l'une des formes équivalentes ci-dessous, que le théorème de Cauchy va nous être le plus utile. La deuxième formulation de ce corollaire implique que l'on peut *déformer continûment le contour d'intégration* d'une fonction holomorphe, à extrémités fixes, ce qui généralise ce que l'on a vu au Théorème 7.12. Une forme plus faible de ce corollaire (elle aussi découlant du Théorème 7.11) consiste à dire

Corollaire : Si f est holomorphe dans un ouvert Ω simplement connexe, on a $\int_{\gamma} f(z)dz = 0$ pour tout chemin fermé dans Ω .

7.3.2 Intégrale de Cauchy

Théorème 7.17 : Soit f holomorphe dans Ω , $a \in \Omega$, et γ un chemin fermé dans Ω ne passant pas par a et homotope à un point dans Ω . Alors

$$\boxed{\frac{1}{2\pi i} \int_{\gamma} \frac{f(z)dz}{z-a} = \text{Ind}_{\gamma}(a) f(a).} \quad (7.24)$$

Voilà un résultat qui va nous être extrêmement utile ! Donnons-en la preuve.

Soit $g(z)$ la fonction définie par

$$g(z) = \begin{cases} \frac{f(z) - f(a)}{z - a} & \text{si } z \neq a \\ f'(a) & \text{si } z = a, \end{cases} \quad (7.25)$$

elle est définie et continue pour tout $z \in \Omega$, et holomorphe pour $z \neq a$. Un théorème de Riemann (voir plus bas Théorème 7.23) assure alors qu'elle est holomorphe sur tout Ω . On a donc par le corollaire du théorème de Cauchy :

$$0 = \int_{\gamma} g(z)dz = \int_{\gamma} \frac{f(z)dz}{z-a} - \int_{\gamma} \frac{f(a)dz}{z-a} = \int_{\gamma} \frac{f(z)dz}{z-a} - 2\pi i \text{Ind}_{\gamma}(a) f(a)$$

ce qui établit (7.24).

Nous verrons au chapitre suivant des applications concrètes de cette formule au calcul d'intégrales. Dans le reste de ce chapitre nous en discutons quelques conséquences fondamentales.

Théorème 7.18 (théorème de la moyenne) : Si f est holomorphe sur Ω , pour tout $a \in \Omega$ et $r > 0$ tel que le disque fermé de centre a et de rayon r soit contenu dans Ω , on a

$$f(a) = \int_{-\pi}^{\pi} f(a + re^{i\theta}) \frac{d\theta}{2\pi}. \quad (7.26)$$

Autrement dit la valeur de f en tout point a est égale à sa moyenne sur un cercle centré en a . La preuve découle simplement de l'intégrale de Cauchy : si γ est le bord du disque orienté dans le sens positif

$$f(a) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(z) dz}{z - a} = \frac{1}{2\pi i} \int_{\gamma} \frac{f(a + re^{i\theta}) ire^{i\theta} d\theta}{re^{i\theta}} = \int_{-\pi}^{\pi} f(a + re^{i\theta}) \frac{d\theta}{2\pi}.$$

7.4 Autres propriétés des fonctions holomorphes

7.4.1 Fonctions holomorphes et fonctions analytiques

Théorème 7.19 : Une fonction est holomorphe dans un ouvert Ω ssi elle est analytique sur Ω .

Résultat remarquable : deux conditions sur les fonctions de variable complexe –dérivabilité et analyticité– ont finalement abouti à la même classe de fonctions !!

Preuve : Soit f une fonction holomorphe dans Ω . Soit $z_0 \in \Omega$, il existe un disque ouvert de centre z_0 et de rayon R contenu dans Ω . Pour $r < R$, soit $z : |z - z_0| < r$ et prenons pour γ un cercle de centre z_0 et de rayon r' , $r < r' < R$. Si $\zeta \in \gamma$, on a $\left| \frac{z-z_0}{\zeta-z_0} \right| = \left| \frac{z-z_0}{r'} \right| < 1$, donc la série géométrique $\sum_{n=0}^{\infty} \frac{(z-z_0)^n}{(\zeta-z_0)^{n+1}}$ converge et a pour somme $(\zeta - z)^{-1}$. Selon (7.24),

$$f(z) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(\zeta)}{\zeta - z} d\zeta = \frac{1}{2\pi i} \int_{\gamma} \sum_{n=0}^{\infty} \frac{(z-z_0)^n}{(\zeta-z_0)^{n+1}} f(\zeta) d\zeta$$

et la convergence uniforme (car normale) de la série pour $|z| \leq r$ et $|\zeta - z_0| = r'$ fait qu'on peut l'intégrer terme à terme (c'est-à-dire permute intégration et sommation)

$$f(z) = \sum_{n=0}^{\infty} (z - z_0)^n \frac{1}{2\pi i} \int_{\gamma} \frac{f(\zeta)}{(\zeta - z_0)^{n+1}} d\zeta = \sum_{n=0}^{\infty} a_n (z - z_0)^n \quad (7.27)$$

ce qui établit bien que f est développable en série entière en z_0 , avec

$$a_n = \frac{1}{2\pi i} \int_{\gamma} \frac{f(\zeta)}{(\zeta - z_0)^{n+1}} d\zeta.$$

Cela étant vrai en tout point $z_0 \in \Omega$, f est bien analytique dans Ω . Et comme (7.27) converge pour tout $r < R$, le rayon de convergence de la série est au moins égal à R . Selon la formule (6.6), on a de plus

$$\frac{1}{n!} f^{(n)}(z_0) = a_n = \frac{1}{2\pi i} \int_{\gamma} \frac{f(\zeta)}{(\zeta - z_0)^{n+1}} d\zeta. \quad (7.28)$$

Réiproquement on veut montrer que si f est analytique dans Ω , elle est indéfiniment dérivable en tout point de Ω , donc holomorphe. Cela a été mentionné au Chap. 6 pour des fonctions analytiques de variable réelle, mais la démonstration doit être reprise pour établir l'holomorphie. Considérons le développement en série de f au point z_0 qu'on prendra égal à 0 sans perte de généralité : $f(z) = \sum_n a_n z^n$ qui converge dans un disque ouvert de rayon R . On a vu au § 6.1.3 que la série dérivée $\sum_n n a_n z^{n-1}$ a aussi R comme rayon de convergence. Pour z tel que $|z| < R$, soient $r : |z| < r < R$ et $h \in \mathbb{C} : 0 \neq h \leq r - |z|$. Formons

$$\frac{f(z+h) - f(z)}{h} - \sum_n n a_n z^{n-1} = \sum_{n \geq 1} u_n(z, h)$$

avec

$$u_n(z, h) = a_n ((z+h)^{n-1} + z(z+h)^{n-2} + \cdots + z^{n-1} - nz^{n-1})$$

Puisque $|z|, |z+h| \leq r$, on a $|u_n(z, h)| \leq 2n|a_n|r^{n-1}$ et comme $r < R$, cette série converge, donc $\forall \epsilon$, il existe n_0 tel que $\sum_{n>n_0} 2n|a_n|r^{n-1} \leq \frac{1}{2}\epsilon$. Quant à la somme finie $\sum_{n \leq n_0} u_n(z, h)$, elle s'annule en $h = 0$ donc peut être rendue $\leq \frac{1}{2}\epsilon$ en prenant $|h| < \eta$. Avec ces choix de n_0 et de η , on a donc rendu $\frac{f(z+h)-f(z)}{h} - \sum_n n a_n z^{n-1}$ arbitrairement petite, et donc établi que $f'(z) = \lim_{h \rightarrow 0} \frac{f(z+h)-f(z)}{h}$, c'est-à-dire l'holomorphie de f .

Complément : Principe de symétrie de Schwarz. Soit Ω un ouvert connexe de \mathbb{C} symétrique par rapport à l'axe réel. Soit Ω' , resp. Ω'' , l'intersection de Ω avec le demi-plan supérieur $y \geq 0$, resp. inférieur $y \leq 0$. Soit une fonction f continue dans Ω' , holomorphe aux points de Ω' tels que $y > 0$ et à valeurs réelles sur l'axe réel. On va montrer que f peut être prolongée de façon unique en une fonction holomorphe sur Ω . Sur Ω'' on définit $g(z) = \overline{f(\bar{z})}$. Elle est continue dans Ω'' et on vérifie qu'elle est holomorphe en tout point non réel de Ω'' . La fonction $h(z)$ égale à $f(z)$ sur Ω' et à $g(z)$ sur Ω'' est continue dans Ω et holomorphe sur $\Omega \setminus \mathbb{R}$. Une réciproque (théorème de Morera) du théorème de Cauchy affirme alors qu'elle est aussi holomorphe en tout point de Ω . Elle prend des valeurs complexes conjuguées en des points symétriques par rapport à l'axe réel. Elle est unique (prolongement analytique).

7.4.2 Fonctions entières. Principe du maximum

Définition 7.5 : On appelle fonction entière une fonction holomorphe dans le plan \mathbb{C} tout entier.

Exemple : la fonction e^z qui est analytique avec un rayon de convergence infini est holomorphe dans tout \mathbb{C} donc entière.

Théorème 7.20 (Liouville) : Une fonction entière bornée est constante.

Autrement dit, une fonction entière non constante doit tendre vers l'infini dans certaines directions du plan complexe. Voir l'exemple de $\exp z$ qui tend vers l'infini quand $|z| \rightarrow \infty$ avec $\Re z > 0$. Inversement ce théorème de Liouville est utile pour établir des identités entre fonctions analytiques : si on peut montrer qu'une combinaison de fonctions analytiques n'a pas de singularité et demeure bornée, c'est une constante. (Exemple, voir TD.)

Preuve. Soit f entière et bornée, $|f(z)| < M$. Calculons $f'(z)$ par Cauchy, ou plutôt par (7.28) : $f'(z) = \frac{1}{2\pi i} \oint_{\gamma} \frac{f(\zeta)}{(\zeta - z)^2} d\zeta$ le long d'un cercle de rayon R autour de z . Donc $|f'(z)| \leq \frac{1}{2\pi} \int_0^{2\pi} \frac{|f(\zeta)|}{|\zeta - z|^2} R d\theta \leq \frac{M}{R}$. Comme R peut être pris arbitrairement grand, $|f'(z)| = 0$ et f est une constante.

Application : théorème de d'Alembert. Montrons que le théorème fondamental de l'algèbre, *tout polynôme P à coefficients complexes et non constant possède au moins une racine complexe*, découle de ce théorème de Liouville. Raisonnons par l'absurde et supposons que P ne s'annule pas. Alors $1/P(z)$ serait holomorphe dans \mathbb{C} et borné. En effet $P(z) = z^n(a_n + \frac{a_{n-1}}{z} + \cdots + \frac{a_0}{z^n}) \rightarrow \infty$ quand $|z| \rightarrow \infty$, donc il existe un disque compact à l'extérieur duquel $1/P(z)$ est borné (puisque $|P| \rightarrow \infty$ à l'infini) et à l'intérieur duquel il est également borné en tant que fonction continue. Donc (Liouville) P serait constant ce qui est contraire à l'hypothèse. q.e.d.

Théorème 7.21 (Principe du maximum) : Soit f continue dans Ω et satisfaisant la propriété de moyenne, (par exemple f holomorphe dans Ω). Si $|f|$ a un maximum local en un point $a \in \Omega$, c'est-à-dire $|f(z)| \leq |f(a)|$ au voisinage de a , alors f est constante dans ce voisinage.

Esquisse de preuve. On peut toujours supposer $f(a)$ réel > 0 , quitte à multiplier la fonction par $\exp -i \arg(f(a))$. Pour $r \geq 0$ assez petit, $M(r) \stackrel{\text{def}}{=} \sup_{\theta} |f(a + re^{i\theta})| \leq f(a)$ par l'hypothèse de maximum. Mais la propriété de moyenne donne $f(a) = \frac{1}{2\pi} \int_0^{2\pi} f(a + re^{i\theta}) d\theta = [\dots] \leq M(r)$. Donc $f(a) = M(r)$. La fonction $g(z) \stackrel{\text{def}}{=} \Re(f(a) - f(z))$ est donc ≥ 0 pour $|z - a| = r$ assez petit [...] ; de plus [...] $g(z) = 0$ ssi $f(z) = f(a)$. Mais $g(z)$, qui est ≥ 0 et continue, ayant une valeur moyenne nulle sur le cercle de centre a et de rayon r est identiquement nulle sur ce cercle. Donc $f(z) = f(a)$ q.e.d. (Exercice : compléter les [...] de cette preuve.)

Noter que ce théorème n'interdit pas un *minimum* local : ainsi la fonction $z \mapsto z$ a un minimum de son module en $z = 0$.

Ce théorème du maximum s'applique aussi aux fonctions harmoniques, de grand intérêt par exemple en électrostatique (potentiel), voir Chap. 8.

7.5 Zéros et singularités des fonctions de variable complexe

7.5.1 Zéros d'une fonction holomorphe

Soit f holomorphe dans un ouvert connexe Ω , et soit $Z(f) \stackrel{\text{def}}{=} \{a \in \Omega : f(a) = 0\}$ l'ensemble des zéros de f dans Ω . Cet ensemble peut être vide, cf la fonction exponentielle.

Théorème 7.22 : (i) Les zéros d'une fonction holomorphe non nulle sont isolés : tout zéro a possède un voisinage sans autre zéro. Autrement dit, si $Z(f)$ possède un point d'accumulation dans Ω , $f = 0$ sur Ω .

(ii) Si $f \neq 0$, $Z(f)$ est fini ou dénombrable. En chaque zéro $a \in Z(f)$, on peut écrire pour tout $\forall z \in \Omega$,

$$f(z) = g(z) (z - a)^{m(a)} \quad (7.29)$$

où g est holomorphe dans Ω et ne s'annule pas au voisinage de a , et l'entier $m(a) \geq 1$ est l'ordre du zéro a .

Exemple, la fonction $f(z) = \sin z$ a des zéros isolés en $z_k = k\pi$. Contre-exemple : la fonction $f(z) = 1 + \exp \frac{1}{z}$ est holomorphe dans $\Omega = \mathbb{C} \setminus \{0\}$ et a des zéros en $z_k = -i((2n+1)\pi)^{-1}$ qui s'accumulent en 0, mais $0 \notin \Omega$.

Éléments de preuve. Soit a un zéro de f analytique dans Ω . Nous admettrons que si toutes les dérivées de f en a s'annulent, alors $f \equiv 0$. Si $f \neq 0$, soit $m \geq 1$ le plus petit entier k tel que $f^{(k)}(a) \neq 0$ et donc dans le développement en série de f en a : $f(z) = \sum_{k \geq m} c_k (z - a)^k$, $c_m = f^{(k)}(a)/m! \neq 0$. La fonction $g(z)$ définie par $\frac{f(z)}{z-a}$ est holomorphe dans Ω sauf peut-être en a , égale à c_m en a , donc selon le même théorème de Riemann déjà invoqué, (voir ci-dessous Th. 7.23), est holomorphe dans tout Ω , ce qui prouve (ii). De plus, $g(z) = a_m + \sum_{k > m} c_k (z - a)^{k-m}$ est non nul dans un voisinage de a , ce qui établit le caractère isolé des zéros de f . Et des points isolés dans $\Omega \subset \mathbb{C}$ forment nécessairement un ensemble fini ou dénombrable (vérifier!).

Application au problème de prolongement : si on connaît une fonction f , supposée holomorphe dans Ω connexe, sur un ensemble E non dénombrable, par exemple $\mathbb{R} \cap \Omega$, elle est (en principe) déterminée dans tout Ω .

Autrement dit son prolongement de E à Ω est unique. En effet si on avait deux fonctions f_1 et f_2 la prolongeant de E dans Ω , leur différence $g = f_1 - f_2$ serait holomorphe dans Ω et y posséderait un ensemble non dénombrable de zéros, donc par le Théorème précédent serait nulle.

Exemple : la fonction $f(z) = \sin^2 z + \cos^2 z - 1$ est entière ; elle s'annule sur \mathbb{R} , elle s'annule donc partout. L'identité $\sin^2 z + \cos^2 z = 1$ est donc vraie sur tout \mathbb{C} .

En revanche, si la fonction f n'est connue que sur les entiers de \mathbb{N} (ou de \mathbb{Z}), son prolongement à \mathbb{C} n'est pas unique, on peut par exemple la multiplier par $e^{2\pi iz}$.

7.5.2 Pôles, singularités essentielles

Supposons qu'une fonction f est holomorphe dans un ouvert privé d'un point a : on parle de *singularité isolée* en a . (On verra plus bas qu'une singularité peut ne pas être isolée : il n'existe pas de disque ouvert pointé $0 < |z - a| < r$ dans lequel la fonction est holomorphe ; c'est le cas des "points de branchement" des fonctions multi-valuées, voir § 8.2³.) Dans le cas d'une

3. C'est le cas aussi pour une fonction telle que $(1 + \exp \frac{1}{z})^{-1}$ dont les pôles en $z_k = -i((2n+1)\pi)^{-1}$ s'accumulent en 0.

singularité isolée, des théorèmes dus à Riemann, Weierstrass, Picard, ... permettent d'affirmer que trois cas sont possibles.

Théorème 7.23 : Soit f holomorphe dans $\Omega \setminus \{a\}$:

1. ou bien f reste bornée au voisinage de a ; alors elle admet une limite en a et on peut la prolonger en une fonction holomorphe sur Ω : on parle de singularité apparente⁴ ;
2. ou bien $\lim_{z \rightarrow a} |f(z)|$ existe (dans $\bar{\mathbb{R}}$) et vaut $+\infty$: on parle de pôle en a ;
3. ou bien, si $\lim_{z \rightarrow a} |f(z)|$ n'existe pas dans $\bar{\mathbb{R}}$, l'image par f de tout disque ouvert pointé $0 < |z - a| < r$ contenu dans Ω est dense dans \mathbb{C} (Weierstrass). En fait (Picard) cette image est le plan \mathbb{C} tout entier ou le plan \mathbb{C} privé d'un point. On parle de singularité essentielle.

Nous nous contenterons d'indications et d'exemples. Un exemple de singularité apparente est fourni par la fonction $f(z) \stackrel{\text{déf}}{=} \frac{g(z)-g(a)}{z-a}$ si $z \neq a$, avec g holomorphe dans Ω . Il est clair que la fonction prolongée par $f(a) = g'(a)$ est holomorphe dans tout Ω , cf § 7.3.2.

Un exemple de singularité essentielle est donné par la fonction $f(z) \stackrel{\text{déf}}{=} \exp \frac{1}{z}$, qui est holomorphe dans $\mathbb{C} \setminus \{0\}$. Exercice : montrer que toute valeur $w \neq 0$ est atteinte (une infinité de fois) par $f(z)$.

On rencontre cette situation en physique dans des problèmes aussi variés que les singularités en température de grandeurs thermodynamiques au voisinage de certaines transitions de phase (modèle XY de rotateurs à deux dimensions, énergie de marche dans la “transition rugueuse”, ... ; ou que les propriétés d’analyticité de certains développements “perturbatifs” en mécanique quantique ou théorie quantique des champs.

Définition 7.6 : Une fonction définie dans Ω qui n'y a que des pôles comme singularités est dite méromorphe sur Ω .

Le cas des pôles des fonctions méromorphes est celui qui va le plus nous occuper dans la suite. Pour ne donner qu'un exemple, toute fraction rationnelle $P(z)/Q(z)$ est une fonction méromorphe, ses pôles sont les zéros (racines) du dénominateur.

7.5.3 Séries de Laurent. Fonctions méromorphes. Résidus

Définition 7.7 : On appelle série de Laurent une expression de la forme

$$\sum_{-\infty}^{\infty} a_n z^n \tag{7.30}$$

4. ou *artificielle*, ou *effaçable*, ...

Quand on discute la convergence d'une telle expression, il faut bien comprendre qu'elle doit être considérée comme la somme de deux séries, l'une portant sur les indices ≥ 0 , l'autre sur les négatifs, et que la convergence de (7.30) signifie la convergence de chacune de ces deux sous-séries. Supposons que $\sum_{n \geq 0} a_n z^n$ a pour rayon de convergence R_1 et que $\sum_{n < 0} a_n \zeta^{-n}$ a pour rayon $1/R_2$ en ζ , donc que $\sum_{n < 0} a_n z^n$ converge pour $|z| > R_2$. Supposons aussi que $R_2 < R_1$. Alors (7.30) converge (uniformément et absolument) dans la couronne $R_2 < |z| < R_1$. Noter que l'on peut avoir $R_2 = 0$ et/ou $R_1 = \infty$.

Théorème 7.24 : Toute fonction f holomorphe dans la couronne $R_2 < |z| < R_1$ y est développable en série de Laurent

$$f(z) = \sum_{-\infty}^{\infty} a_n z^n, \quad (7.31)$$

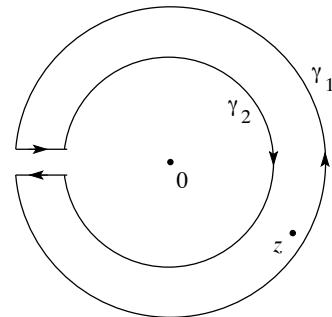
avec des coefficients

$$a_n = \frac{1}{2\pi i} \oint_0 \frac{f(\zeta)}{\zeta^{n+1}} d\zeta, \quad (7.32)$$

le long d'un contour entourant (une fois dans le sens positif) l'origine mais contenu dans la couronne.

Nous avons centré la couronne en 0, mais rien n'interdit de la centrer en un autre point z_0 , avec un développement de Laurent en puissances de $(z - z_0)$, $f(z) = \sum_{-\infty}^{\infty} a_n (z - z_0)^n$.

La preuve de ce théorème est instructive. On considère le contour d'intégration γ représenté sur la figure, complètement contenu dans la couronne d'holomorphie, donc avec des rayons intérieur et extérieur respectivement $r_1 < R_1$ et $r_2 > R_2$, et parcouru dans le sens positif. Les deux petits segments parallèles, choisis dans une direction différente de celle de z , sont distants de ϵ , que l'on fait tendre vers zéro. Dans cette limite, leurs contributions se compensent et seules demeurent celles des deux cercles γ_1 et γ_2 de rayon r_1 et r_2 .



La formule de Cauchy nous donne (dans la limite où $\epsilon \rightarrow 0$)

$$2\pi i f(z) = \int_{\gamma} \frac{f(\zeta)}{\zeta - z} d\zeta = \int_{\gamma_1} \frac{f(\zeta)}{\zeta - z} d\zeta - \int_{\gamma_2} \frac{f(\zeta)}{\zeta - z} d\zeta$$

Pour $\zeta \in \gamma_1$, $|\zeta/z| > 1$ et on développe $(\zeta - z)^{-1} = \sum_{n=0}^{\infty} \frac{z^n}{\zeta^{n+1}}$, et sur γ_2 de même, $(\zeta - z)^{-1} = -z^{-1}(1 - \zeta/z)^{-1} = -\sum_{n=-1}^{-\infty} \frac{z^n}{\zeta^{n+1}}$, ce qui établit le théorème, y compris l'expression de a_n comme intégrale sur le contour $\gamma_1 \sim \gamma_2$ ou tout autre contour homotope contenu dans la couronne.

Les différents cas du Théorème 7.23 se lisent maintenant sur le développement de Laurent au voisinage de z_0 , c'est-à-dire dans une "couronne" $0 < |z - z_0| < r$:

- si le développement de Laurent en z_0 n'a que des termes d'indice ≥ 0 , z_0 n'est pas une

singularité, f est holomorphe en z_0 ;

- s'il n'a qu'un nombre fini de termes d'indice < 0 , et que le plus petit de ces indices vaut m , f a un pôle d'ordre m en z_0 ; si $m = 1$, on parle de *pôle simple*;
- enfin s'il a un nombre infini de termes d'indice < 0 , f a une singularité essentielle en z_0 .

2 exemples de développement de Laurent

1. Fraction rationnelle $f(z) = \frac{1}{(z-a)(z-b)}$ avec par exemple $|a| < |b|$, dont on cherche le développement de Laurent dans la couronne $|a| < |z| < |b|$. On peut soit utiliser directement (7.32), nous y reviendrons plus bas (Exercice 3), soit “décomposer en éléments simples”

$$f(z) = \frac{1}{(z-a)(z-b)} = \frac{1}{a-b} \left(\frac{1}{z-a} - \frac{1}{z-b} \right)$$

et se ramener au développement de Laurent d'un pôle simple $\frac{1}{z-b} = -\sum_{n \geq 0} \frac{z^n}{b^{n+1}}$ pour $|z| < |b|$, ou $\frac{1}{z-a} = \sum_{n \geq 0} \frac{a^n}{z^{n+1}} = \sum_{n < 0} \frac{z^n}{a^{n+1}}$ pour $|z| > |a|$. Au total $f(z) = \sum_{n \in \mathbb{Z}} c_n z^n$ avec

$$c_n = \begin{cases} \frac{1}{b^{n+1}(a-b)} & \text{si } n \geq 0 \\ \frac{1}{a^{n+1}(a-b)} & \text{si } n < 0 \end{cases}. \quad (7.33)$$

Bien entendu, si $|a| \neq 0$, cette même fraction rationnelle $f(z)$ admet aussi un développement en série entière en $z = 0$, valable pour $|z| < |a| < |b|$.

2. Fonction génératrice des fonctions de Bessel $J_n : f(z) = e^{\frac{t}{2}(z-\frac{1}{z})}$, $t \in \mathbb{C}$, admet un développement de Laurent dans le plan pointé \mathbb{C}^* (c'est-à-dire $R_2 = 0, R_1 = \infty$)

$$e^{\frac{t}{2}(z-\frac{1}{z})} = \sum_{n \in \mathbb{Z}} J_n(t) z^n \quad (7.34)$$

avec des coefficients $J_n(t)$ fonctions de t appelés fonctions de Bessel. Selon (7.32),

$$J_n(t) = \frac{1}{2\pi i} \oint_0 \frac{e^{\frac{t}{2}(\zeta-\frac{1}{\zeta})}}{\zeta^{n+1}} d\zeta$$

sur un contour arbitraire entourant (une fois) l'origine, par exemple le cercle unité paramétrisé par $\zeta = e^{i\theta}$, d'où

$$J_n(t) = \frac{1}{2\pi} \int_0^{2\pi} e^{-in\theta} e^{it\sin\theta} d\theta \quad n \in \mathbb{Z}, \quad (7.35)$$

qui est souvent pris comme définition des fonctions de Bessel d'indice entier J_n . Prenant $z = e^{i\alpha}$ dans (7.34) on voit que l'on a obtenu le développement de Fourier de

$$e^{it\sin\alpha} = \sum_{n \in \mathbb{Z}} J_n(t) e^{in\alpha}.$$

Résidus

Soit f une fonction holomorphe dans une couronne $R_2 < |z| < R_1$ centrée à l'origine.

Définition 7.8 : On appelle résidu de f à l'origine le coefficient a_{-1} de (7.31), soit

$$\text{Res}(f, 0) \stackrel{\text{def}}{=} a_{-1} = \frac{1}{2\pi i} \oint_0 f(\zeta) d\zeta. \quad (7.36)$$

Plus généralement le résidu en une (éventuelle) singularité z_0 est $\text{Res}(f, z_0) = \frac{1}{2\pi i} \oint_{z_0} d\zeta f(\zeta)$ avec un contour entourant le point z_0 une seule fois et dans le sens positif.

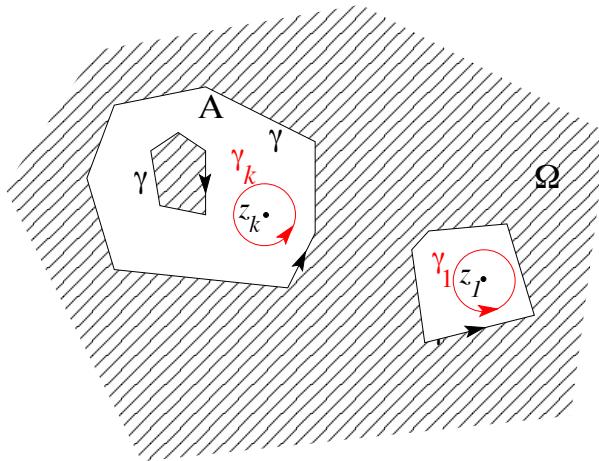


FIGURE 7.4 – La partie hachurée est le complémentaire de A dans Ω . L'intégrale sur le bord orienté γ de A est la somme des intégrales sur les γ_i entourant des singularités.

Théorème 7.25 (théorème des résidus) : Soit f une fonction holomorphe sur Ω , sauf peut-être en des singularités isolées z_k . Soit γ le bord orienté d'un compact A contenu dans Ω , ne passant par aucun des z_k . Alors les z_k contenus dans A sont en nombre fini et

$$\int_{\gamma} f(z) dz = 2\pi i \sum_{\substack{k \\ z_k \in A}} \text{Res}(f, z_k). \quad (7.37)$$

La preuve repose une fois encore sur la déformation du contour γ : avec les hypothèses faites, l'intégrale sur γ est une somme d'intégrales sur des contours $\gamma_1, \dots, \gamma_k$ encerclant les points z_1, \dots, z_k , voir figure 7.4. Chacun de ces intégrales donne lieu au résidu correspondant.

Ce théorème est très utile pour calculer des intégrales comme sommes de résidus, ou inversement des sommes, considérées comme sommes de résidus, comme des intégrales. Cela va être amplement illustré au Chap. 8 et en TD.

Sphère de Riemann. Résidu à l'infini

Il est souvent utile de considérer le plan complexe complété par le point à l'infini : comme la limite $|z| \rightarrow \infty$ ne doit pas dépendre de $\arg z$, ce point est unique ! Cela revient à “compactifier” le plan $\mathbb{C} \simeq \mathbb{R}^2$ en une sphère, la *sphère de Riemann* $\overline{\mathbb{C}} \simeq S^2$.

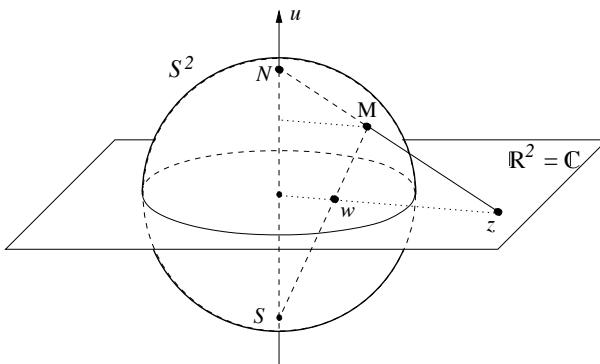


FIGURE 7.5 – Projections stéréographiques d'un point M depuis les pôles Nord N et Sud S .

Ceci peut être vu très explicitement par la projection stéréographique, voir Fig 7.5. Un point M de la sphère unité S^2 de coordonnées (x, y, u) , $x^2 + y^2 + u^2 = 1$, est projeté depuis le pôle Nord N en un point du plan (complexe) d'affixe z . Vérifier que $z = \frac{x+iy}{1-u}$. Le pôle Nord a pour image le point à l'infini dans le plan. La projection depuis le pôle Sud S donne de même $w = \frac{x-iy}{1+u}$, et on a pour un même point M la relation $z.w = 1$. Le pôle Nord a cette fois pour image 0, tandis que S est appliquée à l'infini. Dans le langage de la géométrie différentielle, on dit que l'on a besoin de deux *cartes* $M \mapsto z$ et $M \mapsto w$ pour décrire la sphère. Quand on s'intéresse au voisinage du point z infini, on utilise la coordonnée $w = 1/z$.

On définit sur cette sphère $\overline{\mathbb{C}} = S^2$ les notions d'ouvert, de chemin différentiable, de chemin fermé, de bord orienté d'un compact, etc : à distance finie, ce sont les notions déjà rencontrées, et au voisinage de l'infini, on utilise la variable $w = 1/z$.

On dit alors qu'une fonction $f(z)$ est holomorphe (resp. est méromorphe, a une singularité essentielle) à l'infini si $g(w) \stackrel{\text{déf}}{=} -\frac{1}{w^2} f(z=1/w)$ est holomorphe (resp. est méromorphe, a une singularité essentielle) au voisinage de $w=0$. Pour une fonction $f(z)$ méromorphe à l'infini, le résidu à l'infini se définit par

$$\operatorname{Res}(f, \infty) = -\frac{1}{2\pi i} \oint_0 \frac{1}{w^2} f\left(\frac{1}{w}\right) dw = \operatorname{Res}(g(w), 0) \quad (7.38)$$

avec un contour autour de l'origine, ou encore, si $\sum_n a_n z^n$ est le développement de Laurent de $f(z)$ au voisinage de l'infini, $\operatorname{Res}(f, \infty) = -a_{-1}$. Attention au signe !! Ce signe, qui provient du changement de variable dans $f(z)dz = g(w)dw$, peut aussi être vu comme lié au changement d'orientation d'un contour positif dans le plan complexe quand il est repoussé autour du point à l'infini : il entoure alors ce point à l'infini dans le sens *négatif* (faites l'expérience avec un élastique sur une orange!).

Par exemple, la fonction $f(z) = \frac{1}{z}$ a un pôle simple à l'infini de résidu -1 . En effet $g(w) = -\frac{1}{w}$, donc (7.38) donne $\text{Res}(f, \infty) = \text{Res}(g(w), 0) = -1$.

Théorème des résidus généralisé

Le théorème des résidus admet une généralisation au cas où le contour est dessiné sur la sphère de Riemann. Il s'énonce essentiellement comme le théorème 7.25 :

Théorème 7.26 (théorème des résidus généralisé) : Soit Ω un ouvert de la sphère de Riemann et f une fonction holomorphe sur Ω , sauf peut-être en des singularités isolées z_k . Soit γ le bord orienté d'un compact A de S^2 contenu dans Ω , ne passant par aucun des z_k ni par le point à l'infini. Alors les z_k contenus dans A sont en nombre fini et

$$\int_{\gamma} f(z) dz = 2\pi i \sum_{\substack{k \\ z_k \in A}} \text{Res}(f, z_k) \quad (7.39)$$

où la somme court sur tous les points singuliers $z_k \in A$, y compris éventuellement le point à l'infini.

Preuve : Si $\infty \notin A$, on est dans le cadre du Théorème 7.25. Si $\infty \in A$, soit Γ un contour fermé homotope à un cercle, orienté positivement, contenu dans A , ne passant pas par ∞ et englobant le bord γ de A . On suppose que ∞ est à l'extérieur de ce Γ (au sens de \mathbb{C}) et que toutes les singularités autres que ∞ sont à l'intérieur de Γ , (par exemple, on peut prendre pour Γ un cercle $|z| = R$ suffisamment grand pour entourer toutes les singularités à distance finie, tout en étant dans A .) Voir Fig. 7.6. On peut alors appliquer le théorème des résidus au contour $\gamma \cup \Gamma$. L'intégrale $\oint_{\gamma \cup \Gamma} f dz = 2\pi i \sum_{\substack{z_k \in A \\ z_k \neq \infty}} \text{Res}(f, z_k)$, tandis que $\oint_{\Gamma} f dz = -2\pi i \text{Res}(f, \infty)$. On a donc $\oint_{\gamma} f dz = \sum_k 2\pi i \text{Res}(f, z_k)$, y compris l'éventuel z_k à l'infini, q.e.d.

Plus simplement, on peut aussi arguer qu'il existe toujours un changement de variable $z \mapsto w$ qui ramène à distance finie toutes les singularités contenues dans le compact A (en nombre fini!), voir Exercice 4. On se ramène donc au théorème des résidus ordinaire.

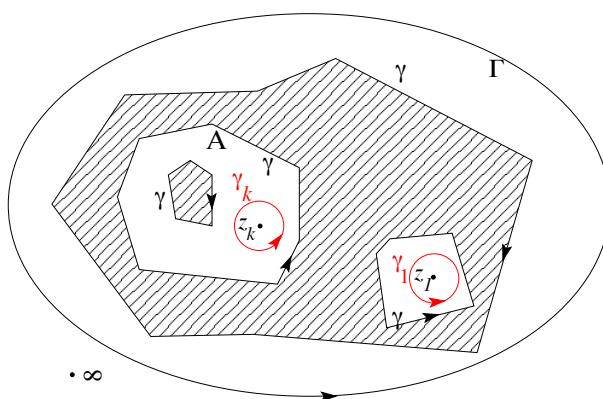


FIGURE 7.6 – La partie hachurée est le complémentaire de A .

Exemple : reprenons la fonction $f(z) = \frac{1}{z}$ et calculons $I = \oint_{\gamma} f(z) dz$ le long d'un contour positif n'entourant ni 0 ni ∞ . Le théorème des résidus ordinaire nous dit que $I = 0$ (pas de

pôle à l'intérieur de γ) ; le théorème des résidus généralisé après déformation du contour en deux contours entourant 0 et ∞ dans le même sens négatif nous dit $I = -2\pi i(\text{Res}(f, 0) + \text{Res}(f, \infty)) = -2\pi i(-1 + 1)$ d'après le calcul ci-dessus. On a bien $I = 0$. En général, par le même argument, montrer que la somme des résidus (y compris à l'infini) d'une fraction rationnelle s'annule.

Lectures complémentaires

Dans ce chapitre, je me suis largement inspiré de [4] et de [1], deux excellentes références qui l'une et l'autre contiennent beaucoup d'informations supplémentaires.

Exercices

1. Justifier par un argument de déformation de contour le calcul de la transformation de Fourier de la fonction gaussienne effectué aux chap.4 et 5.
2. f holomorphe dans Ω , soit $F(z) \stackrel{\text{déf}}{=} \frac{1}{f(z)}$. Montrer que F est méromorphe dans Ω .
3. Armé(e) du théorème des résidus, reprendre le calcul de (7.33) par (7.32). Montrer que selon que $n < 0$ ou $n \geq 0$, on peut refermer le contour soit autour a soit autour de b (et l'infini?).
4. Soit f une fonction holomorphe dans un ouvert Ω avec un nombre fini de singularités en z_k (y compris peut-être à l'infini). Montrer qu'il existe un changement de variable $z = \varphi(w)$ qui applique tous les z_k sur des w_k à distance finie. Soit $\tilde{f} \stackrel{\text{déf}}{=} f \circ \varphi$; comparer les résidus $\text{Res}(f, z_k)$ et $\text{Res}(\tilde{f} \cdot \varphi', w_k)$ et les théorèmes des résidus pour les fonctions f et $\tilde{f} \varphi'$. Montrer que le théorème des résidus généralisé découle alors du théorème usuel.

Appendix D. Système différentiel

Soit une fonction $f : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ (ou \mathbb{C}) satisfaisant le système de deux équations aux dérivées partielles

$$\frac{\partial f(x, y)}{\partial x} = P(x, y) \quad \frac{\partial f(x, y)}{\partial y} = Q(x, y), \quad (\text{D.1})$$

où P et Q sont des fonctions données de classe C^1 . Montrons qu'une condition nécessaire et suffisante pour que le système admette (au moins) une solution locale est que

$$\frac{\partial P(x, y)}{\partial y} = \frac{\partial Q(x, y)}{\partial x}, \quad (\text{D.2})$$

la condition d'intégrabilité du système.

Il est clair que la condition est nécessaire, puisque $\frac{\partial^2 f(x,y)}{\partial x \partial y} = \frac{\partial^2 f(x,y)}{\partial y \partial x}$. Montrons qu'elle est suffisante. Intégrons la première équation (D.1) en $f(x,y) = \int^x P(x',y)dx' + \varphi(y)$ où φ est une fonction de y seulement, de classe C^1 . Avec les hypothèses, on peut dériver sous le signe somme et la deuxième équation (D.1) donne $\frac{\partial f(x,y)}{\partial y} = \int^x \frac{\partial P(x',y)}{\partial y} dx' + \varphi'(y) = Q(x,y)$, qui est bien cohérente car $\psi(x,y) := Q(x,y) - \int^x \frac{\partial P(x',y)}{\partial y} dx'$ ne dépend pas de x : $\frac{\partial \psi(x,y)}{\partial x} = \frac{\partial Q(x,y)}{\partial x} - \frac{\partial P(x,y)}{\partial Y} = 0$ par (D.2). On peut donc déterminer $\varphi(y) = \int^y \left(Q(x,y') - \int^x \frac{\partial P(x',y')}{\partial y} dx' \right) dy' + \text{const.}$ et finalement

$$f(x,y) = \int^x P(x',y)dx' + \int^y \left(Q(x,y') - \int^x \frac{\partial P(x',y')}{\partial y} dx' \right) dy' + \text{const.} \quad (\text{D.3})$$

qui satisfait bien (D.1).

Remarque. Cette construction est une construction *locale*, puisqu'implicitement, on a intégré P ou ψ le long de segments à y ou x constants à partir d'un point de base donné, et que tout point de Ω ne peut pas toujours être atteint ainsi à partir du même point de base. La construction est globale, en revanche, si Ω est un disque ouvert ou un "pavé" $]a, b[\times]c, d[$, pourquoi ?

On a retrouvé le résultat de la Proposition 7.10 : la forme $\omega = Pdx + Qdy$ est fermée ssi (D.2) ; de plus elle est exacte si le domaine est un disque ou un pavé ouvert.

Chapitre 8

Fonctions de variables complexes, applications

8.1 Calcul d'intégrales, de transformées de Fourier, de sommes etc

8.1.1 Calcul pratique des résidus

Par définition, le résidu $\text{Res}(f, z_0)$ d'une fonction en z_0 est le coefficient a_{-1} de son développement de Laurent au voisinage de z_0 . Si f a un pôle simple en z_0

$$\text{Res}(f, z_0) = \lim_{z \rightarrow z_0} (z - z_0)f(z)$$

et cette formule se généralise au cas où le pôle est d'ordre m , comme on s'en convainc aisément

$$\text{Res}(f, z_0) = \lim_{z \rightarrow z_0} \frac{1}{(m-1)!} \frac{d^{m-1}}{dz^{m-1}} ((z - z_0)^m f(z)). \quad (8.1)$$

Si f est une fraction rationnelle $P(z)/Q(z)$, avec un pôle simple en z_0 , zéro simple de Q , il peut être plus aisés de faire appel à la règle de L'Hospital

$$\text{Res}\left(\frac{P}{Q}, z_0\right) = \frac{P(z_0)}{Q'(z_0)}.$$

8.1.2 Lemmes de Jordan

Introduisons des notations commodes :

soit $\mathcal{S}(\theta_1, \theta_2)$ le secteur du plan complexe $\{z = re^{i\theta}, r > 0, 0 \leq \theta_1 \leq \theta \leq \theta_2\}$; et $\gamma(r; \theta_1, \theta_2)$ l'arc de cercle $|z| = r$ dans $\mathcal{S}(\theta_1, \theta_2)$.

Lemmes de Jordan : Soit f une fonction continue $f : \mathcal{S}(\theta_1, \theta_2) \rightarrow \mathbb{C}$.

1. Si $zf(z)$ tend uniformément vers 0 sur l'arc $\gamma(r; \theta_1, \theta_2)$, $\theta_2 \leq 2\pi$, quand $r \rightarrow 0$, resp. $r \rightarrow \infty$, alors

$$\lim_{\substack{r \rightarrow 0^+ \\ \text{resp. } r \rightarrow +\infty}} \int_{\gamma(r; \theta_1, \theta_2)} f(z) dz = 0; \quad (8.2)$$

2. Si $f(z)$ tend uniformément vers 0 sur l'arc $\gamma(r; \theta_1, \theta_2)$, $\theta_2 \leq \pi$, quand $r \rightarrow \infty$, alors

$$\lim_{r \rightarrow \infty} \int_{\gamma(r; \theta_1, \theta_2)} e^{iz} f(z) dz = 0; \quad (8.3)$$

Preuve du 2ème lemme : on prend le secteur $\mathcal{S}(0, \pi)$ (qui est le cas le plus défavorable). Comme $|e^{iz}| = |e^{ire^{i\theta}}| = e^{-r \sin \theta}$, avec $\sin \theta \geq 0$ avec l'hypothèse sur le secteur, on peut majorer

$$\left| \int_{\gamma(r; 0, \pi)} e^{iz} f(z) dz \right| \leq \int_0^\pi |f(re^{i\theta})| re^{-r \sin \theta} d\theta \leq \int_0^{\pi/2} (|f(re^{i\theta})| + |f((re^{i(\pi-\theta)})|) re^{-r \sin \theta} d\theta.$$

Comme $f \rightarrow 0$ uniformément à l'infini, $\forall \epsilon > 0$, $\exists R$ tel que $\forall r > R \ \forall \theta \in [0, \pi]$, $|f(re^{i\theta})| \leq \epsilon$. Donc pour $r > R$

$$\left| \int_{\gamma(r; 0, \pi)} e^{iz} f(z) dz \right| \leq 2\epsilon \int_0^{\pi/2} re^{-r \sin \theta} d\theta$$

Mais si $0 \leq \theta \leq \pi/2$, $0 \leq 2\theta/\pi \leq \sin \theta$ comme on s'en convainc immédiatement. Donc

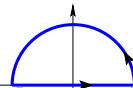
$$\int_0^{\pi/2} re^{-r \sin \theta} d\theta \leq \int_0^{\pi/2} re^{-2r\theta/\pi} d\theta = \frac{\pi}{2}(1 - e^{-r}) < \frac{\pi}{2},$$

d'où il découle que pour tout $r \geq R$, $\left| \int_{\gamma(r; 0, \pi)} e^{iz} f(z) dz \right| \leq \pi\epsilon$, arbitrairement petit, q.e.d.
La preuve du premier lemme, plus facile, est laissée en exercice. Voir aussi TD.

8.1.3 Intégrales sur l'axe réel

Considérons l'intégrale sur l'axe réel d'une fraction rationnelle $I = \int_{-\infty}^{\infty} \frac{P(x)}{Q(x)} dx$. On suppose que Q n'a pas de racine réelle. L'intégrale converge à l'infini si $\deg Q \geq \deg P + 2$, ce que nous supposerons. Procédons selon les étapes suivantes :

- On considère l'intégrale $\int_{-R}^R \frac{P(x)}{Q(x)} dx$ qui converge vers I quand $R \rightarrow \infty$;



- on la remplace par l'intégrale sur le contour $[-R, R] \cup \text{arc}$, où la contribution du cercle tend vers zéro quand $R \rightarrow \infty$, par le lemme de Jordan 1 ;
- le théorème des résidus nous donne alors

$$I = 2\pi i \sum_{\substack{\text{zéros } z_j \text{ de } Q \text{ dans} \\ \text{demi-plan supérieur}}} \text{Res}(P/Q, z_j)$$

On aurait pu aussi refermer le contour dans le demi-plan inférieur, mais attention au signe ! (le contour est alors orienté négativement). Donc aussi bien : $I = -2\pi i \sum_{\substack{\text{zéros } z_j \text{ de } Q \text{ dans} \\ \text{demi-plan inférieur}}} \text{Res}(P/Q, z_j)$. Question : que vaut dans ce cas le résidu à l'infini de P/Q ?

Nous rencontrerons beaucoup d'autres exemples de ce genre de calculs dans la suite.

8.1.4 Intégrales de fractions rationnelles trigonométriques

Considérons l'intégrale

$$I = \int_0^{2\pi} \frac{dt}{a + \sin t}$$

avec a réel > 1 . On écrit $\sin t = (e^{it} - e^{-it})/2i$ et on récrit cette intégrale en termes de $z = e^{it}$ sur le cercle unité. Il vient

$$I = \oint_0 \frac{2dz}{z^2 + 2iaz - 1} = \oint_0 \frac{2dz}{(z - z_+)(z - z_-)} = 2\pi i \text{Res}\left(\frac{2}{(z - z_+)(z - z_-)}, z_+\right) = 2\pi i \frac{2}{z_+ - z_-}$$

où $z_\pm = -ia \pm i\sqrt{a^2 - 1}$ sont les deux pôles, z_+ étant le seul à l'intérieur du cercle. Finalement $I = \frac{2\pi}{\sqrt{a^2 - 1}}$ (qui est positif comme il se doit !).

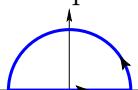
La même méthode s'applique à toute intégrale $\int_0^{2\pi} R(\cos t, \sin t) dt$ où R est une fraction rationnelle.

8.1.5 Transformées de Fourier

On s'intéresse à une intégrale de la forme

$$I = \int_{-\infty}^{\infty} f(x)e^{ikx} dx \quad k \in \mathbb{R}$$

qui est la transformée de Fourier de f . On va à nouveau remplacer cette intégrale sur l'axe réel par une intégrale de contour dans le demi-plan supérieur ou inférieur, selon que le signe de k est > 0 ou < 0 . En effet quand $|z| \rightarrow \infty$, $e^{ikz} f(z)$ tend vers zéro si $\Re e ikz < 0$ et si $f(z)$ croît moins vite qu'une exponentielle (par exemple polynomialement) ; pour cela, si $k > 0$, il faut que $\Im m z > 0$. On peut alors appliquer le 2ème lemme de Jordan, remplacer I par l'intégrale sur le



contour et utiliser le théorème des résidus. À l'inverse si $k < 0$, il faut refermer le contour dans le demi-plan inférieur (en n'ommettant pas le signe venant de l'orientation du contour !).

Prenons comme exemple le cas de la fonction "lorentzienne" rencontrée aux chap. 4 et 5, $f(x) = \frac{1}{x^2 + a^2}$, $a \in \mathbb{R}$, $a > 0$. La fonction $f(z) = (z^2 + a^2)^{-1}$ a deux pôles en $\pm ia$ de résidus

respectifs $\pm \frac{1}{2ia}$. Selon la discussion précédente

$$\int_{-\infty}^{\infty} \frac{1}{x^2 + a^2} e^{ikx} dx = \begin{cases} 2\pi i \operatorname{Res}(f(z)e^{ikz}, ia) = \frac{\pi}{a} e^{-ka} & \text{si } k > 0 \\ -2\pi i \operatorname{Res}(f(z)e^{ikz}, -ia) = \frac{\pi}{a} e^{ka} & \text{si } k < 0 \end{cases}$$

ce que l'on regroupe en une formule unique

$$\int_{-\infty}^{\infty} \frac{1}{x^2 + a^2} e^{ikx} dx = \frac{\pi}{a} e^{-|k|a} \quad (8.4)$$

en accord avec un résultat annoncé au § 4.2.1.

Notons qu'en prenant la partie réelle de I , nous obtenons une autre intégrale non triviale $I = \int_{-\infty}^{\infty} \frac{1}{x^2 + a^2} \cos kx dx = \frac{\pi}{a} e^{-|k|a}$.

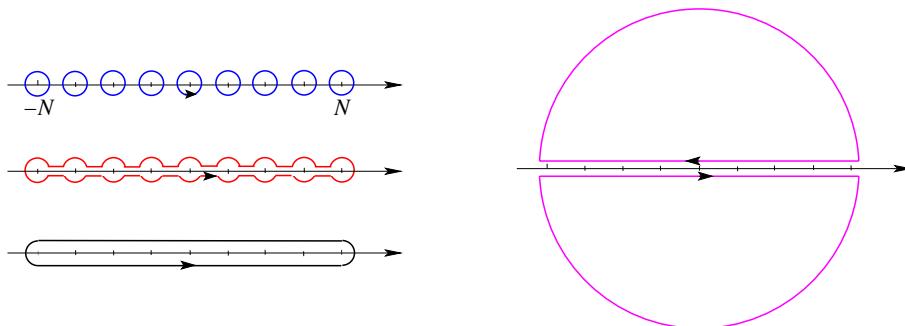
8.1.6 Sommes infinies

Notons en préambule que $\cos \pi z$ est une fonction entière, que $\sin \pi z$ a des zéros en $z = n \in \mathbb{Z}$, et que $\cotan \pi z = \cos \pi z / \sin \pi z$ n'a que des pôles de résidu $1/\pi$ en ces mêmes $z = n$.

Considérons alors la série

$$S = \sum_{n \in \mathbb{Z}} f(n)$$

où f est une fonction méromorphe qui n'a pas de pôle réel et qui tend suffisamment vite vers zéro à l'infini pour assurer la convergence. On peut écrire $f(n) = \operatorname{Res}(\pi f(z)\cotan \pi z, n)$ et la somme S , limitée d'abord à $-N \leq n \leq N$, peut être vue comme venant de l'intégrale de $\frac{1}{2i} f(z)\cotan \pi z$ sur l'un des contours successifs figurés ici



où la contribution des deux demi-cercles tend vers zéro quand $N \rightarrow \infty$ (Jordan 1 !) ; à la limite, le dernier contour entoure les pôles de $f(z)$ en dehors de l'axe réel dans le sens négatif. Donc

$$S = - \sum_{\substack{\text{pôles } z_j \text{ de } f \\ \text{en dehors de l'axe réel}}} \operatorname{Res}(\pi f(z)\cotan \pi z, z_j).$$

Exemple. $S = \sum_{n \in \mathbb{Z}} \frac{1}{n^2 + a^2}$ avec $a > 0$. Les pôles non réels de $\frac{\pi}{z^2 + a^2} \cotan \pi z$ sont en $z = \pm ia$ et ont pour résidu $-\frac{\pi}{2a} \coth \pi a$. Finalement $S = \frac{\pi}{a} \coth \pi a$. On vérifie que quand $a \rightarrow 0$, $S \sim 1/a^2$ comme attendu. Exercice : dans la limite $a \rightarrow 0$, soustraire le terme $n = 0$ et vérifier que l'on retrouve l'expression bien connue $\sum_{n \geq 1} \frac{1}{n^2} = \pi^2/6$.

Par une méthode analogue faisant appel au $1/\sin$ plutôt qu'au \cotan , calculer $\sum_{n \geq 1} \frac{(-1)^n}{n^4}$, réponse dans [1], p. 111.

8.1.7 Intégrales sur un arc. Pôles sur l'axe réel

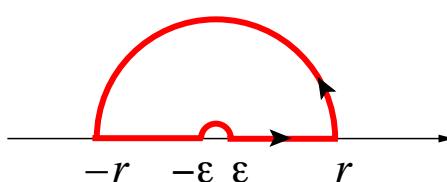
Supposons que f est holomorphe dans un secteur du type défini plus haut au § 8.1.2 : $\mathcal{S} = \{z : |z| > 0, 0 \leq \theta_1 \leq \arg z \leq \theta_2 \leq \pi\}$ avec un pôle simple à l'origine, et qu'on veuille calculer $\lim_{r \rightarrow 0} \int_{\gamma(r; \theta_1, \theta_2)} f(z) dz$. On écrit $f(z) = \frac{a}{z} + g(z)$ où g est holomorphe à l'origine et $a = \text{Res}(f, 0)$. Par Jordan.1, $\lim_{r \rightarrow 0} \int_{\gamma(r)} g(z) dz = 0$ et donc

$$\lim_{r \rightarrow 0} \int_{\gamma(r; \theta_1, \theta_2)} f(z) dz = a \int_{\gamma(r; \theta_1, \theta_2)} \frac{dz}{z} = ia(\theta_2 - \theta_1). \quad (8.5)$$

Application au calcul de $I = \int_0^\infty \frac{\sin x}{x} dx$. On écrit

$$I = \frac{1}{2} \int_{-\infty}^\infty \frac{\sin x}{x} dx = \frac{1}{2i} \lim_{\epsilon \rightarrow 0} \left(\int_{-\infty}^{-\epsilon} \frac{e^{ix}}{x} dx + \int_\epsilon^\infty \frac{e^{ix}}{x} dx \right). \quad (8.6)$$

Mais l'intégrale de $\frac{e^{ix}}{x}$ sur le contour ci-dessous est nulle



puisque le contour n'enclôt aucun pôle. Dans cette intégrale de contour, la contribution sur l'axe réel tend vers $2iI$ quand $r \rightarrow \infty$ et $\epsilon \rightarrow 0$, d'après (8.6) ; la contribution du grand cercle s'annule par Jordan.2 quand $r \rightarrow \infty$; et celle sur le petit cercle est donnée par le calcul précédent (8.5). Il reste dans cette limite

$$2iI = \int_{\gamma(\epsilon, 0, \pi)} \frac{e^{iz}}{z} dz = i\pi$$

d'où finalement $I = \frac{\pi}{2}$.

8.2 Fonctions multivaluées

8.2.1 Points de branchement, feuillets, surface de Riemann

On a déjà rencontré au § 6.2 la notion de fonction multivaluée avec le logarithme complexe, dont les différentes déterminations dans le plan complexe privé de l'origine (plan “pointé” $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$) diffèrent d'un multiple entier de $2\pi i$. Considérons le plan privé d'une demi-droite, par exemple $\mathbb{C} \setminus \mathbb{R}^-$, donc $-\pi < \arg z < \pi$: c'est un domaine simplement connexe, le logarithme y admet une détermination unique, dite détermination principale. Mais on peut très bien choisir une autre demi-droite Δ partant de l'origine, ou toute courbe \mathcal{C} joignant l'origine à l'infini. Dans le plan privé de cette demi-droite ou courbe, le logarithme admet une détermination unique. $\mathbb{C} \setminus \mathcal{C}$ est dit *plan coupé* le long de \mathcal{C} , \mathcal{C} est la “coupure”, l'origine et le point à l'infini en sont les *points de branchement*. (Que l'infini est un point de branchement du log se voit comme toujours par changement de variable $z \mapsto w = \frac{1}{z}$: $\log z = -\log w \dots$)

Le logarithme n'est pas la seule fonction multivaluée. La fonction racine carrée (ou toute autre racine) donne un autre exemple. En effet quand on écrit $z = \rho e^{i\theta}$, on peut penser définir $z^{\frac{1}{2}}$ alias \sqrt{z} par $\sqrt{z} = \sqrt{\rho} e^{i\theta/2}$. Mais θ étant défini à 2π près, $\theta/2$ l'est à π près et \sqrt{z} est défini à un signe près. C'est un problème que nous connaissons bien sur les réels ; mais ici sur les complexes, il prend une nouvelle dimension puisqu'en suivant un chemin continu dans le plan complexe, on peut passer d'une détermination à l'autre. On va à nouveau introduire une coupure du plan complexe partant de l'origine et définir une détermination de la racine carrée dans ce plan coupé. Par exemple, en coupant le long de l'axe réel négatif, donc en prenant à nouveau $-\pi < \arg z < \pi$, on obtient dans $\mathbb{C} \setminus \mathbb{R}^-$ la détermination usuelle de la racine, telle que pour z réel > 0 , $\sqrt{z} > 0$.

Les points de branchement dépendent de la fonction considérée et peuvent être n'importe où dans le plan complexe complété. Par exemple $\sqrt{1-z}$ a un point de branchement en 1 et un à l'infini. Les points de branchement de $\sqrt{1-z^2}$ sont ± 1 ; ceux de $\sqrt{(z^2-a^2)(z^2-b^2)}$ sont en $\pm a$ et $\pm b$, et les coupures peuvent être choisies entre a et $-a$ d'une part, et b et $-b$ de l'autre, (ou entre $-a$ et $-b$ et entre a et b , justifier ce point), etc.

Plutôt que de parler de fonction univaluée dans le plan (coupé), on peut préférer une autre présentation. Étant donnée un choix de coupure(s) et une détermination, on peut associer les autres déterminations à d'autres copies du plan coupé –on parle des différents *feuilles* de la fonction considérée–, puis recoller ces feuilles le long de la (ou des) coupure(s), en suivant par continuité les déterminations de la fonction. L'objet obtenu par ce recollement est appelé surface de Riemann de la fonction considérée. Sur cette surface, qui est un *revêtement* à plusieurs feuilles du plan complexe initial, la fonction est univaluée.

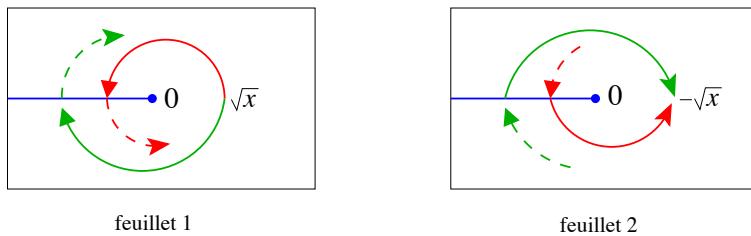


FIGURE 8.1 – Les deux feuillets de la fonction \sqrt{z} : le bord inférieur de la coupure du feuillet 1 est recollé avec le bord supérieur de celle du feuillet 2 et vice versa.

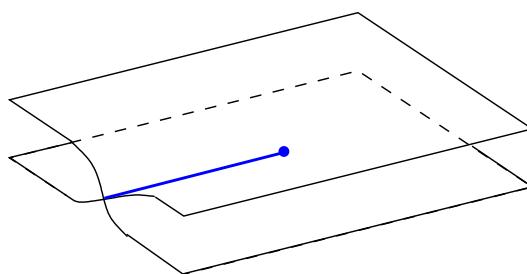


FIGURE 8.2 – La surface de Riemann de la fonction \sqrt{z} : les deux feuillets sont recollés le long du demi-axe réel négatif (en bleu).

Par exemple pour \sqrt{z} dans le plan coupé le long de \mathbb{R}^- , la figure 8.1 représente les deux feuillets, la figure 8.2 montre la surface de Riemann. Noter que (contrairement aux apparences) cette surface de Riemann *est* simplement connexe : un chemin fermé faisant deux fois le tour de l'origine est homotope à zéro, puisqu'il peut être contracté sans rencontrer de singularité (de la fonction). Cela apparaît peut-être plus clairement sur la Fig. 8.3 qui représente le voisinage de l'origine, une fois la surface de Riemann dépliée. La position des coupures est arbitraire et le point O n'est pas singulier. Le même dessin devrait être effectué (dans une autre coordonnée) au voisinage du point à l'infini, avec la même conclusion. En définitive la surface de Riemann de la fonction \sqrt{z} est identifiée à la sphère de Riemann. Elle est donc simplement connexe.¹

La même construction s'applique à toute fonction multivaluée. La surface de Riemann de la racine carrée a deux feuillets, celle du logarithme en a une infinité. Celle de $z^{1/3}$ en a trois, celle de z^α , α réel irrationnel en a une infinité. Celle de la fonction $\sqrt{(z^2 - a^2)(z^2 - b^2)}$ en a quatre etc. On trouve sur le web d'admirables figures de ces surfaces de Riemann ... par exemple http://en.wikipedia.org/wiki/File:Riemann_surface_log.jpg

1. Cela n'est pas le cas général, la surface de Riemann d'une fonction plus compliquée n'est en général pas simplement connexe, elle a un *genre* (= nombre de "poignées") qui dépend du nombre et de l'ordre des points de branchement

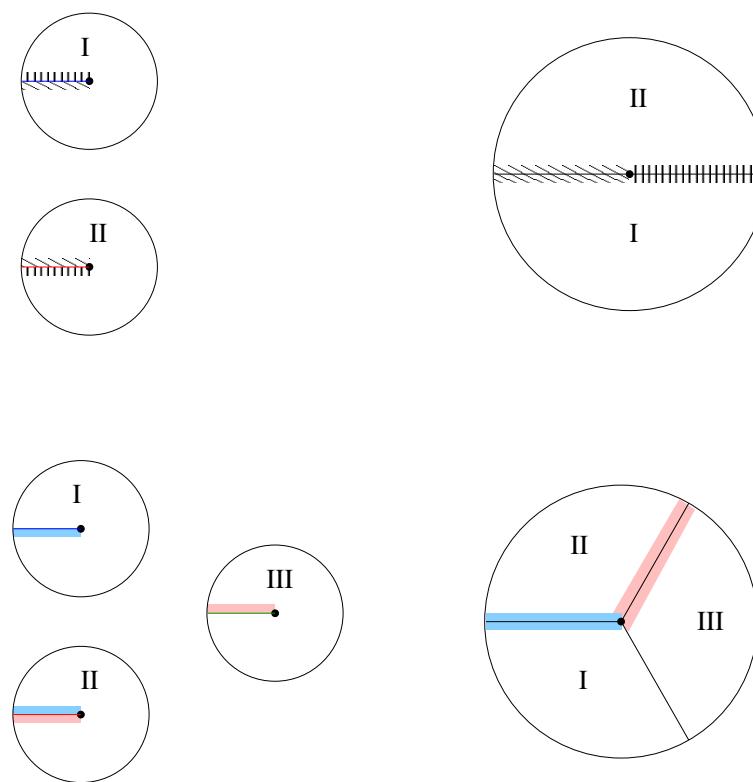


FIGURE 8.3 – La surface de Riemann des fonctions $z^{\frac{1}{2}}$ (en haut) et $z^{\frac{1}{3}}$ (en bas) au voisinage de 0. A gauche, les trois feuillets avec en hâchures ou en couleurs les prescriptions de recollement ; à droite, après dépliage et recollement.

8.2.2 Intégrales de Cauchy de fonctions multivaluées

Une fonction multivaluée comme z^α ou $\log z$ est holomorphe en dehors de ses points de branchement (et de ses éventuelles autres singularités). Rien n'interdit d'utiliser le théorème de Cauchy et ses corollaires le long de chemins fermés évitant ces singularités et continus dans les différents feuillets.

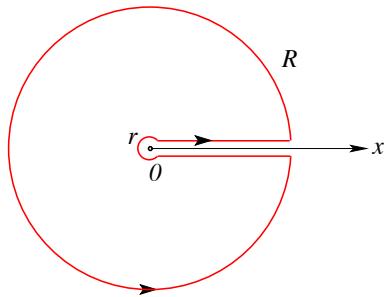


FIGURE 8.4 –

Considérons par exemple l'intégrale $I = \int_0^\infty \frac{F(x)}{x^\alpha} dx$ où $0 < \alpha < 1$ et où F est une fraction rationnelle sans pôle sur le demi-axe $x \geq 0$; on suppose (pour la convergence) que $F \rightarrow 0$ à l'infini (donc au moins aussi vite que $1/z$). Pour étendre $f(z) = F(z)/z^\alpha$ à des valeurs complexes de z , on choisit une détermination de z^α , par exemple dans le plan coupé le long de l'axe réel positif, $0 < \arg z < 2\pi$. Considérons alors l'intégrale $\oint_\gamma \frac{F(z)}{z^\alpha} dz$ le long du contour de la figure 8.4. Comme $zf(z) \rightarrow 0$ quand $|z| \rightarrow 0$ et $\rightarrow \infty$, les deux arcs de cercle $\gamma(r)$ et $\gamma(R)$ ne contribuent pas dans les limites $r \rightarrow 0$ et $R \rightarrow \infty$ (Jordan-1), la détermination de f sur le contour réel supérieur est réelle, celle sur le contour inférieur a un facteur relatif $e^{-2\pi i\alpha}$. On a donc

$$(1 - e^{-2\pi i\alpha})I = 2\pi i \sum \text{Res}(f(z))$$

où la somme court sur tous les pôles de F , réels négatifs ou complexes.

Exemple : $I = \int_0^\infty \frac{dx}{x^{\alpha(1+x)}}$. Le seul pôle est en $-1 = e^{i\pi}$ (choix de la détermination!), son résidu est $e^{-i\pi\alpha}$, d'où $I = \frac{\pi}{\sin \pi\alpha}$.

Autre exemple : $J = \int_0^\infty F(x) \log x dx$ où F est une fraction rationnelle sans pôle sur le demi-axe $x \geq 0$; on suppose (pour la convergence) que $xF(x) \rightarrow 0$ à l'infini. L'astuce est cette fois de considérer $f(z) = F(z)(\log z)^2$ avec un choix de coupure du log le long de l'axe réel *positif*. On intègre sur le même contour de la figure 8.4. À nouveau les deux arcs de cercle ne contribuent pas, les deux déterminations de $\log^2 z$ le long de l'axe réel sont $\log^2 x$ (bord supérieur) et $(\log x + 2\pi i)^2$ (bord inférieur), donc $\int_0^\infty F(x)(\log^2 x - (\log x + 2\pi i)^2)dx = -4\pi i J - 4\pi^2 \int_0^\infty F(x)dx = 2\pi i \sum \text{Res}(F(z)(\log z)^2)$. Il reste à calculer $\int F(x)dx$. Toutefois si F est réel (sur l'axe réel), cette dernière intégrale est réelle et en prenant les parties réelle et imagi-

naire de la relation précédente on obtient $J = \int_0^\infty F(x) \log x \, dx = -\frac{1}{2} \Re e \sum \text{Res}(F(z)(\log z)^2)$, $\int_0^\infty F(x) \, dx = -\frac{1}{2\pi} \Im m \sum \text{Res}(F(z)(\log z)^2)$.

8.3 Fonctions harmoniques

Définition 8.1 : Une fonction de deux variables réelles x et y définie dans un ouvert Ω de \mathbb{R}^2 est dite harmonique dans Ω si elle est de classe C^2 (deux fois continûment différentiable) et satisfait l'équation

$$\Delta f = 0 \quad (8.7)$$

où le laplacien Δ est l'opérateur différentiel

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}. \quad (8.8)$$

En raison de l'apparition du laplacien dans de nombreuses équations de la physique, équation de Poisson (3.26), équation des ondes, équation de la chaleur, ... les fonctions harmoniques y jouent un rôle important.

Introduisant comme au § 7.1.3 les variables $z = x + iy$ et $\bar{z} = x - iy$ et les opérateurs différentiels $\frac{\partial}{\partial z}$ et $\frac{\partial}{\partial \bar{z}}$, on vérifie immédiatement que

$$\Delta = 4 \frac{\partial^2}{\partial z \partial \bar{z}}. \quad (8.9)$$

Une fonction $f(z, \bar{z})$ satisfaisant $\frac{\partial^2 f}{\partial z \partial \bar{z}} = 0$ est donc harmonique.

Théorème 8.1 : Toute fonction holomorphe est harmonique.

En effet si f est holomorphe (dans Ω) elle est indéfiniment dérivable ; on a $\frac{\partial f}{\partial \bar{z}} = 0$, donc aussi $\Delta f = 0$.

En prenant les parties réelle et imaginaire de f il en découle :

Corollaire : La partie réelle et la partie imaginaire d'une fonction holomorphe sont harmoniques.

Exemple : En tout point non nul du plan complexe, $\log z$ est holomorphe, donc sa partie réelle $\log |z|$ est harmonique.

La réciproque n'est vraie que *localement* :

Proposition 8.2 : Toute fonction réelle $g(x, y)$ harmonique dans un ouvert Ω est au voisinage de chaque point de Ω la partie réelle d'une fonction f , holomorphe au voisinage de ce point, déterminée à l'addition d'une constante près.

Cette propriété locale devient globale (vraie en tout point) dans un ouvert Ω simplement connexe. Pour la preuve, voir [4] p. 125 ; voir aussi [1] p.124-125.

On se rappelle (cf Théorème 7.18) qu'une fonction holomorphe satisfait la propriété de moyenne : en tout point, sa valeur au centre d'un petit disque fermé contenu dans Ω égale la moyenne de ses valeurs sur le bord du disque.

Théorème 8.3 : *Toute fonction harmonique satisfait en tout point de Ω la propriété de moyenne : pour r assez petit*

$$u(x_0, y_0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} u(x_0 + r \cos \theta, y_0 + r \sin \theta) d\theta.$$

Cela découle de la Proposition 8.2 : pour toute fonction harmonique réelle u et tout point de Ω il existe une fonction f holomorphe dans un disque centré en ce point, de partie réelle u et qui satisfait la propriété de moyenne ; en prenant la partie réelle on voit que u satisfait la propriété de moyenne.

Au chapitre 7, Théorème 7.21, nous avons montré que toute fonction possédant la propriété de moyenne obéit aussi au principe de maximum. Il en est donc ainsi pour toute fonction harmonique :

Proposition 8.4 : *Toute fonction u réelle harmonique dans Ω est ou bien constante, ou bien n'a en aucun point de Ω un maximum local ou un minimum local.*

Cela s'interprète facilement en termes des dérivées secondes de u . Un maximum (resp. minimum) local en $a \in \Omega$ signifierait que $\frac{\partial^2 u}{\partial x^2}$ et $\frac{\partial^2 u}{\partial y^2}$ sont non nuls et de même signe, en contradiction avec l'équation (8.7).

Ces propriétés trouvent des applications physiques, par exemple en électrostatique, dans l'étude du potentiel créé par une distribution de charges. Comme on l'a rappelé au chapitre 3.6, la loi de Gauss implique que le potentiel électrostatique créé par une distribution de charges est une fonction harmonique en tout point distinct des positions des charges. Une autre application est en aérodynamique dans le calcul de la portance d'une aile d'avion, voir [1], chap. 7. Malheureusement le manque de temps ne permettra pas d'explorer les nombreuses applications des considérations précédentes.

8.4 Méthode du col

On est souvent confronté en physique au problème suivant : on cherche à calculer la limite d'une intégrale de la forme

$$I_\alpha = \int_\gamma e^{-\alpha f(z)} dz \tag{8.10}$$

quand le paramètre α tend vers l'infini. f est supposée holomorphe sur un ouvert Ω indépendant de α , γ est un chemin contenu dans Ω .

Il existe plusieurs versions de ce problème, la méthode du col, la méthode de Laplace, la méthode de la phase stationnaire, etc. Nous nous bornerons à une discussion sommaire et à un exemple.

8.4.1 Méthode du col

L'idée intuitive est de chercher à minimiser la partie réelle de αf tout en empêchant sa partie imaginaire de varier trop rapidement, ce qui risquerait de causer des compensations dans l'intégrale. On va préciser cette idée.

Étant donnée l'intégrale (8.10), où nous prendrons α réel positif, supposons que f est holomorphe dans un ouvert Ω contenant le contour γ et qu'on sache déformer continûment γ en γ' sans rencontrer de singularité de f de telle façon que

- (i) le long de γ' , $\Im m f$ est constante ;
- (ii) il existe le long de γ' un point z_0 , dit *col*, tel que $f'(z_0) = 0$;
- (iii) $\Re e f(z)$ passe par un minimum local en z_0 ;

La condition $\Im m f = \text{const.}$ détermine en général le chemin γ' . La phase de l'exponentielle est alors constante. Développons alors $f(z)$ au voisinage du col. Puisque $f'(z_0) = 0$,

$$f(z) = f(z_0) + \frac{1}{2}(z - z_0)^2 f''(z_0) + o(z - z_0)^2$$

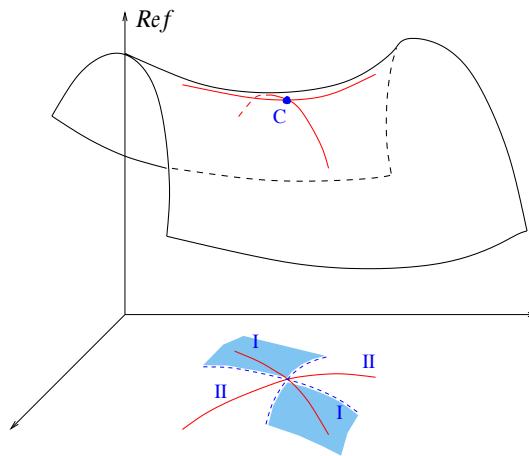
et avec $f''(z_0) = \rho e^{i\theta}$, $z - z_0 = r e^{i\varphi}$,

$$\begin{aligned} \Re e(f(z) - f(z_0)) &= \frac{1}{2}r^2\rho \cos(2\varphi + \theta) + o(r^2) \\ \Im m(f(z) - f(z_0)) &= \frac{1}{2}r^2\rho \sin(2\varphi + \theta) + o(r^2) \end{aligned} \quad (8.11)$$

sur lequel on lit que, pour satisfaire (i), le chemin γ' doit être tel que $\sin(2\varphi + \theta) = 0$ ce qui donne deux directions possibles, orthogonales l'une à l'autre. Les directions médianes, où le cos s'annule, sont celles où la partie réelle change de signe, voir Fig. 8.5. Pour assurer que la partie réelle passe bien par un minimum et non un maximum, c'est le chemin passant dans les zones II qu'on doit choisir. Quand $\alpha \rightarrow \infty$, le minimum de $\Re e(\alpha f)$ est de plus en plus accentué, de plus en plus "piqué", d'où le nom de "steepest descent" donné à la méthode dans la littérature anglo-saxonne. C'est ce qui fait que la contribution dominante à l'intégrale vient de ce chemin. Finalement on s'attend à un comportement asymptotique (et on peut justifier précisément avec des hypothèses adéquates sur f que nous ne détaillerons pas ici)

$$I_\alpha \approx e^{-\alpha f(z_0)} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\alpha f''(z_0)u^2} du$$

$$\approx e^{-\alpha f(z_0)} \left[\frac{2\pi}{\alpha f''(z_0)} \right]^{\frac{1}{2}}. \quad (8.12)$$

FIGURE 8.5 – Col de $\Re f$ et trajectoires de $\Im f$ constant.

8.4.2 Commentaires et variantes

Dans l'intégrale (8.10) le paramètre α prenant de grandes valeurs était explicite et factorisé dans l'exponentielle. Il peut aussi arriver que l'on s'intéresse au comportement d'une intégrale où cette dépendance n'est pas aussi manifeste. Mais l'idée est la même : on cherche les points stationnaires (cols) de la fonction en exponentielle. On en verra un exemple ci-dessous avec la fonction Γ . La méthode du col possède aussi des variantes dont il est utile de connaître l'existence mais que nous ne ferons que mentionner brièvement :

Méthode de Laplace, col réel

C'est la situation où on intègre dans (8.10) une fonction f réelle.

$$J_\alpha = \int_{\mathbb{R}} e^{-\alpha f(x)} dx \quad (8.13)$$

On n'a plus la liberté de déformer le contour d'intégration, mais un résultat similaire s'applique : le comportement dominant de l'intégrale est à nouveau donné par le point x_0 où f est minimale. Comme dans le cas complexe, on développe f au voisinage de ce point

$$f(x) \approx f(x_0) + \frac{1}{2}(x - x_0)^2 f''(x_0) + o(x - x_0)^2$$

et l'intégrale est approximée par une intégrale gaussienne

$$J_\alpha \approx e^{-\alpha f(x_0)} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\alpha f''(x_0)x^2} dx = e^{-\alpha f(x_0)} \sqrt{\frac{2\pi}{\alpha f''(x_0)}}.$$

Voir ci-dessous l'exemple de la fonction Γ .

Méthode de la phase stationnaire

Une autre variante est celle d'une intégrale sur \mathbb{R} d'une exponentielle d'une fonction imaginaire

$$J'_\alpha = \int_{\mathbb{R}} e^{i\alpha f(x)} dx \quad (8.14)$$

où f est réelle. Cette fois encore il s'agit de chercher un extremum ou un point stationnaire de f , où les oscillations de la phase αf sont moins rapides.

On parle de *méthode de phase stationnaire*, et on écrit

$$J'_\alpha \approx e^{i\alpha f(x_0)} \left[\frac{2\pi i}{f''(x_0)} \right]^{\frac{1}{2}}.$$

On rencontrera la méthode du col et ses avatars dans de nombreux domaines de la physique, en Mécanique Statistique (limite thermodynamique, modèles à grand nombre d'états, ...), en Mécanique Quantique et en Optique (vitesse de groupe, limite de courte longueur d'onde...), en théorie des champs (méthode semi-classique, où le paramètre grand est $1/\hbar$), etc.

8.4.3 Fonction Γ . Formule de Stirling

Nous illustrerons la méthode du col (réel) sur le cas de la fonction Gamma d'Euler

$$\Gamma(s) \stackrel{\text{déf}}{=} \int_0^{\infty} u^{s-1} e^{-u} du. \quad (8.15)$$

Observons d'abord que cette intégrale est absolument convergente pour $\Re s > 0$. On calcule aisément $\Gamma(1) = 1$ et, par intégration par parties,

$$\Gamma(s+1) = s\Gamma(s)$$

d'où il découle immédiatement que pour $s = n \geq 1$ entier

$$\Gamma(n) = (n-1)!$$

$\Gamma(s)$ extrapole donc à des valeurs réelles ou complexes la fonction factorielle.

Intéressons-nous à l'asymptotique de la fonction $\Gamma(s+1)$. Selon la méthode précédente, on écrit l'intégrand $e^{f(u)}$ et on cherche les points stationnaires de $f(u) = s \log u - u$, soit ici u_0

solution de $s/u_0 = 1$, donc $u_0 = s$. On développe alors $f(u) = f(u_0) + (u - u_0)f'(u_0) + \frac{1}{2}(u - u_0)^2f''(u_0) + \dots$ mais $f'(u_0) = 0$, et $f''(u_0) = -1/s$, donc

$$\Gamma(s+1) \approx e^{f(u_0)} \int_{-\infty}^{\infty} e^{\frac{1}{2}v^2f''(u_0)} dv = \sqrt{2\pi s} e^{s \log s - s}$$

ou encore

$$\Gamma(s+1) = \left(\frac{s}{e}\right)^s \sqrt{2\pi s} (1 + O(1/s)),$$

où on reconnaît la formule de Stirling pour la factorielle asymptotique... Le lecteur courageux peut-il/elle calculer le terme suivant en $1/s$? Attention, il faut pour cela développer la fonction f jusqu'au terme $(u - u_0)^4$ (et non pas seulement $(u - u_0)^3$, pourquoi?).

Lectures complémentaires

La discussion a suivi H. Cartan [4] et W. Appel [1]. Pour plus de détails sur les subtilités de la méthode du col, consulter C. Aslangul [2], chap. 7.

★

Chapitre 9

Transformation de Laplace

La transformation de Laplace fait passer d'une fonction f à valeurs réelles ou complexes d'une variable t *réelle non négative* : $f : \mathbb{R}^+ \rightarrow \mathbb{R}$ ou \mathbb{C} , à une fonction d'une variable *complexe* p , $\hat{f}(p) = \int_0^\infty e^{-pt} f(t) dt$, avec bien sûr des conditions sur f pour assurer la convergence. Elle est donc apparentée à la transformation de Fourier, mais en diffère par quelques points que nous allons examiner. Comme la transformation de Fourier, la transformation de Laplace est un outil puissant dans l'étude des équations différentielles ou aux dérivées partielles et des systèmes linéaires en physique.

9.1 Définitions et premières propriétés

9.1.1 Abscisse de sommabilité et transformée de Laplace

On rencontre souvent en physique des fonctions définies seulement sur la demi-droite réelle positive \mathbb{R}^+ . C'est par exemple le cas dans un système dynamique soumis au temps $t = 0$ à une excitation (une "source") : on s'intéresse à sa "réponse" $f(t)$ aux temps ultérieurs $t \geq 0$. On peut aussi considérer que la fonction f est définie pour tout $t \in \mathbb{R}$ mais n'est non nulle que pour $t \geq 0$. D'où la

Définition 9.1 : *On appelle fonction causale une fonction $t \mapsto f(t)$ nulle pour $t < 0$.*

Définition 9.2 : *Pour une fonction causale, on définit la transformée de Laplace par*

$$\hat{f}(p) \stackrel{\text{def}}{=} \int_0^\infty e^{-pt} f(t) dt \quad (9.1)$$

mais aussi, compte tenu de l'hypothèse d'annulation à $t < 0$, par $\int_{-\infty}^{\infty} e^{-pt} f(t) dt$.

La transformée de Laplace est notée selon les auteurs (et les circonstances !) $\hat{f}(p)$, $\mathcal{L}[f](p)$, $L_f(p)$, etc, et la transformation parfois $f(t) \square \hat{f}(p)$.

Il reste à préciser les conditions de convergence. On va d'abord supposer que f est localement intégrable (c'est-à-dire intégrable sur tout compact, cf Déf. 3.5), ce qui n'interdit pas à la fonction d'avoir une singularité intégrable à distance finie, comme par exemple $|t - 1|^{-\frac{1}{2}}$. Par ailleurs on observe que si $|f(t)|e^{-st}$ est intégrable pour $s \in \mathbb{R}$, il en est de même de $|f(t)|e^{-s't}$ pour tout $s' > s$. (On rappelle que la fonction f est causale, seule nous intéressera la convergence en $+\infty$.) Cela conduit à la

Définition 9.3 : On appelle abscisse de sommabilité de f la borne inférieure α des $s \in \mathbb{R}$ tels que $f(t)e^{-st}$ est intégrable.

$$\alpha \stackrel{\text{déf}}{=} \inf\{s \in \mathbb{R} : |f(t)|e^{-st} \text{ est intégrable}\}. \quad (9.2)$$

Cette abscisse de sommabilité peut être infinie, voir plus bas.

Plus précisément, si on pose $p = x + i\omega$, l'intégrabilité de $f(t)e^{-pt}$ équivaut à celle de $f(t)e^{-xt}$, et est assurée pour $x > \alpha$. La transformée de Laplace est donc définie pour $x = \Re(p) > \alpha^1$. Inversement l'intégrale ne converge certainement pas pour $x < \alpha$. Pour $x = \alpha$, on peut avoir ou non intégrabilité au sens de Lebesgue ; l'intégrale impropre $\int_{-\infty}^{\infty} e^{-pt} f(t) dt$ peut être non intégrable de Lebesgue mais semi-convergente pour certaines valeurs de p , permettant ainsi d'étendre la transformée de Laplace à ces valeurs : il faut se livrer à une étude cas par cas pour le déterminer. En résumé,

Proposition 9.1 : La transformée de Laplace est définie dans le demi-plan ouvert de sommabilité, $\Re(p) > \alpha$.

Heuristiquement, l'existence d'une abscisse de sommabilité finie ou $-\infty$ signifie que f a une croissance au plus exponentielle à l'infini.

Exemples : a) Une fonction constante a une abscisse de sommabilité nulle ; la transformée de Laplace est définie pour $x = \Re(p) > 0$. Ainsi pour $f = 1$ (pour $t \geq 0$ c'est-à-dire en fait pour la fonction de Heaviside H !).

$$\widehat{H}(p) = \int_{-\infty}^{\infty} H(t)e^{-pt} dt = \int_0^{\infty} e^{-pt} dt = \frac{1}{p}[-e^{-pt}]_0^{\infty}$$

donc $\widehat{H}(p) = \frac{1}{p}$ pour $\Re(p) > 0$. On note que dans ce cas, \widehat{H} peut être étendue (prolongée) à tout $p \neq 0$, en particulier pour $\Re(p) = 0$, $\Im(p) = \omega \neq 0$ avec le résultat $\widehat{H}(i\omega) = \frac{1}{i\omega}$.

1. Certains auteurs appellent *original* une fonction f ayant les propriétés énumérées ci-dessus : causalité, intégrabilité locale, existence d'une abscisse de sommabilité, et *image* sa transformée de Laplace.

b) $f(t) = 1/(1+t^2)$ a aussi une abscisse de sommabilité nulle ; mais la transformée de Laplace est définie pour tout $x = \Re(p) \geq 0$. (Son expression implique des “fonctions spéciales”, le sinus intégral et le cosinus intégral.)

c) La fonction e^{-at^2} pour $a > 0$ a une abscisse de sommabilité $\alpha = -\infty$: la transformée de Laplace existe pour tout p ; à l’inverse pour $a < 0$, f n’a pas d’abscisse de sommabilité (ou si on veut, $\alpha = \infty$) et la transformée de Laplace n’est définie pour aucun p . Nous verrons d’autres exemples au § 9.1.3 ci-dessous.

9.1.2 Holomorphie de \hat{f} , etc

Proposition 9.2 : *La transformée de Laplace $\hat{f}(p)$ est une fonction holomorphe, donc analytique, de la variable p dans le demi-plan ouvert de sommabilité.*

Preuve : Établissons d’abord que l’abscisse de sommabilité α' de $tf(t)$ égale celle de f notée α . Heuristiquement, si $|f|$ a une croissance exponentielle en $e^{\alpha t}$ à l’infini, il en est de même de $t|f(t)|$. Plus précisément, utilisons une évidence utile : si g domine f ($|f| < |g|$) pour t assez grand, l’intégrabilité de ge^{-st} assure celle de fe^{-st} , donc l’abscisse de sommabilité de f est inférieure ou égale à celle de g . Ici, si $t > 1$, $|tf(t)|e^{-st} > |f(t)|e^{-st}$, donc $\alpha \leq \alpha'$; mais par ailleurs comme $t|f| < e^{\epsilon t}|f|$ pour tout $\epsilon > 0$ et t assez grand, $\alpha' \leq \alpha + \epsilon$ quel que soit ϵ , donc finalement $\alpha = \alpha'$, qed. Donc dans le demi-plan ouvert $\Re(p) > \alpha$, $\int_0^\infty tf(t)e^{-pt}$ est absolument convergente, on peut donc dériver $\hat{f}(p)$ sous le signe somme par rapport à $\Re(p)$ ou à $\Im(p)$; on trouve $\frac{\partial \hat{f}(p)}{\partial \Re(p)} = -\int_0^\infty tf(t)e^{-pt} = -i \frac{\partial \hat{f}(p)}{\partial \Im(p)}$, les conditions de Cauchy–Riemann sont satisfaites et on a donc $\frac{d\hat{f}(p)}{dp} = \hat{f}'(p) = -\int_0^\infty tf(t)e^{-pt}$. \hat{f} est dérivable dans le demi-plan, c’est-à-dire holomorphe, donc analytique. Par récurrence la dérivée n -ième vaut $\hat{f}^{(n)}(p) = (-1)^n \int_0^\infty t^n f(t)e^{-pt}$.

Relation avec la transformée de Fourier

On aura noté que la transformée de Fourier d’une *fonction causale* est sa transformée de Laplace à $\Re(p) = 0$, c’est-à-dire sur l’axe imaginaire.

$$\tilde{f}(k) = \mathcal{F}[f](k) = \int_{-\infty}^{\infty} e^{ikt} f(t) dt = \int_0^{\infty} e^{ikt} f(t) dt = \hat{f}(-ik). \quad (9.3)$$

Selon le type de croissance de la fonction f pour $t \rightarrow +\infty$, on peut voir si l’axe imaginaire est ou non dans le domaine de définition de \hat{f} et conclure à l’existence ou non de la transformée de Fourier :

- si f croît moins vite qu’une exponentielle, $\alpha < 0$, donc la transformée de Fourier existe ;
- si f croît plus vite que toute puissance mais au plus comme une exponentielle, on a $\alpha > 0$, l’axe imaginaire n’est pas dans le domaine de définition de \hat{f} et la transformée de Fourier n’existe pas ;
- dans le cas intermédiaire où $\alpha = 0$, la transformée de Fourier n’existe pas toujours au sens des fonctions.

Comportement asymptotique de \hat{f}

Proposition 9.3 : Soit f une fonction causale d'abscisse de sommabilité α et \hat{f} sa transformée de Laplace. Si $x > \alpha$, alors $\hat{f}(x + i\omega) \rightarrow 0$ quand $\omega \rightarrow \pm\infty$

Cela résulte du lemme 4.5 de Riemann–Lebesgue,

$$\lim_{\omega \rightarrow \pm\infty} \int_0^\infty (f(t)e^{-xt}) e^{-i\omega t} dt = 0$$

puisque $f(t)e^{-xt} \in L^1(\mathbb{R})$.

9.1.3 Exemples

a) On a vu la transformée de Laplace de $f = 1$ plus haut. Considérons maintenant celle de $f(t) = t$. Un calcul immédiat, par intégration par parties ou par dérivation sous le signe somme, donne pour $\Re e(p) > 0$

$$\hat{f}(p) = \int_0^\infty te^{-pt} dt = -\frac{d}{dp} \int_0^\infty e^{-pt} dt = \frac{1}{p^2},$$

qui est prolongeable en une fonction méromorphe avec un pôle en 0. Plus généralement, quelle est la transformée de t^n ?

b) $f(t) = \cos t$ a pour abscisse de sommabilité $\alpha = 0$ et pour $\Re e(p) > 0$:

$$\hat{f}(p) = \int_0^\infty e^{-pt} \cos t dt = \frac{1}{2} \int_0^\infty e^{-pt} (e^{it} + e^{-it}) dt = \frac{1}{2} \left[\frac{e^{(i-p)t}}{i-p} - \frac{e^{-(i+p)t}}{i+p} \right]_0^\infty = \frac{p}{p^2 + 1},$$

qui est prolongeable en une fonction méromorphe avec deux pôles en $\pm i$.

c) $f(t) = e^{at}$, avec $a \in \mathbb{C}$ a pour abscisse de sommabilité $\alpha = \Re e(a)$ et pour $\Re e(p) > \Re e(a)$:

$$\hat{f}(p) = \int_0^\infty e^{-(p-a)t} dt = \frac{1}{p-a},$$

à nouveau prolongeable en une fonction méromorphe avec un pôle en a .

d) Au vu de ces exemples, il est tentant de penser que la transformée de Laplace est toujours prolongeable en une fonction méromorphe. Cela est souvent le cas mais n'est pas vrai en général. Ainsi $f(t) = 1/(t+1)$ (fois $H(t)$, fonction causale !) a une transformée de Laplace $e^p \int_p^\infty e^{-u} \frac{du}{u}$ (e^p fois la “fonction Gamma incomplète”) bien définie pour $\Re e(p) > 0$ mais dont on montre qu’elle a un point de branchement en 0.

Cela apparaît aussi sur la transformée de Laplace de la fonction $f(t) = t^\kappa$, κ réel, dont l’abscisse de sommabilité est $\alpha = 0$. Se rappelant la définition de la fonction Γ dans (8.15), on a pour $\kappa > -1$ et p réel d’abord $\hat{f}(p) = \int_0^\infty e^{-pt} t^\kappa dt = p^{-\kappa+1} \int_0^\infty e^{-u} u^\kappa du = \frac{\Gamma(\kappa+1)}{p^{\kappa+1}}$. On invoque alors l’holomorphie pour dire que $\hat{f}(p) = \frac{\Gamma(\kappa+1)}{p^{\kappa+1}}$ reste vrai dans le demi-plan complexe, $\Re e(p) > 0$. Pour $\kappa > -1$ non entier, $\hat{f}(p)$ a un point de branchement en 0.

9.2 Inversion, dérivation, convolution etc

9.2.1 Inversion de la transformation de Laplace

La transformée de Laplace de f en $p = x + i\omega$ peut s'écrire comme une transformée de Fourier d'une fonction reliée à f :

$$\hat{f}(x + i\omega) = \int_{-\infty}^{\infty} H(t)f(t)e^{-xt}e^{-i\omega t} dt = \mathcal{F}[H(t)e^{-xt}f(t)](-\omega).$$

Utilisant alors la formule d'inversion de la transformée de Fourier, pour t un point où $H(t)f(t)$ est continue (cf Théorème 4.6)

$$H(t)f(t)e^{-xt} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(x + i\omega)e^{i\omega t} d\omega,$$

donc

$$H(t)f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{(x+i\omega)t} \hat{f}(x + i\omega) d\omega = \frac{1}{2\pi i} \int_{D_x} \hat{f}(p)e^{pt} dp \quad (9.4)$$

où l'intégration en p est effectuée le long d'une *droite de Bromwich*

$$D_x \stackrel{\text{déf}}{=} \{x + i\omega; \omega \in \mathbb{R}\}. \quad (9.5)$$

Bien noter que cette formule donne un résultat indépendant de $x > \alpha$, domaine où la transformée de Laplace est holomorphe, grâce au théorème de Cauchy. Il convient aussi de s'assurer que cette expression s'annule bien pour $t < 0$. En effet pour $t < 0$, le lemme de Jordan 2 nous dit que l'on peut refermer le contour d'intégration par un grand cercle dans le demi-plan à droite de la droite $\Re(p) = x$, mais le contour est complètement dans le domaine d'holomorphie, donc le résultat est nul, comme attendu pour une fonction causale.

Théorème 9.4 : Soit f une fonction causale d'abscisse de sommabilité α et \hat{f} sa transformée de Laplace. Alors en tout point de continuité de f et avec $x > \alpha$ quelconque

$$f(t) = \frac{1}{2\pi i} \int_{x-i\infty}^{x+i\infty} \hat{f}(p)e^{pt} dp$$

Exemple : Prenons $\hat{f}(p) = \frac{1}{p-a}$ (transformée de Laplace de e^{at} , voir plus haut). Pour $x > \Re(a)$, on calcule

$$f(t) = \frac{1}{2\pi i} \int_{x-i\infty}^{x+i\infty} \hat{f}(p)e^{pt} dp = \lim_{R \rightarrow \infty} \frac{1}{2\pi i} \int_{x-iR}^{x+iR} \frac{1}{p-a} e^{pt} dp$$

Si $t > 0$, on peut refermer le contour par un demi-cercle dans le demi-plan à gauche de la droite $\Re(p) = x$, qui ne contribue pas quand $R \rightarrow \infty$ (lemme de Jordan 2) et qui englobe le pôle en

$p = a$; le théorème des résidus donne alors le résultat e^{at} comme attendu. Si $t < 0$, le résultat est nul, par l'argument précédent.

D'une façon générale, la formule de Laplace inverse combinée avec la formule des résidus fournit le plus souvent le résultat cherché.

9.2.2 Translation

Soit f une fonction causale d'abscisse de sommabilité α , et \hat{f} sa transformée de Laplace. On vérifie alors aisément que la transformée de Laplace de $f(t)e^{-at}$ n'est autre que $\hat{f}(p+a)$

$$\int_0^\infty f(t)e^{-at}e^{-pt}dt = \int_0^\infty f(t)e^{-(p+a)t}dt = \hat{f}(p+a)$$

pour $\Re(p) > \alpha - \Re(a)$. Exemple, de $\mathcal{L}[1](p) = \frac{1}{p}$ on tire $\mathcal{L}[e^{at}](p) = \frac{1}{p-a}$ comme on a vu.

Attention qu'inversement $\hat{f}(p)e^{-\tau p}$ est la transformée de Laplace de $H(t-\tau)f(t-\tau)$ et non de $H(t)f(t-\tau)$!

$$\int_{-\infty}^\infty H(t-\tau)f(t-\tau)e^{-pt}dt = \int_\tau^\infty f(t-\tau)e^{-pt}dt = \int_0^\infty f(t)e^{-p(t+\tau)}dt = \hat{f}(p)e^{-\tau p}.$$

9.2.3 Convolution

Pour deux fonctions causales f et g , le produit de convolution

$$f * g(t) = \int_0^\infty f(s)g(t-s)ds = \int_0^t f(s)g(t-s)ds$$

ne dépend que des valeurs de f et g dans l'intervalle $[0, t]$ pour $t \geq 0$ et s'annule pour $t < 0$. La convolution préserve donc le caractère causal. On démontre alors comme pour la transformation de Fourier le

Théorème 9.5 : Soit f et g deux fonctions causales d'abscisses de sommabilité α et α' , et \hat{f} et \hat{g} leurs transformées de Laplace, définies respectivement dans les demi-plans ouverts $\Re(p) > \alpha$, $\Re(p) > \alpha'$. Alors

$$\mathcal{L}[f * g](p) = \hat{f}(p)\hat{g}(p)$$

est définie pour $\Re(p) > \max(\alpha, \alpha')$. Inversement pour $\Re(p) > \alpha + \alpha'$ et avec $x_0 > \alpha$

$$\mathcal{L}[f.g](p) = \frac{1}{2\pi i} \int_{x_0-i\infty}^{x_0+i\infty} \hat{f}(q)\hat{g}(p-q)dq.$$

La démonstration s'effectue comme au Théorème 4.13.

9.2.4 Opérations de dérivation et intégration

Soit une fonction causale qu'on suppose dérivable (au sens des fonctions!), α l'abscisse de sommabilité de f , on suppose que f' a une abscisse de sommabilité α' . Attention ! par transformée de Laplace de f' , on entend transformée de Laplace de $H(t)f'(t)$, et non de $(Hf)'$! Par intégration par parties, pour $\Re(p) > \max(\alpha, \alpha')$

$$\mathcal{L}[Hf'](p) = \int_0^\infty e^{-pt} f'(t) dt = p \int_0^\infty e^{-pt} f(t) dt + [f(t)e^{-pt}]_0^\infty = p\mathcal{L}[f](p) - f(0^+)$$

car $\lim_{t \rightarrow \infty} e^{-pt} f(t) = 0$ ².

Donc avec les hypothèses ci-dessus (f dérivable, α , resp. α' les abscisses de sommabilité de f et f')

Proposition 9.6 : $\mathcal{L}[Hf'](p) = p\mathcal{L}[f](p) - f(0^+)$ pour $\Re(p) > \max(\alpha, \alpha')$.

La propriété se généralise aux dérivées d'ordre supérieur, cf. [1]

Proposition 9.7 : La transformée de Laplace de $f^{(n)}$ égale $p^n \hat{f}(p) - \sum_{m=0}^{n-1} p^{n-1-m} f^{(m)}(0^+)$.

Comme on l'a observé plus haut, si f a pour abscisse de sommabilité α , pour tout $n \in \mathbb{N}$, $t^n f(t)$ est aussi causale avec la même abscisse de sommabilité et

$$\mathcal{L}[(-t)^n f(t)](p) = \frac{d^n}{dp^n} \mathcal{L}[f](p). \quad (9.6)$$

On peut aussi intégrer

Proposition 9.8 : Si f a pour abscisse de sommabilité α et \hat{f} pour transformée de Laplace, alors pour $\Re(p) > \max(\alpha, 0)$

$$\mathcal{L}\left[\int_0^t f(u) du\right](p) = \frac{\hat{f}(p)}{p}.$$

Inversement, si $\hat{f}(p)$ décroît plus vite que $1/p$ à l'infini et si $\Re(p) > \max(\alpha, 0)$

$$\mathcal{L}\left[\frac{f(t)}{t}\right](p) = \int_p^\infty \hat{f}(z) dz,$$

avec un chemin d'intégration de p à l'infini arbitraire dans le demi-plan de sommabilité, en vertu de l'holomorphie de \hat{f} .

2. On a supposé $\Re(p) > \max(\alpha, \alpha')$, et l'intégrabilité de fe^{-pt} et de sa dérivée implique que $\lim_{t \rightarrow \infty} e^{-pt} f(t) = 0$, selon un argument déjà utilisé dans la transformation de Fourier, cf remarque après (4.13).

9.2.5 Autres exemples

Les exemples de la section 9.1.3 peuvent être retrouvés ou combinés avec les propriétés de dérivation, d'intégration et de linéarité de la transformation de Laplace pour obtenir d'autres formules (où le $H(t)$ est implicite)

$$\begin{aligned}\mathcal{L}[e^{\pm i\omega t}](p) &= \frac{1}{p \mp i\omega} \\ \mathcal{L}[\cos(\omega t)](p) &= \frac{p}{p^2 + \omega^2} \quad , \quad \mathcal{L}[\sin(\omega t)](p) = \frac{\omega}{p^2 + \omega^2} \\ \mathcal{L}[\cosh(\omega t)](p) &= \frac{p}{p^2 - \omega^2} \quad , \quad \mathcal{L}[\sinh(\omega t)](p) = \frac{\omega}{p^2 - \omega^2},\end{aligned}\tag{9.7}$$

toutes formules initialement valables pour $\Re(p) > 0$, puis prolongeables comme on a vu. De même pour $\Re(p) > \Re(a)$,

$$\mathcal{L}[H(t)e^{at}](p) = \frac{1}{p - a}$$

puis par dérivation

$$\mathcal{L}\left[H(t)e^{at} \frac{t^{n-1}}{(n-1)!}\right](p) = \frac{1}{(p-a)^n}.\tag{9.8}$$

En décomposant en pôles simples toute fraction rationnelle, on reconstruit son “original”, c'est-à-dire la fonction dont elle est la transformée de Laplace, etc etc.

9.3 Transformée de Laplace des distributions

On a vu plus haut le cas de la transformée de Laplace de la fonction de Heaviside. Plus généralement pour une fonction causale f localement sommable, la modification sur un ensemble de mesure nulle ne modifie pas l'intégrale de définition de la transformée de Laplace. Cette dernière est donc attachée à la distribution régulière définie par f . Plus généralement encore, pour T une distribution de support contenu dans \mathbb{R}^+ , ($T \in \mathcal{D}'_+$), telle qu'il existe un α réel tel que pour tout $x > \alpha$, $e^{-xt}T \in \mathcal{S}'$, (distribution tempérée), on définit la transformée de Laplace de T par

$$\widehat{T}(p) = \langle T, e^{-pt} \rangle \quad \text{pour } \Re(p) > \alpha.\tag{9.9}$$

En effet si $x = \Re(p) > \alpha$, il existe un $y : \alpha < y < x$ et $\langle e^{-yt}T, e^{-(p-y)t} \rangle$ existe bien et définit $\langle T, e^{-pt} \rangle$ indépendamment de y .

On appelle encore α abscisse de sommabilité de T , et on démontre qu'à nouveau, $\widehat{T}(p)$ est une *fonction* holomorphe de p dans le demi-plan ouvert $\Re(p) > \alpha$. Exemples : La distribution de Dirac a un support dans \mathbb{R}^+ , elle est tempérée et δe^{-xt} l'est aussi pour tout $x : \alpha = -\infty$ et donc pour tout p , $\mathcal{L}[\delta] = 1$, $\mathcal{L}[\delta^{(m)}] = p^m$, $\mathcal{L}[\delta_{t-a}] = e^{-pa}$.

Les propriétés de la transformation de Laplace étudiées plus haut – convolution, translation, relation Laplace–Fourier, etc– s'étendent aux distributions. Le seul point nécessitant un peu d'attention concerne la dérivation. Si $T \in \mathcal{D}'_+$, et si T' est la *distribution dérivée*, alors $\mathcal{L}[T'](p) = p\hat{T}(p)$, sans le terme supplémentaire qui apparaissait à la Prop. 9.6. Cela est dû au fait que la dérivée au sens des distributions de la distribution régulière $T = H(t)f(t)$ a deux termes : $T' = \delta(t)f(0) + H(t)f'(t)$. Le calcul de la Prop. 9.6 ne retenait que le deuxième terme, celui que nous faisons maintenant au sens des distributions les prend en compte tous les deux.

9.4 Applications de la transformée de Laplace

La transformée de Laplace, comme celle de Fourier, a pour effet de transformer les dérivées en multiples de la fonction. Elle simplifie donc considérablement l'étude des équations différentielles ou aux dérivées partielles et permet de prendre aisément en compte les conditions aux limites, comme on va voir. La transformée de Laplace a de plus l'avantage (par rapport à celle de Fourier) de demander moins de régularité à la fonction : intégrabilité locale et comportement exponentiel à $+\infty$, là où Fourier demande l'intégrabilité au sens L^1 .

9.4.1 Équations différentielles, problème de Cauchy

Rappelons d'abord ce qu'on entend par *problème de Cauchy* pour une équation différentielle. Considérons par exemple le cas d'une variable dynamique $f(t)$ satisfaisant une équation différentielle linéaire du second ordre, à coefficients constants, comme on en rencontre couramment en Mécanique, en Électricité, etc. Cette équation est complétée par *deux* conditions au bord (ou conditions initiales), ce nombre étant bien sûr égal à l'ordre de l'équation. On considère donc le système

$$\begin{aligned} a_2 \ddot{f} + a_1 \dot{f} + a_0 f &= g(t) \\ f(t) \Big|_{t=0} &= f_0 \\ \frac{d}{dt} f(t) \Big|_{t=0} &= f_1. \end{aligned} \tag{9.10}$$

Au final, la fonction f satisfaisant (9.10) est unique, et on dit que le problème de Cauchy admet une solution unique. D'un point de vue physique, cette unicité de la solution pour des conditions initiales données est étroitement liée à la question du *déterminisme* de la physique classique : une fois données la position et la vitesse de départ et les équations du mouvement, la dynamique du système est complètement déterminée à tous les temps ultérieurs.

On connaît le principe de résolution : recherche de la solution générale de l'équation “homogène” (sans second membre g), puis recherche d'une solution particulière de l'équation avec

second membre, enfin détermination des constantes d'intégration en utilisant les conditions initiales.

La transformation de Laplace va nous permettre de mener toutes ces opérations simultanément. Soit $\hat{f}(p)$ la transformée de Laplace de f . On a

$$a_2(p^2\hat{f}(p) - pf_0 - f_1) + a_1(p\hat{f}(p) - f_0) + a_0\hat{f}(p) = \hat{g}(p)$$

d'où l'on tire

$$\hat{f}(p) = \frac{\hat{g}(p)}{a_2p^2 + a_1p + a_0} + \frac{a_2(pf_0 + f_1) + a_1f_0}{a_2p^2 + a_1p + a_0}, \quad (9.11)$$

et il ne reste plus qu'à effectuer une transformation de Laplace inverse pour obtenir $f(t)$

$$f(t) = \frac{1}{2\pi i} \int_{B_x} e^{pt} \frac{\hat{g}(p)}{a_2p^2 + a_1p + a_0} dp + \frac{1}{2\pi i} \int_{B_x} e^{pt} \frac{a_2(pf_0 + f_1) + a_1f_0}{a_2p^2 + a_1p + a_0} dp, \quad (9.12)$$

avec une intégration le long d'une droite de Bromwich. Le premier terme peut être considéré comme une solution particulière de l'équation avec second membre (la solution à $f_0 = f_1 = 0$), tandis que le second est la solution générale (si f_0 et f_1 sont considérés comme des paramètres arbitraires) de l'équation sans second membre. L'intérêt de la méthode est son caractère général et systématique : pas besoin de chercher une solution particulière, (9.12) nous la fournit gracieusement ! En pratique, le calcul explicite des intégrales dans (9.12) est mené avec l'aide du théorème des résidus.

Un exemple simple

Soit à résoudre le système

$$\ddot{f} + f = 2 \cos t \quad f(0) = 0 \quad \dot{f}(0) = -1.$$

Selon ce qui précède, $g(t) = 2 \cos t$, $\hat{g}(p) = 2p/(p^2 + 1)$, donc (9.11) donne

$$\hat{f}(p) = \frac{2p}{(p^2 + 1)^2} - \frac{1}{(p^2 + 1)}$$

dont il faut prendre la transformée de Laplace inverse. On peut faire ce calcul par le théorème des résidus, mais il est plus simple d'observer que $\frac{1}{(p^2+1)} = \mathcal{L}[\sin t]$ et $\frac{2p}{(p^2+1)^2} = -\frac{d}{dp} \frac{1}{(p^2+1)} = \mathcal{L}[t \sin t]$ (cf (9.6-9.7)), d'où la solution du problème $f(t) = (t-1) \sin t$, obtenue avec une grande économie de moyens !

Fonction de transfert, susceptibilité, etc

Examinons la structure générale des formules (9.11-9.12). On appelle *fonction de transfert* la fonction (de la variable de Laplace p)

$$Z(p) = a_2 p^2 + a_1 p + a_0$$

(ou plus généralement, pour une équation différentielle linéaire du n -ième ordre à coefficients constants a_q , $Z(p) = \sum_{q=0}^n a_q p^q$). Les zéros z_k de $Z(p)$ donnent des pôles (simples ou multiples) à $\hat{f}(p)$, qui par transformation de Laplace inverse, donnent des exponentielles $e^{z_k t}$, (fois peut-être des puissances de t , cf (9.8)).

Dans le cas où $Z(p)$ n'a que des pôles simples, $Z(p) = \prod_{k=1}^2 (p - z_k)$, la solution a la forme générale $f(t) = \sum_{k=1}^2 \frac{A_k}{Z'(z_k)} e^{z_k t}$ où A_k incorpore les conditions initiales.

Il est aussi d'usage de définir la susceptibilité $\hat{\chi}(p) = \frac{1}{Z(p)}$, qui décrit la réponse du système f à la source g ; par transformée de Laplace inverse, on construit $\chi(t)$, et la dépendance de la réponse $f(t)$ dans la source g est via une intégrale de convolution $f(t) = \int_0^t ds \chi(t-s) g(s) + \dots$, où les points de suspension contiennent la dépendance dans les conditions initiales, cf (9.12).

9.4.2 Exemple : Circuit LRC

Un exemple typique du problème précédent est celui du circuit LRC se chargeant ou se déchargeant, voir Fig. 9.1. La tension u aux bornes du condensateur satisfait

$$LC\ddot{u} + RC\dot{u} + u = v$$

où v est la tension appliquée aux bornes du circuit. Les conditions initiales spécifient les valeurs de $u(0) = u_0$ et de $\dot{u}(0) = u_1 = -\frac{i_0}{C}$, i_0 le courant initial dans le circuit.

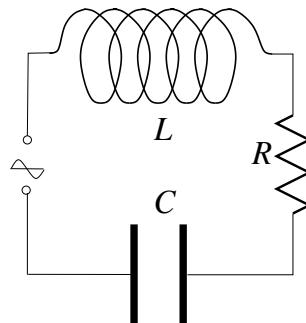


FIGURE 9.1 – Circuit LRC.

On va considérer deux situations différentes.

1. Le circuit n'est pas chargé et le courant y est nul, $u_0 = u_1 = 0$. On le branche au temps

$t = 0$ à un générateur $v = Ve^{i\omega t}$ (comme toujours, on prendra la partie réelle de u à la fin des calculs.)

2. Le circuit est alimenté pour $t \leq 0$ par un générateur $v = V = \text{const}$; au temps $t = 0$, on ouvre le circuit (on déconnecte le générateur). Donc les conditions initiales sont $u_0 = V$, $u_1 = 0$.

La forme générale de la solution (9.11-9.12) fournit, avec $Z(p) = LCp^2 + RCp + 1$

$$\hat{u}(p) = \frac{\hat{v}(p)}{Z(p)} + \frac{LCu_0p + LCu_1 + RCu_0}{Z(p)}. \quad (9.13)$$

Les zéros de $Z(p)$ sont en $z_{\pm} = -\frac{R}{2L} \pm \frac{1}{2L}\sqrt{R^2 - \frac{4L}{C}}$. La susceptibilité du circuit se calcule explicitement, selon le signe de $\Delta = R^2 - 4L/C$ (qu'on supposera non nul)

$$\chi(t) = \frac{1}{2\pi i} \int_{\gamma} \mathrm{d}p e^{pt} \frac{1}{Z(p)} = \sum_{z_{\pm}} \text{Res} \left(\frac{e^{pt}}{Z(p)}, z_{\pm} \right) = H(t) \times \begin{cases} \frac{2}{C\sqrt{\Delta}} e^{-Rt/2L} \sinh(\frac{\sqrt{\Delta}}{2L}t) & \text{si } \Delta > 0 \\ \frac{2}{C\sqrt{-\Delta}} e^{-Rt/2L} \sin(\frac{\sqrt{-\Delta}}{2L}t) & \text{si } \Delta < 0. \end{cases}$$

1. fermeture du circuit : le deuxième terme de la solution (9.13) est nul, il reste

$$u(t) = V \int_0^t \mathrm{d}s \chi(t-s) e^{i\omega s}$$

et on calcule

$$u(t) = \frac{Ve^{i\omega t}}{Z(i\omega)} + \frac{V}{LC(z_+ - z_-)} \left(\frac{e^{z_+ t}}{z_+ - i\omega} - \frac{e^{z_- t}}{z_- - i\omega} \right). \quad (9.14)$$

Aux grands temps, comme $\Re(z_{\pm}) < 0$, seul subsiste le premier terme, proportionnel à la source, tandis que les deux derniers termes décrivent le comportement transitoire du circuit.

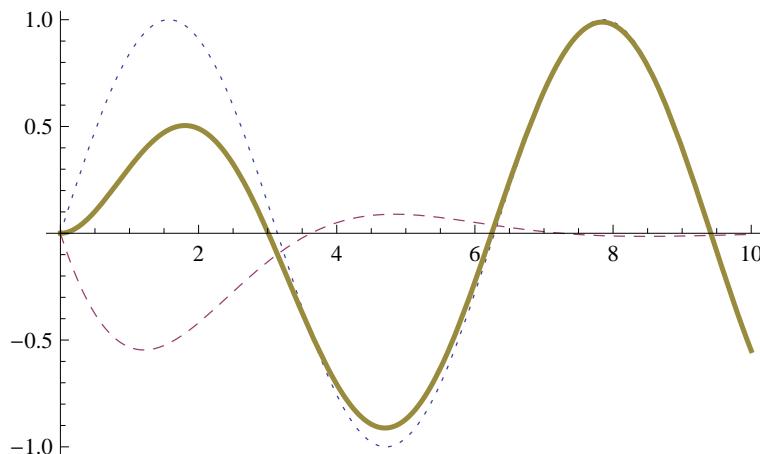


FIGURE 9.2 – Les courbes pointillée, resp. brisée, représentent le premier, resp le second terme de (9.14), et la courbe pleine, la somme, c'est-à-dire la solution $u(t)$. On a pris $R = C = L = \omega = V = 1$.

2. ouverture du circuit : cette fois, l'équation est homogène, seul demeure le deuxième terme de (9.13).

$$\hat{u}(p) = \frac{LCp + RC}{Z(p)}V = \frac{1}{LC(z_+ - z_-)}(LCp + RC) \left(\frac{1}{p - z_+} - \frac{1}{p - z_-} \right) V$$

d'où

$$u(t) = \frac{V}{z_+ - z_-} \left(e^{z_+ t} \left(z_+ + \frac{R}{L} \right) - e^{z_- t} \left(z_- + \frac{R}{L} \right) \right).$$

Là encore, puisque $\Re e(z_{\pm}) < 0$, $u(t) \rightarrow 0$ pour $t \rightarrow \infty$, comme on s'y attend pour ce processus de décharge.

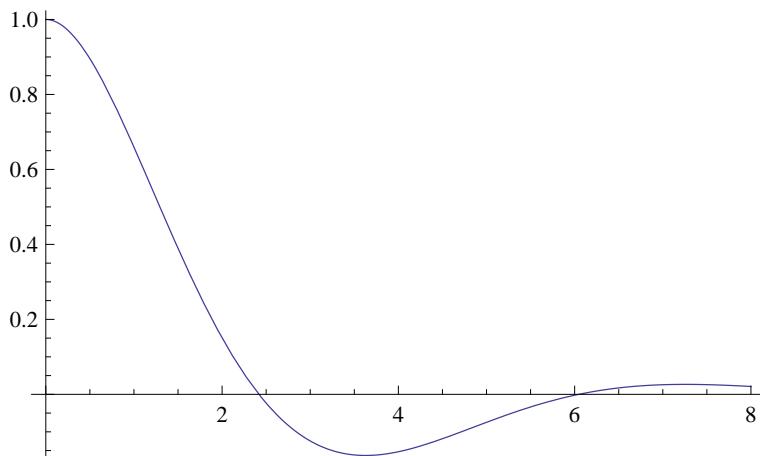


FIGURE 9.3 – La courbe de décharge du circuit, dans le cas (2), toujours avec $R = C = L = \omega = V = 1$.

Il faudrait compléter cette discussion par le cas où $\Delta = 0$, $R^2 = 4LC$. Comme on sait bien et comme on le retrouve ici via la transformation de Laplace, apparaissent alors des fonctions $te^{at}\dots$

9.4.3 Équations linéaires aux dérivées partielles

La méthode faisant appel à la transformation de Laplace peut aussi s'appliquer aussi à des équations différentielles à coefficients non constants, ou à des équations aux dérivées partielles linéaires. Illustrons-le sur le cas de la désintégration radioactive rencontré au chap 5.

On y considérait la fonction génératrice $\phi(x, t) = \sum_{N=0}^{N_0} x^N \mathbf{P}_N(t)$ des probabilités d'avoir N noyaux dans l'état initial au temps t ; on avait montré à l'équation (5.70) que ϕ satisfait l'équation aux dérivées partielles linéaire

$$\frac{\partial}{\partial t} \phi(x, t) = \kappa(1-x) \frac{\partial}{\partial x} \phi(x, t). \quad (5.70)$$

avec des conditions aux limites qu'on va préciser. Considérons la transformée de Laplace $\hat{\phi}(x, p)$ de $H(t)\phi(x, t)$ par rapport à la variable t (ce qui est naturel puisque dans ce problème $t \geq 0$). Elle satisfait l'équation

différentielle ordinaire ($\hat{\phi}'$ est la dérivée par rapport à x)

$$\kappa(1-x)\hat{\phi}'(x,p) = p\hat{\phi}(x,p),$$

équation “à variables séparées” qu’on intègre en

$$\hat{\phi}(x,p) = \hat{\phi}(0,p)(1-x)^{-p/\kappa}. \quad (9.15)$$

C’est le moment de préciser les conditions aux limites. À $x = 0$, $\phi(0,t) = \mathbf{P}_0(t) = (1 - e^{-\kappa t})^{N_0}$ puisque chaque noyau a la probabilité $1 - e^{-\kappa t}$ de s’être désintégré au temps t et que ces noyaux se désintègrent de façon indépendante. Donc $H(t)\phi(0,t) = \sum_{N=0}^{N_0} C_{N_0}^N (-1)^N e^{-N\kappa t} H(t)$ dont la transformée de Laplace est $\hat{\phi}(0,p) = \sum_{N=0}^{N_0} C_{N_0}^N (-1)^N \frac{1}{p + \kappa N}$. Après insertion dans (9.15) et transformation de Laplace inverse, on obtient pour $t \geq 0$ et $\forall a > 0$

$$\begin{aligned} \phi(x,t) &= \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} \sum_{N=0}^{N_0} C_{N_0}^N (-1)^N \frac{1}{p + \kappa N} (1-x)^{-p/\kappa} e^{pt} dp \\ &= \sum_{N=0}^{N_0} C_{N_0}^N (-1)^N (1-x)^N e^{-N\kappa t} \\ &= \left(1 - (1-x)e^{-\kappa t}\right)^{N_0} \end{aligned} \quad (9.16)$$

qui est le résultat obtenu en (5.71), d’où l’on tire la probabilité cherchée $\mathbf{P}_N(t) = C_{N_0}^N (1 - e^{-\kappa t})^{N_0-N} e^{-\kappa t N}$.

On voit que la transformée de Laplace nous a permis de réduire une équation aux dérivées partielles (PDE dans l’acronyme anglo-saxon) en une équation différentielle ordinaire (ODE), et de déduire la solution à x fini de celle à $x = 0$.

Lectures complémentaires

J’ai suivi la discussion de W. Appel [1] complétée par celle de L. Schwartz [8], qu’on pourra consulter pour plus de détails. Le livre de C. Aslangul [2] contient de très nombreuses applications physiques.

Chapitre 10

Aspects de la théorie des groupes

Rôle fondamental de la théorie des groupes en physique : transformations d'un système (par rotations, réflexions, translations...), éventuellement *invariances*. Groupe des transformations, groupe d'invariance d'un système donné.

Différents types de groupes et leur importance en physique

- groupes finis : par ex. groupe d'invariance de rotation d'une molécule ou d'un cristal, sous-groupe du groupe $SO(3)$ des rotations de \mathbb{R}^3
- groupes infinis discrets : par ex. groupe d'invariance d'un cristal infini (rotations, réflexions, translations,...)
- groupes continus *compacts* : groupes $U(n)$, $SU(n)$, $O(n)$, $SO(n), \dots$ ($U(1)$ groupe "de jauge" de l'électrodynamique, groupe $SO(3)$ des rotations, groupes $SU(2)$ et $SU(3)$ en physique des particules, etc)
- groupes continus *non compacts* : groupe de Galilée d'invariance de la Mécanique classique, groupes de Lorentz et de Poincaré de la Relativité restreinte, ...

On va se borner à introduire et étudier sommairement deux concepts fondamentaux :

- *représentations linéaires d'un groupe*

avec la notion de représentation irréductible et le lemme de Schur ;

- *groupes de Lie et algèbres de Lie*

générateurs infinitésimaux, et représentations de l'algèbre de Lie

et les illustrer sur l'exemple de l'

- *algèbre de Lie de $SO(3)$ et ses représentations*

ce qui fera la jonction avec le cours de Mécanique Quantique.

10.1 Représentations linéaires des groupes

10.1.1 Définition

Dans de nombreuses situations en physique ou en mathématiques, on a à considérer l'action d'un groupe G dans un espace vectoriel E . Le groupe agit par des opérations linéaires du groupe linéaire de E , noté $\mathrm{GL}(E)$ (ou plus précisément $\mathrm{GL}(E, \mathbb{R})$ ou $\mathrm{GL}(E, \mathbb{C})$, selon les cas). À tout élément $g \in G$ on associe un opérateur linéaire inversible $D(g) \in \mathrm{GL}(E)$. E peut être l'espace physique, comme \mathbb{R}^3 , ou un espace de tenseurs sur \mathbb{R}^3 , ou être plus général encore. (Il faut bien voir qu'un groupe peut aussi agir sur des ensembles qui ne sont pas des e.v., cf par exemple l'action du groupe des rotations $\mathrm{SO}(3)$ sur la sphère $S^2 \subset \mathbb{R}^3$.)

Il est naturel d'imposer que la loi de composition du groupe est compatible avec celle dans le groupe $\mathrm{GL}(E)$, ou en termes plus mathématiques, que l'application $G \rightarrow \mathrm{GL}(E) : g \mapsto D(g)$ est un *homomorphisme de groupes*

$$\forall g_1, g_2 \in G \quad D(g_1 \cdot g_2) = D(g_1)D(g_2). \quad (10.1)$$

On appelle *représentation du groupe G dans l'espace E* un tel homomorphisme $G \rightarrow \mathrm{GL}(E)$. Noter que (10.1) implique que si e est l'élément neutre de G , $D(e) = I$, l'opérateur identité dans E , et que $D(g)^{-1} = D(g^{-1})$.

L'étude et la classification des représentations des groupes est un domaine important des mathématiques, aux très nombreuses applications physiques.

10.1.2 Représentations réductibles ou irréductibles

Supposons qu'une représentation $G \rightarrow (E, D)$ laisse un sous-espace E_1 de E invariant

$$\forall x \in E_1 \quad \forall g \in G \quad D(g)x \in E_1.$$

Une telle représentation est dite *réductible*. Dans le cas contraire (pas de sous-espace invariant) on dit que la représentation est *irréductible*. Dans les cas qui vont nous occuper (groupes finis ou continus compacts), on peut toujours se ramener à une situation où une représentation réductible D laisse aussi un sous-espace E_2 supplémentaire de E_1 invariant. Autrement dit (si E est de dimension finie), la matrice de $D(g)$ s'écrit dans une certaine base sous forme de blocs

$$D(g) = \begin{pmatrix} D_1(g) & 0 \\ 0 & D_2(g) \end{pmatrix}$$

où D_1 , resp. D_2 , est une représentation de G dans E_1 , resp. E_2 .

Inversement on voit que toute paire de représentations D_1 et D_2 d'un groupe G donné, dans des espaces E_1 et E_2 permet de construire une autre représentation dans l'espace $E_1 \oplus E_2$ somme directe de E_1 et E_2 . Il va suffire de savoir construire (et classifier si possible) les représentations irréductibles pour construire les représentations les plus générales.

10.1.3 Lemme de Schur

Un lemme important que nous admettrons énonce que

Lemme de Schur : *Dans une représentation irréductible de G dans $\mathrm{GL}(E, \mathbb{C})$, tout opérateur commutant avec tous les $D(g)$ est un multiple de l'identité.*

Donc si $\exists V$ tel que $\forall g \in G$, $D(g)V = VD(g)$, alors $V = \lambda I$. Le lemme est en défaut si la représentation n'est pas irréductible, ou si elle est sur les réels. Exemple $G = SO(2)$, $D(\alpha) = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$, cf ci-dessous, groupe commutatif (“abélien”), donc $D(\alpha)D(\beta) = D(\beta)D(\alpha)$ semble un contre-exemple au lemme, mais ici la représentation n'est pas irréductible sur \mathbb{C} , puisque les sous-espaces colinéaires aux vecteurs $\begin{pmatrix} 1 \\ \pm i \end{pmatrix}$ sont invariants par les $D(\alpha)$.

10.2 Groupes et algèbres de Lie

Un *groupe de Lie* est un groupe “topologique”¹ doté de propriétés de différentiabilité. On va donc être naturellement conduit à y considérer des opérations infinitésimalement proches de l'identité. La propriété remarquable d'un groupe de Lie est que la connaissance de ces transformations infinitésimales et de l'*algèbre de Lie* qu'elles forment suffit essentiellement à reconstruire tout le groupe. C'est ce que nous allons illustrer sur trois exemples de complexité croissante.

10.2.1 Groupes R , $U(1)$ et $SO(3)$. Algèbre de Lie

Le plus simple des groupes continus est celui des translations à une dimension : $t(a) : x \mapsto x' = x + a$, avec x et a des variables réelles $x, a \in \mathbb{R}$. Pour toute fonction réelle $f(x)$, définissons la nouvelle fonction $f' = t(a)f$ par $f'(x') = f(x)$ ou de façon équivalente, $f'(x) = f(x - a)$. La loi de composition de ces transformations (loi de groupe) est

$$t(b) \circ t(a) = t(a + b). \quad (10.2)$$

1. Un groupe topologique est un groupe muni d'une topologie, telle que les opérations de produit et de passage à l'inverse soient continues.

La loi est commutative, le groupe est abélien. Ce groupe est en fait \mathbb{R} avec la loi de groupe fournie par l'addition des réels. C'est un groupe de dimension (réelle) égale à 1.

Si a est infinitésimal, on peut écrire $f'(x) \approx f(x) - a \frac{d}{dx} f(x)$ ou encore

$$f'(x) = \left(1 - a \frac{d}{dx}\right) f(x)$$

soit encore, en définissant la variation de la fonction

$$\Delta f(x) = f'(x) - f(x) = -a \left(\frac{d}{dx}\right) f(x) .$$

La variation infinitésimale de la fonction est linéaire dans le paramètre a (on travaille au 1er ordre!) et dans la fonction f , et donnée par l'action du *générateur infinitésimal* $T_a = -aT = -a \left(\frac{d}{dx}\right)$.

Noter que la loi de groupe (10.2) et sa propriété de commutativité se traduisent au niveau infinitésimal par la commutativité des générateurs

$$[T_a, T_b] = T_a T_b - T_b T_a = 0 .$$

Les transformations continues (et différentiables) du groupe impliquent donc l'existence des générateurs infinitésimaux avec des propriétés qui reflètent la loi de groupe. Inversement une fois connu le générateur infinitésimal $T = -\frac{d}{dx}$, il est possible de reconstruire la transformation finie par *action exponentielle*

$$\begin{aligned} \exp(-aT)f(x) &= \left(1 - aT + \frac{1}{2}a^2T^2 + \frac{-a^3}{3!}T^3 + \dots\right)f(x) \\ &= \left(1 - a \left(\frac{d}{dx}\right) + \frac{a^2}{2} \left(\frac{d}{dx}\right)^2 + \frac{(-a)^3}{3!} \left(\frac{d}{dx}\right)^3 + \dots\right)f(x) \\ &= f(x - a) \end{aligned}$$

qui n'est autre que la série de Taylor, que nous supposons convergente (f analytique réelle).

Ces considérations s'étendent sans difficulté à des translations $t(\mathbf{a})$ dans l'espace \mathbb{R}^d et à des fonctions f des d variables (coordonnées) x_1, \dots, x_d . La loi de groupe est toujours additive et donc commutative $t(\mathbf{a}) \circ t(\mathbf{b}) = t(\mathbf{a} + \mathbf{b}) = t(\mathbf{b}) \circ t(\mathbf{a})$, le groupe n'est autre que \mathbb{R}^d avec son addition. L'opérateur infinitésimal est maintenant l'opérateur différentiel $T_{\mathbf{a}} = -\mathbf{a} \cdot \mathbf{T}$ avec \mathbf{T} le vecteur gradient $\mathbf{T} = \vec{\nabla} = \left(\frac{d}{dx_1}, \dots, \frac{d}{dx_d}\right)$, et la formule de Taylor à d variables s'applique.

10.2.2 Groupe de rotation à deux dimensions $\text{SO}(2)$

Considérons maintenant le groupe des rotations à deux dimensions, noté $\text{SO}(2)$. L'élément générique $R(\alpha)$ dépend d'un paramètre continu, l'angle α de rotation défini modulo 2π . La

rotation $R(\alpha)$ agit sur les vecteurs \mathbf{x} du plan par

$$\mathbf{x} \mapsto \mathbf{x}' = R(\alpha)\mathbf{x} \quad (10.3)$$

où la notation doit être comprise comme l'action d'un opérateur linéaire sur le vecteur \mathbf{x} , ou encore, dans un repère orthogonal, comme l'action d'une matrice orthogonale sur les composantes x_1, x_2 de \mathbf{x}

$$\mathbf{x}' = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (10.4)$$

(la matrice est orthogonale par définition de $\text{SO}(2)$, ou encore, géométriquement, parce que les transformations considérées préservent la norme $\|\mathbf{x}'\| = \|\mathbf{x}\|$, or $\mathbf{x}' \cdot \mathbf{x}' = x^T(R^T(\alpha) \cdot R(\alpha))x = \mathbf{x} \cdot \mathbf{x}$ donc $R^T(\alpha) \cdot R(\alpha) = \mathbb{1}$). Ces rotations se composent comme on sait bien

$$R(\alpha)R(\beta) = R(\alpha + \beta) = R(\beta)R(\alpha) \quad (10.5)$$

de façon à nouveau additive (en α) et commutative.

On peut aussi étudier ce groupe des rotations du plan autour de l'origine à l'aide de la coordonnée complexe $z = x_1 + ix_2$. La transformation (10.4) s'exprime par $z' = e^{i\alpha}z$ et on parle alors du groupe $\text{U}(1)$, groupe multiplicatif des nombres de module 1. On a donc l'isomorphisme de groupes $\text{SO}(2) \cong \text{U}(1)$.

Dans le groupe $\text{SO}(2)$, les générateurs infinitésimaux sont des opérateurs, ou des matrices 2×2 . En effet, prenant α infinitésimal et développant au premier ordre, $\cos \alpha \approx 1$, $\sin \alpha \approx \alpha$, donc

$$\alpha \ll 1 \quad R(\alpha) = \mathbb{1} - i\alpha J = \mathbb{1} - i\alpha \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{avec} \quad J = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad (10.6)$$

où l'apparition du i peut paraître curieuse dans ce problème réel, mais est utile si on veut considérer le *générateur infinitésimal* J comme hermitien (plutôt qu'antisymétrique réel donc antihermitien).

Considérons maintenant la loi de groupe (10.5) pour $\beta = d\alpha$ infinitésimal, en faisant usage de (10.6) :

$$R(\alpha + d\alpha) = R(d\alpha)R(\alpha) = (\mathbb{1} - id\alpha J)R(\alpha) = R(\alpha) - id\alpha JR(\alpha)$$

donc $dR(\alpha) = -id\alpha JR(\alpha)$. On obtient donc l'équation différentielle

$$\frac{d}{d\alpha} R(\alpha) = -iJR(\alpha) \quad (10.7)$$

où J est, rappelons le, une matrice fixe, indépendante de α . Cette équation différentielle du premier ordre à coefficients constants et homogène, complétée par la condition initiale $R(0) = \mathbb{1}$, se résout immédiatement en

$$R(\alpha) = \exp -i\alpha J. \quad (10.8)$$

Exercice. Pour $J = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$, vérifier que l'exponentiation de $-i\alpha J$ reproduit bien la matrice de $R(\alpha)$ donnée plus haut. (Indication : calculer les puissances successives J^2, J^3, \dots, J^n et construire $\exp i\alpha J$ par son développement en série.)

L'action (10.3) des rotations de $\text{SO}(2)$ sur les vecteurs $\mathbf{x} \in \mathbb{R}^2$ nous donne un exemple de représentation. Cette représentation est irréductible (sur les réels). D'autres représentations (irréductibles ou non) seraient fournies par l'action de $\text{SO}(2)$ sur des tenseurs.

10.2.3 Générateurs infinitésimaux de $\text{SO}(3)$

Étudions maintenant le groupe $\text{SO}(3)$ des rotations de l'espace euclidien \mathbb{R}^3 et ses générateurs infinitésimaux. Une rotation de $\text{SO}(3)$ est spécifiée par son axe porté par un vecteur unitaire \hat{n} et l'angle de rotation ψ . On la notera $R_{\hat{n}}(\psi)$. Attention que $R_{\hat{n}}(\psi) = R_{-\hat{n}}(-\psi)$, il faut donc choisir $0 \leq \psi \leq \pi$, et \hat{n} quelconque sur la sphère unité, $\hat{n} \in S^2$. Une rotation de \mathbb{R}^3 dépend donc de 3 paramètres réels, et on dit que le groupe $\text{SO}(3)$ est de dimension 3.

Pour un angle $d\psi$ infinitésimal, $R_{\hat{n}}(d\psi)$ diffère de l'opérateur identité par un terme d'ordre $d\psi$, définissant le *générateur infinitésimal* (un opérateur). Comme dans le cas de $\text{SO}(2)$ nous suivons l'usage des physiciens en écrivant les générateurs infinitésimaux des rotations comme des opérateurs hermitiens $J = J^\dagger$. Ainsi on écrit au premier ordre en $d\psi$

$$R_{\hat{n}}(d\psi) = (I - id\psi J_{\hat{n}}) \quad (10.9)$$

où $J_{\hat{n}}$ est le générateur de ces rotations², une matrice 3×3 , qui est hermitique comme conséquence de l'unitarité (ou simplement ici, de l'orthogonalité) de $R_{\hat{n}}$: $R_{\hat{n}} R_{\hat{n}}^\dagger = I$.

Pour trouver l'expression explicite de ces générateurs infinitésimaux, on peut les considérer soit comme des matrices 3×3 agissant sur les composantes des vecteurs \mathbf{x} de \mathbb{R}^3 (comme on a fait pour $\text{SO}(2)$), soit comme des opérateurs différentiels agissant sur les fonctions de \mathbf{x} (cf le cas du groupe \mathbb{R}). Pour une rotation infinitésimale $R_{\hat{n}}(d\psi)$ agissant sur \mathbf{x} , le vecteur transformé \mathbf{x}' est tel que $\mathbf{x}' - \mathbf{x}$ est orthogonal à \mathbf{x} et à \hat{n} , donc s'écrit

$$\mathbf{x}' = \mathbf{x} + d\psi \hat{n} \wedge \mathbf{x}. \quad (10.10)$$

Une fonction scalaire de \mathbf{x} se transforme selon $f'(\mathbf{x}') = f(\mathbf{x})$ soit pour une rotation infinitésimale

$$f'(\mathbf{x}) = f(R^{-1}\mathbf{x}) = f(\mathbf{x} - d\psi \hat{n} \wedge \mathbf{x}) = (1 - d\psi \hat{n} \cdot \mathbf{x} \wedge \vec{\nabla})f(\mathbf{x}) = (1 - id\psi \hat{n} \cdot \mathbf{J})f(\mathbf{x}) \quad (10.11)$$

avec $J_{\hat{n}} = -i\hat{n} \cdot \mathbf{x} \wedge \frac{\partial}{\partial \mathbf{x}}$ soit encore

$$\boxed{J_{\hat{n}} = \hat{n} \cdot \mathbf{J} \quad \text{avec} \quad \mathbf{J} = -i\mathbf{x} \wedge \frac{\partial}{\partial \mathbf{x}}}. \quad (10.12)$$

2. Ne pas confondre $J_{\hat{n}}$ indexé par le vecteur \hat{n} , avec J_k , k ième composante de \mathbf{J} . La relation entre les deux va être donnée ci-dessous.

Remarque. Dans le cours de mécanique quantique, on s'est intéressé à l'opérateur de moment cinétique orbital \mathbf{L} qui diffère du présent générateur \mathbf{J} par un facteur \hbar : $\mathbf{L} = \hbar\mathbf{J}$. Attention ! \mathbf{J} , opérateur de moment angulaire total du cours de mécanique quantique, ne doit pas être confondu avec le présent \mathbf{J} .

Par ailleurs, introduisons les trois matrices de base J_1 , J_2 et J_3 qui décrivent les rotations infinitésimales autour des axes correspondants. De (10.10) on tire

$$J_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad J_2 = \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} \quad J_3 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (10.13)$$

ce qu'on peut exprimer par une formule unique

$$(J_k)_{ij} = -i\epsilon_{ijk} \quad (10.14)$$

à l'aide du tenseur complètement antisymétrique ϵ_{ijk} . Les trois matrices J_i , $i = 1, 2, 3$ satisfont les très importantes relations de commutation suivantes

$$\boxed{[J_i, J_j] = i\epsilon_{ijk}J_k} \quad (10.15)$$

qu'on vérifie aussi sur la forme (10.12). On a utilisé la convention de sommation sur les indices répétés, il y a donc un $\sum_{k=1}^3$ implicite au membre de droite de (10.15). Par exemple (10.12) nous donne $J_1 = J_x = -i(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y})$, $J_2 = J_y = -i(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z})$ d'où on tire $[J_1, J_2] = iJ_3$.

Les générateurs infinitésimaux étant déterminés, comment reconstruit-on les rotations finies ? Les rotations $R_{\hat{n}}(\psi)$ autour d'un axe \hat{n} donné forment un sous-groupe (on parle de *sous-groupe à un paramètre*) ; ces matrices commutent, le sous-groupe est commutatif (ou *abélien*). Par la propriété de groupe,

$$R_{\hat{n}}(\psi + d\psi) = R_{\hat{n}}(d\psi)R_{\hat{n}}(\psi) = (I - id\psi J_{\hat{n}})R_{\hat{n}}(\psi), \quad (10.16)$$

ou encore

$$\frac{\partial R_{\hat{n}}(\psi)}{\partial \psi} = -iJ_{\hat{n}}R_{\hat{n}}(\psi) \quad (10.17)$$

équation différentielle, qui, compte tenu de $R(0) = I$, s'intègre en

$$R_{\hat{n}}(\psi) = e^{-i\psi J_{\hat{n}}}. \quad (10.18)$$

Comme l'a montré (10.12), on peut toujours écrire

$$J_{\hat{n}} = \sum_k n_k J_k \quad \text{donc} \quad R_{\hat{n}}(\psi) = e^{-i\psi \sum_k n_k J_k}. \quad (10.19)$$

Attention ! En raison de la non-commutation des générateurs infinitésimaux J_1 , J_2 et J_3 , cf (10.15), la propriété “bien connue” de la fonction exponentielle $e^a e^b = e^{a+b}$ ne s’applique pas : il ne faut pas écrire $e^{-i\psi \sum_k n_k J_k} = \prod_k e^{-i\psi n_k J_k}$.

Une autre façon d’écrire une rotation finie en termes des générateurs J_i est d’utiliser la paramétrisation des angles d’Euler, cf Fig. 10.1. On a $R(\alpha, \beta, \gamma) = R_z(\alpha)R_y(\beta)R_z(\gamma)$, avec $R_y = e^{-i\beta J_y}$, $R_z(\alpha) = e^{-i\alpha J_z}$.

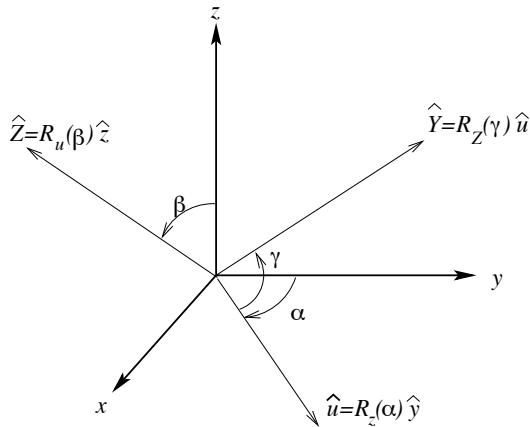


FIGURE 10.1 – Angles d’Euler

10.3 Algèbre de Lie de $\text{SO}(3)$ et ses représentations

10.3.1 Algèbre de Lie $\text{so}(3)$

Récapitulons : nous venons d’introduire l’algèbre de commutation des générateurs infinitésimaux (ou *algèbre de Lie*) du groupe $\text{SO}(3)$, notée $\text{so}(3)$. Elle est définie par les relations (10.15). On utilise aussi beaucoup les trois combinaisons

$$J_z \equiv J_3, \quad J_+ = J_1 + iJ_2, \quad J_- = J_1 - iJ_2 . \quad (10.20)$$

Il est immédiat de calculer

$$[J_3, J_+] = J_+ \quad (10.21a)$$

$$[J_3, J_-] = -J_- \quad (10.21b)$$

$$[J_+, J_-] = 2J_3 . \quad (10.21c)$$

On vérifie aussi que l’opérateur

$$\mathbf{J}^2 = J_1^2 + J_2^2 + J_3^2 = J_3^2 + J_3 + J_- J_+ \quad (10.22)$$

commute avec tous les J

$$[\mathbf{J}^2, J_\cdot] = 0 . \quad (10.23)$$

\mathbf{J}^2 est l'*opérateur de Casimir* de l'algèbre de Lie $\text{so}(3)$.

L'expression des générateurs infinitésimaux a été trouvée dans la représentation (10.13) du groupe $\text{SO}(3)$ sur les vecteurs de \mathbb{R}^3 (la représentation de définition de $\text{SO}(3)$) ou dans celle (10.12) sur les fonctions. On démontre que dans toute représentation du groupe, les générateurs infinitésimaux J_i , $i = 1, 2, 3$, satisfont les mêmes relations de commutation. De même qu'une représentation d'un groupe était définie par un homomorphisme D du groupe G dans un groupe linéaire $\text{GL}(E)$, de même on appelle *représentation de l'algèbre de Lie $\text{so}(3)$* un homomorphisme d de $\text{so}(3)$ dans l'espace $\text{End}(E)$ des opérateurs linéaires sur un certain e.v. E , c'est-à-dire une application de $\text{so}(3)$ dans $\text{End}(E)$ compatible avec les relations de commutation, donc satisfaisant

$$d[J_i, J_j] = [d(J_i), d(J_j)] = i\epsilon_{ijk}d[J_k] \quad (10.24)$$

(toujours avec la sommation sur k implicite). Dans la suite, on omettra le d et on parlera de la représentation de $\text{so}(3)$ par les générateurs infinitésimaux J_i agissant dans l'espace E .

De même que les transformations finies des groupes de Lie $\text{SO}(2)$ et $\text{SO}(3)$ ont été reconstruites par exponentiation des transformations infinitésimales, de même on peut espérer que la connaissance des représentations de l'algèbre de Lie suffit à reconstruire les représentations du groupe. C'est vrai ... à une petite subtilité près que nous allons voir plus bas dans le cas de $\text{SO}(3)$.

En physique, nous sommes surtout intéressés par les "représentations unitaires", où les générateurs J_i , $i = 1, 2, 3$ sont hermitiques, donc

$$J_i^\dagger = J_i, \quad i = 1, 2, 3 \quad J_\pm^\dagger = J_\mp . \quad (10.25)$$

L'importance de l'opérateur de Casimir \mathbf{J}^2 est liée au lemme de Schur : \mathbf{J}^2 commutant avec tous les générateurs infinitésimaux, il commute avec leur exponentielle, donc avec les opérateurs de rotation, et ce, dans toute représentation. Dans une représentation irréductible, \mathbf{J}^2 est donc un multiple de l'identité, $\mathbf{J}^2 = \lambda I$. L'opérateur \mathbf{J}^2 étant semi-défini positif, ce coefficient λ est réel non négatif et on peut l'écrire sous la forme $j(j+1)$, j réel ≥ 0 .

La construction des représentations de l'algèbre $\text{so}(3)$ consiste alors à considérer des vecteurs propres de J_z et à montrer que leur valeur propre m est bornée : $-j \leq m \leq j$, et que $j \pm m$ doivent être des entiers, donc que j et m sont simultanément entiers ou demi-entiers, cf le cours de Mécanique Quantique.

10.3.2 Représentations de l'algèbre $so(3)$ et représentations du groupe $SO(3)$.

L'étude qui a été menée dans le cours de Mécanique Quantique a donc eu pour effet de construire *toutes* les représentations irréductibles de l'algèbre de Lie $so(3)$.

Théorème 10.1 : *Une représentation irréductible de l'algèbre de Lie $so(3)$ est spécifiée par un entier ou demi-entier non négatif, le spin j de la représentation. La représentation de spin j est de dimension $2j + 1$, et une base est fournie par les états propres de J_z notés $|j, m\rangle$ avec $m = -j, -j + 1, \dots, j - 1$. Parmi ces représentations, seules celles de spin j entier sont des représentations du groupe $SO(3)$.*

Le point subtil de cette construction est que toutes ces représentations de l'algèbre de Lie ne fournissent pas par exponentiation une bonne représentation (au sens de la Déf. du 10.1.1) du groupe des rotations $SO(3)$. Cela se voit aisément sur la façon dont une rotation de 2π , en principe l'identité, est représentée : $D(R_z(2\pi)) = \exp -2\pi i J_z = (1)^{2j} I$ qui vaut I si le spin j est entier mais $-I$ s'il est demi-entier. À la rotation identité, ce D associe la matrice $-I$. Ces D ne forment donc pas vraiment une représentation du groupe $SO(3)$ et on parle de représentation au signe près. La cause profonde de cette curiosité, la relation entre les groupes $SO(3)$ et $SU(2)$, le rôle de leur topologie en tant que variétés (simplement connexes ou non), et finalement la raison de l'apparition de ces représentations à un signe près en Mécanique Quantique (théorème de Wigner), autant de sujets passionnants ... que nous ne pourrons pas aborder ici !

★

Lectures supplémentaires

Il existe de nombreux textes d'introduction à la théorie des groupes pour les physiciens. On pourra consulter les notes de l'un des cours que j'ai donnés sur le sujet, sur ma page http://www.lpthe.jussieu.fr/~zuber/Z_Notes.html

Index

abélien, groupe, 189
abscisse de sommabilité, 170
absolument continue, fonction, 79
accumulation, point d', 2
achevée, droite réelle, 3, 13
additivité dénombrable, 21
adhérence, 4
adjacentes, suites, 2
 \aleph_0 (aleph zéro), 11
algèbre, 53
algèbre de Lie $so(3)$, 190
analytique, fonction, 121
argument \arg , fonction, 120

Banach, espace de, 6
Bayes, axiome de, 73
Bessel, fonctions de, 147
Bolzano–Weierstrass, théorème de, 3
bord orienté, 133
Borel, tribu de, 21
borélien, 21
Bromwich, droite de, 173
Buffon, aiguille de, 96

Cantor, ensemble de, 12, 79
Cantor, fonction de, 79
caractéristique, fonction, 84
cardinal, 11
Casimir, opérateur de, 191
Cauchy, intégrale de, 140
Cauchy, partie principale de, 41

Cauchy, règle de, 116
Cauchy, suite de, 1, 6
Cauchy, théorème de, 139
Cauchy–Riemann, conditions de, 128, 130
changement de variable, 32
chemin, 131
circuit LRC, 179
col, méthode du, 163
compact, espace, 5
complet, ensemble, 3
connexe, espace, 5
continue, fonction, 4, 14
contour d'intégration, 137
convergence dominée, théorème de, 27
convergence en probabilité, 100
convergence en loi, 100
convergence normale (CV.N.), 10, 116
convergence presque sûrement, 100
convergence simple (CV.S.), 7, 10
convergence uniforme (CV.U.), 7, 10, 116
convolution de distributions, 49
convolution de fonctions, 48
convolution et transformation de Fourier, 65
coordonnées cylindriques, 33
coordonnées polaires, 33
coordonnées sphériques, 33
corrélation, fonction de, 81
covariance, 76
cumulant, 76

- d'Alembert, théorème de, 143
 décroissance rapide, fonction à, 63
 dénombrable, 11
 dense, ensemble, 3, 4
 densité de probabilité, 80
 densité de probabilité jointe, 81
 désintégration radioactive, 97, 181
 difféomorphisme, 32
 différentiable, fonction, 127
 différentielle, forme, 127
 Dirac, distribution de, 36, 41
 Dirac, masse de, 22
 Dirac, peigne de, 36, 41
 Dirichlet, fonction de, 23
 Dirichlet, théorème de, 57
 disque de convergence, 116
 distribution, 35, 39
 distribution marginale, 81
 distribution régulière, 40
 distribution singulière, 41
 distribution tempérée, 67
 domination, 27, 29
 dual, espace, 39, 53
- écart-type, 75
 équation de la chaleur, 110
 escalier du Diable, 79
 escalier, fonction en, 14, 24
 espérance, 75
 espace des épreuves, 72
 espace topologique (e.t.), 4
 étagée, fonction, 24
 Euler, angles d', 190
 Euler, formule d', 116
 événement atomique, 72
 événement composé, 72
- événement élémentaire, 72
 fermé, ensemble, 3, 4
 fermé, intervalle, 2
 fonction Γ , 125, 166
 fonction causale, 169
 fonction d'erreur complémentaire erfc, 106
 fonction d'erreur erf, 106
 fonction de transfert, 179
 fonction exponentielle, 119
 fonction logarithme, 120
 fonction multivaluée, 120
 fonction saut, 43
 fonction signe, 44
 fonction-test, 38
 forme exacte, 134
 forme linéaire, 39, 53
 Fourier, série de, 57
 Fourier, transformée de, 58
 fraction continue, 1
 frontière, 11
 frontière essentielle, 125
 Fubini, théorème de, 26
- générateur infinitésimal, 186
 générateur infinitésimal des rotations, 188
 génératrice, fonction, 85
 Γ , fonction , 125, 166
 Green, fonction de, 51
 Green–Riemann, formule de, 135
- harmonique, fonction, 51
 Hausdorff, espace de, 4
 Heaviside, fonction de, 43
 Heine–Borel, théorème de, 5
 hermitienne, fonction, 62
 histogramme, 80

- holomorphe, fonction, 128
homomorphisme de groupe, 119, 184
homomorphisme de groupes, 119
homotopes, chemins, 132

impropre, intégrale, 18
indépendants, événements, 74
indicatrice, fonction, 11
indice d'un chemin, 138
intégrable au sens de Lebesgue, fonction, 25
irréductible, représentation, 184

Jacobien, 32

Kolmogorov, axiomes de, 72
kurtosis, 76

Laplace, transformation de, 169
Laplace, transformation inverse, 173
Laurent, série de, 146
Lebesgue, critère de, 30
Lebesgue, intégrale de, 19, 24
Lebesgue, mesure de, 21
Lebesgue, théo. de convergence dominée, 27
Lie, groupe de, 185
limite supérieure, inférieure, 13
Liouville, théorème de, 143
lisse, fonction, 39
localement compact, espace, 5
localement intégrable, 40
logarithme complexe, 120
loi binomiale, 85, 87
loi de probabilité, 74
loi de Poisson, 91
loi des grands nombres, 101
loi forte des grands nombres, 101
loi normale, 88
lois d'échelle, 112

marche aléatoire, 108
maximum, principe du, 143
Maxwell–Boltzmann, distribution de, 97
médiane, 76
méromorphe, 145
mesurable, ensemble, 23
mesurable, espace, 22
mesurable, fonction, 23
mesure, définition, 21
mesuré, espace, 22
mode, 80
moments, 75
Monte-Carlo, 103
mouvement brownien, 108

négligeable, ensemble, 22
norme, 5
normé, espace, 5

ondes, équation des, 55
ordre d'un pôle, 147
ordre d'un zéro, 144
original, 170
ouvert, ensemble, 3, 4
ouvert, intervalle, 2

pôle, 145
pôle simple, 147
Parseval–Plancherel, théorème de, 64
partie principale de Cauchy, 41
partitions, 116
PDF, 80
phase stationnaire, méthode de la, 166
pointé, plan, 120
Poisson, équation de, 51
presque partout, 23
presque sûr, p.s., 73

- primitive, 17
- primitive d'une forme, 134
- principe de symétrie de Schwarz, 142
- probabilité conditionnelle, 73
- prolongement analytique, 123
- propriété de Markov, 110
- réglée, fonction, 17
- répartition, fonction de, 77
- résidu, 148
- résidus, théorème des, 148
- rayon de convergence, 116
- représentation à un signe près, 192
- représentation d'un groupe, 184
- représentation de l'algèbre $so(3)$, 191
- représentation réductible, irréductible, 184
- Riemann, intégrale de, 17
- Riemann, sphère de, 5, 149
- Riemann–Lebesgue, théorème de, 60
- série entière, 115
- série formelle, 115
- Schur, lemme de, 185
- Schwarz, principe de symétrie, 142
- semi-norme, 5
- séparé, espace, 4
- série de Laurent, 146
- série réciproque, 115
- σ -additivité, 21
- simplement connexe, 132
- singularité apparente, 145
- singularité essentielle, 145
- singularité isolée, 145
- sinus cardinal, 65
- $SO(2)$, 186
- $SO(3)$, 188
- sommable, fonction, 25
- sphère de Riemann, 5, 149
- spin d'une représentation, 192
- Stokes, théorème de, 136
- support, 14
- support d'une distribution, 41
- susceptibilité, 179
- Taylor, séries de, 121
- Taylor, théorèmes de, 14
- Tchebychev, inégalité de, 76, 102
- théorème limite central, 103
- théorème des résidus, 148
- tribu, 21
- tribu borélienne, 21
- tribu engendrée par ..., 21
- $U(1)$, 186
- univers, 72
- valeur moyenne, 73, 75
- variable aléatoire (v.a.), 74
- variance, 75
- voisinage, 4
- Weierstrass, théorème de, 14