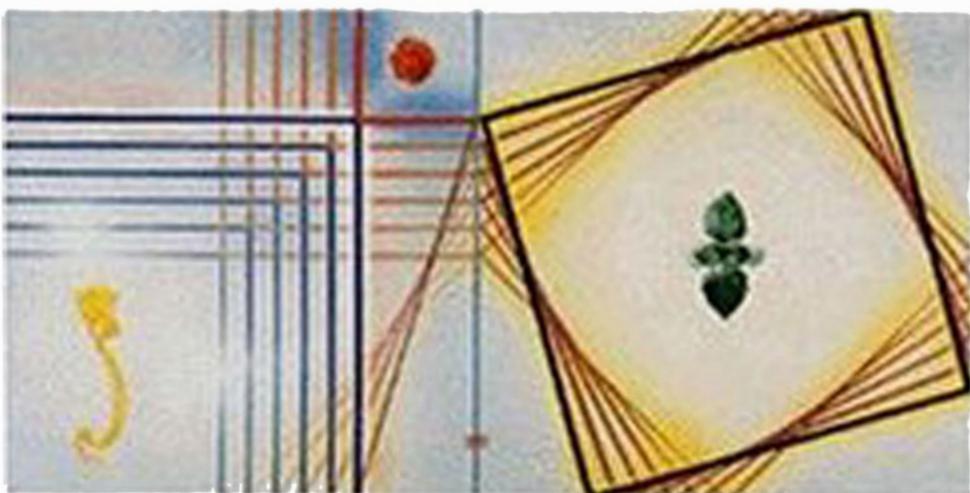


# LE CALCUL INTÉGRAL

Licence de  
Mathématiques

**Henri BUCHWALTER**





Henri BUCHWALTER

Professeur à l'Université Claude Bernard – Lyon 1

# LE CALCUL INTÉGRAL

en

**Licence de Mathématiques**



All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without permission in writing from the Publisher.

La loi du 11 mars 1957 n'autorise que les "copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective". Toute représentation ou reproduction, intégrale ou partielle, faite sans le consentement de l'éditeur, est illicite.

© COPYRIGHT 1991

**EDITION MARKETING**  
EDITEUR DES PREPARATIONS  
GRANDES ECOLES MEDECINE  
32, rue Bargue 75015 PARIS

ISBN 2-7298-4129-6

## Introduction

La théorie de l'intégrale de Riemann a fourni, dès le milieu du XIX<sup>ème</sup> siècle, une première réponse sérieuse au problème de l'existence d'une primitive pour les fonctions continues. Du même coup, elle a permis d'intégrer certaines fonctions discontinues, ouvrant ainsi la voie à diverses extensions. Mais cette théorie s'est avérée insuffisante sur plus d'un point. En particulier elle rencontrait des difficultés quasiment insurmontables pour ramener le calcul des intégrales doubles au cas des intégrales simples successives, dès l'instant où l'on sortait du cadre privilégié de la continuité. Elle ne permettait pas non plus de mettre naturellement en évidence des espaces normés complets, pour les normes liées directement à la notion d'intégrale. Bien entendu ces défauts de complétude n'étaient pas perçus comme tels puisque les notions relatives aux ensembles et aux structures n'étaient pas encore élaborées. Néanmoins la théorie se fermait sur elle-même, sans progresser de façon spectaculaire, malgré l'avancée, située sur un autre plan, due à Stieltjes dans les années 1890 (intégrale de Riemann-Stieltjes).

C'est alors que Borel, en introduisant la mesure des ensembles et les axiomes de dénombrabilité, puis au début du siècle, Lebesgue, en utilisant cette mesure pour résoudre le problème général de la primitive, ont permis de sauter le pas en offrant des outils de calcul, riches de potentialités entièrement nouvelles. Le paradoxe est même que Lebesgue, n'ayant pas complètement résolu le problème de la primitive (puisqu'il a fallu attendre Denjoy en 1915 pour cela) a cependant résolu, ou contribué à résoudre, d'autres problèmes aux implications plus vastes et plus fondamentales. C'est ainsi que la théorie de Lebesgue, et ses généralisations plus ou moins abstraites relevant des mêmes idées, ont envahi, comme outils indispensables la plupart des domaines liés à l'Analyse, que ce soit en Mathématiques pures (espaces  $L^p$  et espaces de Banach, espace  $L^2$  et espaces de Hilbert, Analyse harmonique, théories spectrales, ...), en Mathématiques appliquées (Distributions, espaces de Sobolev, ...) en Physique (espaces de Hilbert et théories quantiques), sans oublier le vaste univers probabiliste.

Tout cela, brièvement résumé, permet de dire que l'intégrale de Lebesgue est maintenant une nécessité contemporaine, et qu'elle doit figurer naturellement assez tôt dans un cursus universitaire, de manière à être utilisée, comme outil banalisé mais performant, tout au long des études ultérieures. C'est pourquoi nous l'avons considérée comme l'un des piliers de la Licence de Mathématiques.

Mais, dans le cadre contraignant de l'enseignement, il a fallu évidemment faire des choix, entre en dire trop ou en dire trop peu. Et pour évaluer ces choix, mieux vaut donc consulter la table des matières des trois chapitres développés. On y verra que, sans entrer dans des développements trop raffinés, nous sommes allés à l'essentiel de la théorie : construction de l'intégrale et de la mesure (Carathéodory), théorèmes de passage à la limite (Beppo Lévi, Fatou, convergence dominée), intégrales doubles et multiples (Tonelli, Fubini, changement de variables), liens avec la topologie (Riesz-Alexandroff et mesures de Radon) et avec l'Analyse fonctionnelle (espaces  $L^1$  et  $L^2$ ).

Quant aux applications elles n'ont été développées que dans deux directions : en offrant tout au long du cours une introduction au langage probabiliste, et en exposant une théorie, succincte mais assez complète, de la transformation de Fourier pour les groupes  $\mathbb{R}^n$ .

Le choix consistant à privilégier la notion d'ensembles, avec les mesures abstraites sur les tribus, se justifie facilement par un argument pédagogique. La théorie peut ainsi démarrer et progresser très vite, en allant cependant assez loin, avec un minimum de prérequis élémentaires de topologie des espaces métriques.

Signalons encore que les résultats énoncés sont complètement démontrés (quand ils ne sont pas triviaux), soit sous forme directe, soit sous forme d'exercices, à l'exception du théorème de Riesz-Alexandroff, dont la preuve (pour la partie "existence"), trop longue et trop technique, est renvoyée à un cours de Maîtrise. En particulier la preuve du théorème de Carathéodory, sur l'extension d'une mesure, et celle du théorème du changement de variables pour la mesure de Borel sur  $\mathbb{R}^n$ , sont toutes deux explicitées de façon complète et détaillée.

Nous espérons donc que, tel qu'il est, ce cours rendra de nombreux services à ceux qui voudront s'initier à la théorie de Lebesgue. Nous souhaitons ainsi leur ouvrir l'accès à des ouvrages plus spécialisés, où cette théorie va alors de soi et n'est employée que comme outil indispensable et puissant, permettant des développements plus approfondis.

# SOMMAIRE

## CHAPITRE 1 : FONCTIONS INTÉGRABLES

<b>1.1</b>	<b>Rappels ensemblistes</b> .....	<b>1</b>
<b>1.2</b>	<b>Arithmétique de <math>\bar{\mathbb{R}}_+ = [0, +\infty]</math></b> .....	<b>3</b>
<b>1.3</b>	<b>Anneaux et tribus</b> .....	<b>3</b>
	Tribu borélienne de $\mathbb{R}^p$ .....	5
	Tribus et partitions .....	5
<b>1.4</b>	<b>Applications mesurables et fonctions mesurables réelles</b> .....	<b>6</b>
	Les fonctions mesurables positives finies ou non .....	7
	Les fonctions mesurables réelles .....	8
	Les fonctions étagées .....	9
	Théorème d'approximation .....	9
<b>1.5</b>	<b>Intégrales et mesures</b> .....	<b>10</b>
	Intégrale (supérieure) et mesure .....	10
	Propriété de Beppo Lévi .....	12
	Mesure image, mesure à densité .....	13
<b>1.6</b>	<b>Construction des mesures</b> .....	<b>14</b>
	Mesures sur un anneau $\mathcal{R}$ .....	14
	Théorème de monotonie décroissante .....	15
	Mesure longueur sur $\mathbb{R}$ ou mesure de Borel .....	15
	Extension d'une mesure. Mesure extérieure .....	16
	Ensembles négligeables .....	17
	Théorème de Carathéodory .....	18
	Problème de l'unicité. Les $\pi$ -classes et les $\lambda$ -classes .....	20
	Théorème de Dynkin .....	20
	Mesures $\sigma$ -finies et unicité du prolongement .....	21
<b>1.7</b>	<b>Mesures sur <math>\mathbb{R}</math> et <math>\mathbb{R}^p</math></b> .....	<b>22</b>
	Mesure de Borel sur $\mathbb{R}$ . Tribu de Lebesgue .....	22
	Mesure de Borel sur $\mathbb{R}^p$ .....	23
	Mesures boréliennes sur $\mathbb{R}$ .....	24
	Mesures diffuses et mesures discrètes .....	28
<b>1.8</b>	<b>Fonctions intégrables</b> .....	<b>28</b>
	Théorème de Fatou .....	28
	Fonctions intégrables .....	30

Théorème de convergence dominée de Lebesgue .....	32
Intégration terme à terme des séries .....	33
Fonctions complexes intégrables .....	33
Intégration sur une partie A .....	33
Intégration par rapport à une mesure image .....	35
Mesure de Borel sur $\mathbb{R}^p$ et homothéties .....	36
Intégration des fonctions sphériques sur $\mathbb{R}^p$ .....	36
Le volume $V_p$ de la boule euclidienne .....	37
<b>1.9 Intégrale de Riemann et intégrale de Lebesgue .....</b>	<b>38</b>
Intégrale sur $[a, b]$ .....	38
Intégrale généralisée sur un intervalle I .....	40
<b>1.10 Fonctions définies par des intégrales .....</b>	<b>42</b>
Théorème de continuité .....	42
Théorème de dérivabilité .....	42
Théorème d'holomorphicité .....	44
Les fonctions B et $\Gamma$ d'Euler .....	47
Théorème d'Artin .....	48
Les formules de Weierstrass et de Gauss .....	50
La formule des compléments .....	51
La formule de Legendre-Gauss .....	52
La formule de Stirling .....	52
La fonction $\Gamma$ dans le champ complexe .....	53
<b>1.11 Le théorème de Riesz-Alexandroff et les mesures de Radon .....</b>	<b>54</b>
Régularité des mesures boréliennes .....	54
Théorème de Riesz-Alexandroff .....	56
Formes linéaires positives sur l'espace $C_0(T)$ .....	56
Remarque historique : Point de vue ensembliste ou fonctionnel .....	57
<b>1.12 Applications au calcul des probabilités .....</b>	<b>58</b>
Indépendance d'événements, indépendance de sous-tribus .....	60
Critère d'indépendance .....	61
Théorème des coalitions .....	61
Tribu asymptotique et loi du zéro-un .....	62
Théorème de Borel-Cantelli .....	63
Les variables aléatoires réelles .....	64
Loi du $\chi^2$ à un degré de liberté .....	64
Variables aléatoires indépendantes .....	65
Vecteurs aléatoires .....	65
Loi du $\chi^2$ à 2, 3, p degrés de liberté .....	65
Espérance mathématique et moments. Variance .....	66
Les inégalités de Markoff et Tchebyscheff .....	69

Produit de deux variables aléatoires réelles intégrables .....	70
Inégalité de Cauchy-Schwarz. Covariance .....	70
Cas d'indépendance .....	70
Règle de Bienaymé-Tchebyscheff .....	71
Loi faible des grands nombres .....	71
Loi faible de Bernoulli et fréquence statistique .....	72
Variance en dimension $p$ .....	73
Matrice de covariance .....	74

**CHAPITRE 2 : INTÉGRATION SUR UN ESPACE PRODUIT**

<b>2.1</b> Produit de deux espaces mesurables .....	77
Cas des tribus boréliennes .....	78
<b>2.2</b> Produit de deux mesures $\sigma$ -finies .....	79
Les fonctions intégrables et le théorème de Fubini .....	83
<b>2.3</b> La formule du changement de variables .....	89
Translations .....	89
Transformations linéaires .....	90
La formule générale du changement de variables .....	94
Passage en coordonnées polaires .....	98
Passage en coordonnées sphériques .....	99
<b>2.4</b> Indépendance et mesure produit .....	103
Vecteurs gaussiens et lois gaussiennes (centrées) .....	103

**CHAPITRE 3 : CONVOLUTION ET TRANSFORMATION DE FOURIER**

<b>3.1</b> Les espaces $L^1$ et $L^2$ .....	109
L'espace $\mathfrak{S}^2 = \mathfrak{S}^2(\Omega, \Sigma, \mu)$ .....	109
Les espaces $L^1$ et $L^2$ .....	110
L'espace de Hilbert $L^2$ .....	111
Théorèmes de densité .....	113

---

<b>3.2</b>	<b>Produit de convolution</b> .....	115
	Produit de convolution de deux mesures bornées .....	117
	Somme de vecteurs aléatoires indépendants .....	118
<b>3.3</b>	<b>Transformation de Fourier et fonction caractéristique</b> .....	120
	Liaison avec le produit de convolution .....	122
	Fonction caractéristique d'un vecteur aléatoire .....	122
	Loi de Poisson .....	123
	Loi binomiale .....	123
	Loi d'Euler .....	123
	Loi de Cauchy .....	124
	Lois gaussiennes sur $\mathbb{R}^n$ .....	124
	Loi de Cauchy sur $\mathbb{R}^n$ .....	126
	Transformation de Fourier et différentiabilité .....	127
	L'espace de Schwartz $\mathcal{S} = \mathcal{S}(\mathbb{R}^n)$ .....	128
	La formule d'inversion pour les fonctions .....	129
	Modules de continuité dans $L^1$ et $L^2$ .....	132
	L'espace de Wiener $\mathcal{W}(\mathbb{R}^n)$ .....	133
	La formule d'inversion pour les mesures bornées .....	135
<b>3.4</b>	<b>Transformation de Fourier et convolution sur <math>L^2</math></b> .....	139
	Calcul de $\mathcal{F}f$ pour $f \in L^2$ .....	141
	Produit de convolution dans $L^2$ .....	142

# CHAPITRE 1 : FONCTIONS INTÉGRABLES

## Introduction

Les propriétés élémentaires de l'intégrale classique des fonctions continues sur  $[a, b]$ ,  $I(f) = \int_a^b f(t) dt$ , se résument à peu de choses : positivité ( $f \geq 0 \Rightarrow I(f) \geq 0$ ), linéarité, principe de monotonie croissante

$$0 \leq f_n \uparrow f \Rightarrow I(f_n) \uparrow I(f)$$

démontré facilement à partir du lemme de Dini :

$$g_n \in C[a, b]; g_n \downarrow 0 \Rightarrow \|g_n\| \downarrow 0$$

en posant  $g_n = f - f_n$ , lemme qui est lui-même une conséquence classique de la compacité de l'intervalle  $[a, b]$ .

Si maintenant on remplace l'intégrale sur  $[a, b]$  par l'intégrale  $\int_{-\infty}^{+\infty} f(t) dt$ ,

prise cette fois sur  $\mathbb{R} = (-\infty, +\infty)$ , pour les fonctions continues et positives (c'est-à-dire telles que  $f \geq 0$ ), et par conséquent à valeurs dans  $[0, +\infty]$ , les mêmes propriétés subsistent sous la forme suivante :

$$\begin{cases} \text{positivité, additivité } I(f + g) = I(f) + I(g) \\ \text{monotonie croissante : } f_n \uparrow f \Rightarrow I(f_n) \uparrow I(f) \end{cases}$$

La question est donc de savoir si ces quelques propriétés peuvent servir de guide pour construire une théorie générale de l'intégration. Nous allons voir que la réponse est positive, une fois précisés la structure de l'ensemble  $\bar{\mathbb{R}}_+ = [0, +\infty]$  et le rôle de la dénombrabilité. Et qu'ainsi l'intégrale, dite de Lebesgue, se construit à partir de données résolument minimales.

## 1.1 Rappels ensemblistes

Étant donné un ensemble  $\Omega$ , on note  $\mathcal{P}(\Omega)$  l'ensemble de toutes ses parties. On rappelle les notions ou propriétés suivantes :

- Réunion  $\bigcup_{i \in I} A_i$  et intersection  $\bigcap_{i \in I} A_i$  d'une famille  $(A_i)_{i \in I}$  de parties de  $\Omega$ , indexée par un ensemble quelconque  $I$ . Distributivité de l'une par rapport à l'autre.

– Complémentaire  $\complement A$  ou  $A^c$  ou  $\Omega \setminus A$ .

Formules de dualité  $(\cup A_i)^c = \cap A_i^c$  et  $(\cap A_i)^c = \cup A_i^c$ .

– Différence ensembliste  $A \setminus B = A \cap B^c$

et différence symétrique  $A \Delta B = (A \setminus B) \cup (B \setminus A)$ .

Formules du type suivant :  $(\cup A_i) \setminus B = \cup (A_i \setminus B)$

$(\cap A_i) \setminus B = \cap (A_i \setminus B)$

$A \setminus (\cup B_j) = \cap (A \setminus B_j)$

$A \setminus (\cap B_j) = \cup (A \setminus B_j)$

– Image réciproque  $f^{-1}(B)$  pour  $f : \Omega \rightarrow X$  et formules de commutation

$f^{-1}(\cap B_j) = \cap f^{-1}(B_j)$

$f^{-1}(\cup B_j) = \cup f^{-1}(B_j)$

$f^{-1}(B^c) = [f^{-1}(B)]^c$

– Fonction indicatrice :

$1_A : \Omega \rightarrow \{0, 1\}$ , définie par  $1_A(\omega) = 1$  si  $\omega \in A$  et  $1_A(\omega) = 0$  si  $\omega \notin A$ .

– Ensembles produits  $X \times Y$ ,  $\prod_{i \in I} X_i$ ,  $X^I = \mathcal{F}(I, X)$  de sorte que

$(X^I)^J = \mathcal{F}(J, \mathcal{F}(I, X)) = \mathcal{F}(J \times I, X) = X^{J \times I}$

– Axiome du choix sous la forme : un produit  $X = \prod_{i \in I} X_i$  d'ensembles  $X_i$  non vides est lui-même non vide.

– Égalité  $\mathcal{P}(\Omega) = \{0, 1\}^\Omega$  dans l'identification  $A \rightarrow f = 1_A$  et  $f \rightarrow A = f^{-1}(1)$ .

– Espace  $\mathbb{R}^{\mathbb{N}}$  des suites réelles.

– **Équipotence et dénombrabilité.** La relation

$X \leq Y \Leftrightarrow$  Il existe une injection  $X \rightarrow Y$

$\Leftrightarrow$  Il existe une surjection  $Y \rightarrow X$

est telle que  $X \leq Y$  et  $Y \leq X \Leftrightarrow X$  et  $Y$  sont équipotents (c'est-à-dire, il existe une bijection  $X \rightarrow Y$ ) mais la preuve est difficile et nécessite l'axiome du choix. On note  $X \leftrightarrow Y$  lorsque  $X$  et  $Y$  sont équipotents.

L'ensemble  $X$  est dit *dénombrable* lorsque  $X \leftrightarrow \mathbb{N}$ . Par exemple  $\mathbb{N} \times \mathbb{N}$  est dénombrable et toute partie infinie de  $\mathbb{N}$  l'est aussi. Donc la condition  $X \leq \mathbb{N}$  signifie que  $X$  est dénombrable ou fini (c'est-à-dire au plus dénombrable). Comme conséquence :

- Une réunion dénombrable d'ensembles dénombrables est dénombrable.
- Un produit fini d'ensembles dénombrables est dénombrable.
- $\mathbb{N}^{\mathbb{P}}$ ,  $\mathbb{Q}$ ,  $\mathbb{Q}^{\mathbb{P}}$ ,  $\mathbb{Z}$ ,  $\mathbb{Z}^{\mathbb{P}}$  sont dénombrables.
- $\mathbb{R} \leftrightarrow \{0, 1\}^{\mathbb{N}} = \mathcal{P}(\mathbb{N})$  n'est pas dénombrable.

- $\mathbb{R} \leftrightarrow (0, 1]^{\mathbb{N}} \leftrightarrow \mathbb{N}^{\mathbb{N}} \leftrightarrow [0, 1]^{\mathbb{N}} \leftrightarrow \mathbb{R}^{\mathbb{N}}$ .
- Dans  $\mathbb{N}^{\mathbb{N}}$  l'ensemble  $\mathbb{N}^{(\mathbb{N})} = \bigcup \mathbb{N}^p$  des suites (à valeurs entières) à support fini est dénombrable.
- Dans  $\mathbb{C}$  l'ensemble des nombres algébriques est dénombrable.

## 1.2 Arithmétique de $\overline{\mathbb{R}}_+ = [0, +\infty]$

L'addition et la multiplication se prolongent à  $\overline{\mathbb{R}}_+$  à l'aide des conventions expresses suivantes, toujours utilisées en théorie de l'intégration :

$$a + \infty = \infty + a = +\infty \text{ pour } a \in \overline{\mathbb{R}}_+$$

$$a \cdot (+\infty) = (+\infty) \cdot a = \begin{cases} +\infty & \text{si } 0 < a \leq +\infty \\ 0 & \text{si } a = 0 \end{cases}$$

Ainsi prolongées l'addition et la multiplication sur  $\overline{\mathbb{R}}_+$  sont commutatives et associatives, et la multiplication est distributive par rapport à l'addition. Il faut toutefois éviter certains pièges :

– ni l'addition ni la multiplication ne sont simplifiables :

$$a + c = b + c \quad \text{n'implique } a = b \text{ que si } c < +\infty$$

$$ac = bc \quad \text{n'implique } a = b \text{ que si } 0 < c < +\infty$$

–  $a_n \rightarrow a$  et  $b_n \rightarrow b$  n'implique pas nécessairement que la suite  $a_n b_n$  soit convergente et, si elle l'est, que sa limite soit  $ab$ .

Par contre, toute suite *croissante* dans  $\overline{\mathbb{R}}_+$  est convergente et si  $a_n \uparrow a$  et  $b_n \uparrow b$  alors  $a_n + b_n \uparrow a + b$  et  $a_n b_n \uparrow ab$ .

Enfin toute série  $\sum u_n$  à termes  $u_n \in \overline{\mathbb{R}}_+$  est convergente dans  $\overline{\mathbb{R}}_+$  et toute modification de l'ordre des termes ou toute sommation par blocs laisse la somme inchangée. Pour que  $\sum u_n = +\infty$ , il faut et il suffit que l'un des  $u_n$  soit  $+\infty$ , ou, s'ils sont tous finis, que la série  $\sum u_n$  soit divergente au sens habituel.

## 1.3 Anneaux et tribus

La théorie de l'intégrale que nous voulons développer est principalement fondée sur la notion de tribu, qui précise et prolonge celle d'anneau. Nous introduisons donc les définitions suivantes :

(1.3.1) *Définition*

- a) On appelle classe toute famille non vide de parties d'un ensemble  $\Omega$ .
- b) On appelle anneau (ou anneau booléen) toute classe  $\mathcal{R}$ , stable par réunion et différence  $A, B \in \mathcal{R} \Rightarrow A \cup B \in \mathcal{R}$  et  $A \setminus B \in \mathcal{R}$
- c) Tout anneau  $\mathcal{R}$  tel que  $\Omega \in \mathcal{R}$  s'appelle une algèbre.
- d) On appelle  $\sigma$ -anneau tout anneau stable pour la réunion dénombrable et  $\sigma$ -algèbre ou *tribu* tout  $\sigma$ -anneau contenant  $\Omega$  comme élément.

(1.3.2) *Proposition*

- a) Tout anneau  $\mathcal{R}$  est stable par réunion, intersection, différence et différence symétrique. De plus,  $\emptyset \in \mathcal{R}$ . Pour que  $\mathcal{R}$  soit une algèbre, il suffit qu'il soit stable par complémentarité.
- b) Tout  $\sigma$ -anneau est stable par réunion dénombrable et intersection dénombrable.
- c) Enfin, toute tribu  $\Sigma$  est stable par réunion dénombrable, intersection dénombrable et complémentarité.

(1.3.3) *Exemples* .- La plus petite tribu sur  $\Omega$  est la classe  $\{\emptyset, \Omega\}$ , et la plus grande est  $\mathcal{P}(\Omega)$ . La classe des parties finies est un anneau, celle des parties finies ou dénombrables est un  $\sigma$ -anneau, noté  $\Delta$ . La classe des parties  $A$  telles que  $A \in \Delta$  ou  $A^c \in \Delta$  est une tribu notée  $\Sigma_\Delta$ .

- Soit  $h : \Omega \rightarrow X$  une application. Alors pour toute tribu  $\mathcal{C}$  sur  $X$ , la classe  $h^{-1}(\mathcal{C}) = \{h^{-1}(B); B \in \mathcal{C}\}$  est une tribu sur  $\Omega$ , comme il résulte aisément des propriétés ensemblistes de l'image réciproque.
- Dans  $\mathbb{R}$  l'ensemble des réunions finies d'intervalles bornés est un anneau, de même que l'ensemble  $\mathcal{J}$  des réunions finies d'intervalles semi-ouverts  $[a, b[$ .
- Dans  $\mathbb{R}^P$  l'ensemble  $\mathcal{J}_p$  des réunions finies de pavés bornés de la forme  $P = \prod_{k=1}^P [a_k, b_k [$  est un anneau.

La plupart des tribus utilisées dans la théorie se construisent par engendrement, une fois vu qu'une intersection quelconque de tribus sur  $\Omega$  est encore une tribu (de même pour les anneaux et les  $\sigma$ -anneaux). D'où

(1.3.4) *Définition*

Soit  $\mathcal{C}$  une classe sur  $\Omega$ . On appelle tribu engendrée par  $\mathcal{C}$ , la plus petite tribu  $\Sigma_{\mathcal{C}}$  contenant la classe  $\mathcal{C}$ .

(1.3.5) *Définition*

On appelle *tribu borélienne* de  $\mathbb{R}$  (resp.  $\mathbb{R}^p$ ) la tribu  $\mathfrak{B}$  (resp.  $\mathfrak{B}_p$ ) engendrée par la classe des parties ouvertes (ou par la classe des parties fermées).

On a alors les premiers résultats :

(1.3.6) *Proposition*

Sur  $\mathbb{R}$  la tribu borélienne  $\mathfrak{B}$  est engendrée par :

- a) la classe des intervalles bornés,
- b) l'anneau  $\mathfrak{J}$  des réunions finies d'intervalles  $[a, b[$ ,
- c) la classe des demi-droites ouvertes  $(-\infty, \alpha[$ ,
- d) la classe des demi-droites fermées  $(-\infty, \alpha]$ .

**Preuve.** Tout résulte du fait qu'un ouvert quelconque de  $\mathbb{R}$  est toujours une réunion finie ou dénombrable d'intervalles ouverts.  $\square$

(1.3.7) *Proposition*

Sur  $\mathbb{R}^p$  la tribu borélienne  $\mathfrak{B}_p$  est engendrée par :

- a) la classe des pavés ouverts  $P = \prod ] a_k, b_k [$
- b) la classe des pavés fermés  $P = \prod [ a_k, b_k ]$
- c) l'anneau  $\mathfrak{J}_p$ .

**Preuve.** Elle est basée sur la densité de  $\mathbb{Q}^p$  dans  $\mathbb{R}^p$ , qui entraîne que tout ouvert  $U$  de  $\mathbb{R}^p$  est toujours une réunion finie ou dénombrable de pavés ouverts.  $\square$

(1.3.8) *Remarque.* On prendra garde que la notion de tribu engendrée, et par conséquent celle de tribu borélienne, est difficile, et délicate à manier correctement. La raison en est que la définition par engendrement ne donne aucun critère "individuel" pour décider si une partie  $A$  appartient ou non à la tribu et nous verrons par la suite que les raisonnements sur les tribus se feront presque toujours de façon "collectiviste" en faisant appel à des classes correctement choisies. Par exemple un ensemble borélien sur  $\mathbb{R}$  ou  $\mathbb{R}^p$  peut être très compliqué puisque la tribu  $\mathfrak{B}_p$  contient évidemment les ouverts, les fermés, les  $G_\delta$  (intersections dénombrables d'ouverts), les  $F_\sigma$  (réunions dénombrables de fermés), les  $G_{\delta\sigma}$ , les  $F_{\sigma\delta}$ , les  $G_{\delta\sigma\delta}$ , les  $F_{\sigma\delta\sigma}$ , etc., c'est-à-dire tous les ensembles construits à partir des ouverts ou des fermés par itérations successives des opérations stables de complémentarité, réunion dénombrable et intersection dénombrable.

**Tribus et partitions.** Étant donné une partie  $A \subset \Omega$  telle que  $\emptyset \neq A$  et  $\Omega \neq A$ , la tribu engendrée par  $A$  est formée des quatre éléments  $(\emptyset, A, A^c, \Omega)$ , et c'est aussi la tribu engendrée par la partition  $\{A, A^c\}$  de  $\Omega$ . Plus généralement, si  $\{A_1, \dots, A_p\}$  est une

partition finie de  $\Omega$  (les  $A_k$  sont donc disjoints, non vides et de réunion  $\Omega$ ) la tribu engendrée par les  $A_k$  est exactement celle des ensembles  $\bigcup_{k \in J} A_k$ , où  $J$  décrit l'ensemble des parties de  $\{1, 2, \dots, p\}$ , de sorte qu'elle a la cardinalité  $2^p$ . Plus généralement encore, si  $\{A_n\}_{n \in \mathbb{N}}$  est une partition dénombrable de  $\Omega$ , la tribu engendrée est exactement celle des parties  $\bigcup_{k \in J} A_k$ , où  $J$  décrit cette fois l'ensemble  $\mathcal{P}(\mathbb{N})$ ; elle a donc la cardinalité de  $\mathcal{P}(\mathbb{N})$ , c'est-à-dire celle de  $\mathbb{R}$ . On remarquera que du point de vue de la cardinalité on passe de  $2^p$  à  $2^{\mathbb{N}} = \mathbb{R}$  en sautant la cardinalité de  $\mathbb{N}$ , c'est-à-dire celle du dénombrable. Et en effet on pourra montrer :

(1.3.9) *Exercice*

Soit  $\Sigma$  une tribu sur  $\Omega$ , supposée au plus dénombrable.

- Démontrer que tout  $\omega \in \Omega$  est contenu dans une plus petite partie de  $\Sigma$ .
- En déduire que  $\Sigma$  est exactement la tribu engendrée par une partition au plus dénombrable.
- Montrer alors que  $\Sigma$  est nécessairement finie, c'est-à-dire qu'il n'existe pas de tribu dénombrable.

## 1.4 Applications mesurables et fonctions mesurables réelles

On appelle *espace mesurable* la donnée d'un couple  $(\Omega, \Sigma)$ , où  $\Sigma$  est une tribu sur  $\Omega$ , et *application mesurable* toute application  $h : (\Omega_1, \Sigma_1) \rightarrow (\Omega_2, \Sigma_2)$  entre deux espaces mesurables, telle que  $h^{-1}(B) \in \Sigma_1$  pour tout  $B \in \Sigma_2$ , autrement dit telle que la tribu  $h^{-1}(\Sigma_2)$  soit une sous-tribu de  $\Sigma_1$ . Il est alors clair que si  $f : \Omega_1 \rightarrow \Omega_2$  et  $g : \Omega_2 \rightarrow \Omega_3$  sont des applications mesurables, il en est de même de  $h = g \circ f : \Omega_1 \rightarrow \Omega_3$ , ce qui fournit la stabilité par composition. Pour tester la mesurabilité d'une application, on a le critère fondamental :

(1.4.1) *Proposition*

Soit  $h : (\Omega_1, \Sigma_1) \rightarrow (\Omega_2, \Sigma_2)$  une application. On suppose que la tribu  $\Sigma_2$  est engendrée par une classe  $\mathcal{C}_2$ . Alors pour que  $h$  soit mesurable il faut et il suffit que  $h^{-1}(\mathcal{C}_2) \subset \Sigma_1$ , autrement dit que l'on ait  $h^{-1}(B) \in \Sigma_1$  seulement pour toute partie  $B \in \mathcal{C}_2$ .

*Preuve.* On introduit la classe  $\mathcal{C}_2$  des parties  $B \subset \Omega_2$  telles que  $h^{-1}(B) \in \Sigma_1$ , en vérifiant de façon évidente que  $\mathcal{C}_2$  est une tribu contenant  $\mathcal{C}_2$ , donc aussi  $\Sigma_2$ , ce qui signifie que  $h^{-1}(\Sigma_2) \subset \Sigma_1$ .  $\square$

Les fonctions mesurables positives finies ou non. Plaçons sur  $\bar{\mathbb{R}}_+ = [0, +\infty]$  sa topologie naturelle, d'ailleurs homéomorphe à celle de l'intervalle compact  $[0, 1]$  dans l'application  $x \rightarrow \frac{x}{1+x}$ . La tribu borélienne de  $\bar{\mathbb{R}}_+$ , qui par définition est celle engendrée par les ouverts, est aussi engendrée par les intervalles ouverts  $[0, \alpha[$  avec  $0 < \alpha < +\infty$ , ou bien par les intervalles fermés  $[0, \alpha]$ . Il suit de là qu'une fonction  $f : (\Omega, \Sigma) \rightarrow \bar{\mathbb{R}}_+$  est mesurable ssi tous les ensembles  $\{f < \alpha\} = \{\omega \in \Omega, f(\omega) < \alpha\}$  sont éléments de  $\Sigma$ , ou bien encore ssi tous les ensembles  $\{f \leq \alpha\}$  sont éléments de  $\Sigma$ . A partir de là et grâce aux préliminaires 1.2 sur  $\bar{\mathbb{R}}_+$ , on a les propriétés *fondamentales* de stabilité suivantes :

(1.4.2) *Théorème*

Soit  $(\Omega, \Sigma)$  un espace mesurable.

- a) Pour qu'une fonction indicatrice  $1_A$  soit mesurable, il faut et il suffit que  $A \in \Sigma$ .
- b) Si  $f$  et  $g$  sont des fonctions  $\Omega \rightarrow \bar{\mathbb{R}}_+$  mesurables, alors  $f + g$  et  $fg$  sont aussi mesurables.
- c) Si  $(f_n)$  est une suite de fonctions  $\Omega \rightarrow \bar{\mathbb{R}}_+$  mesurables, alors les fonctions suivantes sont mesurables :
  - .  $\text{Sup } f_n ; \text{Inf } f_n ; \limsup f_n ; \liminf f_n$ .
  - .  $\lim f_n$  si la suite  $(f_n)$  converge simplement dans  $\bar{\mathbb{R}}_+$ .
  - .  $\sum f_n$ .

*Preuve.* a) est évident car  $1_A^{-1}(1) = A$  et  $(1_A < \alpha)$  est égal à  $A^c$  ou à  $\Omega$  suivant que  $\alpha \leq 1$  ou  $\alpha > 1$ . Pour prouver c) posons  $f = \text{Sup } f_n$  et  $g = \text{Inf } f_n$ . Il est clair que

$$\{f \leq \alpha\} = \bigcap_n \{f_n \leq \alpha\} \in \Sigma$$

$$\{g < \alpha\} = \bigcup_n \{f_n < \alpha\} \in \Sigma$$

donc  $f$  et  $g$  sont mesurables. Ensuite on a :

$$\limsup f_n = \text{Inf}_n \text{Sup}_{k \geq n} f_k$$

$$\liminf f_n = \text{Sup}_n \text{Inf}_{k \geq n} f_k$$

ce qui assure la mesurabilité de  $\limsup f_n$  et  $\liminf f_n$ , donc aussi de  $\lim f_n$  si cette limite existe en tout point. Pour prouver b) il suffit de voir que :

$$\{f + g < \alpha\} = \bigcup_{\substack{r, s \in \mathbb{Q}_+ \\ r+s < \alpha}} \{f < r\} \cap \{g < s\} \in \Sigma$$

$$\{fg < \alpha\} = \bigcup_{\substack{r,s \in \mathbb{Q}_+ \\ rs < \alpha}} \{f < r\} \cap \{g < s\} \in \Sigma$$

Enfin, si  $h = \sum_{n=0}^{\infty} f_n$ , alors  $h = \text{Sup } h_n$  avec  $h_n = \sum_{k=0}^n f_k$  et  $h_n$  est mesurable d'après b),

donc aussi h. □

La conclusion, et le sens profond de (1.4.2), est qu'à condition de ne pas sortir du dénombrable, on peut faire à peu près tout ce qu'on veut avec les fonctions mesurables.

Dans la suite nous noterons  $\overline{\mathfrak{F}}_+(\Sigma)$  le cône des fonctions mesurables  $\Omega \rightarrow \overline{\mathbb{R}}_+$ , en prenant garde que ce n'est pas un espace vectoriel et qu'en particulier la différence  $f - g$  n'y est pas définie correctement.

**Les fonctions mesurables à valeurs réelles.** Il s'agit ici des fonctions mesurables  $\Omega \rightarrow \mathbb{R}$  lorsqu'on place sur  $\mathbb{R}$  sa tribu borélienne. Elles sont caractérisées par la condition  $\{f < \alpha\} \in \Sigma$  pour tout  $\alpha$  réel, ou encore par  $\{f \leq \alpha\} \in \Sigma$  pour tout  $\alpha$  réel. Les propriétés de stabilité sont les suivantes :

(1.4.3) *Théorème*

L'ensemble  $\mathfrak{F}(\Sigma)$  des fonctions mesurables réelles finies sur l'espace  $(\Omega, \Sigma)$  est une algèbre réticulée, et en particulier  $|f|$  est mesurable dès que  $f$  l'est. De plus pour toute suite  $(f_n)$  de fonctions mesurables, les fonctions  $\text{Sup } f_n$ ,  $\text{Inf } f_n$ ,  $\limsup f_n$ ,  $\liminf f_n$ ,  $\lim f_n$ ,  $\sum f_n$  sont mesurables dès qu'elles sont partout définies sur  $\Omega$  (et bien entendu à valeurs finies).

**Preuve.** Si  $f$  est mesurable alors aussi  $-f$  l'est car :

$$\{-f < \alpha\} = \{f > -\alpha\} = \{f \leq -\alpha\}^c \in \Sigma$$

donc aussi  $f + g$ ,  $f - g$  et  $fg = (f^+ - f^-)(g^+ - g^-)$  si  $g$  est aussi mesurable. Le reste n'est qu'une répétition de (1.4.2). □

Comme application immédiate donnons un lemme qui sera utile plus loin :

(1.4.4) *Lemme*

Soit  $f, g \in \overline{\mathfrak{F}}_+(\Sigma)$  telles que  $g \leq f$ . Alors il existe  $h \in \overline{\mathfrak{F}}_+(\Sigma)$  telle que  $f = g + h$ .

**Preuve.** Posons  $A = \{f < +\infty\} \in \Sigma$  et  $B = A^c = \{f = +\infty\}$ . Il suffit de prendre  $h = (1_A f - 1_A g) + (\infty) 1_B$ , car  $1_A f - 1_A g = 1_A (f - g)$  est positive et élément de  $\mathfrak{F}(\Sigma)$ , donc élément de  $\overline{\mathfrak{F}}_+(\Sigma)$ . □

(1.4.5) *Exemple.* En prenant  $(\Omega, \Sigma) = (\mathbb{R}^p, \mathfrak{B}_p)$  on obtient une notion de fonctions mesurables définies sur  $\mathbb{R}^p$ . Ce sont les fonctions dites *boréliennes* ou aussi *Borel-mesurables*. Par exemple toute fonction continue  $f: \mathbb{R}^p \rightarrow \mathbb{R}$  est borélienne puisqu'alors les ensembles  $\{f < \alpha\}$  sont ouverts. Ici apparaît la complexité des fonctions boréliennes car évidemment toute limite simple d'une suite  $(f_n)$  de fonctions continues, à condition qu'elle soit définie en tout point  $x \in \mathbb{R}^p$ , est encore borélienne. Et toute limite simple d'une suite de telles fonctions, etc., etc.

**Les fonctions étagées.** Puisque les fonctions  $1_A, A \in \Sigma$ , sont éléments de  $\mathfrak{L}(\Sigma)$ , il est intéressant d'introduire l'espace vectoriel  $\mathfrak{E}(\Sigma)$  engendré par ces fonctions indicatrices. Il s'agit donc des fonctions de la forme  $g = \sum \alpha_k 1_{A_k}$  avec  $\alpha_k \in \mathbb{R}$  et  $A_k \in \Sigma$ , la sommation étant finie. Une telle fonction est mesurable et ne prend qu'un nombre fini de valeurs, et réciproquement toute fonction mesurable  $g$  qui ne prend que les valeurs  $\lambda_1, \dots, \lambda_p$  est nécessairement égale à  $\sum \lambda_k 1_{A_k}$ , où l'on a posé  $A_k = g^{-1}(\lambda_k) \in \Sigma$ , la famille finie  $(A_k)$  constituant alors une partition finie de  $\Omega$ .

L'intérêt des fonctions étagées est essentiellement qu'elles permettent de reconstituer l'algèbre  $\mathfrak{L}(\Sigma)$  et le cône  $\overline{\mathfrak{L}}_+(\Sigma)$ , ce qui illustre le fait fondamental que l'on peut "remonter" des ensembles aux fonctions en conservant la mesurabilité. En effet :

(1.4.6) **Théorème (d'approximation)**

- a) Toute fonction  $f \in \mathfrak{L}(\Sigma)$  qui est bornée est limite uniforme d'une suite  $(g_n)$  de fonctions étagées.  
 b) Toute fonction  $f \in \overline{\mathfrak{L}}_+(\Sigma)$  est limite simple d'une suite croissante  $(g_n)$  de fonctions étagées positives.

**Preuve.** a) Posons  $A_{k,n} = \left\{ \frac{k}{n} \leq f < \frac{k+1}{n} \right\}$  pour  $n \geq 1$  et  $-n \leq k \leq n$ , après avoir supposé

$\|f\| \leq 1$ . Alors  $A_{k,n} \in \Sigma$  et  $g_n = \sum_k \frac{k}{n} 1_{A_{k,n}} \in \mathfrak{E}(\Sigma)$ . Or  $\|f - g_n\| \leq \frac{1}{n}$  et tout est dit.

b) Pour obtenir la monotonie, il faut être un peu plus subtil. Supposons donc  $0 \leq f \leq +\infty$  et pour chaque  $n \geq 1$  posons

$$A_{k,n} = \{k 2^{-n} \leq f < (k+1) 2^{-n}\} \quad 0 \leq k < n 2^n$$

$$B_n = \{f \geq n\}$$

de sorte que  $A_{k,n} \in \Sigma$  et  $B_n \in \Sigma$ . Alors

$$g_n = \sum_k k 2^{-n} 1_{A_{k,n}} + n 1_{B_n}$$

est étagée et positive. De plus la suite  $(g_n)$  est croissante et  $g_n \uparrow f$  car pour  $f(\omega) = +\infty$  on a  $g_n(\omega) = n \rightarrow +\infty$  et pour  $f(\omega) < +\infty$ , on a  $0 \leq f(\omega) - g_n(\omega) \leq 2^{-n}$  si  $n > f(\omega)$ .  $\square$

**Remarque.** Il résulte de b), par décomposition  $f = f^+ - f^-$ , que toute  $f \in \mathfrak{F}(\Sigma)$  est limite simple d'une suite  $(g_n)$  de fonctions étagées, mais sans qu'on puisse ici garantir la monotonie.

## 1.5 Intégrales et mesures

Donnons immédiatement deux définitions.

### (1.5.1) Définition

On appelle intégrale (ou intégrale supérieure) sur  $(\Omega, \Sigma)$  toute application

$L : \overline{\mathfrak{F}}_+(\Sigma) \rightarrow [0, +\infty]$  vérifiant l'une ou l'autre des conditions équivalentes :

a)  $L(0) = 0$  ;  $L(f + g) = L(f) + L(g)$  ;  $f_n \uparrow f \Rightarrow L(f_n) \uparrow L(f)$

b)  $L(0) = 0$  ;  $L(\sum f_n) = \sum L(f_n)$

Pour une telle  $L$  on a de plus  $L(\lambda f) = \lambda L(f)$  pour  $\lambda \geq 0$  et  $L(f) \leq L(g)$  si  $f \leq g$ .

**Preuve.** On passe aisément de b) à a) avec le lemme (1.4.4) car si  $f_n \uparrow f$ , en écrivant  $f_1 = h_1$  ;  $f_2 = f_1 + h_2$ , ...,  $f_n = f_{n-1} + h_n$ , on obtient  $f = \sum h_n$ , ce qui permet de prouver  $b \Rightarrow a$ .

Réciproquement, si  $(h_n)$  est une suite quelconque de  $\overline{\mathfrak{F}}_+(\Sigma)$ , alors la suite  $f_n = h_1 + \dots + h_n$  est telle que  $f_n \uparrow f = \sum h_n$ , ce qui donne aisément  $a \Rightarrow b$ .

Ensuite par additivité finie on a  $L(pf) = p L(f)$ , donc en changeant  $f$  en  $\frac{f}{q}$ ,  $L\left(\frac{f}{q}\right) = \frac{1}{q} L(f)$ ,

d'où  $L(rf) = r L(f)$  pour tout  $r \in \mathbb{Q}_+$ . Alors si  $\lambda \geq 0$ , on choisit une suite  $r_n \uparrow \lambda$ , donc  $L(\lambda f) = \lim L(r_n f) = \lim r_n L(f) = \lambda L(f)$ .

Enfin, si  $f \leq g$  alors  $L(f) \leq L(g)$  avec le lemme (1.4.4). □

### (1.5.2) Définition

On appelle mesure sur  $\Sigma$  toute application  $\mu : \Sigma \rightarrow [0, +\infty]$  vérifiant l'une ou l'autre des conditions équivalentes :

a)  $\mu \emptyset = 0$  ;  $\mu(A \cup B) = \mu A + \mu B$  si  $A \cap B = \emptyset$  ;  $A_n \uparrow A \Rightarrow \mu A_n \uparrow \mu A$

b)  $\mu \emptyset = 0$  ;  $(A_n)$  disjointe  $\Rightarrow \mu(\cup A_n) = \sum \mu A_n$ .

Or il se trouve que ces deux notions d'intégrale et de mesure n'en font en réalité qu'une, c'est-à-dire qu'il existe une correspondance *bijective*  $L \leftrightarrow \mu$  entre les unes et les autres. C'est ici qu'on touche au point crucial de la théorie, à savoir le passage canonique, et dans les deux sens, entre fonctions et ensembles. La correspondance  $L \rightarrow \mu$  est immédiate.

### (1.5.3) Proposition

Toute intégrale  $L$  définit une mesure  $\mu = \mu_L$  sur  $\Sigma$  par l'égalité  $\mu(A) = L(1_A)$ .

**Preuve.** Avec b) par exemple, car si la suite  $(A_n)$  est disjointe et si  $A = \cup A_n$ , alors  $1_A = \sum 1_{A_n}$ , donc (1.5.1.b)  $\Rightarrow$  (1.5.2.b).  $\square$

La correspondance  $\mu \rightarrow L$  est beaucoup plus délicate car, partant des ensembles, il faut remonter aux fonctions. La méthode consiste, partant d'une mesure  $\mu$ , à construire  $L$  d'abord sur le cône positif  $\mathfrak{E}_+(\Sigma)$  des fonctions étagées positives, puis sur  $\overline{\mathfrak{F}}_+(\Sigma)$ , et à vérifier enfin les propriétés (1.5.1.a).

Étant donné  $f \in \mathfrak{E}_+(\Sigma)$  on l'écrit en décomposition unique, dite *canonique*,  $f = \sum \lambda_k 1_{A_k}$ , où  $(A_k)$  constitue une partition de  $\Omega$  et où les  $\lambda_k$  sont deux à deux distincts, et on pose

$$L(f) = \sum \lambda_k \mu A_k .$$

Si  $f = 1_A$ ,  $A \in \Sigma$ , alors sa décomposition canonique est  $f = 1.1_A + 0.1_{A^c}$ , donc  $L(f) = \mu A = L(1_A)$ , ce qui donne la liaison fondamentale entre  $L$  et  $\mu$ .

Maintenant si  $g = \sum v_j 1_{B_j}$ , en décomposition canonique, alors la décomposition canonique de  $h = f + g$  s'obtient à partir de la décomposition non canonique

$$h = \sum_{kj} (\lambda_k + v_j) 1_{A_k \cap B_j}$$

en rassemblant les  $(\lambda_k + v_j)$  qui sont égaux et en réunissant les  $A_k \cap B_j$  correspondants. L'additivité finie de  $\mu$  permet de vérifier que  $L(f + g) = L(f) + L(g)$ .

Il suit de là que si  $f \in \mathfrak{E}_+(\Sigma)$  s'écrit  $f = \sum \alpha_k 1_{A_k}$  avec  $\alpha_k \geq 0$ , en décomposition canonique ou non, on a toujours  $L(f) = \sum \alpha_k \mu A_k$ . D'où suit aisément :  $f \leq g \Rightarrow L(f) \leq L(g)$

Passons maintenant à la définition de  $L$  sur  $\overline{\mathfrak{F}}_+(\Sigma)$ .

(1.5.4) **Définition**

Pour toute  $f \in \overline{\mathfrak{F}}_+(\Sigma)$  on pose

$$L(f) = \text{Sup} \{ L(g) ; g \in \mathfrak{E}_+(\Sigma), g \leq f \}$$

Évidemment  $L(0) = 0$ , et si  $f \in \mathfrak{E}_+(\Sigma)$ , on retrouve la valeur  $L(f)$  déjà définie, donc en particulier  $L(1_A) = \mu A$ . On a alors

(1.5.5) **Théorème**

La fonctionnelle  $L$  associée à  $\mu$  vérifie les propriétés suivantes :

a)  $L(0) = 0 ; L(1_A) = \mu A, f \leq g \Rightarrow L(f) \leq L(g)$

- b)  $L(f + g) = L(f) + L(g)$   
 c) (Propriété de Beppo Lévi) :  $f_n \uparrow f \Rightarrow L(f_n) \uparrow L(f)$

**Preuve.** a) est évident et b) est une conséquence de c). Car si  $f, g \in \overline{\mathfrak{F}}_+(\Sigma)$ , on sait avec (1.4.6.b), qu'il existe des suites  $(h_n)$  et  $(k_n)$  dans  $\mathfrak{E}_+(\Sigma)$  telles que  $h_n \uparrow f$  et  $k_n \uparrow g$ , d'où  $(h_n + k_n) \uparrow (f + g)$  et avec c)

$$\begin{aligned} L(f + g) &= \lim L(h_n + k_n) \\ &= \lim [L(h_n) + L(k_n)] \\ &= \lim L(h_n) + \lim L(k_n) \\ &= L(f) + L(g) \end{aligned}$$

Prouvons donc c). Pour cela fixons la suite  $(f_n)$  dans  $\overline{\mathfrak{F}}_+(\Sigma)$  telle que  $f_n \uparrow f$ . On a déjà  $L(f_n) \leq L(f)$  et la suite croissante  $L(f_n)$  a une limite  $\ell \leq L(f)$ . Pour voir que  $\ell = L(f)$ , fixons  $h \in \mathfrak{E}_+(\Sigma)$  telle que  $h \leq f$ . Pour tout  $c \in ]0, 1[$ , posons  $B_n = \{f_n \geq ch\} \in \Sigma$ . La suite  $(B_n)$  est croissante et  $\bigcup B_n = \Omega$  car  $h$  est partout finie,  $\lim f_n = f \geq h$  et  $h(\omega) > c h(\omega)$  si  $h(\omega) > 0$ . Par ailleurs on vérifie que  $f_n \geq c h 1_{B_n}$ , donc  $L(f_n) \geq c L(h 1_{B_n})$  et  $\ell \geq c L(h 1_{B_n})$ .

Or si  $h = \sum \alpha_k 1_{A_k}$ , alors  $h 1_{B_n} = \sum \alpha_k 1_{A_k \cap B_n}$  et  $L(h 1_{B_n}) = \sum \alpha_k \mu(A_k \cap B_n)$ . Mais la condition  $B_n \uparrow \Omega$  implique  $A_k \cap B_n \uparrow A_k$  et  $\mu(A_k \cap B_n) \uparrow \mu A_k$  d'après (1.5.2.a), d'où à la limite  $\ell \geq c L(h)$ , puis  $\ell \geq L(h)$  avec  $c \uparrow 1$ , et enfin  $\ell \geq L(f)$ , ce qui termine toute la preuve.  $\square$

En résumé on a :

(1.5.6) **Théorème**

Il y a correspondance bijective entre mesures et intégrales sur l'espace mesurable  $(\Omega, \Sigma)$ , la liaison  $\mu \leftrightarrow L$  étant donnée par la formule fondamentale

$$\mu A = L(1_A)$$

pour toute partie  $A \in \Sigma$ .

On remarquera que cette liaison définit complètement  $L$  si l'on sait déjà que les propriétés de (1.5.1.a) sont vérifiées. Car alors *nécessairement*  $L$  se détermine sur  $\mathfrak{E}_+(\Sigma)$  par

$$L(g) = \sum \alpha_k \mu A_k \quad \text{si} \quad g = \sum \alpha_k 1_{A_k}$$

et sur  $\overline{\mathfrak{F}}_+(\Sigma)$  par

$$L(f) = \lim L(g_n) \quad \text{si} \quad g_n \uparrow f \quad \text{avec} \quad g_n \in \mathfrak{E}_+(\Sigma).$$

Le travail fait plus haut dans la preuve de (1.5.5) a consisté surtout à garantir l'*existence* de l'intégrale  $L$ .

**Notations.** Dans la suite,  $\mu$  étant supposée donnée, on notera  $L(f)$  sous la forme d'une intégrale selon

$$L(f) = \int f \, d\mu = \int f(t) \, d\mu(t)$$

(1.5.7) *Exemples*

- **Mesure de Dirac  $\delta_a$ .** Ici  $a \in \Omega$  et  $\Sigma = \mathcal{P}(\Omega)$ . On définit  $\delta_a$  par  $\delta_a(A) = \begin{cases} 1 & \text{si } a \in A \\ 0 & \text{si } a \notin A \end{cases}$ , soit  $\delta_a(A) = 1_A(a)$ . Il en résulte que l'intégrale associée est  $L_a(f) = f(a)$ .
- **Mesures discrètes.** On fixe une suite  $a_k \in \Omega$  et une suite  $\alpha_k > 0$  et soit  $\mu = \sum \alpha_k \delta_{a_k}$ , qui correspond à l'intégrale  $L(f) = \sum \alpha_k f(a_k)$ .
- **Mesure dénombrement sur  $\mathbb{N}$ .** Ici  $\Omega = \mathbb{N}$  et  $\Sigma = \mathcal{P}(\mathbb{N})$ . Avec  $\mu = \sum \delta_k$ , on obtient :

$$\mu A = \begin{cases} \text{card } A & \text{si } A \text{ est fini} \\ +\infty & \text{si } A \text{ est infini} \end{cases}$$

Le cône  $\overline{\mathcal{F}}_+(\Sigma)$  s'identifie à  $(\overline{\mathbb{R}}_+)^{\mathbb{N}}$ , donc à l'ensemble des suites  $x = (x_n)$  avec  $x_n \in [0, +\infty]$ , et l'intégrale associée à  $\mu$  n'est autre que  $L(x) = \sum x_n$ .

- **Mesure image.** Étant donné  $(\Omega, \Sigma)$  et  $(X, \mathcal{C})$ , et une application mesurable  $h : \Omega \rightarrow X$ , la donnée d'une mesure  $\mu$  sur  $\Sigma$  détermine aussi une mesure  $\nu$  sur  $\mathcal{C}$ . En effet  $\mu$  est associée à  $L_\mu$  et l'application  $L$ , définie sur  $\overline{\mathcal{F}}_+(\mathcal{C})$  par  $L(g) = L_\mu(g \circ h)$ , vérifie les conditions (1.5.1). C'est donc une intégrale, dont la mesure associée  $\nu$  est définie par  $\nu(B) = L_\mu(1_B \circ h) = L_\mu(1_{h^{-1}(B)})$ , soit  $\nu(B) = \mu(h^{-1}(B))$  pour toute  $B \in \mathcal{C}$ . On dit que  $\nu$  est la mesure image de  $\mu$  par  $h$ , notée  $\nu = h(\mu)$ . On remarquera que si  $\mu = \delta_a$ , alors  $\nu = h(\mu) = \delta_{h(a)}$ , ce qui signifie que l'application  $\mu \rightarrow \nu = h(\mu)$ , définie sur les mesures, prolonge l'application  $h : a \rightarrow h(a)$ , définie sur les points.
- **Mesure à densité.** On fixe ici  $(\Omega, \Sigma)$ , une mesure  $\mu$  sur  $\Sigma$  et une fonction  $h \in \overline{\mathcal{F}}_+(\Sigma)$ .

Alors l'application  $L$ , définie sur  $\overline{\mathcal{F}}_+(\Sigma)$  par

$$L(f) = L_\mu(fh) = \int f h \, d\mu$$

est une intégrale, dont la mesure associée  $\nu$  sur  $\Sigma$  est donnée par

$$\nu(A) = \int 1_A h \, d\mu, \quad A \in \Sigma$$

ce que l'on note plutôt  $\nu(A) = \int_A h \, d\mu$

La construction par mesures images ou par mesures à densité permet donc d'obtenir de nouvelles mesures lorsqu'on en connaît une; il suffit d'avoir à sa disposition l'application mesurable ou la fonction mesurable  $h$ .

## 1.6 Construction des mesures

En pratique on ne dispose pas, a priori, d'une mesure sur une tribu  $\Sigma$  mais seulement d'une mesure sur un anneau  $\mathcal{R}$  (telle qu'elle va être définie) et le problème est de savoir si l'on peut *construire* une extension à la tribu  $\Sigma_{\mathcal{R}}$  engendrée par  $\mathcal{R}$ , extension qui, bien entendu, doit rester une mesure. Et du même coup un second problème se greffe sur le premier, à savoir *l'unicité* de l'extension.

### (1.6.1) Définition

Soit  $\mathcal{R}$  un anneau sur l'ensemble  $\Omega$ . On appelle mesure sur  $\mathcal{R}$  toute application  $\mu : \mathcal{R} \rightarrow [0, +\infty]$  vérifiant l'une ou l'autre des deux conditions équivalentes :

$$(M) \begin{cases} a/ \mu \emptyset = 0 \\ b/ \mu(A \cup B) = \mu A + \mu B \text{ si } A, B \in \mathcal{R} \text{ et } A \cap B = \emptyset \\ c/ B_n \in \mathcal{R}, B_n \uparrow B \text{ et } B \in \mathcal{R} \Rightarrow \mu(B) = \lim \mu B_n = \text{Sup } \mu B_n \end{cases}$$

$$(D) \begin{cases} a/ \mu \emptyset = 0 \\ b/ \text{Pour toute suite disjointe } (A_n) \text{ de } \mathcal{R}, \text{ telle que} \\ \quad A = \cup A_n \in \mathcal{R}, \text{ on a } \mu(A) = \sum \mu A_n \end{cases}$$

La condition  $M_b$  est celle d'additivité finie, la condition  $M_c$  celle de monotonie croissante, et la condition  $D_b$  celle d'additivité dénombrable. On passe de (M) à (D) en posant  $B_n = A_1 \cup A_2 \dots \cup A_n$  lorsque les  $A_k$  sont donnés et  $A_1 = B_1, A_2 = B_2 \setminus B_1, \dots, A_n = B_n \setminus B_{n-1}$ , lorsque les  $B_k$  sont donnés.

Bien entendu si  $\mathcal{R}$  est déjà une tribu  $\Sigma$  on retrouve la définition (1.5.2), de sorte que toute mesure  $\mu$  sur une tribu  $\Sigma$  induit une mesure sur tout anneau  $\mathcal{R}$  contenu dans  $\Sigma$ .

Les propriétés les plus immédiates des mesures sont rassemblées dans les deux énoncés suivants :

### (1.6.2) Proposition

Soit  $\mu$  une mesure sur l'anneau  $\mathcal{R}$

- $\mu$  est croissante :  $A \subset B \Rightarrow \mu A \leq \mu B$ .
- $A, B \in \mathcal{R}, B \subset A$  et  $\mu B < +\infty \Rightarrow \mu(A \setminus B) = \mu A - \mu B$ .
- $\mu$  est dénombrablement sous-additive au sens suivant :  
 $A_n \in \mathcal{R}, A \in \mathcal{R}$  et  $A \subset \cup A_n \Rightarrow \mu A \leq \sum \mu A_n$ .

**Preuve.** a) et b) sont immédiats avec  $M_b$ . Pour prouver c), remplaçons  $A_n$  par  $A'_n = A \cap A_n \in \mathcal{R}$  de sorte que  $A = \cup A'_n$ . Posons  $B_n = A'_1 \cup \dots \cup A'_n \uparrow A$ , de sorte que

$\mu A = \lim \mu B_n$ . Il reste à voir que  $\mu B_n \leq \sum_1^n \mu A'_k \leq \sum_1^n \mu A_k$ , ce qui résulte de la sous-additivité finie de  $\mu$ , conséquence de  $\mu(A \cup B) + \mu(A \cap B) = \mu A + \mu B$  pour  $A, B \in \mathcal{R}$ .  $\square$

En couplant b) avec la condition  $(M_b)$  de monotonie croissante, on obtient le résultat suivant, *fondamental* pour les applications.

(1.6.3) **Théorème (de monotonie décroissante)**

Soit  $\mu$  une mesure sur l'anneau  $\mathcal{R}$ . Pour toute suite  $(A_n)$  de  $\mathcal{R}$ , décroissante et telle que  $A = \bigcap A_n \in \mathcal{R}$  (donc  $A_n \downarrow A$ ) on a

$$\mu(A) = \lim \mu A_n = \inf \mu A_n$$

chaque fois qu'il existe un entier  $p$  tel que  $\mu A_p < +\infty$ .

En particulier sous cette condition, on a

$$A_n \downarrow \emptyset \Rightarrow \mu A_n \downarrow 0$$

**Remarque.** Si l'on abandonne la condition  $\mu A_p < +\infty$ , c'est-à-dire si l'on suppose  $\mu A_n = +\infty$  pour tout  $n$ , le résultat devient faux car on peut avoir  $\mu A = 0$ . Par exemple avec  $\Omega = \mathbb{N}$ ,  $\mathcal{R} = \mathcal{P}(\Omega)$ ,  $\mu = \sum \delta_n$ , mesure dénombrement, et  $A_n = [n, +\infty) \subset \mathbb{N}$ .

(1.6.4) **Exemple important : la mesure longueur sur  $\mathbb{R}$ .** Désignons par  $\mathcal{R}$  l'anneau engendré par les intervalles  $(a, b)$  de  $\mathbb{R}$ , sans qu'on précise si ces intervalles sont ouverts, fermés, ou bornés. Il est facile de vérifier que  $\mathcal{R}$  est exactement l'ensemble des réunions finies d'intervalles *disjoints*. On a alors :

Sur  $\mathcal{R}$ , il n'existe qu'une seule mesure  $\lambda$ , dite mesure de Borel, ou mesure longueur, telle que  $\lambda(a, b) = b - a$ .

L'unicité est évidente. Réciproquement, si  $A$ , réunion finie *disjointe*  $\cup (a_k, b_k)$ , est élément de  $\mathcal{R}$ , on pose  $\lambda A = \sum (b_k - a_k)$ , ce qui donne immédiatement les propriétés  $(M_a)$ ,  $(M_b)$  de (1.6.1). Mais il reste à prouver  $(M_c)$  ou, ce qui revient au même,  $(D_b)$ . Pour prouver  $(D_b)$  on se ramène au cas où l'intervalle  $(a, b)$  est exactement égal à une réunion  $\cup J_n$  d'intervalles *disjoints*  $J_n = (a_n, b_n)$ , et à démontrer que  $b - a = \sum \lambda(J_n)$ .

Or déjà on a  $\sum \lambda(J_n) \leq \lambda(a, b) = b - a$  car toute réunion finie (disjointe) des  $J_n$  est contenue dans  $(a, b)$ . Pour l'inégalité contraire, fixons un intervalle compact  $[\alpha, \beta] = K$  dans  $(a, b)$  et, pour  $\varepsilon > 0$  fixé, posons  $U_n = ] a_n - \varepsilon 2^{-n}, b_n + \varepsilon 2^{-n} [$ . Alors  $K$  est recouvert par la suite d'ouverts  $(U_n)$ , donc par Borel-Lebesgue, il suffit d'une famille finie pour recouvrir  $K$ ,

autrement dit il existe un entier  $p$  tel que  $K \subset \bigcup_{n=1}^p U_n$ . Par sous-additivité finie de  $\lambda$

(conséquence immédiate de l'additivité finie), on a  $\lambda K \leq \sum_{n=1}^p \lambda U_n \leq \sum_{n=1}^{\infty} \lambda U_n$ , soit

$\beta - \alpha \leq \sum \lambda (J_n) + \varepsilon \sum_{n=1}^{\infty} 2^{-n}$ . Avec  $\varepsilon \downarrow 0$ , on obtient  $\beta - \alpha \leq \sum \lambda (J_n)$  et avec  $\alpha \downarrow a$  et  $\beta \uparrow b$ , on a enfin  $b - a \leq \sum \lambda (J_n)$ .  $\square$

Ce qui est remarquable dans cette affaire est surtout que, du point de vue historique, l'axiome de Borel-Lebesgue (à la base de toute la théorie de la compacité) a été obtenu précisément pour prouver l'additivité dénombrable de la mesure longueur sur  $\mathbb{R}$ . Ce n'est donc pas un hasard si la théorie de la compacité (elle-même à la base de la topologie générale) porte le nom des initiateurs de la théorie de la mesure !

**Le problème de l'extension.** Il s'énonce très simplement : étant donné une mesure  $\mu$  sur un anneau  $\mathfrak{R}$ , peut-on la prolonger en une mesure sur la tribu engendrée par  $\mathfrak{R}$  ? Et si la réponse est positive, y a-t-il unicité de ce prolongement ? Nous allons voir que la réponse à la question de l'existence est toujours positive, mais celle relative à l'unicité est en général négative; toutefois elle est positive dans de nombreux cas qui couvrent le domaine des applications pratiques. Mais pour étudier ces questions il faut introduire de nouvelles notions.

**La mesure extérieure.** Étant donné une mesure  $\mu$  sur l'anneau  $\mathfrak{R}$ , on introduit une application  $\mu^* : \mathcal{P}(\Omega) \rightarrow [0, +\infty]$ , définie pour *tous* les ensembles de  $\Omega$  :

(1.6.5) *Définition*

On appelle mesure extérieure associée à  $\mu$  l'application  $\mu^*$  définie sur la tribu  $\mathcal{P}(\Omega)$  selon

$$\mu^*(E) = \begin{cases} +\infty & \text{si } E \text{ ne peut être recouvert par une suite d'éléments de } \mathfrak{R} \\ \text{Inf} \{ \sum \mu A_n, E \subset \bigcup A_n, A_n \in \mathfrak{R} \} & \text{sinon} \end{cases}$$

En particulier si  $\mathfrak{R}$  est une tribu  $\Sigma$  alors on a plus simplement

$$\mu^*(E) = \text{Inf} \{ \mu B, E \subset B \text{ et } B \in \Sigma \}$$

A partir de  $\mu$  on a donc une application  $\mu^*$  définie sur un ensemble  $\mathcal{P}(\Omega)$  beaucoup plus grand que  $\mathfrak{R}$ , mais le prix à payer est qu'en général  $\mu^*$  n'est pas additive sur  $\mathcal{P}(\Omega)$ . Les propriétés qui restent sont les suivantes :

(1.6.6) *Proposition*

Pour toute mesure  $\mu$  sur  $\mathfrak{R}$  la mesure extérieure  $\mu^*$  vérifie :

- a)  $\mu^*$  prolonge  $\mu$ , c'est-à-dire que  $\mu^* A = \mu A$  pour tout  $A \in \mathfrak{R}$ , et en particulier  $\mu^* \emptyset = 0$
- b)  $E \subset F \Rightarrow \mu^* E \leq \mu^* F$ , c'est-à-dire que  $\mu^*$  est croissante
- c)  $\mu^*$  est dénombrablement sous-additive

$$\mu^* \left( \bigcup E_n \right) \leq \sum \mu^* E_n$$

**Preuve.** a) Si  $A \in \mathfrak{R}$  on a déjà  $\mu^* A \leq \mu A$  avec  $A \subset A$  et  $\mu A \leq \mu^* A$  avec (1.6.2.c). Puisque b) est évident prouvons c) et pour cela il suffit de supposer  $\sum \mu^* E_n < +\infty$ . Alors pour

$\varepsilon > 0$  fixé, il existe, pour tout  $n \geq 1$ , un recouvrement  $E_n \subset \bigcup_k A_{n,k}$  avec  $A_{n,k} \in \mathfrak{R}$  et

$\sum_k \mu A_{n,k} \leq \mu^* E_n + \varepsilon 2^{-n}$ . La condition  $E \subset \bigcup_{k,n} A_{n,k}$  fournit donc

$$\mu^* E \leq \sum_{k,n} \mu A_{n,k} = \sum_n \sum_k \leq \varepsilon + \sum_n \mu^* E_n$$

ce qui donne c) avec  $\varepsilon \downarrow 0$ . □

### (1.6.7) Définition

Les parties  $N \subset \Omega$  telles que  $\mu^* N = 0$  sont appelées *négligeables* pour  $\mu$ . Soit  $\mathcal{N}$  leur classe.

Et évidemment

### (1.6.8) Proposition

- a)  $N \in \mathcal{N}$  et  $M \subset N \Rightarrow M \in \mathcal{N}$
- b) Toute réunion dénombrable d'ensembles négligeables est encore négligeable.
- c) Si l'anneau  $\mathfrak{R}$  est une tribu  $\Sigma$ , pour que l'ensemble  $N$  soit négligeable il faut et il suffit qu'il existe  $A \in \Sigma$  tel que  $N \subset A$  et  $\mu A = 0$ .

### (1.6.9) Exemples

**Ex 1.** Soit  $\Omega = I$  un ensemble non dénombrable (mais infini) et  $\Delta$  l'anneau des parties dénombrables ou finies de  $I$ . Alors la tribu  $\Sigma = \Sigma_\Delta$  engendrée est formée des parties  $A$  telles que  $A \in \Delta$  ou  $A^c \in \Delta$  (et on ne peut avoir les deux !). La mesure  $\mu = 0$  sur  $\Delta$  donne la mesure extérieure  $\mu^*$  sur  $\mathcal{P}(I)$ , définie par  $\mu^* E = 0$  si  $E \in \Delta$  et  $\mu^* E = +\infty$  si  $E \notin \Delta$ . La mesure  $\lambda$  sur  $\Sigma$  définie par  $\lambda A = 0$  si  $A \in \Delta$  et  $\lambda A = 1$  si  $A^c \in \Delta$  (on vérifiera que c'est bien une mesure !) donne comme mesure extérieure  $\lambda^* E = 0$  si  $E \in \Delta$  et  $\lambda^* E = 1$  si  $E \notin \Delta$ . Ici  $\mu^*$  est une mesure, mais  $\lambda^*$  n'est même pas additive. Par ailleurs les mesures  $\lambda$  sur  $\Sigma$  et  $\mu^* \upharpoonright_\Sigma$  sont différentes sur  $\Sigma$ , mais leurs restrictions coïncident sur l'anneau  $\Delta$ .

**Ex 2.** Pour la mesure longueur sur  $\mathbb{R}$  les ensembles négligeables sont les parties  $N$  telles que, pour tout  $\varepsilon > 0$ , il existe une suite  $(J_n)$  d'intervalles telle que  $N \subset \bigcup J_n$  et  $\sum \ell(J_n) \leq \varepsilon$ . Par exemple tout singleton  $\{a\}$  est négligeable, et par conséquent tout ensemble fini ou dénombrable. Ainsi  $\mathbb{Q}$  est négligeable dans  $\mathbb{R}$ , quoique partout dense. Dans l'autre sens, on peut trouver des ensembles négligeables dans  $\mathbb{R}$  qui soient équipotents à  $\mathbb{R}$  (donc non dénombrables), par exemple l'ensemble de Cantor  $K \subset [0, 1]$ .

Revenons maintenant au cas général d'une mesure  $\mu$  sur un anneau  $\mathfrak{R}$  et à sa mesure extérieure associée  $\mu^*$ . Puisque  $\mu^*|_{\mathfrak{R}} = \mu$ , il est clair que  $\mu^*$  est dénombrablement additive sur  $\mathfrak{R}$  et la question qui peut se poser est de savoir s'il existe des classes plus grandes que  $\mathfrak{R}$ , et en particulier une tribu, sur lesquelles  $\mu^*$  reste dénombrablement additive (et d'ailleurs additive suffirait grâce à la sous-additivité dénombrable (1.6.6.c)). Or il se trouve que les éléments  $A \in \mathfrak{R}$  possèdent, vis-à-vis de  $\mu^*$ , une propriété d'additivité bien particulière :

(1.6.10) *Lemme*

Soit  $A \in \mathfrak{R}$  une partie fixée. Alors  $A$  partitionne additivement (pour  $\mu^*$ ) toute partie  $G \subset \Omega$ , c'est-à-dire que l'on a

$$\mu^* G = \mu^*(G \cap A) + \mu^*(G \setminus A)$$

On peut dire encore que, pour toute  $E \subset A$  et toute  $F \subset A^c$ , on a

$$\mu^*(E \cup F) = \mu^* E + \mu^* F$$

**Preuve.** Il suffit de prouver l'inégalité  $\mu^* E + \mu^* F \leq \mu^*(E \cup F)$  de sorte qu'on peut supposer  $\mu^*(E \cup F) < +\infty$ . Alors soit  $(C_n)$  une suite de  $\mathfrak{R}$  telle que  $E \cup F \subset \bigcup C_n$ . En posant  $A_n = A \cap C_n$  et  $B_n = C_n \setminus A$  on a  $E \subset \bigcup A_n$  et  $F \subset \bigcup B_n$ , donc aussi

$$\mu^* E + \mu^* F \leq \sum \mu A_n + \sum \mu B_n = \sum (\mu A_n + \mu B_n) = \sum \mu C_n$$

d'où  $\mu^* E + \mu^* F \leq \mu^*(E \cup F)$  avec (1.6.5). □

Maintenant on peut se demander si les parties  $A \in \mathfrak{R}$  sont les seules à vérifier la propriété de partition du lemme. La réponse constitue le résultat-clé qui résout le problème de l'extension de la mesure  $\mu$ .

(1.6.11) *Théorème (Carathéodory)*

Soit  $\mu$  une mesure quelconque sur l'anneau  $\mathfrak{R}$ . Alors la classe  $\Sigma_\mu$  des parties  $A \subset \Omega$  vérifiant la propriété :

$$(C) \quad \mu^*(E \cup F) = \mu^* E + \mu^* F \text{ pour toutes } E \subset A \text{ et } F \subset A^c$$

est une tribu contenant l'anneau  $\mathfrak{R}$  (donc aussi la tribu  $\Sigma = \Sigma_{\mathfrak{R}}$  engendrée) et la classe des parties négligeables.

De plus, la mesure extérieure  $\mu^*$  est dénombrablement additive sur la tribu  $\Sigma_\mu$ .

**Preuve.** Déjà  $\Sigma_\mu$  contient  $\mathcal{R}$  (c'est le lemme) et si  $A = N \in \mathcal{N}$  alors

$$\mu^* E + \mu^* F = \mu^* F \leq \mu^* (E \cup F),$$

donc  $N \in \Sigma_\mu$ . La classe  $\Sigma_\mu$  est évidemment stable par passage au complémentaire et si  $A, B \in \Sigma_\mu$  alors pour  $E \subset A \cap B$  et  $F \subset (A \cap B)^c$ , posons  $F_A = F \cap A$ ,  $F_B = F \cap B$  et  $G = F \setminus (A \cup B)$ , de sorte que  $F = F_A \cup F_B \cup G$  et  $E \cup F = (E \cup F_A) \cup (G \cup F_B)$ . D'où

$$\begin{aligned} \mu^* (E \cup F) &= \mu^* (E \cup F_A) + \mu^* (G \cup F_B) && \text{car } A \in \Sigma_\mu \\ &= \mu^* E + \mu^* F_A + \mu^* G + \mu^* F_B && \text{car } B \in \Sigma_\mu \\ &\geq \mu^* E + \mu^* F && \text{car } F = F_A \cup F_B \cup G \end{aligned}$$

Ainsi  $\Sigma_\mu$  est stable par intersection, donc aussi par réunion. Pour voir que c'est une tribu, il suffit de prouver la stabilité par réunion dénombrable disjointe. Soit donc  $(A_n)$  une suite disjointe de  $\Sigma_\mu$ ,  $A = \bigcup A_n$ ,  $E \subset A$  et  $F \subset A^c$ ; posons  $E_n = E \cap A_n$ . Alors pour tout  $n$  fixé

$$\begin{aligned} \mu^* (E \cup F) &\geq \mu^* (E_1 \cup \dots \cup E_n \cup F) \\ &= \mu^* E_1 + \mu^* (E_2 \cup \dots \cup E_n \cup F) && \text{car } A_1 \in \Sigma_\mu \\ &= \mu^* E_1 + \mu^* E_2 + \mu^* (E_3 \cup \dots \cup E_n \cup F) && \text{car } A_2 \in \Sigma_\mu \\ &\vdots && \vdots \\ &= \mu^* E_1 + \mu^* E_2 + \dots + \mu^* E_n + \mu^* F && \text{car } A_n \in \Sigma_\mu \end{aligned}$$

Donc quand  $n \rightarrow +\infty$

$$\begin{aligned} \mu^* (E \cup F) &\geq \sum_1^\infty \mu^* E_k + \mu^* F \\ &\geq \mu^* E + \mu^* F && \text{car } E = \bigcup E_n \end{aligned}$$

Ainsi  $A \in \Sigma_\mu$  et  $\Sigma_\mu$  est une tribu. Il reste à prouver que  $\mu^*$  est une mesure sur  $\Sigma_\mu$  mais cela est évident car  $\mu^* \emptyset = \mu \emptyset = 0$  puisque  $\emptyset \in \mathcal{R}$  et si  $(A_n)$  est une suite disjointe de  $\Sigma_\mu$ , alors l'inégalité précédente

$$\mu^* (E \cup F) \geq \sum \mu^* E_k + \mu^* F \geq \mu^* E + \mu^* F$$

appliquée à  $E = A = \bigcup A_n$  et  $F = \emptyset$  donne

$$\mu^* (\bigcup A_n) = \sum \mu^* A_n$$

ce qui termine tout. □

### (1.6.12) Corollaire

Soit  $\mathcal{R}$  un anneau,  $\Sigma = \Sigma_{\mathcal{R}}$  la tribu engendrée. Alors pour toute mesure  $\mu$  sur  $\mathcal{R}$ , la restriction  $\bar{\mu} = \mu^*|_\Sigma$  est une mesure sur la tribu  $\Sigma$ , qui prolonge  $\mu$ . Et c'est même la plus grande mesure sur  $\Sigma$  qui prolonge  $\mu$ .

**Preuve.** Il est clair que  $\bar{\mu}$  est une mesure prolongeant  $\mu$ , ce qui résout complètement le problème de l'existence. Si  $\nu$  est une autre mesure sur  $\Sigma$  telle que  $\nu|_{\mathcal{R}} = \mu$ , prouvons

$\nu E \leq \bar{\mu} E = \mu^* E$  pour toute partie  $E \in \Sigma$ . On peut supposer  $\mu^* E < +\infty$ , et alors si  $E \subset \bigcup A_n$ , avec  $A_n \in \mathcal{R}$ , on a

$$\nu E \leq \sum \nu A_n = \sum \mu A_n$$

d'où  $\nu E \leq \mu^* E$  par passage à la borne inférieure.  $\square$

L'exemple 1 de (1.6.9) assure effectivement qu'il n'y a pas en général unicité du prolongement.

**Le problème de l'unicité.** Pour aborder cette question, il est nécessaire d'introduire encore de nouvelles notions, qui seront d'ailleurs utiles dans d'autres contextes. Il s'agit des classes de Dynkin.

(1.6.13) *Définition*

a) On appelle  $\pi$ -classe toute classe  $\mathcal{P}$  de parties de  $\Omega$  telle que

$$A, B \in \mathcal{P} \Rightarrow A \cap B \in \mathcal{P}$$

b) On appelle  $\lambda$ -classe toute classe  $\mathcal{L}$  de parties de  $\Omega$  telle que :

- $\Omega \in \mathcal{L}$

- $\mathcal{L}$  est stable par différence propre

$$A, B \in \mathcal{L}, B \subset A \Rightarrow A \setminus B \in \mathcal{L}$$

- $\mathcal{L}$  est stable par réunion dénombrable croissante

$$A_n \in \mathcal{L}, A_n \uparrow A \Rightarrow A \in \mathcal{L}$$

L'intérêt des  $\pi$ -classes et  $\lambda$ -classes est évidemment relatif aux tribus car une classe  $\Sigma$  est une tribu si et seulement si c'est à la fois une  $\pi$ -classe et une  $\lambda$ -classe. Et dans cet esprit on a le résultat fondamental :

(1.6.14) *Théorème (Dynkin)*

Soit  $\mathcal{P}$  une  $\pi$ -classe et  $\mathcal{L}$  une  $\lambda$ -classe telles que  $\mathcal{P} \subset \mathcal{L}$ . Alors la tribu  $\Sigma_{\mathcal{P}}$  engendrée par  $\mathcal{P}$  est contenue dans  $\mathcal{L}$

$$\mathcal{P} \subset \mathcal{L} \Rightarrow \Sigma_{\mathcal{P}} \subset \mathcal{L}$$

**Preuve.** Désignons par  $\mathcal{M}$  la  $\lambda$ -classe engendrée par  $\mathcal{P}$ , c'est-à-dire la plus petite  $\lambda$ -classe contenant  $\mathcal{P}$ . On a évidemment  $\mathcal{M} \subset \mathcal{L}$ , et si l'on démontre que  $\mathcal{M}$  est aussi une  $\pi$ -classe, alors ce sera une tribu, d'où  $\Sigma_{\mathcal{P}} \subset \mathcal{M} \subset \mathcal{L}$ . Pour prouver que  $A \in \mathcal{M}$  et  $B \in \mathcal{M} \Rightarrow A \cap B \in \mathcal{M}$ , opérons en deux temps. Fixons d'abord  $A \in \mathcal{P}$  et considérons la classe

$$\mathcal{D}_A = \{B \subset \Omega; A \cap B \in \mathcal{M}\}$$

Il est facile de voir que  $\mathfrak{D}_A$  est une  $\lambda$ -classe telle que  $\mathfrak{P} \subset \mathfrak{D}_A$ , puisque  $\mathfrak{P}$  est une  $\pi$ -classe, donc  $\mathcal{M} \subset \mathfrak{D}_A$  et ainsi  $A \in \mathfrak{P}$  et  $B \in \mathcal{M} \Rightarrow A \cap B \in \mathcal{M}$ .

Ensuite fixons  $B \in \mathcal{M}$  et considérons la classe

$$\mathfrak{D}_B = \{ A \subset \Omega ; A \cap B \in \mathcal{M} \}$$

qui est, de la même façon, une  $\lambda$ -classe contenant  $\mathfrak{P}$ , donc aussi  $\mathcal{M}$ .

Ainsi  $A, B \in \mathcal{M} \Rightarrow A \cap B \in \mathcal{M}$  et tout est dit.  $\square$

(1.6.15) *Définition*

Soit  $\mathfrak{R}$  un anneau et  $\mu$  une mesure sur  $\mathfrak{R}$ .

- a) On dit que  $\mu$  est finie (ou bornée) lorsque  $\Omega \in \mathfrak{R}$  et  $\mu\Omega < +\infty$ . En particulier  $\mu$  est une probabilité si  $\mu\Omega = 1$ .
- b) On dit que  $\mu$  est  $\sigma$ -finie lorsqu'il existe une suite croissante  $\Omega_n \in \mathfrak{R}$  telle que  $\Omega = \bigcup \Omega_n$ , et  $\mu\Omega_n < +\infty$  pour tout  $n$ .

Nous avons maintenant tous les ingrédients pour établir le résultat d'unicité utile pour les applications.

(1.6.16) *Théorème (d'unicité)*

- a) Soit  $\Sigma$  une tribu et  $\mathfrak{P}$  une  $\pi$ -classe engendrant  $\Sigma$ . Alors deux probabilités sur  $\Sigma$ , qui coïncident sur  $\mathfrak{P}$ , sont nécessairement égales.
- b) Soit  $\mathfrak{R}$  un anneau,  $\Sigma = \Sigma_{\mathfrak{R}}$  sa tribu engendrée. Alors deux mesures sur  $\Sigma$ , qui sont  $\sigma$ -finies sur  $\mathfrak{R}$ , et qui coïncident sur  $\mathfrak{R}$ , sont nécessairement égales.

*Preuve.* a) Soit  $\mu, \nu$  les deux probabilités. Alors la classe

$$\mathfrak{L} = \{ A \in \Sigma, \mu A = \nu A \}$$

est trivialement une  $\lambda$ -classe contenant  $\mathfrak{P}$ , donc  $\mathfrak{L} = \Sigma$  et  $\mu = \nu$ .

b) Soit  $\mu, \nu$  les deux mesures. Il existe une suite  $\Omega_n \in \mathfrak{R}$  telle que  $\Omega_n \uparrow \Omega$  et  $\mu\Omega_n = \nu\Omega_n < +\infty$ . Pour chaque  $k$  la classe

$$\mathfrak{L}_k = \{ A \in \Sigma ; \mu(A \cap \Omega_k) = \nu(A \cap \Omega_k) \}$$

est une  $\lambda$ -classe contenant  $\mathfrak{R}$ , donc  $\mathfrak{L}_k = \Sigma$  puisque  $\mathfrak{R}$  est une  $\pi$ -classe. Alors pour toute  $A \in \Sigma$  on a  $\mu(A \cap \Omega_k) = \nu(A \cap \Omega_k)$ , et par monotonie croissante on obtient  $\mu A = \nu A$ , donc  $\mu = \nu$ .  $\square$

En rassemblant (1.6.11) et 1.6.16) on obtient le résultat définitif important.

(1.6.17) *Théorème*

Soit  $\mu$  une mesure  $\sigma$ -finie sur l'anneau  $\mathfrak{R}$ . Alors  $\mu$  possède une unique mesure prolongement à la tribu  $\Sigma = \Sigma_{\mathfrak{R}}$  engendrée par  $\mathfrak{R}$ , coïncidant avec la mesure extérieure  $\mu^*$  sur  $\Sigma$ .

De plus la tribu  $\Sigma_{\mu}$  des parties, dites  $\mu$ -mesurables, vérifiant la condition (C) de (1.6.11), est exactement la tribu engendrée par  $\Sigma$  et la classe  $\mathcal{N}$  des parties négligeables, formée d'ailleurs des ensembles  $E$  tels qu'il existe une partie  $B \in \Sigma$  vérifiant  $E \Delta B \in \mathcal{N}$ .

**Preuve.** Désignons par  $\mathfrak{C}$  la tribu engendrée par  $\Sigma$  et  $\mathcal{N}$ , telle que  $\mathfrak{C} \subset \Sigma_{\mu}$ , contenant déjà comme éléments toutes les parties  $E$  du type précédent. Réciproquement soit  $E \in \Sigma_{\mu}$ . Si  $(\Omega_n)$  est une suite de  $\mathfrak{R}$  telle que  $\mu \Omega_n < +\infty$  et  $\Omega_n \uparrow \Omega$ , alors  $E \cap \Omega_n = E_n \in \Sigma_{\mu}$  et  $\mu^* E_n = \bar{\mu} E_n < +\infty$ . Il en résulte (facile) qu'il existe  $B_n \in \Sigma$  telle que  $E_n \subset B_n$  et  $\mu^* E_n = \mu^* B_n = \mu B_n$ , donc  $\mu^*(B_n \setminus E_n) = 0$  par additivité de  $\mu^*$  sur  $\Sigma_{\mu}$ . Alors  $E \subset B$  et  $N = B \setminus E$  vérifie  $N \subset \bigcup (B_n \setminus E_n)$ , donc  $N \in \mathcal{N}$ . Et évidemment  $E \Delta B = N \in \mathcal{N}$ .  $\square$

**Notations.** Dans le cas de  $\sigma$ -finitude, il n'y a donc aucune différence entre mesures sur un anneau  $\mathfrak{R}$  et mesures sur une tribu  $\Sigma$ . On convient alors de noter encore  $\mu$  l'extension à  $\Sigma$  et de réserver la notation  $\bar{\mu}$  à la mesure  $\mu^*|_{\Sigma_{\mu}}$  sur la tribu des parties  $\mu$ -mesurables. Bien entendu la mesure extérieure  $\mu^*$  reste définie sur  $\mathfrak{P}(\Omega)$ , mais ce n'est pas une mesure en général, puisqu'elle n'est même pas additive sur  $\mathfrak{P}(\Omega)$ .

Revenons maintenant au cas  $\Omega = \mathbb{R}$  ou  $\Omega = \mathbb{R}^p$ , avant d'aborder l'étude des fonctions intégrables.

## 1.7 Mesures sur $\mathbb{R}$ et $\mathbb{R}^p$

**Mesure de Borel sur  $\mathbb{R}$ .** On a vu en (1.6.4) que la fonctionnelle "longueur" est en réalité une mesure  $\lambda$  sur l'anneau  $\mathfrak{R}$  engendré par les intervalles  $(a, b)$  de  $\mathbb{R}$ , et cette mesure est évidemment  $\sigma$ -finie. Elle se prolonge donc de façon unique en une mesure toujours dite de Borel, sur la tribu borélienne  $\mathfrak{B}$  de  $\mathbb{R}$ , qui est exactement la tribu  $\Sigma_{\mathfrak{R}}$  engendrée par  $\mathfrak{R}$ .

Dans ce contexte la tribu  $\Sigma_{\lambda}$  porte le nom de *tribu de Lebesgue* et la mesure associée  $\bar{\lambda}$  s'appelle aussi mesure de Lebesgue.

Mais alors toute fonction continue  $f$  sur  $\mathbb{R}$ , qui est positive (c'est-à-dire telle que  $f \geq 0$ ), est un élément de  $\bar{\mathfrak{F}}_+$ , donc admet une intégrale  $\int f d\lambda$  par rapport à  $\lambda$ . Et la question est

posée de comparer cette intégrale à l'intégrale classique  $\int_{-\infty}^{+\infty} f(x) dx = \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow +\infty}} \int_a^b f(x) dx$ ,

positive finie ou non.

C'est ici que nous rencontrons le premier résultat important de cohérence :

(1.7.1) **Proposition**

Pour toute fonction  $f$ , continue et positive sur  $\mathbb{R}$ , on a

$$\int f \, d\lambda = \int_{-\infty}^{+\infty} f(x) \, dx \leq +\infty$$

**Preuve.** Fixons  $n$  et pour  $-n \cdot 2^n \leq k < n \cdot 2^n$  considérons l'intervalle  $I_{n,k} = [k \cdot 2^{-n}, (k+1)2^{-n} [$ .

Alors  $\bigcup_k I_{n,k} = J_n = ]-n, +n [$ . Définissons maintenant la fonction "en escaliers"  $\varphi_n$  par

$$\varphi_n = \begin{cases} 0 & \text{en dehors de } J_n \\ \text{Inf } \{f(t), t \in I_{n,k}\} & \text{sur } I_{n,k} \end{cases}$$

Un peu d'attention montre que la suite  $(\varphi_n)$  est croissante et que  $\varphi_n(x) \uparrow f(x)$  en tout point  $x$ , puisque  $f$  est continue. Par monotonie croissante on a donc  $\int f \, d\lambda = \lim \int \varphi_n \, d\lambda$ .

$$\text{Or } \int \varphi_n \, d\lambda = \int_{-n}^n \varphi_n(x) \, dx = \int_{-\infty}^{+\infty} \varphi_n(x) \, dx \text{ et il reste à voir que } \int_{-\infty}^{+\infty} \varphi_n(x) \, dx \uparrow \int_{-\infty}^{+\infty} f(x) \, dx .$$

$$\text{Or, par } \varphi_n \leq f, \text{ on a déjà } \int_{-\infty}^{+\infty} \varphi_n(x) \, dx \leq \int_{-\infty}^{+\infty} f(x) \, dx \text{ et } \int f \, d\lambda \leq \int_{-\infty}^{+\infty} f(x) \, dx .$$

Réciproquement la continuité uniforme de  $f$  sur chaque intervalle  $[a, b]$  compact, assure que  $\varphi_n \uparrow f$  *uniformément* sur  $[a, b]$ , d'où suit

$$\int_a^b f(x) \, dx = \lim \int_a^b \varphi_n \, dx \leq \lim \int_{-\infty}^{+\infty} \varphi_n(x) \, dx = \int f \, d\lambda$$

$$\text{et ainsi } \int_{-\infty}^{+\infty} f(x) \, dx \leq \int f \, d\lambda \text{ avec } a \downarrow -\infty \text{ et } b \uparrow \infty . \quad \square$$

**Mesure de Borel sur  $\mathbb{R}^p$ .** On remplace ici les intervalles  $(a, b)$  par les pavés

$P = \prod (a_k, b_k)$  et la longueur  $b - a$  par le "volume"  $\lambda_p(P) = \prod_k (b_k - a_k)$ . Alors sur l'anneau  $\mathcal{R} = \mathcal{R}_p$  des réunions finies de pavés (que l'on peut supposer disjointes) la fonctionnelle  $\lambda_p$  est une mesure. On le voit en démontrant d'abord que  $\lambda_p$  est additive, puis qu'elle est dénombrablement additive, de la même façon qu'en (1.6.4), en utilisant la compacité des pavés fermés bornés  $\prod [\alpha_k, \beta_k]$ . Et la même preuve que (1.7.1) donne aussi :

(1.7.2) *Proposition*

Pour toute fonction  $f$ , continue et positive sur  $\mathbb{R}^p$ , on a

$$\int f d\lambda_p = \int_{\mathbb{R}^p} f(x) dx = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x) dx_1 \dots dx_p$$

**Mesures boréliennes sur  $\mathbb{R}$ .** Revenons au cas  $p = 1$  pour déterminer *toutes* les mesures sur la tribu borélienne  $\mathfrak{B}$  de  $\mathbb{R}$ , qui sont finies sur les compacts, et qu'on appelle plus sommairement mesures boréliennes. Donnons tout d'abord une condition nécessaire :

(1.7.3) *Proposition*

Soit  $\mu$  une mesure borélienne sur  $\mathbb{R}$ . Pour tout  $a \in \mathbb{R}$  on définit la fonction  $F_a$  selon

$$F_a(t) = \begin{cases} \mu([a, t[) & \text{si } t \geq a \\ -\mu([t, a[) & \text{si } t \leq a \end{cases}$$

On a alors :

- a)  $F_a$  est croissante et continue à gauche, et  $F_a(a) = 0$
- b)  $F_a(b) = -F_b(a)$
- c)  $F_a(t) = F_a(b) + F_b(t)$  (relation de Chasles).

**Preuve.** Seule la continuité à gauche mérite le détail. Or si  $t_n \uparrow t$  alors  $F_a(t_n) \uparrow F_a(t)$  car :

- si  $t \leq a$  on a  $[t_n, a[ \downarrow [t, a[$  et la monotonie décroissante
- si  $t > a$  on a  $[a, t_n[ \uparrow [a, t[$  et la monotonie croissante. □

On voit donc que les fonctions  $F_a$  se déduisent de l'une d'entre elles par addition d'une constante. On peut donc considérer, en faisant  $a = 0$  par exemple, que  $\mu$  est définie sur l'anneau  $\mathfrak{R}$  engendré par les intervalles  $(a, b)$  par la donnée d'une fonction  $F$ , croissante et continue à gauche, et il est alors facile de vérifier les formules :

$$(*) \quad \begin{cases} \mu([s, t[) = F(t) - F(s) & s \leq t \\ \mu([s, t]) = F(t+0) - F(s) & s \leq t \\ \mu(]s, t]) = F(t+0) - F(s+0) & s \leq t \\ \mu(]s, t[) = F(t) - F(s+0) & s < t \end{cases}$$

La réciproque est vraie.

(1.7.4) *Proposition*

Soit  $F$  une fonction croissante et continue à gauche sur  $\mathbb{R}$ . Il lui correspond une unique mesure borélienne  $\mu = \mu_F$  sur  $\mathbb{R}$  vérifiant les conditions (\*). On dit que  $\mu_F$  est la mesure de Borel-Stieltjes associée à  $F$ .

**Preuve.** On utilise ici l'anneau  $\mathcal{J}$  des réunions finies d'intervalles  $[s, t[$ , introduit en (1.3.6), et qui engendre la tribu borélienne  $\mathfrak{B}$ , en définissant  $\mu_F$  par additivité finie à partir de  $\mu_F([s, t[) = F(t) - F(s)$  pour  $s \leq t$ .

Et la seule chose intéressante à prouver est l'additivité dénombrable, à savoir que si  $[s, t[ = \bigcup_n J_n$ , où la réunion est disjointe, avec  $J_n = [s_n, t_n[$ , alors  $\mu_F([s, t[) = \sum \mu_F(J_n)$ . Par additivité finie on a déjà  $\sum \mu_F(J_n) \leq \mu_F([s, t[)$ . Pour prouver  $\mu_F([s, t[) \leq \sum \mu_F(J_n)$ , on peut supposer  $s < t$  et si  $\tau \in [s, t[$  est fixé, ainsi que  $\varepsilon > 0$ , on commence, grâce à la continuité à gauche de  $F$ , à introduire  $\sigma_n < s_n$  tel que  $F(s_n) - F(\sigma_n) \leq \varepsilon 2^{-n}$  pour tout  $n$ . Avec  $K = [s, \tau]$  et  $U_n = ]\sigma_n, t_n[$ , on obtient un recouvrement ouvert  $K \subset \bigcup U_n$ , de sorte qu'un recouvrement fini suffit, et ainsi il existe  $p$  tel que

$$[s, \tau[ \subset K \subset \bigcup_{n=1}^p U_n \subset \bigcup_{n=1}^p [s_n, t_n[$$

Par additivité finie, et sous-additivité finie, on a donc

$$\begin{aligned} F(\tau) - F(s) &\leq \sum_{n=1}^p [F(t_n) - F(\sigma_n)] \\ &\leq \sum_{n=1}^p [F(t_n) - F(s_n) + \varepsilon 2^{-n}] \\ &\leq \varepsilon + \sum_{n=1}^{\infty} [F(t_n) - F(s_n)] \end{aligned}$$

et avec  $\varepsilon \downarrow 0$  d'abord, puis  $\tau \uparrow t$  ensuite (et la continuité à gauche de  $F$ ) on obtient finalement

$$\mu_F[s, t[) = F(t) - F(s) \leq \sum_{n=1}^{\infty} [F(t_n) - F(s_n)] = \sum \mu_F(J_n)$$

ce qui suffit.  $\square$

(1.7.5) **Remarque.** Lorsque  $\mu$  est une probabilité sur  $\mathbb{R}$ , on lui associe volontiers la fonction  $F$  définie par  $F(t) = \mu(-\infty, t[$ , appelée *fonction de répartition* de  $\mu$ . Elle est telle que  $F(t) \rightarrow 0$  quand  $t \downarrow -\infty$  et  $F(t) \uparrow 1$  quand  $t \uparrow +\infty$ , et bien entendu croissante et continue à gauche.

L'importance de (1.7.3) et (1.7.4), qui décrivent *toutes* les mesures boréliennes sur  $\mathbb{R}$ , n'apparaîtra que plus loin quand nous parlerons des variables aléatoires et de leurs lois. Cette importance justifie que l'on s'attarde un peu en décrivant de façon plus détaillée les deux sortes de mesures qui interviennent : à savoir les mesures *diffuses* et les mesures *discrètes*. Remarquons déjà qu'il résulte des formules (\*) que l'on a

$$\mu(\{a\}) = F(a + 0) - F(a) = \Delta_F(a)$$

où  $\Delta_F(a)$  est le saut de  $F$  au point  $a$ . Ainsi

(1.7.6) *Proposition*

On dit que  $\mu$  est diffuse lorsque  $\mu(\{a\}) = 0$  pour tout  $a \in \mathbb{R}$ . Pour cela il est nécessaire et suffisant que la fonction  $F$  associée soit continue sur  $\mathbb{R}$ .

(1.7.7) *Exemples*

Ex 1. La mesure de Borel sur  $\mathbb{R}$ , soit  $\lambda$ , associée à  $F(t) = t$ .

Ex 2. La probabilité uniforme sur  $[0, 1]$  définie par la mesure  $\mu$  telle que  $\mu(a, b) = \lambda((a, b) \cap [0, 1])$ , définie par la fonction

$$F(t) = \begin{cases} 0 & \text{si } t \leq 0 \\ t & \text{si } 0 \leq t \leq 1 \\ 1 & \text{si } t \geq 1 \end{cases} = \text{Sup} \{ 0, \text{Inf}(t, 1) \}$$

Ex 3. La donnée d'une fonction continue  $h \geq 0$  sur  $\mathbb{R}$  détermine la mesure  $\mu$ , de

densité  $h$  par rapport à la mesure de Borel et telle que  $\mu(a, b) = \int_a^b h(t) dt$ . La fonction  $F$  associée est donc une primitive de  $h$ . Pour que  $\mu$  soit une probabilité, il faut et il suffit

$$\text{que } \int_{-\infty}^{+\infty} h(t) dt = 1.$$

Ex 4. La probabilité de Cauchy, de paramètre  $a > 0$ , définie par la densité  $h(t) = \frac{1}{\pi} \frac{a}{a^2 + t^2}$  et la fonction de répartition  $F(t) = \frac{1}{2} + \frac{1}{\pi} \text{Arc tg } \frac{t}{a}$ .

Ex 5. La probabilité de Laplace-Gauss réduite, notée  $\gamma$  ou LG (0, 1), définie par la densité  $h(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)$  et la fonction de répartition  $F = \Phi$ , dite *fonction d'erreur*

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{t^2}{2}\right) dt$$

sachant bien entendu que l'intégrale de Gauss  $\int_{-\infty}^{\infty} \exp\left(-\frac{t^2}{2}\right) dt$  est égale à  $\sqrt{2\pi}$ .

Ex 6. La probabilité de Laplace-Gauss LG ( $m, \sigma$ ) de paramètres  $m \in \mathbb{R}$  et  $\sigma > 0$ , associée à la densité

$$h(t) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left[-\frac{(t-m)^2}{2\sigma^2}\right]$$

et à la fonction de répartition  $F(x) = \Phi\left(\frac{x-m}{\sigma}\right)$ .

**Ex 7. La probabilité générale d'Euler.** On rappelle rapidement que la fonction  $\Gamma$  d'Euler est définie, pour  $x > 0$ , par

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$$

Elle vérifie la relation fondamentale  $\Gamma(x+1) = x \Gamma(x)$ , d'où l'on tire, avec  $\Gamma(1) = 1$ , l'égalité  $\Gamma(n+1) = n!$  pour  $n$  entier.

Étant donné deux paramètres  $\alpha > 0$  et  $\beta > 0$ , on leur associe la probabilité, dite d'Euler, définie par la densité

$$h(t) = \begin{cases} 0 & \text{si } t \leq 0 \\ \frac{1}{\beta^\alpha \Gamma(\alpha)} t^{\alpha-1} \exp\left(-\frac{t}{\beta}\right) & \text{si } t > 0 \end{cases}$$

**Ex 8.** Soit  $\gamma = \text{LG}(0, 1)$  et soit  $g(x) = |x|^\alpha$  avec  $\alpha > 0$ . On désigne par  $\mu$  la mesure image de  $\gamma$  par la fonction  $g$ . Déterminer sa densité  $h$ . Pour quelle valeur de  $\alpha$  obtient-on une probabilité d'Euler ?

(Réponse :  $h(t) = 0$  si  $t < 0$ ,  $h(t) = \frac{2}{\alpha \sqrt{2\pi}} t^{\frac{1}{\alpha}-1} \exp\left(-\frac{t^{2/\alpha}}{2}\right)$  si  $t > 0$  et  $\mu$  est la probabilité d'Euler si  $\alpha = 2$ , de paramètres  $\alpha = 2$  et  $\beta = \frac{1}{2}$ ).

**Ex 9.** Soit  $\gamma = \text{LG}(0, 1)$  et  $g(x) = e^x$ . On désigne par  $\mu$  la probabilité image de  $\gamma$  par la fonction  $g$ , dite probabilité log-normale. Donner sa densité.

(Réponse :  $h(t) = 0$  si  $t < 0$ ,  $h(t) = \frac{1}{\sqrt{2\pi}} \frac{1}{t} \exp\left[-\frac{1}{2} (\ln t)^2\right]$  si  $t > 0$ ).

Passons maintenant au cas des mesures discrètes, dont le prototype est la mesure de Dirac  $\delta_a$  au point  $a \in \mathbb{R}$ . Le cas général s'obtient en fixant une suite  $(a_n)$  de points de  $\mathbb{R}$  et une suite  $(\alpha_n)$  de scalaires positifs ou nuls, donnant  $\mu = \sum \alpha_n \delta_{a_n}$ . On dit quelquefois que  $\mu$  s'obtient en fixant les masses  $\alpha_n$  aux points  $a_n$ . Lorsque  $\sum \alpha_n = 1$  on obtient une probabilité.

**Ex 10. Probabilité de Bernoulli de paramètre  $p$ .** On fixe  $0 < p < 1$  et  $q = 1 - p$ . Alors  $\mu = p \delta_1 + q \delta_0$ .

**Ex 11. Probabilité binomiale de paramètres  $n$  et  $p$ .** Avec  $n$  entier  $\geq 1$ ,  $0 < p < 1$  et  $q = 1 - p$ , on construit

$$\mu = \sum_{k=0}^n C_n^k p^k q^{n-k} \delta_k$$

Pour  $n = 1$  on retrouve la probabilité de Bernoulli.

**Ex 12. Probabilité de Poisson de paramètre  $\lambda > 0$ ,** définie par

$$\mu = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \delta_k$$

**Ex 13.** Probabilité géométrique de paramètre  $p$ . Avec  $0 < p < 1$  et  $q = 1 - p$ , on pose

$$\mu = p \sum_{k=0}^{\infty} q^k \delta_{k+1}$$

On constate qu'une mesure discrète  $\mu$  est totalement déterminée par les valeurs  $\mu(\{a\})$ , alors que ces valeurs, qui sont nulles, ne donnent aucune information pour les mesures diffuses. En un sens mesures discrètes et mesures diffuses apparaissent comme complémentaires. Ce point de vue est précisé par l'exercice théorique suivant :

(1.7.8) *Exercice*

Soit  $\mu = \mu_F$  une mesure de Borel-Stieltjes quelconque. Démontrer les assertions :

- a) L'ensemble  $D$  des points  $a \in \mathbb{R}$  tels que  $\mu(\{a\}) > 0$  est fini (éventuellement vide) ou dénombrable.
- b) La mesure discrète  $\lambda = \sum_{t \in D} \mu(\{t\}) \delta_t$  est telle que  $\lambda \leq \mu$ .
- c) La mesure  $\mu$  peut se décomposer en  $\mu = \lambda + \nu$ , où  $\nu$  est une mesure diffuse.
- d) La décomposition de  $\mu$  en somme d'une mesure discrète et d'une mesure diffuse est unique.

## 1.8 Fonctions intégrables

Revenons maintenant à la situation générale d'un espace mesurable  $(\Omega, \Sigma)$ ,  $\Sigma$  étant donc une tribu, sur lequel existe une mesure  $\mu$ , associée évidemment à une intégrale  $f \rightarrow \int f d\mu$ , définie sur  $\bar{\mathcal{F}}_+$  et à valeurs dans  $[0, +\infty]$ .

Avant d'aborder la notion de fonction intégrable, nous allons compléter certains résultats relatifs à  $\bar{\mathcal{F}}_+$  et à l'intégrale  $\int f d\mu$ , intéressants surtout par les propriétés de finitude.

Déjà la propriété de monotonie croissante, dite de Beppo Lévi, conduit à un résultat moins précis, mais où la monotonie a disparu.

(1.8.1) *Théorème (Fatou)*

Soit  $(f_n)$  une suite quelconque de  $\bar{\mathcal{F}}_+$ . On a alors

$$\int (\liminf f_n) d\mu \leq \liminf \int f_n d\mu$$

En particulier si  $f_n \rightarrow f$  simplement et si  $\int f_n d\mu \leq M$  alors  $\int f d\mu \leq M$ .

**Preuve.** Déjà  $\liminf f_n = \sup_n \inf_{k \geq n} f_k$ , et puisque la suite  $h_n = \inf_{k \geq n} f_k$  est monotone croissante, on a par Beppo Lévi

$$\int (\liminf f_n) d\mu = \sup_n \int h_n d\mu$$

Mais pour  $k \geq n$  on a  $h_n \leq f_k$ , donc  $\int h_n d\mu \leq \int f_k d\mu$  et aussi  $\int h_n d\mu \leq \inf_{k \geq n} \int f_k d\mu$ , d'où le résultat voulu, à savoir

$$\begin{aligned} \int (\liminf f_n) d\mu &\leq \sup_n \inf_{k \geq n} \int f_k d\mu \\ &\leq \liminf \int f_n d\mu \end{aligned} \quad \square$$

**Remarque.** On pourra vérifier sur des exemples simples que même si  $f_n \rightarrow f$ , la suite des intégrales  $\int f_n d\mu$  peut ne pas avoir de limite, et si elle en a une cette limite peut être différente de  $\int f d\mu$ . Le théorème de Fatou ne constitue donc qu'une approche assez grossière au véritable théorème de passage à la limite dans les intégrales qui sera donné plus loin sous le nom de "théorème de convergence dominée de Lebesgue".

Il importe maintenant de savoir dans quelles conditions on a  $\int f d\mu < +\infty$  ou  $\int f d\mu = 0$  lorsque  $f \in \bar{\mathfrak{L}}_+$  est fixée. Pour cela nous avons besoin de dégager la notion de propriété vraie  $\mu$ -presque partout. Étant donné une propriété  $P(x)$ , dépendant du point  $x \in \Omega$ , on dit que " $P$  est vraie  $\mu$  pp" ou " $P(x)$  est vraie pour  $\mu$ -presque tout  $x$ " lorsque l'ensemble  $\{x; P(x) \text{ n'est pas vérifiée}\}$  est négligeable.

Alors :

(1.8.2) **Proposition**

Soit  $f \in \bar{\mathfrak{L}}_+$ . Pour que  $\int f d\mu = 0$  il faut et il suffit que  $f = 0$   $\mu$ pp.

**Preuve.** Soit  $A = \{x; f(x) > 0\} = \{f > 0\}$ . Alors les suites croissantes  $(nf)$  et  $(n1_A)$  ont la même limite, qui est la fonction  $(+\infty)1_A$ . Par Beppo Lévi on en déduit que

$$\lim n \int f d\mu = \lim n \mu A$$

d'où résulte aisément que  $\int f d\mu = 0 \Leftrightarrow \mu A = 0$ . □

(1.8.3) **Corollaire**

Soit  $f, g \in \bar{\mathfrak{L}}_+$ .

a) Si  $f \leq g$   $\mu$ pp alors  $\int f d\mu \leq \int g d\mu$

b) Si  $f = g$   $\mu$ pp alors  $\int f d\mu = \int g d\mu$

**Preuve.** Il suffit de prouver a). Or posons  $B = \{f > g\}$ ; on obtient un élément de  $\Sigma$  tel que  $\mu B = 0$ , et tel aussi que  $f 1_{B^c} \leq g 1_{B^c}$  (en tout point). Écrivons  $f = f 1_{B^c} + f 1_B$  et  $g = g 1_{B^c} + g 1_B$ , et remarquons que  $f 1_B = 0$   $\mu$ pp et  $g 1_B = 0$   $\mu$ pp. Alors, avec la proposition on a

$$\int f d\mu = \int f 1_{B^c} d\mu \leq \int g 1_{B^c} d\mu = \int g d\mu. \quad \square$$

**Remarque.** Il importe de bien voir, réciproquement que l'égalité  $\int f d\mu = \int g d\mu$  n'implique aucune relation intéressante entre  $f$  et  $g$  dans le cas général.

(1.8.4) **Proposition**

Soit  $f \in \bar{\mathfrak{L}}_+$  telle que  $\int f d\mu < +\infty$ .

Alors on a  $f < +\infty$   $\mu$ pp.

**Preuve.** Soit  $A = \{f = +\infty\} \in \Sigma$ . On a donc  $n 1_A \leq f$  pour tout  $n$ , donc  $n \mu A \leq \int f d\mu < +\infty$  et  $\mu A = 0$ .  $\square$

Cette proposition presque triviale (connaissant les propriétés d'une mesure) a cependant comme conséquences des résultats extrêmement importants pour les applications.

(1.8.5) **Corollaire**

a) Soit  $(f_n)$  une suite croissante de  $\bar{\mathfrak{L}}_+$ . On suppose que la suite  $(\int f_n d\mu)$  reste bornée dans  $\mathbb{R}$ . Alors la suite  $(f_n)$  est  $\mu$ pp convergente dans  $\mathbb{R}$ .

b) Soit  $(f_n)$  une suite quelconque de  $\bar{\mathfrak{L}}_+$ . On suppose que la série  $\sum \int f_n d\mu$  est convergente dans  $\mathbb{R}$ . Alors la série  $\sum f_n$  est elle-même  $\mu$ pp convergente dans  $\mathbb{R}$ , et en particulier on a  $f_n \rightarrow 0$   $\mu$ pp.

On rencontre ici un résultat typique de la théorie de Lebesgue. A savoir qu'une information numérique (une suite de nombres est bornée, ou bien une série est convergente) donne une information fonctionnelle (une suite de fonctions converge ponctuellement, ou bien une série de fonctions converge ponctuellement), le prix à payer étant l'intervention du "négligeable", rassemblant les points où "ça ne passe pas". On voit donc que les ensembles négligeables de  $\Sigma$  s'introduisent naturellement dans le jeu.

**Fonctions intégrables.** Rappelons les espaces  $\mathfrak{L} = \mathfrak{L}(\Sigma)$  et  $\mathfrak{E} = \mathfrak{E}(\Sigma)$  regroupant les fonctions réelles (finies) mesurables et les fonctions étagées mesurables, ainsi que leurs cônes positifs  $\mathfrak{L}_+$  et  $\mathfrak{E}_+$ , contenus dans  $\bar{\mathfrak{L}}_+$ . Pour chaque  $f \in \mathfrak{L}$  on pose

$$f^+ = \text{Sup}(f, 0); \quad f^- = \text{Sup}(-f, 0)$$

qui sont des fonctions de  $\mathfrak{L}_+$  vérifiant les égalités

$$\begin{cases} f = f^+ - f^- \\ |f| = f^+ + f^- \end{cases} \quad \begin{cases} f^+ = \frac{1}{2}(|f| + f) \\ f^- = \frac{1}{2}(|f| - f) \end{cases}$$

(1.8.6) *Définition*

Soit  $f$  une fonction mesurable réelle.

- a) On dit que  $f$  est intégrable pour  $\mu$  lorsque l'on a  $\int |f| \, d\mu < +\infty$ .
- b) Dans ce cas le nombre réel fini

$$I(f) = \int f^+ \, d\mu - \int f^- \, d\mu$$

est appelé l'intégrale de  $f$  par rapport à  $\mu$ , que l'on note encore  $\int f \, d\mu$  ou  $\int f(t) \, d\mu(t)$ , ou même  $\int f(t) \, \mu(dt)$ .

On a alors facilement :

(1.8.7) *Proposition*

- a) Toute fonction mesurable, qui est  $\mu$ pp nulle, est intégrable et d'intégrale nulle.
- b) Si  $f$  est une fonction mesurable telle que  $|f| \leq g$ , où  $g$  est intégrable, alors  $f$  est intégrable.
- c) Si  $f$  est intégrable et  $g$  mesurable et bornée, alors  $fg$  est intégrable.

Désignons désormais par  $\mathfrak{L}^1(\Omega, \Sigma, \mu)$ , ou seulement  $\mathfrak{L}^1(\mu)$  ou même  $\mathfrak{L}^1$  l'ensemble des fonctions intégrables. Les propriétés les plus simples de  $\mathfrak{L}^1$  sont :

(1.8.8) *Théorème*

- a)  $\mathfrak{L}^1$  est un sous-espace vectoriel réticulé de l'algèbre  $\mathfrak{L}$ .
- b) L'intégrale  $I(\cdot)$  est une forme linéaire positive sur  $\mathfrak{L}^1$ . En particulier, si  $f \in \mathfrak{L}^1$  on a l'inégalité, dite de la moyenne

$$\left| \int f \, d\mu \right| \leq \int |f| \, d\mu$$

**Le théorème de convergence dominée de Lebesgue.** On aborde maintenant l'un des aspects les plus célèbres de la théorie, qui va donner un outil puissant pour les calculs

d'intégrales, en explicitant un cas très général où l'on peut intervertir l'intégration et le passage à la limite simple pour les suites.

(1.8.9) **Théorème (Convergence dominée de Lebesgue)**

Soit  $(f_n)$  une suite de fonctions intégrables telles que :

- a)  $f_n$  converge  $\mu$ pp vers une fonction mesurable  $f$ .
- b) Il existe une fonction positive intégrable  $g$  telle que
 
$$|f_n| \leq g \quad \mu\text{pp pour chaque } n$$
 (condition de domination).

Alors  $f$  est intégrable et  $\int f_n d\mu \rightarrow \int f d\mu$ .

**Preuve.** Les ensembles d'exception  $N = \{x, f_n(x) \not\rightarrow f(x)\}$ ,  $M_n = \{|f_n| > g\}$ ,  $M = \bigcup M_n$ , sont tous éléments de  $\Sigma$ , et négligeables, donc  $B = N \cup M$  est négligeable. En modifiant toutes les fonctions en les remplaçant par 0 sur  $B$ , c'est-à-dire en les multipliant par  $1_{B^c}$ , on ne change ni leur mesurabilité, ni leur intégrabilité, ni leurs intégrales. On peut donc supposer que  $f_n \rightarrow f$  partout et  $|f_n| \leq g$  partout. Soit alors

$$u_n = \inf_{k \geq n} f_k, \quad v_n = \sup_{k \geq n} f_k, \quad w_n = v_n - u_n$$

de sorte que  $u_n \uparrow f$ ,  $v_n \downarrow f$  et  $w_n \downarrow 0$ . De plus la condition  $u_n \leq f_k \leq v_n$  pour  $k \geq n$  garantit que  $u_n \leq f \leq v_n$ , donc aussi  $|f_n - f| \leq w_n$ , et puisque  $|f| \leq g$ , la fonction  $f$  est intégrable et

$$\left| \int f_n d\mu - \int f d\mu \right| \leq \int |f_n - f| d\mu \leq \int w_n d\mu$$

Tout sera démontré si l'on prouve que  $\int w_n d\mu \downarrow 0$ . Or  $w_n \leq w_1 \leq 2g$  assure que  $w_1 - w_n \uparrow w_1$  et  $\int w_1 d\mu < +\infty$ . Par Beppo Lévi on a donc  $\int (w_1 - w_n) d\mu \rightarrow \int w_1 d\mu$ , d'où  $\int w_n d\mu \rightarrow 0$  par différence.  $\square$

**Remarque.** 1) Si la suite  $(f_n)$  converge partout vers  $f$ , alors  $f$  est automatiquement mesurable. Mais si  $f_n \rightarrow f$  seulement  $\mu$ pp, alors la mesurabilité de  $f$  doit être supposée, sauf si toutefois la tribu de départ  $\Sigma$  contient les ensembles négligeables (auquel cas on dit que la mesure  $\mu$  est *complète* sur  $\Sigma$ ).

2) On donnera quelques exemples simples montrant que la propriété  $\int f_n d\mu \rightarrow \int f d\mu$  devient fausse si l'on abandonne la condition de convergence dominée, même si on la remplace par la condition notoirement plus faible  $\int |f_n| d\mu \leq M$ , ce qui ramène au théorème de Fatou (1.8.1).

La première conséquence du théorème de Lebesgue est essentielle pour les applications puisqu'elle donne le résultat le plus puissant sur la permutation du signe  $\int$  d'intégration et du signe  $\Sigma$  de sommation.

(1.8.10) **Théorème (Intégration terme à terme des séries)**

Soit  $(f_n)$  une suite de fonctions intégrables telles que

$$\sum_n \int |f_n| d\mu < +\infty$$

Alors la série  $\sum f_n$  est  $\mu$ pp absolument convergente dans  $\mathbb{R}$ , et il existe une fonction intégrable  $F$  telle que

$$F = \sum f_n \quad \mu\text{pp} \quad \text{et} \quad \int F d\mu = \sum \int f_n d\mu$$

*Preuve.* Soit  $h = \sum |f_n| \in \bar{\mathcal{F}}_+$ . Alors  $\int h d\mu = \sum \int |f_n| d\mu < +\infty$ , donc  $h$  est  $\mu$ pp finie. L'ensemble  $B = \{h < +\infty\} \in \Sigma$  est donc tel que  $\mu B^c = 0$ , de sorte que la fonction  $g = h 1_B$  est

partout finie et intégrable. Par ailleurs les sommes partielles  $F_n = \sum_{k=0}^n f_k$  sont telles que

$|F_n| \leq h$ , donc  $|F_n| \leq g$   $\mu$ pp et  $F_n \rightarrow F = \sum_0^\infty f_n$  sur  $B$ , puisque la convergence absolue implique la convergence. Il n'y a plus qu'à appliquer le théorème de Lebesgue en définissant  $F$  partout par  $F = \sum f_n 1_B$ .  $\square$

**Fonctions complexes intégrables.** Si  $f = g + ih$  est une fonction complexe  $\Omega \rightarrow \mathbb{C}$ , on dit qu'elle est mesurable lorsqu'elle est mesurable par rapport à la tribu borélienne  $\mathcal{B}_2$  de  $\mathbb{C} = \mathbb{R}^2$  et à la tribu  $\Sigma$  sur  $\Omega$ . Puisque  $\mathcal{B}_2$  est engendrée par les ensembles quadrants  $\{x < \alpha, y < \beta\}$ , on voit que  $f$  est mesurable ssi ses parties réelle  $g = \text{Re } f$  et imaginaire  $h = \text{Im } f$  sont mesurables. Comme par ailleurs

$$|f| \leq |g| + |h| \leq 2|f|$$

on voit aussi que  $|f|$  est intégrable ssi  $g$  et  $h$  sont intégrables. Ce dernier fait légitime la définition suivante :

(1.8.11) **Définition**

Une fonction complexe mesurable  $f = g + ih$  est dite intégrable (par rapport à la mesure  $\mu$ ) lorsque  $|f|$  est elle-même intégrable. Pour une telle fonction la quantité

$$I(f) = \int f d\mu = \int g d\mu + i \int h d\mu$$

est appelée l'intégrale de  $f$  par rapport à  $\mu$ .

L'espace  $\mathfrak{L}^1 = \mathfrak{L}^1(\mu, \mathbb{C})$  de ces fonctions complexes intégrables est évidemment un espace vectoriel sur lequel la fonctionnelle intégrale est linéaire. De plus, on a encore validité de la formule de la moyenne (1.8.8) sous la forme

(1.8.12) *Proposition (Formule de la moyenne)*

$$\left| \int f \, d\mu \right| \leq \int |f| \, d\mu$$

*Preuve.* Posons  $\int f \, d\mu = \rho e^{i\theta}$  avec  $\rho = \left| \int f \, d\mu \right|$  et ainsi

$$\begin{aligned} \rho &= \int e^{-i\theta} (g + ih) \, d\mu && \text{si } f = g + ih \\ &= \int (g \cos \theta + h \sin \theta) \, d\mu && \text{car } \rho \text{ est réel} \\ &\leq \int |g \cos \theta + h \sin \theta| \, d\mu \leq \int |f| \, d\mu \end{aligned}$$

car pour  $z = x + iy = r e^{i\varphi}$  on a

$$|x \cos \theta + y \sin \theta| = |r \cos(\theta - \varphi)| \leq r. \quad \square$$

**Intégration sur une partie A.** Étant donné une fonction intégrable  $f$ , alors pour toute partie  $A \in \Sigma$ , la fonction  $f 1_A$  est encore intégrable. On convient donc de poser, en accord avec les notations de (1.5.7) relatives aux mesures à densité

$$\int_A f \, d\mu = \int f 1_A \, d\mu$$

On a alors la propriété suivante.

(1.8.13) *Proposition*

$$\left| \begin{array}{l} \text{Soit } f \text{ une fonction intégrable (éventuellement complexe). Pour que l'on ait} \\ f = 0 \text{ } \mu\text{pp, il faut et il suffit que } \int_A f \, d\mu = 0 \text{ pour toute partie } A \in \Sigma. \end{array} \right.$$

*Preuve.* On se ramène immédiatement au cas réel, et seule la suffisance de la condition est à prouver. Or posons  $A = \{f \geq 0\} \in \Sigma$  et  $B = A^c = \{f < 0\}$ , ce qui donne

$$|f| = f 1_A - f 1_B$$

d'où  $\int |f| \, d\mu = \int_A f \, d\mu - \int_B f \, d\mu = 0$  et  $|f| = 0 \text{ } \mu\text{pp}$  avec (1.8.2).  $\square$

On remarquera bien la différence entre (1.8.2) et (1.8.13). Dans le premier cas la fonction  $f$  est positive et  $\int f \, d\mu = 0$  suffit pour impliquer  $f = 0 \text{ } \mu\text{pp}$ ; dans le second cas  $f$  est quelconque et bien entendu la nullité de l'intégrale  $\int f \, d\mu$  sur tout l'espace ne suffit plus; il y faut la nullité de *toutes* les intégrales  $\int_A f \, d\mu$  pour  $A \in \Sigma$ .

Lorsque  $h$  est une fonction intégrable et positive, alors l'application  $A \rightarrow \int_A h \, d\mu$ , définie sur  $\Sigma$ , est une mesure  $\nu$ , notée en général  $h \cdot \mu$  ou  $h\mu$ , ou  $d\nu = h d\mu$ . Et c'est même exactement la mesure de densité  $h$  par rapport à  $\mu$ . Mais ici  $\nu$  est en plus une mesure finie sur  $\Sigma$  puisque

$$\nu(\Omega) = \int h \, d\mu < +\infty$$

Avec (1.8.13) on voit que si  $\nu$  admet deux densités  $h_1$  et  $h_2$  par rapport à  $\mu$ , alors

$\int_A h_1 \, d\mu = \int_A h_2 \, d\mu$  pour toute partie  $A \in \Sigma$ , donc  $h_1 = h_2$   $\mu$ pp. On a donc unicité de la densité modulo l'égalité  $\mu$ pp. En résumé :

(1.8.14) **Proposition**

Pour toute fonction  $h$  intégrable et positive, la fonctionnelle  $\nu$  définie sur  $\Sigma$  par

$$\nu(A) = \int_A h \, d\mu$$

est une mesure finie, et c'est la mesure de densité  $h$ . De plus on a

- a) Toute partie  $A \in \Sigma$  qui est  $\mu$ -négligeable est nécessairement  $\nu$ -négligeable.
- b) Pour qu'une fonction mesurable  $f : \Omega \rightarrow \mathbb{R}$  soit intégrable par rapport à la mesure  $\nu$ , il faut et il suffit que la fonction  $f \cdot h$  soit intégrable par rapport à  $\mu$ , et dans ce cas

$$\int h \, d\nu = \int f \, h \, d\mu$$

**Preuve.** L'assertion b) provient de (1.5.7) puisque l'intégrale associée à  $\nu$  a été précisément définie par la relation  $\int f \, d\nu = L_\nu(f) = \dot{L}_\mu(fh) = \int f \, h \, d\mu$  pour toute  $f \in \bar{\mathcal{F}}_+$ . Il suffit donc de remplacer la fonction intégrable  $f$  par les fonctions  $|f|$ ,  $f^+$ ,  $f^-$  successives pour conclure.  $\square$

**Remarque.** La condition a) donne une condition *nécessaire* pour qu'une mesure finie soit *densitable* par rapport à  $\mu$ , qui peut être utile dans certains cas. Par exemple si  $\mu = \lambda$  est la mesure de Borel sur  $\mathbb{R}$  et si  $\nu = \delta_a$  est une mesure de Dirac, ou même une mesure discrète  $\nu = \sum \alpha_n \delta_{a_n}$  avec  $0 < \sum \alpha_n < +\infty$  de façon que  $\nu$  soit finie et non nulle, alors l'ensemble au plus dénombrable  $D = \{a\}$ , ou  $D = \{a_1, a_2, \dots, a_n, \dots\}$  est tel que  $\lambda(D) = 0$  et  $\nu(D) = \nu(\Omega)$ . La mesure  $\nu$  n'est donc pas densitable par rapport à la mesure  $\lambda$ .

**Intégration par rapport à une mesure image.** Fixons une application mesurable  $h : (\Omega, \Sigma) \rightarrow (X, \mathcal{C})$  entre deux espaces mesurables et désignons, conformément à (1.5.7), par  $\nu = h(\mu)$  la mesure image par  $h$  de toute mesure  $\mu$  sur  $\Sigma$ . On rappelle que pour  $B \in \mathcal{C}$  et  $f : X \rightarrow \bar{\mathbb{R}}_+$  mesurable on a

$$\nu(B) = \mu [ h^{-1}(B) ]$$

$$\int f \, d\nu = \int (f \circ h) \, d\mu = \int f [h(t)] \, d\mu(t)$$

ce qui fournit, de la même manière qu'en (1.8.14), le résultat :

(1.8.15) **Proposition**

Pour qu'une fonction mesurable  $f : X \rightarrow \mathbb{R}$  (ou  $X \rightarrow \mathbb{C}$ ) soit intégrable par rapport à la mesure image  $\nu = h(\mu)$ , il faut et il suffit que la fonction  $f \circ h : \Omega \rightarrow \mathbb{R}$  (ou  $\Omega \rightarrow \mathbb{C}$ ) soit intégrable pour la mesure  $\mu$ . De plus, dans ce cas

$$\int f \, d\nu = \int (f \circ h) \, d\mu = \int f [h(t)] \, d\mu(t)$$

**Preuve.** Tout provient de  $|f \circ h| = |f| \circ h$ , qui ramène au cas  $f \geq 0$ . □

Donnons deux exemples d'application, le second étant déjà assez élaboré.

(1.8.16) **Exemple 1. Mesure de Borel sur  $\mathbb{R}^p$  et homothéties.** Fixons une homothétie  $H = H(O, k)$ , de centre  $O$  et de rapport  $k > 0$ , et considérons la mesure image  $\mu = H(\lambda_p)$ , où  $\lambda_p$  est la mesure de Borel sur  $\mathbb{R}^p$ . Alors pour tout pavé  $P$  de  $\mathbb{R}^p$  on a évidemment

$$\mu(P) = \lambda_p (H^{-1}(P)) = \frac{1}{k^p} \lambda_p (P)$$

de sorte que  $\mu$  et  $\frac{1}{k^p} \lambda_p$  sont deux mesures  $\sigma$ -finies sur la tribu borélienne  $\mathfrak{B}_p$  de  $\mathbb{R}^p$ , qui coïncident sur la classe des pavés, donc sur l'anneau  $\mathfrak{R}$  engendré (qui est formé, rappelons-le, des réunions finies disjointes de pavés). Comme  $\mu$  et  $\frac{1}{k^p} \lambda_p$  sont en fait  $\sigma$ -finies sur  $\mathfrak{R}$ , on peut appliquer le théorème d'unicité (1.6.16.b) et obtenir que  $\mu = \frac{1}{k^p} \lambda_p$ . Or  $H^{-1} = H\left(O, \frac{1}{k}\right)$ , de sorte qu'en changeant  $k$  en  $\frac{1}{k}$  on obtient :

Pour toute partie borélienne  $A$  de  $\mathbb{R}^p$ , désignons par  $kA = H_k(A)$  l'image de  $A$  dans l'homothétie  $H_k = H(O, k)$  de centre l'origine et de rapport  $k > 0$ . Alors

$$\lambda_p (kA) = k^p \lambda_p (A)$$

Pour  $p = 1, 2, 3$  on retrouve un résultat classique lorsqu'on particularise l'ensemble  $A$ .

(1.8.16) **Exemple 2. Intégration des fonctions sphériques sur  $\mathbb{R}^p$ .** Désignons par  $\|x\| = (x_1^2 + \dots + x_p^2)^{1/2}$  la norme euclidienne sur  $\mathbb{R}^p$ . On dit qu'une fonction  $g : \mathbb{R}^p \rightarrow \mathbb{R}$  ou

$\mathbb{C}$  est sphérique, ou radiale, si elle ne dépend que de  $\|x\|$ , autrement dit s'il existe une fonction  $f : [0, +\infty) \rightarrow \mathbb{R}$  ou  $\mathbb{C}$ , telle que  $g(x) = f(\|x\|)$ . Il est clair alors, si  $f$  est

borélienne, que l'intégrale  $\int g d\lambda_p = \int g(x) dx$ , si elle existe, va pouvoir se calculer aisément sous la forme

$$\int g(x) dx = \int f(\|x\|) dx = \int f(t) d\mu(t)$$

en introduisant la mesure image  $\mu = h(\lambda_p)$ , avec  $h(x) = \|x\|$ . Il faut donc déterminer  $\mu$  et pour cela, en remarquant que  $\mu$  est une mesure sur  $[0, +\infty)$ , calculer sa fonction de répartition

$$F_\mu(t) = \mu[0, t[ = \lambda_p \{ x ; \|x\| < t \}$$

pour  $t > 0$ . Or la boule ouverte  $\{ \| \bullet \| < t \}$  est l'homothétique de la boule unité ouverte  $\overset{\circ}{B}_p = \{ x, \|x\| < 1 \}$  dans l'homothétie de rapport  $t$ . Il en résulte que si l'on désigne par  $V_p = \lambda_p(\overset{\circ}{B}_p)$  le  $p$ -volume de cette boule unité, on a nécessairement

$F_\mu(t) = V_p t^p = \int_0^t p V_p u^{p-1} du$ , ce qui prouve que  $\mu$  est définie par la densité  $p V_p u^{p-1}$  sur  $[0, +\infty)$ . Et puisque  $\mu$  est diffuse, on voit que  $V_p$  est aussi le volume  $\lambda_p(B_p)$  de la boule unité fermée  $B_p$ . En résumé, et ne connaissant pas encore  $V_p$ , on a :

(1.8.17) **Proposition**

Soit  $V_p = \lambda_p(B_p)$  le volume  $p$ -dimensionnel de la boule unité  $B_p$  de  $\mathbb{R}^p$ . On a l'égalité

$$\int_{\mathbb{R}^p} f(\|x\|) dx = p V_p \int_0^\infty t^{p-1} f(t) dt$$

chaque fois que  $f$  est borélienne et positive sur  $[0, +\infty)$ , ou bien telle que  $\int_0^\infty t^{p-1} |f(t)| dt < +\infty$ .

**Calcul de  $V_p$ .** En particulierisant  $f$  on doit pouvoir calculer le volume  $V_p$ . Choisissons en effet  $f(t) = \exp\left(-\frac{t^2}{2}\right)$  et désignons par

$$I = \int_{-\infty}^\infty \exp\left(-\frac{t^2}{2}\right) dt = 2 \int_0^\infty \exp\left(-\frac{t^2}{2}\right) dt$$

l'intégrale de Gauss (égale à  $\sqrt{2\pi}$ , mais dont nous allons justement retrouver la valeur).

Puisque  $\int_{\mathbb{R}^p} \exp\left[-\frac{\|x\|^2}{2}\right] dx = I^p$ , on a déjà

$$I^p = p V_p \int_0^{\infty} t^{p-1} \exp\left(-\frac{t^2}{2}\right) dt$$

Or l'intégrale  $J_p = \int_0^{\infty} t^p \exp\left(-\frac{t^2}{2}\right) dt$  est telle que

$$J_{p+1} = \int_0^{\infty} t^p d\left[-\exp\left(-\frac{t^2}{2}\right)\right] = p J_{p-1}$$

d'où l'on tire

$$\begin{cases} J_{2n+1} = 2.4 \dots 2n J_1 = 2^n n! & \text{car } J_1 = 1 \\ J_{2n} = 1.3 \dots (2n-1) \frac{1}{2} & \text{car } J_0 = \frac{1}{2} \end{cases}$$

de sorte que la formule  $I^p = p V_p J_{p-1}$  donne

$$V_{2n} = \frac{1^{2n}}{2^n n!} \qquad V_{2n+1} = \frac{2 I^{2n}}{1.3 \dots (2n+1)}$$

Or l'égalité  $V_2 = \pi$  fournit  $I^2 = 2\pi$ , donc  $I = \sqrt{2\pi}$ , et par suite on obtient le résultat définitif

$$\left| \begin{array}{l} V_{2n} = \frac{\pi^n}{n!} \\ V_{2n+1} = \frac{2^{n+1} \pi^n}{1.3 \dots (2n+1)} \end{array} \right.$$

## 1.9 Intégrale de Riemann et intégrale de Lebesgue

Il s'agit ici de comparer l'intégrale de Riemann sur  $\mathbb{R}$ , qui généralise l'intégrale des fonctions continues, avec l'intégrale que nous venons de construire, associée à la mesure de Borel "longueur"  $\lambda$ .

Il convient toutefois de bien séparer le cas de l'intégrale de Riemann sur  $[a, b]$  du cas de l'intégrale généralisée (ou impropre) sur  $\mathbb{R}$ , qui n'est pas, à proprement parler, une intégrale de Riemann.

**Intégrale sur  $[a, b]$ .** Soit  $f$  une fonction réelle bornée sur  $[a, b]$ . A chaque subdivision  $\sigma : a = x_0 < x_1 < \dots < x_{n+1} = b$ , on associe les valeurs

$$m_k = \text{Inf} \{ f(t) ; t \in [x_k, x_{k+1}] \}$$

$$M_k = \text{Sup} \{ f(t) ; t \in [x_k, x_{k+1}] \}$$

puis les sommes de Darboux

$$s(\sigma) = \sum m_k (x_{k+1} - x_k) ; \qquad S(\sigma) = \sum M_k (x_{k+1} - x_k)$$

On rappelle que  $f$  est dite Riemann-intégrable sur  $[a, b]$ , s'il existe un nombre réel  $I$ , appelé intégrale de  $f$ , tel que les sommes  $s(\sigma)$  et  $S(\sigma)$  tendent vers  $I$  quand le pas  $\delta(\sigma) = \text{Max}(x_{k+1} - x_k)$  tend vers 0. On sait alors (théorème de Riemann) qu'il faut et il suffit que l'on ait

$$\Omega(\sigma) = S(\sigma) - s(\sigma) \rightarrow 0 \quad \text{quand} \quad \delta(\sigma) \rightarrow 0$$

Associant à chaque subdivision  $\sigma$  les fonctions étagées  $\varphi_\sigma = \sum m_k 1_{J_k}$  et  $\psi_\sigma = \sum M_k 1_{J_k}$ , avec  $J_k = [x_k, x_{k+1}]$  [ pour  $0 \leq k < n$  et  $J_n = [x_n, x_{n+1}]$  ] (de façon à partitionner  $[a, b]$  et non pas seulement  $[a, b[$ ), on voit que

$$\varphi_\sigma \leq f \leq \psi_\sigma, \quad s(\sigma) = \int \varphi_\sigma d\lambda, \quad S(\sigma) = \int \psi_\sigma d\lambda.$$

Choisissons maintenant une suite de subdivisions emboîtées en prenant pour  $\sigma_n$  la subdivision définie par un découpage équidistant de  $[a, b]$ , de pas  $\delta_n = 2^{-n}(b-a)$ , soit :

$$x_k = a + \frac{k}{2^n}(b-a) \quad k = 0, 1, \dots, 2^n$$

et notons  $\varphi_n$  et  $\psi_n$  les fonctions étagées associées. Alors la suite  $(\varphi_n)$  est croissante, la suite  $(\psi_n)$  est décroissante et

$$\int (\psi_n - \varphi_n) d\lambda = S(\sigma_n) - s(\sigma_n) = \Omega(\sigma_n).$$

En posant  $\varphi = \lim \varphi_n$  et  $\psi = \lim \psi_n$ , on obtient l'encadrement  $\varphi \leq f \leq \psi$  et la condition de Riemann donne, par monotonie décroissante

$$\int (\psi - \varphi) d\lambda = \lim \Omega(\sigma_n) = 0$$

d'où l'on déduit  $\psi = \varphi$  presque partout, donc aussi  $f = \varphi = \psi$  presque partout, et bien entendu

$$\int_a^b f(x) dx = \int \varphi d\lambda = \int \psi d\lambda.$$

Maintenant il est clair que les fonctions  $\varphi$  et  $\psi$  sont Borel-mesurables (car limites simples de fonctions étagées), mais  $f$  ne l'est pas nécessairement. En fait, avec (1.6.17) et le début de (1.7), on voit que  $f$  est en réalité mesurable par rapport à la tribu  $\Sigma_\lambda = \mathfrak{B}_\lambda$  de Lebesgue, engendrée par  $\mathfrak{B}$  et les ensembles négligeables, donc aussi intégrable par rapport à la mesure de Lebesgue  $\bar{\lambda}$  sur  $\mathfrak{B}_\lambda$ , et telle enfin que  $\int f d\bar{\lambda} = \int \varphi d\bar{\lambda} = \int_a^b f(x) dx$ .

En résumé, on a :

(1.9.1) **Théorème**

Toute fonction  $f$  Riemann-intégrable sur  $[a, b]$  est Lebesgue-mesurable et Lebesgue-intégrable sur  $[a, b]$  (mais pas nécessairement Borel-mesurable) et

$$\int_a^b f(x) dx = \int_{[a,b]} f d\bar{\lambda}$$

On voit donc que l'intégrale de Lebesgue est beaucoup plus générale que l'intégrale de Riemann. Elle permet d'ailleurs d'intégrer des fonctions non bornées, ce que ne peut faire l'intégrale de Riemann. Mais même dans le champ des fonctions bornées, elle présente un avantage (et nous en verrons beaucoup d'autres). Par exemple, la fonction  $g$  de Dirichlet, égale à la fonction indicatrice de  $\mathbb{Q} \cap [a, b]$ , est Lebesgue-intégrable (et Borel-mesurable) et d'intégrale nulle puisque  $g = 0$  p.p ; elle est *partout* discontinue, et pour chaque subdivision  $\sigma$  on a  $m_k = 0$  et  $M_k = 1$ , donc

$$S(\sigma) - s(\sigma) = \Omega(\sigma) = b - a$$

ce qui assure que  $g$  n'est pas Riemann-intégrable.

Pour compléter cet état comparatif rappelons qu'une fonction  $h$  est dite "*en escaliers*" s'il existe une partition finie de  $[a, b]$  formée d'intervalles, sur chacun desquels  $h$  est constante. Toute fonction en escaliers est donc étagée, mais la réciproque est fautive évidemment. On dit encore qu'une fonction  $f$  est *réglée*, lorsqu'elle est limite uniforme sur  $[a, b]$  de fonctions en escaliers : c'est le cas des fonctions continues et des fonctions monotones. On fait d'ailleurs souvent la théorie élémentaire de l'intégration sur  $[a, b]$  avec les fonctions réglées. Et les comparaisons avec l'intégrale de Riemann et l'intégrale de Lebesgue, du point de vue de la puissance de la théorie, se font avec l'énoncé suivant :

- a) L'ensemble des points de discontinuité d'une fonction réglée  $f$  est au plus dénombrable et  $f$  admet en chacun de ces points une limite à droite et une limite à gauche.
- b) L'ensemble des points de discontinuité d'une fonction Riemann-intégrable est négligeable (pour la mesure de Borel  $\lambda$ )
- c) Une fonction Lebesgue-intégrable peut être partout discontinue.

**Intégrale généralisée sur un intervalle  $I$ .** Ici on part avec un intervalle  $I = (\alpha, \beta)$  supposé a priori non compact (donc non borné, ou alors borné mais non fermé), et une fonction  $f$  définie sur  $I$  et supposée seulement localement intégrable au sens de Riemann, c'est-à-dire Riemann-intégrable sur tout sous-intervalle compact  $[a, b]$  de  $I$ . L'intégrale généralisée (ou impropre) de  $f$  sur  $I$  est définie par

$$\int_I f(x) dx = \lim_{a \downarrow \alpha, b \uparrow \beta} \int_a^b f(x) dx$$

lorsque cette limite existe dans  $\mathbb{R}$ . On dit encore que l'intégrale  $\int_I f(x) dx$  est *convergente*, et qu'elle est *divergente* dans le cas contraire.

D'après (1.9.1)  $f$  est déjà localement Lebesgue-intégrable sur  $I$ . Mais elle peut ne pas être Lebesgue-intégrable sur  $I$ . En effet :

(1.9.2) *Théorème*

Soit  $f$  une fonction localement Riemann-intégrable sur  $I$ . Pour que  $f$  soit Lebesgue-intégrable, il faut et il suffit que l'intégrale impropre  $\int_I f(x) dx$  soit absolument convergente, c'est-à-dire que l'intégrale  $\int_I |f(x)| dx$  soit convergente (donc finie). On a alors dans ce cas l'égalité

$$\int_I f(x) dx = \int_I f d\bar{\lambda}$$

*Preuve.* On fixe les suites  $(a_n)$  et  $(b_n)$  telles que  $a_n \downarrow \alpha$  et  $b_n \uparrow \beta$  et soit  $\varphi_n = 1_{[a_n, b_n]}$ .

Si  $\int_I |f(x)| dx < +\infty$  alors

$$\int \varphi_n |f| d\bar{\lambda} = \int_{a_n}^{b_n} |f| dx \uparrow \int_I |f(x)| dx$$

et ainsi la suite  $\varphi_n |f|$  croissant vers  $|f|$ , le théorème de Beppo Lévi donne  $\int_I |f| d\bar{\lambda} = \int_I |f(x)| dx < +\infty$ , donc  $f$  est Lebesgue-intégrable sur  $I$ . Alors le théorème de convergence dominée de Lebesgue s'applique à la suite  $\varphi_n f$ , de sorte que

$$\int_I f d\bar{\lambda} = \lim \int \varphi_n f d\bar{\lambda} = \lim \int_{a_n}^{b_n} f(x) dx = \int_I f(x) dx$$

Réciproquement si  $f$  est Lebesgue-intégrable sur  $I$ , alors aussi  $|f|$ , donc pour  $[a, b] \subset I$ , on a

$$\int_a^b |f(x)| dx = \int_{[a,b]} |f| d\bar{\lambda} \leq \int_I |f| d\bar{\lambda} < +\infty$$

ce qui prouve la convergence absolue de l'intégrale  $\int_I f(x) dx$ .  $\square$

**Remarque 1.** Si  $f$  est continue et positive sur  $\mathbb{R}$ , alors avec  $I = \mathbb{R}$ , on retrouve la

proposition (1.7.1), dans le cas où l'intégrale  $\int_{-\infty}^{\infty} f(x) dx$  est finie. Mais la preuve de (1.7.1) ressemble étrangement à la preuve de (1.9.1)!

**Remarque 2.** On sait que l'intégrale  $\int_0^{\infty} \frac{\sin x}{x} dx$  est convergente (et de valeur  $\frac{\pi}{2}$ ), sans être absolument convergente. La fonction  $\frac{\sin x}{x}$  possède donc sur  $I = [0, +\infty)$  une intégrale

généralisée sans être Lebesgue-intégrable (bien qu'elle soit continue). Ce fait réhabilite en quelque sorte la théorie de l'intégrale généralisée, qui dans les cas de semi-convergence, peut fournir des résultats inaccessibles par la théorie de l'intégrale de Lebesgue.

## 1.10 Fonctions définies par des intégrales

On fixe une mesure  $\mu$  sur  $(\Omega, \Sigma)$ . En considérant un autre espace  $X$  et une fonction  $f : \Omega \times X \rightarrow \mathbb{R}$  ou  $\mathbb{C}$ , on peut introduire une nouvelle fonction

$$F(x) = \int f(\omega, x) d\mu(\omega) \quad x \in X$$

si l'on suppose que pour chaque  $x \in X$  la fonction partielle  $\omega \rightarrow f(\omega, x)$  est intégrable. La question naturelle qui se pose est de se demander dans quelles conditions les propriétés de régularité des fonctions  $x \rightarrow f(\omega, x)$  sur  $X$  vont se transmettre à la fonction  $F$ . Par exemple si  $X = T$  est un espace métrique, ou  $X = I$  un intervalle de  $\mathbb{R}$ , ou  $X = U$  un ouvert du plan complexe, peut-on garantir que  $F$  soit continue sur  $T$ , ou dérivable sur  $I$ , ou holomorphe sur  $U$  ?

Examinons ces questions une à une, la réponse étant toujours sous-tendue par le théorème de convergence dominée de Lebesgue.

### (1.10.1) *Théorème (de continuité)*

Soit  $T$  un espace métrique et  $f$  une fonction réelle ou complexe, définie sur  $\Omega \times T$ . On suppose que :

- Pour tout  $x \in T$  la fonction  $\omega \rightarrow f(\omega, x)$  est intégrable.
- Pour  $\mu$ -presque tout  $\omega \in \Omega$ , la fonction  $x \rightarrow f(\omega, x)$  est continue au point  $x_0 \in T$  (resp. est continue sur  $T$ ).
- Il existe une fonction  $g$  intégrable telle que, pour tout  $x \in T$ , on ait  $|f(\omega, x)| \leq g(\omega)$   $\mu$ pp sur  $\Omega$ . (condition CD de convergence dominée).

Alors la fonction  $F : x \rightarrow \int f(\omega, x) d\mu(\omega)$  est continue au point  $x_0$  (resp. sur  $T$ ).

*Preuve.* Fixons dans  $T$  la suite  $(x_n)$  telle que  $x_n \rightarrow x_0$ . Pour voir que  $F(x_n) \rightarrow F(x_0)$ , il suffit alors d'appliquer le théorème de Lebesgue (1.8.9) à la suite  $h_n : \omega \rightarrow f(\omega, x_n)$ , qui tend  $\mu$ pp vers la fonction  $h : \omega \rightarrow f(\omega, x_0)$ .  $\square$

### (1.10.2) *Théorème (de dérivabilité)*

Soit  $I$  un intervalle de  $\mathbb{R}$  et soit  $f$  une fonction, réelle ou complexe, définie sur  $\Omega \times I$ . On suppose que :

- Pour tout  $x \in I$  la fonction  $\omega \rightarrow f(\omega, x)$  est intégrable.

- b) Pour tout  $\omega \in \Omega$ , la fonction  $x \rightarrow f(\omega, x)$  est dérivable sur  $I$   
 c) Il existe une fonction  $g$  intégrable telle que l'on ait, pour tout  $x \in I$ , et tout  $\omega \in \Omega$ ,

$$\left| \frac{\partial}{\partial x} f(\omega, x) \right| \leq g(\omega)$$

(condition CD de convergence dominée)

Alors la fonction  $F : x \rightarrow \int f(\omega, x) d\mu(\omega)$  est continue et dérivable sur l'intervalle  $I$ . Pour tout  $x \in I$ , la fonction  $\omega \rightarrow \frac{\partial}{\partial x} f(\omega, x)$  est intégrable sur  $\Omega$ , et on a la formule de dérivation sous le signe d'intégration

$$F'(x) = \int \frac{\partial}{\partial x} f(\omega, x) d\mu(\omega)$$

*Preuve.* Supposons  $f$  réelle et fixons  $x \in I$  et une suite  $x_n \rightarrow x$ , avec  $x_n \in I$  et  $x_n \neq x$ . La formule des accroissements finis montre que la suite des fonctions

$$k_n : \omega \rightarrow \frac{f(\omega, x) - f(\omega, x_n)}{x - x_n}$$

est telle que  $|k_n| \leq g$  partout, et  $k_n \rightarrow \frac{\partial}{\partial x} f(\omega, x)$ .

Ici encore le théorème de Lebesgue (1.8.9) donne le résultat.  $\square$

**Remarque 1.** On constate qu'il est inutile de chercher d'abord à obtenir la continuité de  $F$ . En particulier l'hypothèse (1.10.1c) de convergence dominée est inutile, mais elle doit être remplacée par l'hypothèse CD relative à la dérivée. Si l'on veut donc démontrer que  $F$  est  $p$  fois dérivable, il suffira de prouver l'intégrabilité de chacune des fonctions  $\omega \rightarrow \frac{\partial^k}{\partial x^k} f(\omega, x)$  pour  $x \in I$  et  $0 \leq k \leq p - 1$  en s'assurant que *la condition CD*

*est satisfaite au niveau de la dernière dérivée*  $\frac{\partial^p f}{\partial x^p}$ .

**Remarque 2.** La dérivabilité étant une propriété locale, il suffit pour la validité de (1.10.2) que la condition de convergence dominée c) soit vérifiée localement, c'est-à-dire sur chaque sous-intervalle compact  $J = [a, b]$  de  $I$ , autrement dit qu'il existe, pour chaque tel  $J$ , une fonction intégrable  $g_J$  telle que  $\left| \frac{\partial}{\partial x} f(\omega, x) \right| \leq g_J(\omega)$  pour tout  $x \in J$  et tout  $\omega \in \Omega$ .

**Remarque 3.** On observera aussi la différence entre les énoncés b) et c) de (1.10.1) et (1.10.2). Dans le second cas, les égalités ou inégalités ont lieu partout et non pas seulement  $\mu$ pp. Cela tient à l'utilisation de la formule des accroissements finis

$$k_n(\omega) = \frac{\partial}{\partial x} f(\omega, y_n)$$

avec  $y_n \in ]x, x_n[$ , *dépendant a priori de  $\omega$* , sans que l'on sache comment.

**Remarque 4.** Choissant  $\Omega = \mathbb{N}$  avec pour  $\mu$  la mesure dénombrement, on obtient du même coup des théorèmes de continuité et de dérivabilité pour des fonctions définies par des séries.

**Remarque 5.** On mesurera enfin l'efficacité de ces théorèmes en les comparant, sur des exemples concrets, aux énoncés analogues relatifs à la convergence uniforme (et à la mesure de Borel sur  $\mathbb{R}$ ). En particulier le théorème de continuité ne nécessite ici aucune hypothèse *globale* de continuité de la fonction  $f(\omega, x)$  *par rapport au couple*  $(\omega, x)$ .

Le dernier résultat que nous donnerons est relatif à l'holomorphie et peut-être sauté dans une première lecture. Il convient de remarquer qu'il est d'une simplicité toute particulière, quoiqu'extrêmement puissant en pratique.

### (1.10.3) Théorème (d'holomorphie)

Soit  $U$  un ouvert et soit  $f$  une fonction complexe, définie sur  $\Omega \times U$ . On suppose que :

- Pour tout  $z \in U$  la fonction  $\omega \rightarrow f(\omega, z)$  est intégrable sur  $\Omega$ .
- Pour tout  $\omega \in \Omega$ , la fonction  $z \rightarrow f(\omega, z)$  est holomorphe sur  $U$ .
- Pour tout compact  $K$  de  $U$ , il existe une fonction intégrable  $g_K$  telle que, pour tout  $z \in K$ , on ait  $|f(\omega, z)| \leq g_K(\omega)$   $\mu$ -pp sur  $\Omega$  (condition CD de convergence dominée).

Alors la fonction  $F : z \rightarrow \int f(\omega, z) d\mu(\omega)$  est holomorphe sur  $U$ . De plus, pour tout entier  $k$  et tout  $z \in U$ , la fonction  $\omega \rightarrow \frac{\partial^k}{\partial z^k} f(\omega, z)$  est intégrable sur  $\Omega$  et

$$F^{(k)}(z) = \int \frac{\partial^k}{\partial z^k} f(\omega, z) d\mu(\omega)$$

**Preuve.** Soit  $K$  un compact de  $U$ ,  $F = \bigcup U$  et  $d = d(K, F) > 0$  la distance de  $K$  au fermé disjoint  $F$ . Posons  $\rho = \frac{d}{2} = \rho_K$ . Déjà le théorème (1.10.1) garantit la continuité de  $F$  sur  $K$ , donc en fait la continuité de  $F$  sur  $U$ .

Fixons  $x_0 \in K$  et le disque  $D(z_0, \rho)$  de frontière le cercle  $\Gamma = \{z; |z - z_0| = \rho\}$ . Alors  $D(z_0, \rho) \subset L$ , où  $L$  est le compact de  $U$  défini par  $L = L_K = \{z; d(z, K) \leq \rho\}$ . Par holomorphie de la fonction  $z \rightarrow f(\omega, z)$  on a les formules de Cauchy :

$$(1) \quad \begin{cases} f(\omega, z) = \frac{1}{2i\pi} \int_{\Gamma} \frac{f(\omega, \xi)}{\xi - z} d\xi \\ \frac{\partial f}{\partial z}(\omega, z) = \frac{1}{2i\pi} \int_{\Gamma} \frac{f(\omega, \xi)}{(\xi - z)^2} d\xi \end{cases} \quad \text{si } |z - z_0| < \rho$$

En paramétrant  $\Gamma$  selon  $\xi = z_0 + \rho e^{i\theta}$ , et en faisant intervenir les points  $\xi_n = z_0 + \rho e^{2i\pi \frac{k}{n}}$ , on approche les intégrales précédentes par les sommes de Riemann

$$(2) \quad \begin{cases} h_n(\omega) = \frac{\rho}{n} \sum_{k=1}^n \frac{f(\omega, \xi_k)}{\xi_k - z} e^{2i\pi \frac{k}{n}} \rightarrow f(\omega, z) \\ k_n(\omega) = \frac{\rho}{n} \sum_{k=1}^n \frac{f(\omega, \xi_k)}{(\xi_k - z)^2} e^{2i\pi \frac{k}{n}} \rightarrow \frac{\partial f}{\partial z}(\omega, z) \end{cases}$$

Or on a les conditions de domination

$$(3) \quad \begin{cases} |h_n(\omega)| \leq \frac{\rho \|f\|_L}{\rho - |z - z_0|} \leq \frac{\rho g_L(\omega)}{\rho - |z - z_0|} \\ |k_n(\omega)| \leq \frac{\rho \|f\|_L}{(\rho - |z - z_0|)^2} \leq \frac{\rho g_L(\omega)}{(\rho - |z - z_0|)^2} \end{cases}$$

ce qui donne aussi, avec  $z = z_0$ , en faisant varier  $z_0$  dans  $K$ , la condition

$$(4) \quad \left| \frac{\partial f}{\partial z}(\omega, z) \right| \leq \frac{1}{\rho_K} g_L(\omega) \quad \text{pour tout } z \in K$$

et bien entendu garanti, avec le théorème de Lebesgue, les égalités

$$(5) \quad \begin{cases} F(z) = \lim \frac{\rho}{n} \sum_{k=1}^n \frac{F(\xi_k)}{\xi_k - z} e^{2i\pi \frac{k}{n}} = \frac{1}{2i\pi} \int_{\Gamma} \frac{f(\xi)}{\xi - z} d\xi \\ \int \frac{\partial f}{\partial z}(\omega, z) d\mu(\omega) = \frac{1}{2i\pi} \int_{\Gamma} \frac{f(\xi)}{(\xi - z)^2} d\xi \end{cases}$$

grâce à la continuité de  $F$ , permettant de passer à la limite pour les sommes de Riemann introduites.

En supposant  $z_0 = 0$  pour simplifier, donc  $|z| < \rho$ , on voit que  $F$  est développable en série entière convergente pour  $|z| < \rho$ , ce qui assure que  $F$  est holomorphe sur  $U$ . De plus, par dérivation des séries entières, ou par application de la formule de Cauchy, on a

$$F'(z) = \frac{1}{2i\pi} \int_{\Gamma} \frac{f(\xi)}{(\xi - z)^2} d\xi = \int \frac{\partial f}{\partial z}(\omega, z) d\mu(\omega)$$

pour  $|z - z_0| < \rho$ , donc en particulier sur  $K$ , donc en fait sur  $U$ . Il n'y a plus alors qu'à repartir avec  $F'$  à la place de  $F$  et la fonction  $\frac{\partial f}{\partial z}$  à la place de  $f$ , une fois vu que la condition (4) donne exactement la condition de convergence dominée c) pour  $\frac{\partial f}{\partial z}$ .  $\square$

#### (1.10.4) Exercices

##### Ex 1.

On pose  $G(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ixt} \exp\left(-\frac{t^2}{2}\right) dt$  pour  $x$  réel.

a) Calculer  $G$  par dérivation en montrant que  $G'(x) = -x G(x)$ .

b) Déterminer la fonction  $H(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{xt} \exp\left(-\frac{t^2}{2}\right) dt$  pour  $x$  réel, puis démontrer que  $H$  est la restriction à l'axe réel d'une fonction  $H(z)$  entière, c'est-à-dire holomorphe dans tout le plan complexe. En déduire que  $H(z) = \exp\left(\frac{z^2}{2}\right)$  et retrouver ainsi la valeur de  $G(x)$ .

##### Ex 2.

On considère la fonction

$$K(x) = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \exp\left[-\frac{1}{2}\left(\frac{x^2}{t^2} + t^2\right)\right] dt \text{ pour tout } x \text{ réel.}$$

a) Prouver que  $K$  est continue et bornée.

b) Prouver que  $K$  est dérivable pour  $x \neq 0$ .

c) En faisant le changement de variable  $u = \frac{x}{t}$  dans l'expression intégrale de  $K'(x)$ , prouver que  $K'(x) = -K(x)$  pour  $x > 0$ .

d) En déduire l'égalité  $K(x) = e^{-|x|}$  pour tout  $x$ .

## Ex 3.

On considère les fonctions, définies pour tout  $x$  réel

$$F(x) = \int_0^{\infty} \frac{\cos tx}{1+t^2} dt \qquad G(x) = \int_0^{\infty} \frac{1 - \cos tx}{t^2} \frac{dt}{1+t^2}$$

- Établir que  $F$  et  $G$  sont continues et calculer  $F(0)$  et  $G(0)$ .
- Prouver que  $F(0) - F(x) + G(x) = C|x|$  avec  $C = \int_0^{\infty} \frac{\sin^2 t}{t^2} dt$ .
- Démontrer que  $G$  est deux fois continûment dérivable sur  $\mathbb{R}$  et telle que  $G'' = F$ . En déduire l'expression de  $F(x)$  pour  $x > 0$ , puis pour  $x$  quelconque. La fonction  $F$  est-elle dérivable au point  $x = 0$  ?
- Donner la valeur de la constante  $C$ .
- Prouver enfin les égalités suivantes pour  $a > 0$

$$\int_0^{\infty} \frac{\sin t}{t} \frac{dt}{a^2 + t^2} = \frac{\pi}{2a^2} (1 - e^{-a})$$

$$\int_0^{\infty} \frac{\sin^2 t}{t^2} \frac{dt}{a^2 + t^2} = \frac{\pi}{4a^3} (e^{-a} - 1 + 2a)$$

Quelle égalité obtient-on par multiplication par  $a^2$  en faisant tendre  $a$  vers  $+\infty$  ?

**Les fonctions  $B$  et  $\Gamma$  d'Euler.** Pour illustrer les théorèmes de continuité, de dérivabilité ou d'holomorphicité, revenons à la fonction  $\Gamma$  d'Euler, déjà introduite en (1.7.7., Ex. 7), et dont l'utilisation est essentielle pour le calcul de nombreuses intégrales en rapport avec la mesure de Laplace-Gauss et les mesures d'Euler. Il s'agit ici d'en esquisser une théorie faisant apparaître ses principales propriétés, et pour cela, recourons pour simplifier aux exercices. Rappelons tout d'abord que, pour une mesure  $\mu$  quelconque sur un espace  $(\Omega, \Sigma)$ , et pour des fonctions  $f, g$  mesurables et *positives*, on a toujours l'inégalité, dite de Cauchy-Schwarz

$$\left( \int fg \, d\mu \right)^2 \leq \int f^2 \, d\mu \cdot \int g^2 \, d\mu$$

que les intégrales soient finies ou non. Pour démontrer cette inégalité on se ramène au cas où les deux intégrales  $\int f^2 \, d\mu$  et  $\int g^2 \, d\mu$  sont finies et strictement positives, et on considère l'intégrale  $T(\lambda) = \int (f + \lambda g)^2 \, d\mu$  pour tout  $\lambda$  réel, ce qui définit un trinôme positif ou nul sur  $\mathbb{R}$ , dont le discriminant  $\Delta$  est négatif ou nul par conséquent, ce qui donne l'inégalité cherchée.

(1.10.5) *Exercice*

Soit  $\mu$  une mesure borélienne sur l'intervalle  $]0, +\infty)$ , telle que la fonction  $t^{\alpha-1}$  soit intégrable pour tout  $\alpha > 0$ . On suppose  $\mu$  non nulle.

- Montrer que les fonctions  $t^{\alpha-1} |\ln t|^p$  sont intégrables pour  $\mu$ , pour tout  $\alpha > 0$  et tout entier  $p \geq 0$ .
- On pose  $\Phi(x) = \int t^{x-1} d\mu(t)$  pour tout  $x > 0$ . Montrer que  $\Phi$  est une fonction strictement positive, indéfiniment dérivable sur  $]0, +\infty)$ . Expliciter  $\Phi'$  et  $\Phi''$ , puis démontrer l'inégalité  $\Phi'^2 \leq \Phi \Phi''$  comme conséquence de l'inégalité de Cauchy-Schwarz. En déduire que la fonction  $\ln \Phi$  est convexe, c'est-à-dire que la fonction  $\frac{\Phi'}{\Phi}$  est croissante. On dit alors que  $\Phi$  est *logarithmiquement convexe*.

(1.10.6) *Exercice*

On choisit pour  $\mu$  la probabilité définie par la densité  $e^{-t}$  par rapport à la mesure de Borel sur  $[0, +\infty)$ . On désigne par  $\Gamma$  la fonction  $\Phi$  associée, de sorte que

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt \quad x > 0$$

- Montrer que  $\Gamma(n+1) = n!$  pour tout entier  $n \geq 0$ , ce qui justifie le terme de "fonction factorielle" que l'on attribue quelquefois à la fonction  $\Gamma$  d'Euler.
- Montrer que  $\Gamma$  vérifie les trois conditions suivantes :
  - $\Gamma(1) = 1$
  - $\Gamma(x+1) = x \Gamma(x)$  pour tout  $x > 0$
  - $\Gamma$  est logarithmiquement convexe sur  $]0, +\infty)$ .

La propriété la plus remarquable de la fonction  $\Gamma$  est qu'elle est exactement caractérisée par l'ensemble de ces trois conditions. En effet :

(1.10.7) *Théorème (Artin)*

Soit  $\Phi$  une fonction définie sur  $]0, +\infty)$ , strictement positive et dérivable. On suppose que  $\Phi(x+1) = x \Phi(x)$  pour tout  $x > 0$  et que  $\Phi$  est logarithmiquement convexe. On a alors :

$$\Phi(x) = \Gamma(x) \Gamma(x)$$

pour tout  $x > 0$ .

*Preuve.* En posant  $H = \frac{\Phi}{\Gamma}$ , on introduit une fonction qui est manifestement 1-périodique et telle que  $H(1) = \Phi(1)$ . Fixons  $0 < x \leq 1$  et considérons le point  $n+x$  tel que

$n \leq n+x \leq n+1$ . Puisque les fonctions  $\frac{\Phi'}{\Phi}$  et  $\frac{\Gamma'}{\Gamma}$  sont croissantes et vérifient toutes deux l'égalité fonctionnelle  $f(x+1) = f(x) + \frac{1}{x}$ , obtenue par dérivation logarithmique à partir de la condition  $\beta$ ), on a avec  $\frac{H'}{H} = \frac{\Phi'}{\Phi} - \frac{\Gamma'}{\Gamma}$ .

$$\frac{\Phi'(n)}{\Phi(n)} - \frac{\Gamma'(n+1)}{\Gamma(n+1)} \leq \frac{H'(x+n)}{H(x+n)} \leq \frac{\Phi'(n+1)}{\Phi(n+1)} - \frac{\Gamma'(n)}{\Gamma(n)}$$

soit encore

$$\frac{H'(n)}{H(n)} - \frac{1}{n} \leq \frac{H'(x)}{H(x)} \leq \frac{H'(n)}{H(n)} + \frac{1}{n}$$

Or  $\frac{H'(n)}{H(n)} = \frac{H'(1)}{H(1)}$  par périodicité, donc par passage à la limite  $n \rightarrow +\infty$ , on obtient  $\frac{H'(x)}{H(x)} = \frac{H'(1)}{H(1)} = \alpha$  pour tout  $x \in ]0, 1[$ , donc aussi pour tout  $x > 0$ . Il suit de là que  $H(x) = C e^{\alpha x}$ , et la périodicité de  $H$  garantit que  $\alpha = 0$ , donc  $H(x) = C = H(1) = \Phi(1)$  et  $\Phi(x) = \Phi(1) \Gamma(x)$ .  $\square$

La fonction  $B$  (lire béta), dite fonction eulérienne de première espèce, est définie pour  $x > 0$  et  $y > 0$  par

$$B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt$$

On remarquera qu'en fixant  $y$  et en considérant la mesure  $\mu$  sur  $]0, 1[$  définie par la densité  $(1-t)^{y-1}$  on a, avec les notations de (1.10.5), l'égalité  $B(x, y) = \Phi_\mu(x)$ .

(1.10.8) *Exercice*

a) Établir, pour  $x > 0$  et  $y > 0$ , les égalités

$$\begin{cases} B(x, y) = B(y, x) \\ B(x, y+1) = B(x, y) - B(x+1, y) = \frac{y}{x} B(x+1, y) \\ B(x+1, y) = \frac{x}{x+y} B(x, y) \end{cases}$$

b) On fixe  $y > 0$ . Montrer que la fonction  $\Phi(x) = B(x, y) \Gamma(x+y)$  vérifie les conditions du théorème d'Artin. En déduire l'égalité fondamentale

$$B(x, y) = \frac{\Gamma(x) \Gamma(y)}{\Gamma(x+y)}$$

c) Retrouver, à partir de là, la valeur de l'intégrale de Gauss

$$I = \int_{-\infty}^{\infty} \exp\left(-\frac{t^2}{2}\right) dt \text{ par le calcul de } \Gamma\left(\frac{1}{2}\right).$$

(1.10.9) *Exercice*

En revenant à l'intégration des fonctions sphériques vue en (1.8.17), établir que le volume  $V_p$  de la boule euclidienne  $B_p$  de  $\mathbb{R}^p$  est donné, que  $p$  soit pair ou impair, par la formule

$$V_p = \frac{\pi^{p/2}}{\Gamma\left(\frac{p}{2} + 1\right)}$$

(1.10.10) *Exercice*

a) Montrer que l'intégrale  $J(\alpha, \beta) = \int_0^\infty \frac{t^{\alpha-1}}{(1+t^2)^\beta} dt$  est finie pour  $2\beta > \alpha > 0$ , et exprimer dans ce cas sa valeur à l'aide de la fonction  $\Gamma$ , en faisant les changements de variables  $u = t^2$  et  $v = \frac{1}{1+u}$ .

b) En revenant aux fonctions sphériques montrer alors l'égalité

$$\int_{\mathbb{R}^p} \frac{dx}{[1 + \|x\|^2]^{\frac{p+1}{2}}} = \frac{\pi^{\frac{p+1}{2}}}{\Gamma\left(\frac{p+1}{2}\right)}$$

où  $\|x\|$  est la norme euclidienne sur  $\mathbb{R}^p$ .

**Les formules de Weierstrass et de Gauss.** On désigne ici par  $\gamma$  la constante d'Euler  $\gamma = \lim_{n \rightarrow \infty} \left[ 1 + \frac{1}{2} + \dots + \frac{1}{n} - \ln n \right] = 0,577\dots$  (on rappelle qu'on ne connaît pas la nature arithmétique de  $\gamma$ , en particulier on ignore si  $\gamma$  est un nombre rationnel ou non) et l'on introduit la fonction  $H$ , définie pour  $x > 0$  par

$$H(x) = -\frac{1}{x} + \sum_{n=1}^{\infty} \left( \frac{1}{n} - \frac{1}{x+n} \right)$$

(1.10.11) *Exercice*

a) Montrer que  $H$  est indéfiniment dérivable et calculer  $H'(x)$  et  $H^{(k)}(x)$  sous forme de séries.

Montrer que  $H(1) = 0$  et que  $H(x+1) = H(x) + \frac{1}{x}$ .

En déduire alors que la fonction  $M(x) = \int_1^x H(t) dt$  est telle que

$M(x+1) = M(x) + \ln x + \gamma$  et que la fonction  $\Phi(x) = \exp[-\gamma x + M(x)]$  vérifie les conditions du théorème d'Artin.

b) Établir la formule de Weierstrass

$$\frac{1}{\Gamma(x)} = x e^{\gamma x} \prod_{n=1}^{\infty} \left(1 + \frac{x}{n}\right) e^{-\frac{x}{n}} \quad x > 0$$

c) En déduire que  $\Gamma(x) = -\gamma = \int_0^{\infty} e^{-t} \ln t \, dt$  et expliciter la relation existant entre  $H$  et  $\frac{\Gamma}{\Gamma}$ .

d) Établir enfin la formule de Gauss

$$\Gamma(x) = \lim_{n \rightarrow \infty} \frac{n^x n!}{x(x+1) \dots (x+n)} \quad x > 0$$

### La formule des compléments

#### (1.10.12) Exercice

a) Montrer que l'intégrale  $I(\alpha) = \int_0^{\infty} \frac{t^{\alpha-1}}{1+t} \, dt$  est finie pour  $0 < \alpha < 1$ , et prouver que  $I(\alpha) = \frac{\pi}{\sin \pi \alpha}$  par la méthode des résidus. On peut aussi obtenir la valeur de  $I(\alpha)$  en prouvant la continuité de la fonction  $I$  sur  $]0, 1[$  et en calculant  $I(\alpha)$  pour  $\alpha = \frac{p}{q}$  rationnel, par le changement de variable  $t = u^q$ .

b) En déduire la formule des compléments

$$\Gamma(x) \Gamma(1-x) = \frac{\pi}{\sin \pi x} \quad 0 < x < 1$$

Qu'obtient-on pour  $x = \frac{1}{2}$  ?

c) En déduire encore, avec la formule de Weierstrass, les développements eulériens de  $\sin x$  et  $\cotg x$ .

$$\sin x = x \prod_{n=1}^{\infty} \left(1 - \frac{x^2}{n^2 \pi^2}\right) \quad x \text{ réel}$$

$$\cotg x = \frac{1}{x} + \sum_{n=1}^{\infty} \frac{2x}{x^2 - n^2 \pi^2} \quad x \text{ réel, } x \notin \pi \mathbb{Z}$$

Comment retrouver, à partir de là, l'égalité classique

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} \quad ?$$

## La formule de Legendre-Gauss

(1.10.13) *Exercice*

- a) Utiliser encore une fois le théorème d'Artin pour démontrer la formule de duplication de Legendre

$$2^{x-1} \Gamma\left(\frac{x}{2}\right) \Gamma\left(\frac{x+1}{2}\right) = \sqrt{\pi} \Gamma(x) \quad x > 0$$

- b) En déduire l'égalité  $\int_0^1 \ln \Gamma(t) dt = \ln \sqrt{2\pi}$ .

- c) Plus généralement, montrer que, pour tout entier  $p \geq 2$ , la fonction  $p^x \Gamma\left(\frac{x}{2}\right) \Gamma\left(\frac{x+1}{2}\right) \dots \Gamma\left(\frac{x+p-1}{2}\right)$  est proportionnelle à  $\Gamma(x)$ . Déterminer la constante de proportionnalité avec b), et obtenir la formule de Legendre-Gauss

$$p^x \Gamma\left(\frac{x}{2}\right) \Gamma\left(\frac{x+1}{2}\right) \dots \Gamma\left(\frac{x+p-1}{2}\right) = (2\pi)^{\frac{p-1}{2}} \Gamma(x)$$

La formule de Stirling. On désigne par  $\varphi$  la fonction (discontinue) de période 1, égale

à  $t - \frac{1}{2}$  pour  $0 \leq t < 1$ . On commencera par montrer que l'intégrale  $\int_0^\infty \frac{\varphi(t)}{x+t} dt$  existe, pour tout  $x > 0$ , en tant qu'intégrale impropre de Riemann semi-convergente et que

$$\lim_{x \rightarrow \infty} \int_0^\infty \frac{\varphi(t)}{x+t} dt = 0.$$

(1.10.14) *Exercice*

- a) Utiliser les résultats de (1.10.11) pour obtenir l'égalité

$$\frac{\Gamma'(x)}{\Gamma(x)} = \ln x - \frac{1}{2x} + \int_0^\infty \frac{\varphi(t)}{(x+t)^2} dt$$

- b) En déduire, avec (1.10.13.b) l'égalité

$$\ln \Gamma(x) = \left(x - \frac{1}{2}\right) \ln x - x + \ln \sqrt{2\pi} - \int_0^\infty \frac{\varphi(t)}{x+t} dt$$

c) Obtenir alors les formules de Stirling

$$\Gamma(x) \sim x^{x-\frac{1}{2}} e^{-x} \sqrt{2\pi} \quad x \rightarrow +\infty$$

$$n! \sim n^n e^{-n} \sqrt{2\pi n} \quad n \rightarrow +\infty$$

et en déduire que, pour tout  $a > 0$ , on a  $\Gamma(x+a) \sim x^a \Gamma(x)$  quand  $x \rightarrow +\infty$  (ce qui offre une petite généralisation de la formule  $\Gamma(x+1) = x \Gamma(x)$ ).

d) Montrer avec a) l'inégalité  $\frac{\Gamma'(x)}{\Gamma(x)} < \ln x$ . En déduire que la fonction

$\left(\frac{e}{x}\right)^x \Gamma(x)$  est décroissante pour  $x > 0$ , puis démontrer que cette propriété,

jointe aux deux conditions  $\Gamma(1) = 1$  et  $\Gamma(x+1) = x \Gamma(x)$ , caractérise complètement la fonction  $\Gamma$ .

**La fonction  $\Gamma$  dans le champ complexe.** Il est facile de vérifier que le produit infini (1.10.11.b) de la formule de Weierstrass est convergent pour tout  $x = z$  complexe, et qu'il définit donc une fonction entière (c'est-à-dire holomorphe dans tout le plan complexe) notée encore

$$\frac{1}{\Gamma(z)} = z e^{z\gamma} \prod_{n=1}^{\infty} \left(1 + \frac{z}{n}\right) e^{-\frac{z}{n}}$$

qui admet pour zéros simples les points  $z = 0, -1, \dots, -n, \dots$ . La fonction  $\Gamma(z)$  ainsi définie, qui prolonge analytiquement sur  $\mathbb{C}$  la fonction  $\Gamma(x)$  définie pour  $x > 0$ , est donc *méromorphe* et admet pour pôles simples les points  $z = 0, -1, \dots, -n, \dots$ . La théorie du prolongement analytique permet alors d'affirmer que  $\Gamma$  vérifie les formules

$$\begin{cases} \Gamma(z+1) = z \Gamma(z) & \text{si } z \neq 0, -1, \dots, -n, \dots \\ \Gamma(z) \Gamma(1-z) = \frac{\pi}{\sin \pi z} & \text{si } z \notin \mathbb{Z} \end{cases}$$

(1.10.15) *Exercice*

a) Étudier l'intégrale  $\Phi(z) = \int_0^{\infty} t^{z-1} e^{-t} dt$  pour  $z$  complexe. Montrer qu'elle

est définie pour  $x = \Re z > 0$  et que, dans ce demi-plan, on a  $\Phi(z) = \Gamma(z)$ .

b) Établir la relation

$$\Gamma(z) = e^{\frac{iz}{2}} \int_0^{\infty} u^{z-1} e^{-iu} du$$

valable pour  $0 < \Re z < 1$ , en intégrant le long d'un contour convenablement choisi. En déduire les formules

$$\int_0^{\infty} t^{z-1} \cos t \, dt = \Gamma(z) \cos \frac{\pi z}{2} \quad 0 < \Re z < 1$$

$$\int_0^{\infty} t^{z-1} \sin t \, dt = \Gamma(z) \sin \frac{\pi z}{2} \quad 0 < \Re z < 1$$

c) Soit  $I = \int_0^{\infty} \cos(t^2) \, dt$  et  $J = \int_0^{\infty} \sin(t^2) \, dt$  les intégrales de Fresnel.

Montrer que

$$I = J = \sqrt{\frac{\pi}{8}}$$

## 1.11 Le théorème de Riesz-Alexandroff et les mesures de Radon

Rappelons qu'un espace métrique  $T$  est dit localement compact lorsque tout point  $x \in T$  admet un voisinage compact (et il admet alors une base de voisinages compacts). On dit que  $T$  est dénombrable à l'infini lorsqu'il est réunion dénombrable de parties compactes. Par exemple l'espace  $\mathbb{R}^p$  est évidemment de ce type, de même que tout espace compact.

Sur un espace métrique  $T$ , localement compact et dénombrable à l'infini, toute mesure  $\mu$  borélienne, c'est-à-dire définie sur la tribu borélienne de  $T$  (engendrée par les ouverts, ou les fermés, ou les compacts), qui est finie sur les compacts (et on ne considère que celles-ci) est donc  $\sigma$ -finie. Si l'on désigne par  $\mathcal{K}(T)$  l'espace vectoriel des fonctions continues réelles, qui sont à support compact (c'est-à-dire nulles en dehors d'un compact), on voit que toute  $f \in \mathcal{K}(T)$  est intégrable pour toute mesure borélienne  $\mu$ .

On peut donc considérer la fonctionnelle  $L = L_{\mu}$ , définie sur  $\mathcal{K}(T)$  par  $L(f) = \int f \, d\mu$ , qui est toujours à valeurs réelles finies, et qui possède en outre les propriétés évidentes :

- elle est linéaire sur  $\mathcal{K}(T)$
- elle est positive sur  $\mathcal{K}(T)$  :  $f \geq 0 \Rightarrow L(f) \geq 0$

Or il se trouve que  $L$ , qui est déterminée par  $\mu$ , détermine à son tour la mesure  $\mu$ . Mais cette détermination ne peut pas s'exprimer de façon directe par la formule  $\mu B = L(1_B)$ , puisque la fonction indicatrice  $1_B$  n'est élément de  $\mathcal{K}(T)$  que si  $B$  est à la fois ouvert et compact. Il faut donc ruser et opérer de façon indirecte, en décrivant ce qu'on appelle les propriétés de *régularité* de la mesure  $\mu$ .

### (1.11.1) Proposition (Régularité des mesures boréliennes)

a) Pour tout compact  $K$  de  $T$  on a

$$\mu K = \text{Inf} \left\{ \int f \, d\mu ; f \in \mathcal{K}(T), 1_K \leq f \right\}$$

$$\left| \begin{array}{l} \text{b) Pour tout borélien } B \text{ de } T \text{ on a} \\ \mu B = \text{Sup} \{ \mu K, K \text{ compact}, K \subset B \} \end{array} \right.$$

**Preuve.** a) Dans  $T$  (espace métrique) tout fermé est un  $G_\delta$ , c'est-à-dire une intersection dénombrable d'ouverts. Il suit de là facilement qu'il existe une suite  $(U_n)$  d'ouverts relativement compacts (c'est-à-dire tels que  $\overline{U_n}$  soit compact) tels que  $U_n \downarrow K$ , de sorte que  $\mu U_n \downarrow \mu K$ . Avec le théorème d'Urysohn (évident pour les espaces métriques) on peut construire, pour tout  $n$ , une fonction continue  $f_n$  telle que  $0 \leq f_n \leq 1$ ,  $f_n = 1$  sur  $K$  et  $f_n = 0$  sur  $U_n^c$ . On a alors  $1_K \leq f_n \leq 1_{U_n}$ ,  $\int f_n d\mu \rightarrow \mu K$  et  $f_n \in \mathcal{K}(T)$ , d'où a).

b) Supposons tout d'abord  $\mu$  finie, c'est-à-dire  $\mu T < +\infty$ , et posons pour tout borélien  $B$ ,  $\tilde{\mu} B = \text{Sup} \{ \mu K, K \text{ compact}, K \subset B \} \leq \mu B$

Introduisons maintenant la classe  $\mathcal{C}$  des boréliens  $B$  tels que  $\tilde{\mu} B = \mu B$  et  $\tilde{\mu} B^c = \mu B^c$ , évidemment stable par passage au complémentaire. Montrons que  $\mathcal{C}$  est stable par réunion dénombrable. Soit  $B_n \in \mathcal{C}$  et  $B = \bigcup_{n \geq 1} B_n$ . Pour chaque  $n$ , il existe un compact  $K_n \subset B_n$  et un compact  $L_n \subset B_n^c$  tels que

$$\mu(B_n \setminus K_n) \leq \varepsilon 2^{-(n+1)}$$

$$\mu(B_n^c \setminus L_n) \leq \varepsilon 2^{-n}$$

Posons  $L = \bigcap L_n$  et  $H = \bigcup K_n$ . Alors les inclusions  $B \setminus H \subset \bigcup (B_n \setminus K_n)$  et  $B^c \setminus L \subset \bigcup (B_n^c \setminus L_n)$  donnent les conditions

$$\mu(B \setminus H) \leq \frac{\varepsilon}{2} \quad \text{et} \quad \mu(B^c \setminus L) \leq \varepsilon$$

Or  $L$  est compact, mais  $H$  ne l'est pas. Soit donc  $K'_N = \bigcup_{n \leq N} K_n$ , tel que  $K'_N \uparrow H$  et  $\mu(H \setminus K'_N) \downarrow 0$ . En choisissant  $N$  tel que  $\mu(H \setminus K'_N) \leq \frac{\varepsilon}{2}$  on obtient bien  $\mu(B \setminus K'_N) \leq \varepsilon$ , et ainsi  $B$  appartient à la classe  $\mathcal{C}$ .

Alors  $\mathcal{C}$  est donc une tribu. Et cette tribu contient les ouverts, car dans  $T$  tout fermé est un  $K_\sigma$  (réunion dénombrable de compacts), et tout ouvert un  $F_\sigma$ . Donc tout ouvert est aussi un  $K_\sigma$ , de même que tout fermé. Il en résulte que pour  $U$  ouvert, on a nécessairement  $\tilde{\mu} U = \mu U$  et  $\tilde{\mu} U^c = \mu U^c$ , donc  $U \in \mathcal{C}$ . Ainsi  $\mathcal{C}$  est exactement la tribu borélienne de  $T$ , ce qui prouve b).

Si maintenant  $\mu T = +\infty$ , fixons une suite de compacts  $H_p$  tels que  $\mu H_p < +\infty$  et  $H_p \uparrow T$ , et posons  $\mu_p(B) = \mu(B \cap H_p)$ . Alors  $\mu_p$  est une mesure finie sur  $T$  et  $\mu B = \text{Sup}_p \mu_p(B)$ . D'où, pour tout borélien  $B$

$$\mu B = \text{Sup}_p \mu_p(B) = \text{Sup}_p \text{Sup}_{K \subset B} \mu_p(K) = \text{Sup}_K \text{Sup}_p \mu_p(K) = \text{Sup}_K \mu K . \quad \square$$

Ainsi l'application  $\mu \rightarrow L_\mu$  est injective. Le théorème de Riesz-Alexandroff, que nous énonçons ici sans en donner la preuve, longue et délicate, assure que l'application  $\mu \rightarrow L_\mu$  est aussi surjective.

(1.11.2) **Théorème (Riesz-Alexandroff)**

Soit  $T$  un espace métrique localement compact et dénombrable à l'infini. Pour toute forme linéaire positive  $L$  sur l'espace vectoriel  $\mathcal{K}(T)$ , il existe une mesure borélienne unique  $\mu$  telle que  $L(f) = \int f d\mu$  pour toute  $f \in \mathcal{K}(T)$ .

Une conséquence intéressante concerne l'espace de Banach  $C_0(T)$  des fonctions réelles continues sur  $T$ , qui tendent vers zéro à l'infini, c'est-à-dire qui sont telles que, pour tout  $\varepsilon > 0$ , il existe un compact  $K$  tel que  $|f| \leq \varepsilon$  sur  $K^c$ , muni de sa norme uniforme. Il est immédiat que  $\mathcal{K}(T)$  est un sous-espace vectoriel *dense* de  $C_0(T)$ , ce qu'on obtient avec Urysohn en construisant une suite  $\varphi_n \in \mathcal{K}(T)$  telle que  $0 \leq \varphi_n \leq 1$  et  $\varphi_n \uparrow 1$ , donc telle aussi que  $\varphi_n \uparrow 1$  uniformément sur tout compact  $K$  de  $T$  (avec Dini), ce qui implique  $\varphi_n f \rightarrow f$  uniformément sur  $T$ . Alors

(1.11.3) **Corollaire**

Soit  $T$  un espace métrique localement compact et dénombrable à l'infini.

a) Toute mesure borélienne  $\mu$ , finie sur  $T$ , détermine une forme linéaire positive  $L_\mu$  sur l'espace  $C_0(T)$  selon

$$L_\mu(f) = \int f d\mu$$

b) Réciproquement toute forme linéaire positive  $L$  sur l'espace  $C_0(T)$  est définie par une unique mesure borélienne finie  $\mu$  selon  $L(f) = \int f d\mu$ .

**Preuve.** Il suffit de prouver b). Pour cela montrons d'abord que  $L$  est continue pour la norme uniforme de  $C_0(T)$ , en introduisant l'ensemble

$$\Delta = \{ f \in C_0(T), 0 \leq f \leq 1 \}.$$

En effet, il suffit de montrer que  $L$  est majorée par une constante (finie)  $M$  sur  $\Delta$ , car alors pour  $\|f\| \leq 1$  on aura, avec l'encadrement  $-|f| \leq f \leq |f|$  et la positivité de  $L$ , l'inégalité

$$|L(f)| \leq L(|f|) \leq M$$

Or, si l'on avait  $\sup_{f \in \Delta} L(f) = +\infty$ , il existerait une suite  $f_n \in \Delta$  telle que  $L(f_n) \geq 2^n$  et la

fonction  $f = \sum_1^N 2^{-n} f_n$  serait un élément de  $C_0(T)$  pour lequel on aurait  $f \geq \sum_1^N 2^{-n} f_n$ , donc  $L(f) \geq N$  pour tout  $N$ , donc  $L(f) = +\infty$  ce qui est absurde.

Maintenant, la restriction de  $L$  à l'espace  $\mathcal{K}(T)$  est, par le théorème de Riesz-Alexandroff, définie par une mesure borélienne  $\mu$ , de sorte qu'on a  $0 \leq \int f \, d\mu = L(f) \leq M$  pour  $f \in \Delta_{\mathcal{K}} = \Delta \cap \mathcal{K}(T)$ . Avec une suite  $\varphi_n \in \mathcal{K}(T)$  telle que  $\varphi_n \uparrow 1$ , on en déduit que  $\int d\mu \leq M$ , et  $\mu$  est une mesure finie sur  $T$ . Alors les deux formes linéaires  $L$  et  $L_{\mu}$ , qui sont continues sur  $C_0(T)$  et qui coïncident sur le sous-espace vectoriel dense  $\mathcal{K}(T)$ , sont évidemment identiques.  $\square$

(1.11.4) **Remarque historique.** Le premier résultat dans cette voie a été obtenu par F. Riesz dès 1909. Il décrivait toute forme linéaire positive sur l'espace  $C[0, 1]$  comme une intégrale (de Stieltjes) associée à une mesure  $\mu = \mu_f$  sur l'intervalle compact  $[0, 1]$ . Et ceci, bien entendu, en profitant des travaux de Stieltjes qui, vers 1890-91, avait généralisé l'intégrale de Riemann. Ensuite les extensions successives du théorème de F. Riesz ont permis de se dégager peu à peu de la notion restrictive d'intégrale de Stieltjes pour aboutir à la notion générale de mesure. C'est ainsi que Fréchet, en 1910, passait au cas des intégrales doubles de Stieltjes sur le carré  $[0, 1]^2$ , puis Radon, en 1913, au cas, déjà très élaboré, d'un compact quelconque de  $\mathbb{R}^p$ . Après de multiples travaux, diverses formes de l'énoncé général ont été obtenues par Markoff en 1938 (on appelle aussi (1.11.2) le théorème de Riesz-Markoff) puis par Alexandroff en 1940.

La question historique à ce sujet n'est pas sans importance. Car le théorème montre que la théorie des mesures peut se construire de deux points de vue, a priori éloignés l'un de l'autre :

- le point de vue *ensembliste*, que nous avons choisi d'exposer, qui privilégie les parties d'un ensemble et axe tout son développement sur la notion de tribu;
- le point de vue *fonctionnel*, où partant d'un espace localement compact quelconque  $T$  (non nécessairement métrique) on appelle *mesure de Radon* toute forme linéaire positive sur l'espace  $\mathcal{K}(T)$ .

En fait, les deux points de vue ont chacun leurs avantages et leurs inconvénients. Le théorème (1.11.2) montre qu'il conduisent à la même théorie, au moins pour les espaces (métriques) localement compacts dénombrables à l'infini. Mais, bien que le choix fonctionnel mette l'accent sur la régularité intérieure dans le cas général (toute mesure de Radon sur  $T$  localement compact quelconque détermine une mesure borélienne intérieurement régulière) il apparaît comme trop difficile à exposer de façon élémentaire, et aussi comme trop restrictif pour l'utilisation en calcul des probabilités. En effet :

- il oblige à n'utiliser que des espaces topologiques (métriques) localement compacts, nécessitant des préliminaires topologiques (compacité, Urysohn) pas forcément élémentaires;

- il rend difficilement compte de la théorie des mesures boréliennes (même régulières) dans le cas topologique (même métrique) non localement compact. Or les espaces métriques non localement compacts interviennent fréquemment en probabilités, par exemple  $T = \mathbb{R}^{\mathbb{N}}$  ou  $T = H$ , espace de Hilbert;
- enfin, même pour le cas d'un espace localement compact, il oblige pratiquement à ne s'intéresser qu'à la seule tribu borélienne (encore qu'une certaine tribu de Baire interviennent aussi). Ceci est bien entendu une exigence absurde pour un probabiliste, qui pour parler d'indépendance, de conditionnement ou de martingales, doit nécessairement faire appel à de multiples sous-tribus.

## 1.12 Applications au calcul des probabilités

A la différence de la théorie de la mesure (ou celle de l'intégration) qui constitue un outil de calcul, la théorie des probabilités peut être considérée comme une théorie "physique", dont le but est la description la plus précise possible (au sens quantitatif) des "lois du hasard". Celles-ci existent en effet, et sont suffisamment explicites pour avoir été, historiquement, découvertes de façon expérimentale. Mais elles apparaissent très souvent comme contraires à l'intuition première que l'on se fait d'un phénomène, d'où la nécessité, pour la clarté de la situation, d'une axiomatisation (ou d'une modélisation) très poussée.

Considérons à titre d'illustration le jeu de pile ou face, où deux joueurs A et B jettent une pièce de monnaie à tour de rôle, A donnant 1 franc à B si le côté pile sort, ou recevant 1 franc de B si le côté face sort. Au bout de  $n$  épreuves le gain de A est noté  $G_n$ ; c'est un nombre entier élément de  $\mathbb{Z}$ , et une question intéressante est d'en donner une estimation raisonnable. On peut alors démontrer que la moyenne  $\frac{G_n}{n}$  des gains à chaque jet de pièce tend vers zéro quand  $n \rightarrow +\infty$ , ce qui paraît normal et intuitif si le jeu est équilibré, c'est-à-dire si la pièce n'est pas truquée. Mais ce qui est plus paradoxal, et quelque peu rebelle à l'intuition, est qu'on peut montrer aussi que pour tout entier  $N \geq 1$  fixé, le gain  $G_n$  peut prendre, pour  $n$  assez grand, des valeurs supérieures à  $N$  et d'autres inférieures à  $-N$ . Autrement dit, la fonction  $G(n) = G_n$  subit des oscillations aussi grandes que l'on veut. On peut encore montrer qu'une "courbe de sécurité" est de la forme  $y = \lambda [x \ln(\ln x)]^{1/2}$ , c'est-à-dire que l'on a, presque toujours en un sens à préciser,  $|G_n| \leq \lambda [n \ln(\ln n)]^{1/2}$ .

Introduisons maintenant le formalisme nécessaire au développement de la théorie. Une expérience aléatoire est, par définition, une expérience dont le résultat est indéterminé, et dépend du hasard, à l'intérieur d'une famille de résultats possibles. Par exemple l'observation de la durée de vie d'un individu dans une population, ou celle d'une particule atomique, du nombre d'appels passant chaque journée (ou chaque heure) dans un standard téléphonique, du résultat du jet d'un ou plusieurs dés, sont des expériences aléatoires.

L'expérience se décrit donc par la donnée de tous ses résultats possibles, ce qui fait apparaître un ensemble  $\Omega$ , rassemblant tous ces résultats. Pour des appels téléphoniques on aura  $\Omega = \mathbb{N}$ , pour des durées de vie  $\Omega = \mathbb{R}_+$ , ou  $\Omega = \mathbb{N}$  si on choisit une unité de temps minimum (siècle, année, seconde, nanoseconde). Pour un jet de deux dés, on aura pour  $\Omega$  l'ensemble des couples  $(m, n)$  d'entiers tels que  $1 \leq m, n \leq 6$ .

La complexité de l'expérience détermine donc partiellement celle de  $\Omega$ , en notant bien que l'ensemble  $\Omega$  à lui seul ne rend pas compte de la manière dont l'expérience est faite, mais seulement de son résultat.

Maintenant si l'on répète  $n$  fois une expérience aléatoire associée à  $\Omega$ , l'espace décrivant ces nouveaux résultats sera  $\Omega^n$ . Si l'on répète une infinité de fois, ce sera  $\Omega^{\mathbb{N}}$ , de sorte qu'on voit naturellement apparaître les espaces produits, finis ou infinis, et plus tard apparaîtront aussi les probabilités produits.

La notion importante d'événements est celle des parties de  $\Omega$  que l'on désire considérer. Par exemple si l'on jette deux dés, le fait que le total de leurs points est inférieur ou égal à 10 est l'événement aléatoire  $\{(m, n) ; m + n \leq 10\}$ . Pour des raisons historiques, le calcul des probabilités datant du 17ème siècle alors que la théorie des ensembles date de la fin du 19ème siècle, la terminologie utilisée en calcul des probabilités diffère de celle utilisée en théorie des ensembles. Le tableau de correspondance suivant donne quelques exemples essentiels :

Terminologie ensembliste	Terminologie probabiliste
référentiel	espace d'épreuves
partie	événement
partie complémentaire	événement contraire
partie vide, partie pleine	événement impossible, événement certain
inclusion	implication
intersection	et
parties disjointes	événements incompatibles
union	ou (inclusif)
différence symétrique	ou (exclusif)
presque partout	presque sûrement

Ainsi une expérience aléatoire se décrit mathématiquement par la donnée d'un ensemble  $\Omega$  et d'une classe  $\Sigma$  de parties de  $\Omega$ , qui rassemble les événements que l'on désire étudier ou considérer. Comme on exige rapidement que  $\Sigma$  soit stable par passage au complémentaire, par réunion et intersection finies ou dénombrables, on voit que  $\Sigma$  doit être une tribu sur  $\Omega$ .

La notion de probabilité sur  $\Sigma$  s'est dégagée à partir de la notion de fréquence statistique. Fixons un événement  $A \in \Sigma$  et répétons  $N$  fois l'expérience en notant  $N(A)$  le nombre de fois où  $A$  s'est réalisé. On a évidemment  $0 \leq N(A) \leq N$  et l'on constate expérimentalement (au moins sur des cas simples) que le quotient  $\frac{N(A)}{N}$  admet une

limite, notée  $P(A)$ , lorsque  $N \rightarrow +\infty$ . C'est la loi empirique des grands nombres, et on a évidemment  $P(A) \in [0, 1]$ ,  $P(\Omega) = 1$  et  $P(\emptyset) = 0$

La fonctionnelle ensembliste  $P : \Sigma \rightarrow [0, 1]$  ainsi construite est manifestement additive sur la tribu  $\Sigma$ , car  $N(A \cup B) = N(A) + N(B)$  si  $A$  et  $B$  sont des événements incompatibles. Mais pour les besoins de la théorie on exige qu'elle soit dénombrablement additive, ce qui en fait une probabilité sur  $\Sigma$ , c'est-à-dire une mesure finie de masse totale 1. Cette exigence supplémentaire n'a apporté aucune contradiction expérimentale dans le maniement de la loi empirique des grands nombres, mais a, au contraire, permis la prévision de résultats vérifiés par l'expérience.

En résumé une expérience aléatoire se décrit fondamentalement par la donnée d'un espace probabilisé  $(\Omega, \Sigma, P)$ , de sorte qu'à première vue la théorie probabiliste apparaît comme une illustration de la théorie générale des mesures, mais ceci n'est que partiellement vrai. Car il existe en effet une notion très particulière, que l'on appelle l'indépendance, traduisant mathématiquement le fait que les résultats associés à un événement  $A$  n'influent pas sur ceux associés à un autre événement  $B$ , indépendance qui donne toute son autonomie à la théorie probabiliste. Abordons donc la question.

**Indépendance d'événements, indépendance de sous-tribus.** L'espace probabilisé  $(\Omega, \Sigma, P)$  étant fixé, on dit que deux événements  $A$  et  $B$  sont *indépendants* lorsque l'on a

$$P(A \cap B) = P(A)P(B)$$

Il est alors immédiat que  $A$  et  $B^c$  sont indépendants, de même que  $A^c$  et  $B$ , ou  $A^c$  et  $B^c$ . Lorsque l'on considère trois événements  $A, B, C$  on constate, sur des exemples, que l'indépendance deux à deux n'implique pas la relation  $P(A \cap B \cap C) = P(A)P(B)P(C)$ , et réciproquement que cette relation n'implique pas l'indépendance deux à deux. On est donc amené à définir l'indépendance (globale) de  $A, B, C$  en imposant la formule de multiplicativité pour 3 et aussi pour 2. Alors en généralisant on a :

(1.12.1) *Définition*

On dit qu'une famille non vide  $(A_i)_{i \in I}$  d'événements est indépendante (globalement) lorsque pour toute partie finie  $J \subset I$ , non vide, on a

$$P\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} P(A_i)$$

(1.12.2) *Définition*

On dit qu'une famille non vide  $(\mathcal{G}_i)_{i \in I}$  de classes sur  $\Omega$ , contenues dans la tribu  $\Sigma$ , est indépendante (globalement) lorsque, pour toute choix des  $A_i \in \mathcal{G}_i$ , la famille d'événements  $(A_i)_{i \in I}$  est indépendante.

Lorsque les classes  $\mathcal{G}_i$  sont des sous-tribus  $\Sigma_i$  de  $\Sigma$ , on obtient la notion importante d'indépendance de sous-tribus. Et si les tribus  $\Sigma_i$  sont respectivement engendrées par des

classes  $\mathcal{C}_i$ , on conçoit qu'il est plus facile de tester l'indépendance de la famille  $(\mathcal{C}_i)$  que celle de la famille  $(\Sigma_i)$ . Mais alors dans quels cas a-t-on le même résultat ?

(1.12.3) **Théorème (Critère d'indépendance)**

Soit  $(\Sigma_i)_{i \in I}$  une famille non vide de sous-tribus de  $\Sigma$ , chaque  $\Sigma_i$  étant engendrée par une classe  $\mathcal{C}_i$ . On suppose que les  $\mathcal{C}_i$  sont des  $\pi$ -classes, c'est-à-dire qu'elles sont stables par intersection. Alors pour que la famille  $(\Sigma_i)_{i \in I}$  soit indépendante, il suffit que la famille  $(\mathcal{C}_i)_{i \in I}$  le soit.

**Preuve.** On se ramène à supposer  $I$  fini, soit  $I = \{1, 2, \dots, p\}$ , et à prouver que  $P(B_1 \cap \dots \cap B_p) = P(B_1) \dots P(B_p)$  pour tout choix des  $B_i \in \Sigma_i$ . Or considérons dans  $\Sigma_1$  la classe  $\mathcal{C}_1$  des parties  $B_1$  telles que

$$P(B_1 \cap A_2 \cap \dots \cap A_p) = P(B_1) P(A_2) \dots P(A_p)$$

pour tout choix des  $A_k \in \mathcal{C}_k$ ,  $k \geq 2$ . Il est immédiat que  $\mathcal{C}_1$  est stable par différence propre et réunion dénombrable croissante; de plus  $\Omega \in \mathcal{C}_1$  et  $\mathcal{C}_1 \subset \Sigma_1$  puisque la famille  $(\mathcal{C}_i)_{i \in I}$  est indépendante. Alors  $\mathcal{C}_1$  est en fait une  $\lambda$ -classe contenant la  $\pi$ -classe  $\mathcal{C}_1$ , donc aussi la tribu engendrée  $\Sigma_1$  d'après (1.6.14) (théorème de Dynkin), et ainsi

$$P(B_1 \cap A_2 \cap \dots \cap A_p) = P(B_1) P(A_2) \dots P(A_p)$$

pour tout  $B_1 \in \Sigma_1$  et tous  $A_k \in \mathcal{C}_k$ ,  $k \geq 2$ . Il suffit alors d'opérer de proche en proche en considérant maintenant la classe  $\mathcal{C}_2$  des parties  $B_2 \in \Sigma_2$  telles que

$$P(B_1 \cap B_2 \cap A_3 \cap \dots \cap A_p) = P(B_1) P(B_2) P(A_3) \dots P(A_p)$$

pour tout  $B_1 \in \Sigma_1$  et tous  $A_k \in \mathcal{C}_k$ ,  $k \geq 3$ , pour voir que  $\mathcal{C}_2 = \Sigma_2$ . Et ainsi de suite, jusqu'à remplacer tous les  $A_k \in \mathcal{C}_k$  par des  $B_k \in \Sigma_k$ .  $\square$

Le critère précédent est absolument nécessaire pour la preuve du résultat qui suit, essentiel pour tout ce qui concerne l'indépendance, et traduisant le fait que l'indépendance se conserve par regroupements disjoints.

(1.12.4) **Théorème (des coalitions)**

Soit  $(\Sigma_i)_{i \in I}$  une famille indépendante de sous-tribus de  $\Sigma$ . On fixe une partition  $(I_\alpha)_{\alpha \in \mathfrak{A}}$  de l'ensemble  $I$  des indices, et pour chaque  $\alpha \in \mathfrak{A}$  on désigne par  $\Sigma'_\alpha$  la tribu engendrée par la classe  $\bigcup_{i \in I_\alpha} \Sigma_i$ . Alors la famille  $(\Sigma'_\alpha)_{\alpha \in \mathfrak{A}}$  des blocs ainsi constitués est indépendante.

**Preuve.** Désignons par  $\mathcal{C}_\alpha$  la classe des parties  $B \in \Sigma$  de la forme  $B = \bigcap_{i \in I_\alpha} A_i$ , avec  $A_i \in \Sigma_i$  et  $A_i = \Omega$  sauf pour un nombre fini d'indices. Alors  $\mathcal{C}_\alpha$  est une  $\pi$ -classe contenant toutes les  $\Sigma_i$  pour  $i \in I_\alpha$ , contenue dans la tribu  $\Sigma'_\alpha$ , et par conséquent engendrant  $\Sigma'_\alpha$ . Et comme

il est immédiat que l'indépendance de la famille  $(\Sigma_i)_{i \in I}$  implique l'indépendance de la famille  $(\mathcal{E}_\alpha)_{\alpha \in \mathfrak{A}}$ , le résultat est une simple conséquence de (1.12.3).  $\square$

**Tribu asymptotique et loi du zéro-un.** Le théorème des coalitions va permettre d'obtenir un premier résultat asymptotique du calcul des probabilités.

(1.12.5) *Définition*

Soit  $(\Sigma_n)_{n \geq 1}$  une suite de sous-tribus de  $\Sigma$ . On appelle tribu des événements asymptotiques la tribu  $\Sigma_\infty$  égale à l'intersection des tribus  $\Sigma'_n$  engendrées par les  $\Sigma_k$ ,  $k \geq n$ .

Considérons par exemple une suite d'événements  $A_n \in \Sigma_n$ . Alors les événements

$$A = \limsup A_n = \bigcap_n \bigcup_{k \geq n} A_k$$

$$B = \liminf A_n = \bigcup_n \bigcap_{k \geq n} A_k$$

sont tous deux asymptotiques. On remarquera que la réalisation de A signifie qu'une infinité d'événements  $A_k$  sont réalisés, et que la réalisation de B signifie que tous les  $A_k$  sauf un nombre fini sont réalisés. On a évidemment  $B \subset A$ .

L'intérêt principal de la tribu asymptotique réside dans l'énoncé suivant :

(1.12.6) *Théorème (Loi du zéro-un)*

Soit  $(\Omega, \Sigma, P)$  un espace probabilisé et soit  $(\Sigma_n)_{n \geq 1}$  une suite indépendante de sous-tribus de  $\Sigma$ . Pour tout événement asymptotique  $A \in \Sigma_\infty$  on a  $P(A) = 0$  ou  $P(A) = 1$ .

**Preuve.** Soit  $\Sigma''_n$  la tribu engendrée par les  $\Sigma_k$  pour  $k < n$ . Alors  $\Sigma''_n$  et  $\Sigma'_n$  sont indépendantes par (1.12.4), et puisque  $\Sigma_\infty \subset \Sigma'_n$  on voit que  $\Sigma''_n$  et  $\Sigma_\infty$  sont indépendantes. La suite  $(\Sigma''_n)$  est croissante, et sa réunion  $\mathfrak{R}$  est un anneau, et a fortiori une  $\pi$ -classe, qui engendre la tribu  $\Sigma''$  engendrée par toutes les  $\Sigma_n$ . Il résulte du critère (1.12.3) que  $\Sigma''$  et  $\Sigma_\infty$  sont indépendantes. Comme  $\Sigma_\infty \subset \Sigma''$ , on en déduit que  $\Sigma_\infty$  est indépendante d'elle-même (ce qui n'a aucun caractère de contradiction) et alors pour  $A \in \Sigma_\infty$  on a

$$P(A) = P(A \cap A) = P(A)^2$$

donc  $P(A) = 0$  ou  $P(A) = 1$ .  $\square$

En particulier soit  $(A_n)$  une suite indépendante d'événements. Pour chaque  $n$ , soit  $\Sigma_n = \{ \emptyset, \Omega, A_n, A_n^c \}$  la tribu engendrée par  $A_n$ . En appliquant le critère (1.12.3) aux classes  $\mathcal{E}_n = \{A_n\}$  on voit que la suite  $(\Sigma_n)$  est indépendante, de sorte que l'événement

$\limsup A_n$  est asymptotique et donc  $P(\limsup A_n) = 0$  ou  $P(\limsup A_n) = 1$ . Pour distinguer les deux cas on a :

(1.12.7) **Théorème (Borel-Cantelli)**

Soit  $(\Omega, \Sigma, P)$  un espace probabilisé et soit  $(A_n)$  une suite d'événements.

- a) Si  $\sum P(A_n) < +\infty$  alors  $P(\limsup A_n) = 0$  (même si la suite  $(A_n)$  n'est pas indépendante).  
 b) Si la suite  $(A_n)$  est indépendante et si  $\sum P(A_n) = +\infty$ , alors  $P(\limsup A_n) = 1$ .

**Preuve.** Soit  $B_n = \bigcup_{k \geq n} A_k$  et  $A = \bigcap B_n = \limsup A_n$ . Dans le cas a) on a

$$P(A) \leq P(B_n) \leq \sum_{k \geq n} P(A_k) = \rho_n, \text{ où } \rho_n \text{ est le reste d'une série convergente}$$

(dans  $\mathbb{R}_+$ ), donc  $\rho_n \downarrow 0$  et  $P(A) = 0$ . Dans le cas b) remarquons que  $A^c = \bigcup_n B_n^c$ , avec la

monotonie  $B_n^c \uparrow A^c$ , puisque  $B_n \downarrow A$ . Par ailleurs, pour  $n$  fixé,  $B_n^c = \bigcap_{k \geq n} A_k^c$ , donc  $B_n^c$  est

l'intersection de la suite décroissante  $C_p = \bigcap_{n \leq k \leq p} A_k^c$ , avec  $p \geq n$ . Or par indépendance,

on a

$$P(C_p) = \prod_{k=n}^p [1 - P(A_k)]$$

La série  $\sum P(A_k)$  étant divergente, il en est de même de la série  $\sum \ln [1 - P(A_k)]$ , de sorte que  $P(C_p) \downarrow 0$  quand  $p \rightarrow +\infty$ . Alors  $P(B_n^c) = 0$  pour tout  $n$ , et par monotonie croissante,  $P(A^c) = 0$  donc  $P(A) = 1$ .  $\square$

(1.12.8) **Exercice**

Soit  $(A_n)$  une suite indépendante d'événements et  $B = \liminf A_n$ . En remarquant que  $B^c = \limsup A_n^c$ , donner les conditions sur les  $P(A_n)$  pour que  $P(B) = 0$  ou  $P(B) = 1$ .

(1.12.9) **Exercice (Le singe dactylographe)**

Un singe tape à la machine à écrire au hasard sur les touches. Sachant que cette machine possède  $m$  signes (lettres, ponctuations, ...), quelle est la probabilité pour que le singe tape une œuvre littéraire donnée comportant  $N$

signes ? On désigne par  $A_k$  l'événement consistant en la frappe de l'œuvre précédente du  $[(k-1)N+1]^{\text{ème}}$  coup au  $(kN)^{\text{ème}}$ . Montrer que la suite  $(A_k)$  est indépendante et que  $P(A_k) = p$  est constant. Conclusion : l'œuvre sera-t-elle tapée au moins une fois, une infinité de fois ? (On fait bien entendu abstraction du temps passé, de l'usure de la machine, et aussi de celle du singe !).

**Les variables aléatoires réelles.** Sur un espace  $(\Omega, \Sigma, P)$  probabilisé, on désigne classiquement par "variable aléatoire réelle" toute fonction mesurable réelle, à valeurs finies, que l'on note bien souvent sous la forme  $X : \Omega \rightarrow \mathbb{R}$ . La probabilité  $P$  n'intervient pas directement dans cette définition, mais elle intervient aussitôt en permettant de considérer la mesure image  $X(P)$ , transportant la probabilité  $P$  sur la droite numérique  $\mathbb{R}$ .

On note souvent  $\mu_X = X(P)$  et on dit que  $\mu_X$  est la *loi*, ou la *distribution* de la variable aléatoire réelle  $X$ . C'est évidemment une probabilité sur  $\mathbb{R}$ , que l'on peut définir aussi par sa *fonction de répartition*, notée  $F_X$  :

$$F_X(t) = \mu_X((-\infty, t]) = P\{X < t\}$$

en notant, comme il est classique, par  $\{X < t\}$  l'événement

$$\{\omega; X(\omega) < t\} = X^{-1}((-\infty, t]) .$$

On transportera très fréquemment la classification des probabilités sur  $\mathbb{R}$ , esquissée dans la suite des exemples (1.7.7). C'est ainsi qu'on dira qu'une variable aléatoire réelle  $X$  est uniforme sur  $[0, 1]$ , ou de Cauchy, ou de Laplace-Gauss  $LG(0, 1)$ , ou d'Euler, ou binomiale, ou de Poisson, etc. si sa loi a la même dénomination.

(1.12.10) **Exemple 1. Loi du  $\chi^2$  à un degré de liberté.** Étant donné une variable aléatoire réelle  $X$  suivant une loi de Laplace-Gauss réduite (c'est-à-dire  $LG(0, 1)$ ), on appelle loi du  $\chi^2$  à un degré de liberté la loi de la variable aléatoire réelle  $S = X^2$ . Il est facile de voir que  $S$  est définie par la fonction de répartition

$$F(t) = \begin{cases} 0 & \text{si } t \leq 0 \\ \frac{1}{\sqrt{2\pi}} \int_0^t e^{-\frac{u}{2}} \frac{du}{\sqrt{u}} & \text{si } t > 0 \end{cases}$$

donc aussi par la densité  $h(t)$ , nulle pour  $t < 0$ , et égale à  $\frac{1}{\sqrt{2\pi} t} \exp\left(-\frac{t}{2}\right)$  si  $t > 0$ .

On retrouve une loi d'Euler de paramètres  $\alpha = \frac{1}{2}$  et  $\beta = 2$ .

(1.12.11) **Exemple 2. Loi log-normale de paramètre  $a$ .** Soit encore  $X$  une variable aléatoire réelle suivant une loi  $LG(0, 1)$ . Alors pour tout  $a > 0$  les variables aléatoires réelles positives  $Y = e^{aX}$  et  $Z = e^{-aX}$  suivent une même loi, dite log-normale de paramètre  $a$ , définie sur  $[0, +\infty)$  par la densité

$$h_a(t) = \frac{1}{\sqrt{2\pi t}} \frac{1}{at} \exp \left[ -\frac{1}{2} (\ln t)^2 \right]$$

On pourra vérifier que la loi log-normale  $\mu_a$  précédente est sa propre image dans la transformation  $x \rightarrow \frac{1}{x}$  de  $\mathbb{R}_+^*$  en lui-même.

**Variabes aléatoires indépendantes.** Si  $X$  est une variable aléatoire réelle, la tribu  $\mathcal{X}^{-1}(\mathcal{B})$ , où  $\mathcal{B}$  est la tribu borélienne de  $\mathbb{R}$ , est évidemment une sous-tribu  $\Sigma_X$  de  $\Sigma$ . On dit alors qu'une famille  $(X_i)_{i \in I}$  de variables aléatoires réelles est indépendante lorsque la famille correspondante des sous-tribus  $\Sigma_{X_i}$  est elle-même indépendante. Le critère (1.12.3) permet d'affirmer que la famille  $(X_i)_{i \in I}$  est indépendante ssi, pour toute partie finie  $J \subset I$  et tous réels  $\alpha_i, i \in J$ , on a

$$P \{X_i < \alpha_i \quad \forall i \in J\} = \prod_{i \in J} P \{X_i < \alpha_i\}$$

On voit apparaître ici clairement l'intérêt du théorème des coalitions qui autorise les regroupements. Par exemple si la suite finie  $(X_1, X_2, X_3, X_4)$  est indépendante, il en est de même de la suite  $(X_1^2, X_2, X_3, X_4)$ , ou de la suite  $(X_1, X_2, X_3, X_4)$ , etc.

**Vecteurs aléatoires.** On désigne sous ce vocable toute application mesurable  $V : \Omega \rightarrow \mathbb{R}^p$ . En posant  $V = (X_1, \dots, X_p)$ , on associe à  $V$  ses composantes  $X_k : \Omega \rightarrow \mathbb{R}$  qui sont évidemment des variables aléatoires réelles. Mais réciproquement si les  $X_k$  sont des variables aléatoires réelles, alors  $V$  est aussi un vecteur aléatoire, comme on voit en remarquant que la tribu borélienne de  $\mathbb{R}^p$  est engendrée par les ensembles de la forme  $\{x_1 < \alpha_1, x_2 < \alpha_2, \dots, x_p < \alpha_p\}$ . Par analogie avec le cas  $p = 1$  on dit que la mesure image  $V(P)$ , qui est une probabilité sur  $\mathbb{R}^p$ , est la loi de  $V$ .

(1.12.12) *Exercice (Loi du  $\chi^2$  à deux degrés de liberté)*

Soit  $V = (X, Y)$  un vecteur aléatoire de dimension 2 suivant la loi de Laplace-Gauss sur  $\mathbb{R}^2$ , de densité

$$g_2(x, y) = \frac{1}{2\pi} \exp \left[ -\frac{x^2 + y^2}{2} \right]$$

- Montrer que  $X$  et  $Y$  sont deux variables aléatoires réelles de même loi LG  $(0, 1)$ , qui sont de plus indépendantes.
- On appelle loi du  $\chi^2$  à deux degrés de liberté, la loi de la variable aléatoire réelle positive  $S_2 = \|V\|^2 = X^2 + Y^2$ . Établir, par un calcul en coordonnées polaires, que la loi de  $S_2$  est définie par la densité

$$h_2(t) = \frac{1}{2} e^{-\frac{t}{2}}$$

sur  $[0, +\infty)$ , autrement dit est la loi d'Euler de paramètres  $\alpha = 1$  et  $\beta = 2$ .

(1.12.13) **Exercice (Loi du  $\chi^2$  à trois degrés de liberté)**

Soit  $V = (X, Y, Z)$  un vecteur aléatoire de dimension 3 suivant la loi de Laplace-Gauss sur  $\mathbb{R}^3$ , de densité

$$g_3(x, y, z) = \frac{1}{(2\pi)^{3/2}} \exp \left[ -\frac{1}{2}(x^2 + y^2 + z^2) \right]$$

- Montrer que les variables aléatoires réelles  $X, Y, Z$  sont indépendantes et suivent la même loi LG  $(0, 1)$ .
- Par un calcul en coordonnées sphériques dans  $\mathbb{R}^3$ , établir que la variable aléatoire réelle positive  $S_3 = \|V\|^2 = X^2 + Y^2 + Z^2$  suit la loi (dite du  $\chi^2$  à trois degrés de liberté) définie par la densité

$$h_3(t) = \frac{1}{\sqrt{2\pi}} \sqrt{t} e^{-\frac{t}{2}} \quad \text{sur } [0, +\infty).$$

De quelle loi d'Euler s'agit-il ?

(1.12.14) **Exercice (Loi du  $\chi^2$  à  $p$  degrés de liberté)**

Soit  $V = (X_1, X_2, \dots, X_p)$  un vecteur aléatoire  $p$ -dimensionnel dont la loi admet la densité sur  $\mathbb{R}^p$

$$g(x) = \frac{1}{(2\pi)^{p/2}} \exp \left[ -\frac{\|x\|^2}{2} \right]$$

- Vérifier que la suite finie  $(X_1, X_2, \dots, X_p)$  est indépendante et que chaque variable aléatoire réelle  $X_k$  suit la loi LG  $(0, 1)$ .
- En utilisant l'intégration des fonctions sphériques vue en (1.8.17), ainsi que l'exercice (1.10.9), prouver que la variable aléatoire réelle positive  $S_p = \|V\|^2 = X_1^2 + \dots + X_p^2$  suit la loi (dite du  $\chi^2$  à  $p$  degrés de liberté) définie par la densité

$$h_p(t) = \frac{1}{2^{p/2}} \frac{1}{\Gamma\left(\frac{p}{2}\right)} t^{\frac{p}{2}-1} e^{-\frac{t}{2}} \quad \text{sur } [0, +\infty).$$

Vérifier qu'il s'agit de la loi d'Euler de paramètres  $\alpha = \frac{p}{2}$  et  $\beta = 2$ .

**Espérance mathématique et moments d'une variable aléatoire.** Lorsqu'une variable aléatoire réelle  $X: \Omega \rightarrow \mathbb{R}$  est intégrable, on introduit évidemment son intégrale

$\int X dP$ . Et alors :

(1.12.15) *Définition*

Soit  $X$  une variable aléatoire réelle intégrable sur l'espace probabilisé  $(\Omega, \Sigma, P)$

a) On appelle *espérance mathématique* (ou moyenne) de  $X$  le nombre

$$m = E(X) = \int X dP = \int t d\mu_X(t)$$

b) On dit que  $X$  est une *variable centrée* lorsque  $E(X) = 0$ .

En remplaçant  $X$  par la variable aléatoire réelle  $X^n$  on obtient la notion de *moment*.

(1.12.16) *Définition*

Soit  $n \geq 1$  un entier. On dit que la variable aléatoire réelle  $X$  admet un moment d'ordre  $n$  lorsque la variable aléatoire réelle  $X^n$  est intégrable. On appelle alors *moment d'ordre  $n$*  le nombre  $m_n = E(X^n)$ , et *moment absolu* le nombre  $\tilde{m}_n = E(|X|^n)$  si  $n$  est impair.

L'inégalité  $|X|^p \leq 1 + |X|^n$  valable pour  $1 \leq p \leq n$ , et vérifiée en distinguant le cas  $|X(\omega)| \leq 1$  et le cas  $|X(\omega)| > 1$ , montre que  $X$  admet un moment d'ordre  $p \leq n$ , dès qu'elle admet un moment d'ordre  $n$ . En particulier la fonction  $(X - E(X))^n$  est alors intégrable, d'où la définition :

(1.12.17) *Définition*

Soit  $X$  une variable aléatoire admettant un moment d'ordre  $n \geq 2$ . On appelle *moment centré* d'ordre  $n$  la quantité

$$m'_n = \int [X - E(X)]^n dP = \int [t - E(X)]^n d\mu_X(t)$$

Lorsque  $n = 2$ , le moment centré  $m'_2$  prend le nom de *variance* de  $X$ , notée  $\text{var}(X)$ , la quantité  $\sigma(X) = \text{var}(X)^{1/2}$  prenant le nom d'*écart-type*.

**Remarque.** La moyenne  $E(X)$  s'apparente à un centre de gravité dans la représentation d'une probabilité comme distribution de masses, tandis que la variance s'apparente à un moment d'inertie par rapport à ce centre de gravité. Plus la variance est petite et plus la variable aléatoire réelle  $X$  est concentrée autour de sa valeur moyenne, de sorte que cette variance offre une mesure de la *dispersion* de  $X$ . Le cas limite  $\text{var}(X) = \sigma^2 = 0$  correspond d'ailleurs au cas où  $X = m$   $P$ -presque partout, c'est-à-dire presque sûrement. On a alors  $\mu_X = \delta_m$ , masse de Dirac localisée au point  $m$ .

En développant  $[X - E(X)]^2$  on obtient :

(1.12.18) *Proposition*

Soit  $X$  une variable aléatoire admettant un moment d'ordre 2. On a alors

$$\text{var}(X) = \sigma^2(X) = E(X^2) - E(X)^2$$

En particulier on a l'inégalité  $E(X)^2 \leq E(X^2)$ .

Lorsque  $\sigma(X) > 0$  on associe à  $X$  la variable aléatoire réelle

$$Y = \frac{X - E(X)}{\sigma(X)}$$

telle que  $E(Y) = 0$  et  $\sigma(Y) = 1$ . La loi de  $Y$  est alors dite loi réduite associée à la loi de  $X$ ; on dit aussi que  $Y$  est la variable réduite associée à  $X$ .

(1.12.19) *Exemples.* Donnons sous forme de tableau les valeurs des espérances et des moments de quelques variables aléatoires déjà introduites par la donnée de leur loi. Pour les lois discrètes les intégrales sont des séries qu'il faut sommer, ce qui est plus ou moins facile. Pour les lois densitables, les intégrales sont plus classiques et se calculent souvent par l'intermédiaire de la fonction  $\Gamma$  d'Euler.

Loi de $X$	$E(X)$	$\sigma^2(X)$
Bernoulli de paramètre $p$	$p$	$pq$
Binomiale de paramètres $n, p$	$np$	$npq$
Poisson de paramètre $\lambda$	$\lambda$	$\lambda$
Géométrique de paramètre $p$	$\frac{1}{p}$	$\frac{q}{p^2}$

Loi de $X$	$E(X)$	$\sigma^2(X)$	$m_n$ ou $m'_n$
Uniforme sur $[0, 1]$	$\frac{1}{2}$	$\frac{1}{12}$	$m'_n = 0$ si $n$ impair $m'_n = \frac{2^{-n}}{n+1}$ si $n$ pair
Cauchy de paramètre $a$	<del>XXXX</del>	<del>XXXX</del>	<del>XXXX</del>
Laplace-Gauss réduite	0	1	$m_{2p+1} = 0$ $m_{2p} = 1.3.5 \dots (2p-1)$
LG ( $m, \sigma$ )	$m$	$\sigma^2$	$m'_{2p+1} = 0$ $m'_{2p} = 1.3.5 \dots (2p-1) \sigma^{2p}$
Euler de paramètres $\alpha, \beta$	$\alpha\beta$	$\alpha \beta^2$	$m_n = \alpha(\alpha+1) \dots (\alpha+n-1) \beta^n$
$\chi_2$ à 1 degré de liberté	1	2	$m_n = 1.3.5 \dots (2n-1)$

$\chi_2$ à 2 degrés de liberté	2	4	$m_n = 2.4.6. \dots 2n$
$\chi_2$ à 3 degrés de liberté	3	6	$m_n = 3.5.7. \dots (2n-1)$
$\chi_2$ à p degrés de liberté	p	2p	$m_n = p(p+2) \dots (p+2n-2)$
log-normale de paramètre a	$\exp\left(\frac{a^2}{2}\right)$	$\exp(2a^2) - \exp(a^2)$	$m_n = \exp\left(\frac{a^2 n^2}{2}\right)$

**Les inégalités de Markoff et Tchebyscheff.** On a dit plus haut que les moments d'une variable aléatoire réelle  $X$  pouvaient servir de mesure de la dispersion de  $X$ . Les inégalités qui suivent vont préciser quantitativement cette question. On remarquera qu'elles ont toutes en commun pour but de montrer que la probabilité pour que  $|X|$ , ou  $|X - E(X)|$ , prenne de grandes valeurs est suffisamment petite.

(1.12.20) *Proposition (Inégalité de Markoff)*

Soit  $X$  une variable aléatoire réelle positive intégrable. Pour tout  $r > 0$ , on a

$$P\{X \geq r\} \leq \frac{E(X)}{r}$$

**Preuve.** Car si  $A = \{X \geq r\}$ , alors  $E(X) \geq \int_A X dP \geq rP(A)$ . □

On en déduit aisément :

(1.12.21) *Corollaire*

Soit  $\alpha > 0$ . Pour toute variable aléatoire réelle  $X$  telle que  $E(|X|^\alpha) < +\infty$ , et pour tout  $r > 0$ , on a

$$P\{|X| \geq r\} \leq \frac{E(|X|^\alpha)}{r^\alpha}$$

En supposant  $\alpha = 2$  et en remplaçant  $X$  par  $X - E(X)$ , on obtient l'inégalité classique :

(1.12.22) *Proposition (Inégalité de Tchebyscheff)*

Soit  $X$  une variable aléatoire réelle admettant un moment d'ordre 2. Pour tout  $r > 0$ , on a

$$P\{|X - E(X)| \geq r\} \leq \frac{\sigma^2(X)}{r^2}$$

On voit bien ici que  $\sigma(X)$  apparaît comme une unité d'écart par rapport à la moyenne  $m = E(X)$ , d'où le nom d'écart-type.

**Produit de deux variables aléatoires réelles intégrables.** Lorsque  $X$  et  $Y$  sont deux variables aléatoires réelles intégrables le produit  $XY$  n'est pas en général intégrable. Par exemple sur  $\Omega = [0, 1]$ , avec pour probabilité  $P$  la mesure de Borel, la fonction  $X(t) = \frac{1}{\sqrt{t}}$  est intégrable mais  $X^2$  ne l'est pas. Par contre, lorsque  $X$  et  $Y$  ont des moments

d'ordre 2, l'inégalité

$$2 |XY| \leq X^2 + Y^2$$

garantit que le produit  $XY$  est intégrable. De plus on a l'inégalité de Cauchy-Schwarz, déjà rencontrée lors de l'étude de la fonction  $\Gamma$ , et que nous retrouverons plus loin dans un cadre plus général :

(1.12.23) *Proposition (Inégalité de Cauchy-Schwarz)*

Soit  $X$  et  $Y$  deux variables admettant des moments d'ordre 2. Alors la variable aléatoire réelle  $XY$  est intégrable et

$$E(XY) \leq [E(X^2)]^{\frac{1}{2}} [E(Y^2)]^{\frac{1}{2}}$$

*Preuve.* Toujours avec le trinôme

$$T(\lambda) = E[(X + \lambda Y)^2] = E(X^2) + 2\lambda E(XY) + \lambda^2 E(Y^2)$$

qui, étant positif ou nul, a son discriminant négatif ou nul. □

A partir de là, et sous les mêmes hypothèses, on introduit la *covariance de  $X$  et  $Y$*  selon

$$\text{cov}(X, Y) = E\{(X - E(X))(Y - E(Y))\} = E(XY) - E(X)E(Y)$$

qui vérifie donc l'inégalité, autre forme de Cauchy-Schwarz

$$|\text{cov}(X, Y)| \leq \sigma(X) \sigma(Y)$$

Revenons aux variables  $X$  et  $Y$  supposées intégrables. Il est un cas où leur produit est encore intégrable : c'est celui où  $X$  et  $Y$  sont indépendantes. En effet :

(1.12.24) *Proposition*

Soit  $X, Y$  deux variables intégrables et indépendantes. Alors  $XY$  est intégrable, et on a

$$E(XY) = E(X)E(Y)$$

$$\text{cov}(X, Y) = 0$$

*Preuve.* Il suffit de prouver la première formule dans le cas  $X \geq 0$  et  $Y \geq 0$ , car ensuite on remplace  $X$  et  $Y$  par  $|X|$  et  $|Y|$  pour vérifier que

$$E(|XY|) = E(|X|)E(|Y|) < +\infty,$$

puis dans le cas général on décompose en  $X = X^+ - X^-$  et  $Y = Y^+ - Y^-$ . Or pour  $X \geq 0$  et  $Y \geq 0$ , il existe une suite  $(\varphi_n)$  et une suite  $(\psi_n)$  de fonctions étagées telles que  $\varphi_n \uparrow X$  et  $\psi_n \uparrow Y$ , donc aussi  $\varphi_n \psi_n \uparrow XY$ . De plus d'après la preuve de (1.4.6), on peut choisir  $\varphi_n$  mesurable par rapport à la tribu  $\Sigma_X$  et  $\psi_n$  mesurable par rapport à  $\Sigma_Y$ , de sorte que  $\varphi_n$  et  $\psi_n$  sont aussi indépendantes entre elles. Alors il ne reste plus qu'à prouver que

$$E(\varphi_n \psi_n) = E(\varphi_n) E(\psi_n),$$

ce qui ramène au cas où  $X$  et  $Y$  sont étagées, et toujours indépendantes. Or dans ce cas on a

$$X = \sum a_i 1_{A_i} \quad \text{et} \quad Y = \sum b_j 1_{B_j}$$

avec  $A_i \in \Sigma_X$  et  $B_j \in \Sigma_Y$ , donc chaque  $A_i$  est indépendant de chaque  $B_j$ . Alors

$$XY = \sum_{i,j} a_i b_j 1_{A_i \cap B_j} \text{ et}$$

$$E(XY) = \sum_{i,j} a_i b_j P(A_i \cap B_j)$$

$$= \sum_{i,j} a_i b_j P(A_i) P(B_j)$$

$$= \left[ \sum_i a_i P(A_i) \right] \left[ \sum_j b_j P(B_j) \right] = E(X) E(Y). \quad \square$$

(1.12.25) *Corollaire (Règle de Bienaymé-Tchebyscheff)*

Soit  $X_1, X_2, \dots, X_p$  des variables aléatoires réelles d'ordre 2, supposées 2 à 2 indépendantes. On pose  $S = X_1 + X_2 + \dots + X_p$ . Alors  $S$  est d'ordre 2 et

$$\text{var}(S) = \text{var}(X_1) + \dots + \text{var}(X_p).$$

**Preuve.** Soit  $m_i = E(X_i)$  et  $m = E(S) = \sum m_i$ , de sorte que  $\tilde{S} = S - E(S) = \sum \tilde{X}_i$  avec  $\tilde{X}_i = X_i - E(X_i)$ . En calculant  $\tilde{S}^2$  on a donc

$$\text{var}(S) = \sum_i \text{var}(X_i) + \sum_{\substack{i,j \\ i \neq j}} \text{cov}(X_i, X_j)$$

et le résultat provient des égalités  $\text{cov}(X_i, X_j) = 0$  si  $i \neq j$ , conséquences de l'indépendance 2 à 2.  $\square$

**Loi faible des grands nombres.** Donnons comme application de la règle précédente et de l'inégalité de Tchebyscheff la preuve de la loi faible des grands nombres qui, historiquement, a joué un rôle fondamental dans la mise au point des principes mathématiques qui régissent le calcul des probabilités.

(1.12.26) *Théorème (Loi faible des grands nombres)*

Soit  $(X_n)_{n \geq 1}$  une suite de variables réelles d'ordre 2, supposées toutes de même loi, et deux à deux indépendantes. On pose  $m = E(X_1)$  et  $\sigma^2 = \text{var}(X_1)$ , puis  $S_n = X_1 + \dots + X_n$ . On a alors, pour tout  $\varepsilon > 0$  :

$$P \left\{ \left| \frac{S_n}{n} - m \right| \geq \varepsilon \right\} \leq \frac{\sigma^2}{n \varepsilon^2}$$

**Preuve.** Puisque  $E(X_k) = m$  et  $\text{var}(X_k) = \sigma^2$ , car les variables aléatoires réelles  $X_k$  ont même loi, on a  $E(S_n) = nm$ , donc  $E\left(\frac{S_n}{n}\right) = m$ , et tout provient de l'inégalité de Tchebscheff puisque l'on a, avec (1.12.25),

$$\text{var} \left( \frac{S_n}{n} - m \right) = \text{var} \left( \frac{S_n}{n} \right) = \frac{1}{n^2} \text{var}(S_n) = \frac{1}{n^2} \sum_{k=1}^n \text{var}(X_k) = \frac{\sigma^2}{n}. \quad \square$$

Comme corollaire, retrouvons les tout premiers résultats de Bernoulli en calcul des probabilités, obtenus dès le 18<sup>ème</sup> siècle, et justifiant l'introduction de la probabilité d'un événement à partir de la fréquence de ses réalisations.

(1.12.27) *Corollaire (Loi faible de Bernoulli)*

Soit  $(A_n)_{n \geq 1}$  une suite d'événements, tous de même probabilité  $p$ , et supposés deux à deux indépendants. Soit  $N_n$  le nombre de réalisations des  $A_k$  pour  $k \leq n$ , et  $\varphi_n = \frac{N_n}{n}$  la fréquence de réalisation des  $A_x$  pour  $k \leq n$ . On a alors pour tout  $\varepsilon > 0$

$$P \{ |\varphi_n - p| \geq \varepsilon \} \leq \frac{1}{4n \varepsilon^2}$$

**Preuve.** Pour chaque épreuve  $\omega$  on a  $N_n(\omega) = \sum_{k=1}^n 1_{A_k}(\omega)$ , de sorte qu'il suffit de poser  $X_k = 1_{A_k}$  et de constater que la loi de chaque  $X_k$  est exactement la probabilité de Bernoulli  $\mu = p \delta_1 + q \delta_0$  avec  $q = 1 - p$ . Donc les variables aléatoires réelles  $X_k$  ont même loi, et de plus  $E(X_k) = p$  et  $\text{var}(X_k) = pq = p(1 - p)$ , dont le maximum est atteint pour  $p = \frac{1}{2}$ .  $\square$

En lisant la propriété obtenue sous la forme

$$P \{ |\varphi_n - p| < \varepsilon \} \geq 1 - \frac{1}{4n \varepsilon^2}$$

on voit donc que pour tout voisinage  $J$  du point  $p$ , on aura  $\varphi_n \in J$  pour  $n$  assez grand, avec une probabilité aussi voisine de 1 que l'on voudra. C'est en ce sens que l'on dit que  $\varphi_n \rightarrow p$ , la notion de convergence introduite s'appelant convergence en probabilité.

**Variance en dimension p.** Rappelons tout d'abord que l'espace  $\mathbb{R}^P$  est muni de son produit scalaire euclidien

$$\langle x, y \rangle = \sum_{k=1}^P x_k y_k \quad \text{si } x = (x_k) \text{ et } y = (y_k)$$

et que sa base canonique est notée  $(e_1, e_2, \dots, e_p)$ .

On se fixe un vecteur aléatoire  $X = (X_1, X_2, \dots, X_p)$  défini sur l'espace probabilisé  $(\Omega, \Sigma, P)$ . Puisque sur  $\mathbb{R}^P$  on a

$$\text{Max } |x_k| \leq \|x\| \leq |x_1| + \dots + |x_p|$$

on voit que  $\|X\|$  est intégrable ssi chacune des composantes  $X_k$  est intégrable. En posant  $m_k = E(X_k)$  et  $m = (m_k)$ , on obtient un vecteur  $m \in \mathbb{R}^P$ , appelé vecteur *espérance mathématique* de  $X$ , ou encore *moyenne* de  $X$ , noté  $m = E(X)$ , qui est caractérisé par la condition que l'on ait, pour tout  $a \in \mathbb{R}^P$ ,

$$\langle m, a \rangle = E(\langle X, a \rangle) = \int \langle X(\omega), a \rangle dP(\omega)$$

Bien entendu, on dit que  $X$  est un vecteur centré lorsque  $E(X) = 0$ , c'est-à-dire lorsque chaque composante  $X_k$  est une variable aléatoire réelle centrée.

On dit maintenant que  $X$  est d'ordre 2, ou admet des moments d'ordre 2, lorsque la variable aléatoire réelle  $\|X\|^2$  est intégrable, c'est-à-dire lorsque les composantes  $X_k$  sont toutes d'ordre 2. On peut alors évidemment introduire le moment normique

$$E(\|X\|^2) = \sum_{k=1}^P E(X_k^2) \quad \text{et le moment normique centré } \sigma^2 = m'_2 = E(\|X - E(X)\|^2) =$$

$\sum_{k=1}^P \text{var}(X_k)$  mais ce moment est trop "globalisant" pour donner une information "directionnelle" sur le comportement de  $X$ .

On préfère introduire, pour chaque  $x \in \mathbb{R}^P$ , la quantité

$$V(x) = \int \langle X(\omega) - m, x \rangle^2 dP(\omega)$$

$$V(x) = E(\langle X - m, x \rangle^2)$$

qui est donc la variance de la variable aléatoire réelle  $\langle X, x \rangle$ . Il est clair que cette variance va alors mesurer la concentration ou la dispersion de  $X$  dans la direction définie par  $x$ . La fonction  $V$  est une forme quadratique sur  $\mathbb{R}^P$ , qui est de plus positive ou nulle, associée à la forme bilinéaire.

$$L(x, y) = E(\langle X - m, x \rangle \langle X - m, y \rangle)$$

selon  $V(x) = L(x, x)$ . On dit que  $L$ , qui est donc une forme bilinéaire symétrique de type positif, est la forme bilinéaire de *covariance* du vecteur aléatoire  $X$ . En revenant aux composantes  $(X_k)$  on voit que  $L$  est complètement déterminée par les coefficients

$$\begin{cases} c_{ij} = L(e_i, e_j) = E((X_i - m_i)(X_j - m_j)) = \text{cov}(X_i, X_j) & \text{si } i \neq j \\ c_{kk} = L(e_k, e_k) = V(e_k) = \text{var}(X_k) & \text{si } i = j = k \end{cases}$$

ce qui incite à introduire la matrice symétrique

$$\Gamma = [c_{ij}]$$

dite *matrice de covariance* de  $X$  et alors

$$\begin{cases} L(x, y) = \langle \Gamma x, y \rangle = \langle x, \Gamma y \rangle = \sum_i \sum_j c_{ij} x_i y_j \\ V(x) = L(x, x) = \langle \Gamma x, x \rangle = \langle x, \Gamma x \rangle = \sum_i \sum_j c_{ij} x_i x_j \end{cases}$$

La liaison entre  $L$  et  $V$  se précise encore par le fait que l'on a l'inégalité de Cauchy-Schwarz

$$L(x, y)^2 \leq V(x) V(y)$$

conséquence de la positivité de  $V(x)$ . Par ailleurs la formule suivante est fréquemment utile

$$V(x + y) = V(x) + 2L(x, y) + V(y)$$

Lorsque  $\Gamma = I$ , matrice identité de  $\mathbb{R}^p$ , on dit que  $X$  est un vecteur réduit; chaque  $X_k$  est une variable aléatoire réelle réduite et de plus  $\text{cov}(X_i, X_j) = 0$  si  $i \neq j$ . Les relations entre  $V$  et  $L$  se réduisent alors aux relations habituelles entre la norme et le produit scalaire sur  $\mathbb{R}^p$ .

Pour aller plus loin, il convient de rappeler qu'une matrice symétrique est diagonalisable par une matrice orthogonale, ce qui traduit un changement de base orthonormale dans  $\mathbb{R}^p$ . La matrice  $\Gamma$  étant ici de type positif, ses valeurs propres  $\lambda_k$  sont

telles que  $\lambda_k \geq 0$ . Enfin  $E(\|X\|^2) = \sum_n c_{kk} = \text{tr } \Gamma$ , trace de la matrice  $\Gamma$ .

(1.12.28) **Exercice**

- a) Établir avec Cauchy-Schwarz que  $\langle \Gamma x, x \rangle = 0$  équivaut à  $\Gamma x = 0$ .
- b) On dit que  $X$  est non dégénéré lorsque sa matrice de covariance  $\Gamma$  est inversible. Montrer que les assertions suivantes sont équivalentes :
  - $X$  est non dégénéré;
  - la forme quadratique  $V$  est de type défini positif, c'est-à-dire que l'on a  $V(x) > 0$  pour tout  $x \neq 0$ ;
  - les valeurs propres de  $\Gamma$  sont strictement positives.

(1.12.29) *Exercice*

Soit  $X = (X_1, \dots, X_p)$  un vecteur aléatoire dégénéré, dont la matrice de covariance  $\Gamma$  n'est donc pas inversible. On pose  $S = \text{Ker } \Gamma = \{x \in \mathbb{R}^p, V(x) = 0\}$  et  $Q = S^\perp$ .

- Prouver que  $\Gamma(x) \in Q$  pour tout  $x \in \mathbb{R}^p$ , puis que  $\Gamma$  est injective sur  $Q$ . En déduire que  $\text{Im } \Gamma = \Gamma(Q) = Q$ .
- En remarquant que  $\int \langle X(\omega) - m, s \rangle^2 dP(\omega) = 0$  pour tout  $s \in S$ , prouver que l'on a  $X(\omega) \in m + Q$  presque sûrement. En déduire que le sous-espace affine  $Q_m = m + Q$  porte la distribution  $\mu_X$  de  $X$ , c'est-à-dire que l'on a  $\mu_X(Q_m) = 1$ .

(1.12.30) *Exercice*

Soit  $X = (X_1, \dots, X_p)$  un vecteur aléatoire centré non dégénéré,  $\Gamma$  sa matrice de covariance,  $\lambda_k$  les valeurs propres de  $\Gamma$ ,  $\Delta$  la matrice diagonale ( $\lambda_k$ ). On fixe une matrice orthogonale  $U$  telle que  $\Gamma = U \Delta U^{-1}$ , et on pose encore  $\Lambda = \Gamma^{-1} = U \Delta^{-1} U^{-1}$ .

- On pose  $Y = U^{-1} X$ , ce qui donne un nouveau vecteur aléatoire centré. Établir que sa matrice de covariance est la matrice  $\Delta$ .
- Prouver que  $\langle \Lambda X, X \rangle = \langle \Delta^{-1} Y, Y \rangle$  et en déduire que  $E(\langle \Lambda X, X \rangle) = p$ .
- Soit  $B_r$  la boule euclidienne de rayon  $r$ , et soit  $E_r$  l'ellipsoïde  $E_r = \{x \in \mathbb{R}^p, \langle \Lambda x, x \rangle \leq r^2\}$ . Prouver les inégalités

$$P\{X \notin B_r\} \leq \frac{\text{tr } \Gamma}{r^2}$$

$$P\{X \notin E_r\} \leq \frac{p}{r^2}$$

- Montrer que si  $p = 1$ , chacune de ces inégalités se réduit à l'inégalité de Tchebyscheff.



## CHAPITRE 2 : INTÉGRATION SUR UN ESPACE PRODUIT

Dans ce chapitre, il s'agit de confronter la théorie de l'intégration jusqu'ici exposée (tribus, mesures, fonctions mesurables, fonctions intégrables) à une situation nouvelle, qui est celle où l'espace de base  $\Omega$  est un produit cartésien  $\Omega = \Omega_1 \times \Omega_2$ .

Précisons tout d'abord quelques notations.

- On appelle *pavé* dans  $\Omega$  tout produit  $A = A_1 \times A_2$  avec  $A_1 \subset \Omega_1$  et  $A_2 \subset \Omega_2$ .
- Pour toute partie  $A \subset \Omega$ , tout  $x_1 \in \Omega_1$  et tout  $x_2 \in \Omega_2$  on désigne par :
  - $A(x_1, \bullet) = \{x_2 \in \Omega_2, (x_1, x_2) \in A\}$  la coupe de  $A$  au point  $x_1$  et par
  - $A(\bullet, x_2) = \{x_1 \in \Omega_1, (x_1, x_2) \in A\}$  la coupe de  $A$  au point  $x_2$ .
- Pour toute fonction  $f : \Omega \rightarrow X$ , où  $X$  est un ensemble quelconque, on désigne par  $f(x_1, \bullet) : \Omega_2 \rightarrow X$  et  $f(\bullet, x_2) : \Omega_1 \rightarrow X$  les fonctions partielles définies par  $x_2 \rightarrow f(x_1, x_2)$  et  $x_1 \rightarrow f(x_1, x_2)$  respectivement.
- Pour toutes fonctions numériques (réelles ou complexes)  $f_1$  et  $f_2$  définies respectivement sur  $\Omega_1$  et  $\Omega_2$ , on note  $f_1 \otimes f_2$  la fonction définie sur  $\Omega$  par  $(x_1, x_2) \rightarrow f_1(x_1) f_2(x_2)$ .
- Enfin, on note  $\pi_1 : \Omega \rightarrow \Omega_1$  et  $\pi_2 : \Omega \rightarrow \Omega_2$  les projections canoniques de  $\Omega$ .

### 2.1 Produit de deux espaces mesurables

On suppose que  $\Omega_1$  et  $\Omega_2$  sont munis respectivement de tribus  $\Sigma_1$  et  $\Sigma_2$ . On dit alors qu'un pavé  $A = A_1 \times A_2$  avec  $A_1 \in \Sigma_1$  et  $A_2 \in \Sigma_2$  est un *pavé mesurable*. Il est alors immédiat que deux pavés mesurables  $A = A_1 \times A_2$  et  $B = B_1 \times B_2$  sont tels que  $A \cap B = (A_1 \cap B_1) \times (A_2 \cap B_2)$  est encore un pavé et que la différence  $A \setminus B = [(A_1 \setminus B_1) \times A_2] \cup [(A_1 \cap B_1) \times (A_2 \setminus B_2)]$  est *réunion disjointe* de deux pavés mesurables. Il suit de là aisément qu'une réunion finie de pavés mesurables peut toujours s'écrire comme réunion finie disjointe. On a par conséquent, en remarquant que  $\Omega = \Omega_1 \times \Omega_2$  est un pavé mesurable, l'énoncé :

#### (2.1.1) Proposition

L'algèbre  $\mathfrak{A}$  engendrée par les pavés mesurables est exactement formée des réunions finies (disjointes) de pavés mesurables. C'est aussi l'algèbre engendrée par l'ensemble des bandes  $A_1 \times \Omega_2$  et  $\Omega_1 \times A_2$ , avec  $A_1 \in \Sigma_1$  et  $A_2 \in \Sigma_2$ .

Il est donc naturel d'introduire la tribu  $\Sigma$  engendrée par l'algèbre  $\mathfrak{A}$ . D'où la définition :

(2.1.2) *Définition*

On appelle tribu produit de  $\Sigma_1$  et  $\Sigma_2$ , et on note  $\Sigma = \Sigma_1 \otimes \Sigma_2$  la tribu sur  $\Omega = \Omega_1 \times \Omega_2$  engendrée par l'algèbre  $\mathfrak{A}$ , ou par l'ensemble des pavés mesurables, ou par l'ensemble des bandes mesurables.

On peut caractériser  $\Sigma$  autrement. En effet :

(2.1.3) *Proposition*

- a)  $\Sigma = \Sigma_1 \otimes \Sigma_2$  est aussi la plus petite tribu rendant mesurables les projections  $\pi_1$  et  $\pi_2$ .
- b) C'est encore la plus petite tribu  $\mathfrak{R}$  rendant mesurables toutes les fonctions  $f_1 \otimes f_2$  où  $f_1 : \Omega_1 \rightarrow \mathbb{R}$  et  $f_2 : \Omega_2 \rightarrow \mathbb{R}$  sont mesurables.

*Preuve.* a) est conséquence du fait que  $\Sigma$  est engendrée par les bandes mesurables, c'est-à-dire par les deux tribus  $\pi_1^{-1}(\Sigma_1)$  et  $\pi_2^{-1}(\Sigma_2)$ . Pour prouver b) soit  $f = f_1 \otimes f_2$ , avec  $f_1$  et  $f_2$  mesurables. Alors  $f = (f_1 \circ \pi_1) \cdot (f_2 \circ \pi_2)$  est  $\Sigma$ -mesurable, d'où  $\mathfrak{R} \subset \Sigma$ . Réciproquement on a  $\Sigma \subset \mathfrak{R}$  car tout pavé mesurable  $A = A_1 \times A_2$  s'écrit  $A = (1_{A_1} \otimes 1_{A_2})^{-1}(\{1\})$ .  $\square$

Les définitions précédentes ayant toutes recours à l'engendrement, on ne connaît pas explicitement les éléments  $A$  de la tribu  $\Sigma$ , ni les fonctions  $\Sigma$ -mesurables  $f : \Omega \rightarrow \mathbb{R}$ . Les conditions nécessaires suivantes sont donc d'une grande importance :

(2.1.4) *Proposition*

Soit  $A \in \Sigma$  et  $f : \Omega \rightarrow \mathbb{R}$  une fonction  $\Sigma$ -mesurable. Alors pour tout  $x_1 \in \Omega_1$  et tout  $x_2 \in \Omega_2$  on a :

- a)  $A(x_1, \bullet) \in \Sigma_2$  et  $A(\bullet, x_2) \in \Sigma_1$ .
- b) La fonction partielle  $f(x_1, \bullet)$  est  $\Sigma_2$ -mesurable et la fonction partielle  $f(\bullet, x_2)$  est  $\Sigma_1$ -mesurable.

*Preuve.* On remarque que la classe des parties  $A \subset \Omega$  vérifiant la propriété a) est une tribu sur  $\Omega$  contenant tous les pavés mesurables, et que, pour tout borélien  $B$  de  $\mathbb{R}$ , on a  $f(x_1, \bullet)^{-1}(B) = f^{-1}(B)(x_1, \bullet)$ .  $\square$

**Cas des tribus boréliennes.** Considérons l'espace  $\mathbb{R}^{p+q}$  comme le produit  $\mathbb{R}^p \times \mathbb{R}^q$ . On voit alors que pour tous boréliens  $B_1 \subset \mathbb{R}^p$  et  $B_2 \subset \mathbb{R}^q$  l'ensemble  $B_1 \times B_2$  est un borélien de  $\mathbb{R}^{p+q}$  puisque les projections canoniques  $\pi_1$  et  $\pi_2$  sont continues. Réciproquement, la tribu borélienne  $\mathfrak{B}_{p+q}$  de  $\mathbb{R}^{p+q}$  est engendrée par les pavés compacts, produits de  $(p+q)$  intervalles compacts, qui sont donc des pavés mesurables très particuliers. D'où le résultat :

(2.1.5) *Proposition*

La tribu borélienne  $\mathfrak{B}_{p+q}$  de  $\mathbb{R}^{p+q}$  est exactement le produit  $\mathfrak{B}_p \otimes \mathfrak{B}_q$  des tribus boréliennes de  $\mathbb{R}^p$  et  $\mathbb{R}^q$ .

Plus généralement on a, en rappelant qu'un espace métrique  $(T, d)$  est dit *séparable* lorsqu'il existe dans  $T$  une partie *dénombrable et partout dense* (ce qui est le cas des espaces  $\mathbb{R}^p$ ):

(2.1.6) *Exercice*

Soit  $(T_1, d_1)$  et  $(T_2, d_2)$  deux espaces métriques séparables,  $T = T_1 \times T_2$  leur produit muni de la distance

$$d(x, y) = d_1(x_1, y_1) + d_2(x_2, y_2).$$

- a) Montrer que  $(T, d)$  est séparable.
- b) En déduire que dans  $T$  tout ouvert est réunion dénombrable de produits  $B_1 \times B_2$  de boules ouvertes dans  $T_1$  et  $T_2$ .
- c) Démontrer enfin que la tribu borélienne  $\mathfrak{B}$  de  $T$  est exactement la tribu produit  $\mathfrak{B}_1 \otimes \mathfrak{B}_2$  des tribus boréliennes de  $T_1$  et  $T_2$ .

## 2.2 Produit de deux mesures $\sigma$ -finies

Plaçons sur  $(\Omega_1, \Sigma_1)$  et  $(\Omega_2, \Sigma_2)$  respectivement des mesures  $\sigma$ -finies  $\mu_1$  et  $\mu_2$ . Avec

(2.1.4.b) on voit donc que pour toute  $f \in \overline{\mathfrak{F}}_+(\Sigma)$ , les fonctions partielles  $f(x_1, \bullet)$  et  $f(\bullet, x_2)$  sont respectivement éléments de  $\overline{\mathfrak{F}}_+(\Sigma_2)$  et  $\overline{\mathfrak{F}}_+(\Sigma_1)$ , de sorte qu'elles peuvent être intégrées. On a alors :

(2.2.1) *Lemme 1*

Soit  $f \in \overline{\mathfrak{F}}_+(\Sigma)$  fixée. Alors la fonction

$$\psi : x_2 \rightarrow \int f(\bullet, x_2) d\mu_1$$

est  $\Sigma_2$ -mesurable et la fonction

$$\varphi : x_1 \rightarrow \int f(x_1, \bullet) d\mu_2$$

est  $\Sigma_1$ -mesurable.

**Preuve.** D'après le théorème d'approximation (1.4.6), il existe une suite  $f_n \in \mathfrak{F}^+(\Sigma)$  telle que  $f_n \uparrow f$ , d'où l'on tire aisément, avec Beppo Lévi, que  $\psi_n \uparrow \psi$ . Il suffit donc de prouver

que  $\psi_n$  est  $\Sigma_2$ -mesurable, ce qui ramène finalement à supposer  $f = 1_B$ , avec  $B \in \Sigma$ , et à prouver que la fonction  $\psi_B$  est élément de  $\overline{\mathfrak{F}}_+(\Sigma_2)$ .

Or supposons d'abord  $\mu_1$  finie, et dans ce cas désignons par  $\mathfrak{L}$  la classe des parties  $B \in \Sigma$  telles que  $\psi_B \in \overline{\mathfrak{F}}_+(\Sigma_2)$ . On a évidemment  $\Omega \in \mathfrak{L}$ , et on vérifie sans difficulté que  $\mathfrak{L}$  est stable par différence propre (ce qui utilise la finitude de  $\mu_1$ ), et par réunion dénombrable croissante (toujours avec Beppo Lévi). Autrement dit  $\mathfrak{L}$  est une  $\lambda$ -classe d'après (1.6.13). Or si  $B = A_1 \times A_2$  avec  $A_1 \in \Sigma_1$ , alors  $\psi_B = \mu_1 A_1 \cdot 1_{A_2}$ , donc  $B \in \mathfrak{L}$ , de sorte que  $\mathfrak{L}$  contient la  $\pi$ -classe formée des pavés mesurables. Par le théorème de Dynkin (1.6.14), on voit alors que  $\mathfrak{L} = \Sigma$ .

Si maintenant  $\mu_1$  est  $\sigma$ -finie, alors il existe une suite  $\Omega_1^n \in \Sigma_1$  telle que  $\Omega_1^n \uparrow \Omega_1$  et

$$\mu_1 \Omega_1^n < +\infty. \text{ En posant, pour } B \in \Sigma, \psi_B^n(x_2) = \int_{\Omega_1^n} 1_B(., x_2) d\mu_1 \text{ on voit que } \psi_B^n \in \overline{\mathfrak{F}}_+(\Sigma_2)$$

d'après le cas  $\mu_1$  finie, et la condition évidente  $\psi_B^n \uparrow \psi_B$  fournit le résultat général. Avec la condition  $\mu_2$   $\sigma$ -finie on obtient de même le résultat sur la fonction  $\varphi$ , ce qui termine tout. □

(2.2.2) *Lemme 2*

La fonctionnelle  $f \rightarrow L(f) = \int \psi d\mu_2$  est une intégrale sur  $(\Omega, \Sigma)$  au sens de (1.5.1.).

*Preuve.* Car  $L(0) = 0$  et, par Beppo Lévi,  $f_n \uparrow f$  implique  $\psi_n \uparrow \psi$  puis  $L(f_n) \uparrow L(f)$ . □

La liaison entre intégrales et mesures étudiée en (1.5) donne alors le résultat fondamental suivant :

(2.2.3) *Théorème*

Soit  $(\Omega_1, \Sigma_1, \mu_1)$  et  $(\Omega_2, \Sigma_2, \mu_2)$  deux espaces mesurés  $\sigma$ -finis. Il existe une unique mesure  $\mu$  sur l'espace mesurable produit  $(\Omega, \Sigma)$  telle que l'on ait

$$(*) \quad \mu(A_1 \times A_2) = \mu_1 A_1 \cdot \mu_2 A_2$$

pour tous  $A_1 \in \Sigma_1$  et  $A_2 \in \Sigma_2$ . Cette mesure  $\mu$  est  $\sigma$ -finie : elle est appelée mesure produit (ou produit direct, ou produit tensoriel) de  $\mu_1$  et  $\mu_2$ , et notée  $\mu = \mu_1 \otimes \mu_2$ .

*Preuve.* Toute mesure  $\mu$  sur  $\Sigma$  vérifiant la condition (\*) est évidemment  $\sigma$ -finie. Alors si  $\mu$  et  $\nu$  vérifient (\*), elles coïncident nécessairement sur l'algèbre  $\mathfrak{A}$  engendrée par les

pavés mesurables d'après (2.1.1.), donc aussi sur la tribu  $\Sigma$  d'après (1.6.16), et ainsi  $\mu = \nu$ , d'où la clause d'unicité.

Pour prouver l'existence, il suffit de considérer la mesure  $\mu$  associée canoniquement par (1.5.3.) à l'intégrale L. Alors

$$\mu(A_1 \times A_2) = L(1_B) = \int \psi_B d\mu_2 \quad \text{pour } B = A_1 \times A_2 .$$

Or  $\psi_B = \mu_1 A_1 \cdot 1_{A_2}$  et  $\int \psi_B d\mu_2 = \mu_1 A_1 \cdot \mu_2 A_2$ , et tout est dit.  $\square$

On en déduit pour le calcul de  $\mu B$ ,  $B \in \Sigma$ , les deux formules de balayage :

(2.2.4) **Corollaire 1**

Soit  $B \in \Sigma$ . Alors

$$\mu B = \int \mu_1 [ B( \cdot , x_2 ) ] d\mu_2 (x_2)$$

$$\mu B = \int \mu_2 [ B(x_1, \cdot ) ] d\mu_1 (x_1)$$

**Preuve.** La première formule est  $\mu B = \int \psi_B d\mu_2$  déjà vue. Pour obtenir la seconde, il suffit de permuter le rôle de  $\mu_1$  et  $\mu_2$ , ce qui fournit une autre mesure  $\tilde{\mu}$  par  $\tilde{\mu} B = \int \psi_B d\mu_1$ . Mais  $\tilde{\mu}$  vérifie aussi (\*) donc la clause d'unicité garantit l'égalité  $\mu = \tilde{\mu}$  et la validité de la seconde formule.  $\square$

(2.2.5) **Corollaire 2**

Si  $\mu_1$  et  $\mu_2$  sont des probabilités alors  $\mu = \mu_1 \otimes \mu_2$  est aussi une probabilité.

(2.2.6) **Corollaire 3**

Pour qu'une partie  $B \in \Sigma$  soit négligeable pour  $\mu = \mu_1 \otimes \mu_2$ , il faut qu'elle satisfasse aux deux conditions :

a) On a  $\mu_2 [ B(x_1, \cdot ) ] = 0$  pour  $\mu_1$ -presque tout  $x_1 \in \Omega_1$ .

b) On a  $\mu_1 [ B( \cdot , x_2 ) ] = 0$  pour  $\mu_2$ -presque tout  $x_2 \in \Omega_2$ .

et il suffit qu'elle vérifie l'une d'entre elles.

(2.2.7) **Corollaire 4**

Les mesures de Borel  $\lambda_n$  sur  $\mathbb{R}^n$  sont telles que  $\lambda_{p+q} = \lambda_p \otimes \lambda_q$  et  $\lambda_n = \lambda_1^{\otimes n}$ .

L'intégration des fonctions  $f$ , définies sur  $\Omega$ ,  $\Sigma$ -mesurables à valeurs positives finies ou non (donc  $f \in \bar{\mathcal{F}}_+(\Sigma)$ ), nous ramène à l'intégrale  $L$  de (2.2.2). En échangeant le rôle de  $\mu_1$  et  $\mu_2$ , on obtient donc le résultat important rattaché aux intégrations successives, ou à la permutation des intégrales :

(2.2.8) **Théorème (Tonelli)**

Pour toute  $f \in \bar{\mathcal{F}}_+(\Sigma)$  on a :  $\int f d(\mu_1 \otimes \mu_2) =$

$$\int \left[ \int f(\cdot, x_2) d\mu_1 \right] d\mu_2(x_2) = \int \left[ \int f(x_1, \cdot) d\mu_2 \right] d\mu_1(x_1)$$

Pour rappeler le calcul par intégrales superposées, on note très souvent  $\int f d\mu$  sous forme d'une intégrale "double" :

$$\int f d\mu = \int \int f d(\mu_1 \otimes \mu_2) = \int \int f(x_1, x_2) d\mu_1(x_1) d\mu_2(x_2)$$

d'où l'apparition des intégrales "multiples" sur les espaces produits finis.

En particulierisant  $f$ , on a aussi

(2.2.9) **Corollaire 1**

Soit  $f_1 : \Omega_1 \rightarrow [0, +\infty]$  et  $f_2 : \Omega_2 \rightarrow [0, +\infty]$  mesurables. Alors  $f = f_1 \otimes f_2$  est  $\Sigma$ -mesurable, à valeurs positives finies ou non, et :

$$\int (f_1 \otimes f_2) d(\mu_1 \otimes \mu_2) = \int f_1 d\mu_1 \cdot \int f_2 d\mu_2$$

(2.2.10) **Corollaire 2**

Soit  $h_1$  et  $h_2$  des fonctions positives ou nulles, à valeurs finies, intégrables respectivement par rapport aux mesures  $\mu_1$  et  $\mu_2$ , auxquelles on associe les mesures à densité  $dv_1 = h_1 d\mu_1$  et  $dv_2 = h_2 d\mu_2$ . Alors la mesure  $\nu = \nu_1 \otimes \nu_2$  admet la densité  $h = h_1 \otimes h_2$  par rapport à la mesure  $\mu = \mu_1 \otimes \mu_2$ .

*Les fonctions intégrables et le théorème de Fubini.*- En revenant à (2.2.9) on voit encore que si  $f_1 : \Omega_1 \rightarrow \mathbb{R}$  est supposée intégrable par rapport à  $\mu_1$ , alors la finitude des

intégrales, et la décomposition  $f_1 = f_1^+ - f_1^-$ , jointes à la condition  $f_2 \geq 0$  et  $\int f_2 d\mu_2 < +\infty$ ,

garantissent la validité de la formule du corollaire 1. On étend ensuite au cas où  $f_2 : \Omega_2 \rightarrow \mathbb{R}$  est intégrable par rapport à  $\mu_2$ , pour obtenir

(2.2.11) *Proposition*

Pour toutes fonctions  $f_1 : \Omega_1 \rightarrow \mathbb{R}$  et  $f_2 : \Omega_2 \rightarrow \mathbb{R}$ , intégrables respectivement par rapport à  $\mu_1$  et  $\mu_2$ , la fonction  $f_1 \otimes f_2$  est intégrable par rapport à  $\mu = \mu_1 \otimes \mu_2$  et

$$\int (f_1 \otimes f_2) d(\mu_1 \otimes \mu_2) = \int f_1 d\mu_1 \cdot \int f_2 d\mu_2$$

Examinons l'autre aspect du problème en partant d'une fonction  $f : \Omega \rightarrow \mathbb{R}$ , supposée intégrable par rapport à  $\mu$ , et posons

$$\tilde{\psi}(x_2) = \int |f(\cdot, x_2)| d\mu_1 \leq +\infty.$$

Avec le théorème de Tonelli, on a donc  $\int \tilde{\psi}(x_2) d\mu_2(x_2) = \int |f| d\mu < +\infty$  ce qui

assure que la fonction  $\tilde{\psi}$  est  $\mu_2$ -presque partout finie, de sorte que l'ensemble  $N_2 = \{x_2, \tilde{\psi}(x_2) = +\infty\}$  est élément de  $\Sigma_2$  et négligeable. De même, si

$$\tilde{\varphi}(x_1) = \int |f(x_1, \cdot)| d\mu_2 \leq +\infty,$$

l'ensemble  $N_1 = \{x_1, \tilde{\varphi}(x_1) = +\infty\}$  est élément de  $\Sigma_1$  et  $\mu_1 N_1 = 0$ , et ainsi l'ensemble  $M = N_1^c \times N_2^c \in \Sigma$  est tel que  $\mu M^c = 0$ , puisque  $M^c = (N_1 \times \Omega_2) \cup (\Omega_1 \times N_2)$ . Remplaçant

$f$  par  $g = 1_M f$ , on obtient alors une fonction intégrable par rapport à  $\mu$ , telle que les fonctions partielles  $g(\cdot, x_2)$  et  $g(x_1, \cdot)$  soient respectivement intégrables par rapport à  $\mu_1$  et  $\mu_2$ , pour tous  $x_1$  et  $x_2$ . La finitude des intégrales garantit donc, avec Tonelli, et la

décomposition  $f = f^+ - f^-$ , ou  $g = g^+ - g^-$ , que  $\int g d\mu = \int \psi_g d\mu_2 = \int \varphi_g d\mu_1$  avec

$$\psi_g(x_2) = \int g(\cdot, x_2) d\mu_1 \text{ et } \varphi_g(x_1) = \int g(x_1, \cdot) d\mu_2.$$

Or  $\int g \, d\mu = \int f \, d\mu$  et pour  $x_2 \notin N_2$ ,  $\psi_g(x_2) = \int f(\cdot, x_2) \, d\mu_1$ , de même que

$\varphi_g(x_1) = \int f(x_1, \cdot) \, d\mu_2$  pour tout  $x_1 \notin N_1$ . On aboutit à ceci que les fonctions

$$\varphi_f(x_1) = \int f(x_1, \cdot) \, d\mu_2 \quad x_1 \in N_1^c$$

$$\psi_f(x_2) = \int f(\cdot, x_2) \, d\mu_1 \quad x_2 \in N_2^c$$

qui ne sont que *presque partout définies*, sont en réalité presque partout égales à des fonctions intégrables, que l'on appelle *versions intégrables* de  $\varphi_f$  et  $\psi_f$ .

Convenons alors de *définir*  $\int \varphi_f \, d\mu_1$  en remplaçant  $\varphi_f$  par l'une ou l'autre de ses

versions intégrables, et de même pour  $\int \psi_f \, d\mu_2$ , de sorte que l'on obtient le résultat définitif.

(2.2.12) *Théorème (Fubini)*

Soit  $f : \Omega \rightarrow \mathbb{R}$  une fonction intégrable pour  $\mu = \mu_1 \otimes \mu_2$ . Alors :

a) La fonction  $f(x_1, \cdot)$  est intégrable par rapport à  $\mu_2$  pour  $\mu_1$ -presque tout  $x_1$  et la fonction  $f(\cdot, x_2)$  est intégrable par rapport à  $\mu_1$  pour  $\mu_2$ -presque tout  $x_2$ .

b) Les fonctions

$$\varphi_f : x_1 \rightarrow \int f(x_1, \cdot) \, d\mu_2 \quad \text{et} \quad \psi_f : x_2 \rightarrow \int f(\cdot, x_2) \, d\mu_1$$

sont définies presque partout respectivement sur  $\Omega_1$  et  $\Omega_2$ .

c) On a de plus la permutation des intégrales

$$\begin{aligned} \int f \, d(\mu_1 \otimes \mu_2) &= \int \left[ \int f(x_1, \cdot) \, d\mu_2 \right] d\mu_1(x_1) \\ &= \int \left[ \int f(\cdot, x_2) \, d\mu_1 \right] d\mu_2(x_2) \end{aligned}$$

**Remarque.** Les formules de (2.2.11) et (2.2.12), établies pour  $f$  réelle, sont évidemment valables aussi pour  $f$  complexe, et la formule c) du théorème de Fubini est fréquemment utilisée pour calculer des intégrales (ou des séries) successives par permutation des variables. On prendra garde toutefois qu'il peut arriver, dans des cas très simples, qu'une fonction  $f$  soit  $\Sigma$ -mesurable, telle que les fonctions  $f(x_1, \cdot)$  et  $f(\cdot, x_2)$  soient

toujours intégrables, telle encore que les fonctions  $\varphi : x_1 \rightarrow \int f(x_1, \cdot) d\mu_2$  et

$\psi : x_2 \rightarrow \int f(\cdot, x_2) d\mu_1$  soient elles aussi intégrables, mais que l'on ait

$$I = \int \varphi d\mu_1 \neq \int \psi d\mu_2 = J$$

C'est par exemple le cas, lorsque  $\mu_1$  et  $\mu_2$  sont toutes deux égales à la mesure de Borel sur  $[0, 1]$ , avec la fonction  $f(x, y) = \frac{x^2 - y^2}{(x^2 + y^2)^2}$  pour laquelle  $I = \frac{\pi}{4}$  et  $J = -\frac{\pi}{4}$ . Bien entendu  $f$  n'est pas intégrable sur le carré  $[0, 1] \times [0, 1]$ .

Ce que disent donc le théorème de Tonelli et le théorème de Fubini est que cela ne peut pas se produire si  $f$  est positive ou bien si  $f$  vérifie les conditions du corollaire suivant, qui explicite le critère d'application du théorème de Fubini.

(2.2.13) **Corollaire (Critère de Fubini)**

Soit  $f$  une fonction définie sur  $\Omega = \Omega_1 \times \Omega_2$ , supposée  $\Sigma$ -mesurable. Si l'on a

$$\int d\mu_1(x_1) \int |f(x_1, \cdot)| d\mu_2 < +\infty$$

ou bien  $\int d\mu_2(x_2) \int |f(\cdot, x_2)| d\mu_1 < +\infty$

alors  $f$  est intégrable par rapport à  $\mu = \mu_1 \otimes \mu_2$  et les intégrales du théorème de Fubini sont finies et égales à  $\int f d\mu$ .

En pratique, on commencera par majorer  $|f(x_1, x_2)|$ , puis on testera le critère précédent en choisissant la plus accessible des deux formules à vérifier.

Donnons pour terminer quelques exemples sous forme d'exercices.

(2.2.14) *Exercice 1*

a) Calculer l'intégrale  $J(a) = \int_0^{\infty} \frac{dx}{(a^2 + x^2)^\alpha}$  pour  $a > 0$  et  $\alpha > \frac{1}{2}$ .

b) En déduire que l'intégrale  $I = \int_0^{\infty} \int_0^{\infty} \frac{dx dy}{(1 + x^2 + y^2)^\alpha}$  n'est finie que si  $\alpha > 1$ , et qu'alors  $I = \frac{\pi}{4} \frac{1}{\alpha - 1}$ .

(2.2.15) *Exercice 2*

a) Établir que pour  $0 < \alpha < 2$ , les deux intégrales

$$\int_0^{\infty} \frac{1 - \cos x}{x^{\alpha+1}} dx \quad \text{et} \quad \frac{1}{\alpha} \int_0^{\infty} \frac{\sin x}{x^\alpha} dx$$

existent et sont égales. On désignera par  $J(\alpha)$  leur valeur commune.

b) A toute probabilité  $\mu$  sur  $\mathbb{R}$  on associe la fonction

$$F(x) = \int_{\mathbb{R}} \cos tx \, d\mu(t) \quad x \in \mathbb{R}$$

Démontrer que  $F$  est continue et bornée, puis établir la formule, valable pour  $0 < \alpha < 2$ ,

$$(*) \quad \int_0^{\infty} \frac{1 - F(x)}{x^{\alpha+1}} dx = J(\alpha) \int_{\mathbb{R}} |t|^\alpha d\mu(t) \leq +\infty$$

c) On choisit pour  $\mu$  la probabilité d'Euler, de densité  $e^{-t}$  sur  $[0, +\infty)$ . Déterminer la fonction  $F$  associée, puis en utilisant (1.10.12) prouver l'égalité

$$J(\alpha) = \frac{\pi}{2} \frac{1}{\Gamma(\alpha + 1)} \frac{1}{\sin \frac{\pi\alpha}{2}}$$

d) On définit les intégrales  $C(\alpha) = \int_0^{\infty} \frac{\cos x}{x^\alpha} dx$  pour  $0 < \alpha < 1$  et

$$S(\alpha) = \int_0^{\infty} \frac{\sin x}{x^\alpha} dx \quad \text{pour } 0 < \alpha < 2. \text{ Prouver que}$$

$$C(\alpha) = \frac{\pi}{2} \frac{1}{\Gamma(\alpha)} \frac{1}{\cos \frac{\pi\alpha}{2}} \quad 0 < \alpha < 1$$

$$S(\alpha) = \frac{\pi}{2} \frac{1}{\Gamma(\alpha)} \frac{1}{\sin \frac{\pi\alpha}{2}} \quad 0 < \alpha < 2$$

e) Montrer que les intégrales de Fresnel  $I = \int_0^{\infty} \cos(x^2) dx$  et  $J = \int_0^{\infty} \sin(x^2) dx$

sont toutes deux égales à  $\sqrt{\frac{\pi}{8}}$ .

f) Comparer avec l'exercice (1.10.15).

(2.2.16) *Exercice 3*

Pour  $\alpha > 0$ , on pose :  $f_\alpha(t) = \frac{t^{\alpha-1}}{\Gamma(\alpha)} e^{-t}$  pour  $t > 0$

et  $\varphi_\alpha(x) = \int_0^{\infty} e^{itx} f_\alpha(t) dt$  pour  $x \in \mathbb{R}$ .

a) Prouver que  $\int_0^t f_\alpha(t-s) f_\beta(s) ds = f_{\alpha+\beta}(t)$  pour  $t > 0$ .

b) En déduire que  $\varphi_{\alpha+\beta} = \varphi_\alpha \varphi_\beta$ .

c) Montrer que la fonction  $\varphi_\alpha$  est dérivable et solution de l'équation différentielle  $y' = \frac{i \alpha y}{1 - ix}$ . En déduire que

$$\varphi_\alpha(x) = (1+x^2)^{-\frac{\alpha}{2}} \exp[i \alpha \operatorname{Arc} \operatorname{tg} x]$$

et retrouver ainsi b).

(2.2.17) *Exercice 4*

Dans l'espace  $\mathbb{R}^2$  on considère le quadrant  $P = \{x \geq 0, y \geq 0\}$  et pour  $a > 0$  et  $b > 0$ , on considère la mesure  $\mu$  sur  $P$  de densité  $x^{a-1} y^{b-1}$  par rapport à la mesure de Borel, donc  $d\mu = x^{a-1} y^{b-1} dx dy$ . A chaque fonction borélienne positive  $f$ , définie sur  $[0, +\infty)$  on associe l'intégrale finie ou non

$$D(a, b) = \int \int_P x^{a-1} y^{b-1} f(x+y) dx dy$$

- a) Montrer que  $D(a, b) = \int_0^{\infty} f(t) d\nu(t)$ , où  $\nu$  est la mesure image sur  $[0, +\infty)$  de la mesure  $\mu$  par l'application  $s(x, y) = x + y$ .
- b) Déterminer la mesure  $\nu$  en calculant sa fonction de répartition  $\nu[0, t[$ .
- c) En déduire que  $D(a, b) = B(a, b) \int_0^{\infty} t^{a+b-1} f(t) dt$  où  $B(a, b)$  est la fonction bêta d'Euler.
- d) Quel résultat obtient-on en choisissant  $f(t) = e^{-t}$  ?
- e) On remplace maintenant l'espace  $\mathbb{R}^2$  par l'espace  $\mathbb{R}^n$ ,  $n \geq 2$ , et le quadrant  $P$  par le cône positif  $P_n = (\mathbb{R}_+)^n$ . Déduire de c), avec les mêmes hypothèses sur  $f$ , l'égalité

$$\int_{P_n} x_1^{a_1-1} x_2^{a_2-1} \dots x_n^{a_n-1} f(x_1 + \dots + x_n) dx_1 \dots dx_n =$$

$$\frac{\Gamma(a_1) \dots \Gamma(a_n)}{\Gamma(a_1 + \dots + a_n)} \int_0^{\infty} t^{a_1 + \dots + a_n - 1} f(t) dt$$

pour tous nombres  $a_1 > 0, a_2 > 0, \dots, a_n > 0$ .

- f) Soit  $S_n$  le simplexe unité de  $\mathbb{R}^n$ ,  $S_n = P_n \cap \{x_1 + \dots + x_n \leq 1\}$ . Montrer que son volume est donné par  $W_n = \lambda_n S_n = \frac{1}{n!}$ .

(2.2.18) *Exercice 5*

Dans l'espace  $T = \mathbb{R}^n \times [0, +\infty)$ , de coordonnées  $(x, r)$ , on place la mesure produit  $dx dr$ , et soit  $P \subset T$  le cône défini par

$$P = \{(x, r) ; \|x\| \leq r\}$$

où  $\|x\|$  est la norme euclidienne.

- a) En calculant de deux façons différentes l'intégrale

$$I = \int \int_P r \exp\left(-\frac{r^2}{2}\right) dx dr$$

retrouver le volume  $V_n$  de la boule euclidienne de  $\mathbb{R}^n$ .

- b) On remplace la norme euclidienne sur  $\mathbb{R}^n$  par la fonctionnelle  $\|x\|_\alpha = (|x_1|^\alpha + \dots + |x_n|^\alpha)^{\frac{1}{\alpha}}$ , avec  $\alpha > 0$ , et la boule euclidienne par l'ensemble  $B_\alpha = \{\|x\|_\alpha \leq 1\}$ . En remplaçant  $P$  par le cône  $P_\alpha$  correspondant, et l'intégrale  $I$  par l'intégrale

$$I_\alpha = \int \int_{P_\alpha} r^{\alpha-1} \exp\left(-\frac{r^\alpha}{2}\right) dx dr$$

établir que le volume  $V_n^{(\alpha)}$  de  $B_\alpha$  est donné par la formule générale

$$V_n^{(\alpha)} = \left(\frac{2}{\alpha}\right)^n \frac{\Gamma\left(\frac{1}{\alpha}\right)^n}{\Gamma\left(\frac{n}{\alpha} + 1\right)}$$

- c) Établir que le cas  $\alpha = 1$  permet de retrouver le volume  $W_n = \lambda_n S_n$  du simplexe calculé en (2.2.17.f).

### 2.3 La formule du changement de variables

Nous nous limiterons ici à la mesure de Borel  $\lambda_n$  sur  $\mathbb{R}^n$ , pour prouver correctement (ce qui est délicat) la formule classique du changement de variables dans une intégrale multiple, et d'en donner quelques applications. Étudions tout d'abord l'influence d'une transformation affine.

*Translations.* La formule  $\lambda_n(P) = \prod (b_k - a_k)$  donnant le volume d'un pavé  $P = \prod [a_k, b_k]$  montre immédiatement que pour toute translation de vecteur  $u \in \mathbb{R}^n$ , le volume  $\lambda_n(P + u)$  du pavé translaté de  $P$  reste égal à  $\lambda_n(P)$ . Il s'ensuit que pour tout borélien  $B$  de  $\mathbb{R}^n$  on a  $\lambda_n(B + u) = \lambda_n(B)$ .

On a d'ailleurs là une caractérisation de la mesure de Borel sur  $\mathbb{R}^n$ . Car :

#### (2.3.1) Proposition

Soit  $\mu$  une mesure borélienne (localement finie) sur  $\mathbb{R}^n$ , invariante par les translations de  $\mathbb{R}^n$ . Alors  $\mu = k \lambda_n$  est proportionnelle à la mesure de Borel et la constante  $k$  est la mesure  $\mu(I^n)$  du cube unité  $I^n = [0, 1]^n$  de  $\mathbb{R}^n$ .

*Preuve.* Introduisons l'anneau  $\mathfrak{R}$  des parties qui sont réunions finies (disjointes) de pavés  $P = \prod [a_k, b_k]$ , où les  $a_k$  et les  $b_k$  sont *rationnels*, anneau qui engendre la tribu borélienne. Soit  $J = [0, 1]^n$  et  $k = \mu(J)$ . Remarquons ensuite qu'il existe un entier  $q \geq 1$  et des entiers  $p_k \geq 1$  tels que  $b_k - a_k = \frac{p_k}{q}$ , de sorte qu'on peut diviser  $P$  en une réunion disjointe de  $N = p_1 p_2 \dots p_n$  petits pavés cubiques de côtés  $\frac{1}{q}$ , qui sont tous de même mesure  $\alpha_q$  pour  $\mu$ . Et de même, on peut diviser  $J$  en une réunion disjointe de  $q^n$  petits pavés cubiques de même mesure  $\alpha_q$  pour  $\mu$ . Il suit de là que  $k = \mu(J) = q^n \alpha_q$ , et que

$\mu(P) = N \alpha_q = k \cdot \frac{P_1 P_2 \cdots P_n}{q^n} = k \lambda_n(P)$ , d'où l'on tire  $\mu = k \lambda_n$  avec (1.6.16.b), de sorte que  $k = \mu(I) = \mu(I^n)$ . □

**Transformations linéaires.** Soit  $T$  une transformation linéaire de  $\mathbb{R}^n$  dans lui-même, supposée bijective, donc de déterminant  $\det T$  non nul. Dans ce cas  $T = S^{-1}$ , avec  $S = T^{-1}$ , de sorte que  $T$  transforme tout borélien  $B$  en un borélien  $T(B) = S^{-1}(B)$ , et que l'application  $B \rightarrow \lambda_n [T(B)]$  est une mesure borélienne  $\mu = S(\lambda_n)$ , image linéaire de la mesure de Borel.

Pour déterminer  $\mu$  commençons par remarquer que  $\mu$  est invariante par translation, puisque

$$\mu(B + u) = \lambda_n [T(B) + T(u)] = \lambda_n [T(B)] = \mu(B)$$

de sorte qu'il existe une constante  $k_T$  telle que  $\mu = k_T \lambda_n$ .

Et ainsi

$$(*) \quad \lambda_n [T(B)] = k_T \lambda_n(B) \text{ pour tout borélien } B.$$

Pour obtenir la valeur de  $k_T$  remarquons encore que si  $T = T_1 T_2$  alors  $k_T = k_{T_1} k_{T_2}$  de manière évidente à partir de la formule précédente. Autrement dit, la fonction  $T \rightarrow k_T$  est une fonctionnelle multiplicative sur le groupe  $GL(n, \mathbb{R})$  des transformations linéaires bijectives de  $\mathbb{R}^n$ . Ce qui implique en particulier que  $k_I = 1$ , où  $I$  est l'opérateur identité, et que  $k_{T^{-1}} = (k_T)^{-1}$ . Alors :

(2.3.2) **Théorème**

Pour toute transformation  $T \in GL(n, \mathbb{R})$  on a  $k_T = |\det T|$ . Autrement dit pour tout borélien  $B$  de  $\mathbb{R}^n$ , on a

$$\lambda_n [T(B)] = |\det T| \lambda_n(B)$$

**Preuve.** La multiplicativité de  $k_T$  et celle de  $|\det T|$  garantissent que si le résultat est vrai pour  $T_1$  et pour  $T_2$ , il est vrai aussi pour  $T = T_1 T_2$ . Or

a) *Il est vrai pour  $T$  diagonale*, car si  $(e_1, e_2, \dots, e_n)$  est la base canonique de  $\mathbb{R}^n$  et si  $T e_k = \lambda_k e_k$ , avec  $\lambda_k \neq 0$ , alors l'image  $T(I^n)$  est le paralléloèpe  $P = \prod [0, \lambda_k]$ , d'où  $k_T = \prod |\lambda_k| = |\det T|$ .

b) *Il est vrai pour  $T$  hermitienne* (ou symétrique), car si  $T^*$  est l'adjointe de  $T$  (ou la transposée) et si  $T = T^*$ , on sait alors que  $T$  est diagonalisable par une matrice orthogonale et que ses valeurs propres sont réelles, d'où  $T = U^{-1} \Delta U$  et  $k_T = k_\Delta = |\det \Delta| = |\det T|$ .

c) *Il est vrai pour  $T$  orthogonale*, car si  $T = U$  est une isométrie de  $\mathbb{R}^n$ , nécessairement surjective, alors  $k_T = 1$  comme on voit avec la formule (\*) appliquée en prenant pour  $B$  la boule unité euclidienne  $B_n$  de  $\mathbb{R}^n$ . Mais  $U U^{-1} = I = U U^*$  donne aussi  $1 = (\det U)^2$ , et  $|\det U| = 1$ .

d) Il est vrai pour  $T$  quelconque, car la transformation  $T^*T$  est hermitienne, bijective, et ses valeurs propres sont strictement positives. On a donc  $T^*T = V^{-1} \Delta V$ , où  $V$  est orthogonale et où la transformation diagonale  $\Delta$  admet de façon évidente une racine carrée  $\Delta^{1/2}$ . Posons  $|T| = V^{-1} \Delta^{1/2} V$ , ce qui donne un opérateur hermitien tel que  $T^*T = |T|^2$ , et remarquons que

$$\begin{aligned} \| |T|x \|^2 &= \langle |T|x, |T|x \rangle = \langle |T|^2 x, x \rangle = \langle T^*T x, x \rangle \\ &= \langle T x, T x \rangle = \|T x\|^2 \end{aligned}$$

pour tout  $x \in \mathbb{R}^n$ . La bijectivité de  $|T|$  garantit donc que l'opérateur  $U = T|T|^{-1}$  est une isométrie, puisque si  $y \in \mathbb{R}^n$  et si  $x = |T|^{-1}y$  alors  $Uy = Tx$  et  $\|Uy\| = \|y\|$ . Or  $T = U|T|$ , et le résultat est conséquence de b) et c).  $\square$

### (2.3.3) Corollaire 1

La mesure de Borel sur  $\mathbb{R}^n$  est invariante par les translations, les symétries et les isométries de  $\mathbb{R}^n$  (donc par les déplacements et les anti-déplacements). Si  $H$  est une homothétie de rapport  $k \neq 0$ , alors

$$\lambda_n [H(B)] = |k|^n \lambda_n (B)$$

pour tout borélien  $B$ .

### (2.3.4) Corollaire 2

Dans  $\mathbb{R}^n$ , tout hyperplan, et plus généralement toute variété affine de dimension  $p \leq n - 1$ , est négligeable pour la mesure de Borel.

### (2.3.5) Corollaire 3

Soit  $(x_1, x_2, \dots, x_n)$  un système de  $n$  vecteurs de  $\mathbb{R}^n$ . Alors le volume du parallélotope (oblique) construit sur ces vecteurs est égal à la valeur absolue de leur déterminant.

**Preuve.** Si le système est lié, alors les  $x_k$  sont éléments d'un même hyperplan, qui est de mesure nulle d'après le corollaire 2, et le déterminant est nul aussi. Si le système est libre, on définit la transformation linéaire bijective  $T$  par  $Te_k = x_k$ . La matrice  $M_T$  représentant  $T$  sur la base canonique  $(e_1, \dots, e_n)$  admet donc les  $x_k$  comme vecteurs colonnes, d'où  $\det T = \det(x_1, x_2, \dots, x_n)$ . Mais  $k_T = \lambda_n [T(I^n)]$  et on reconnaît en  $T(I^n)$  le parallélotope construit sur les  $x_k$ .  $\square$

### (2.3.6) Corollaire 4

Soit  $T$  une transformation linéaire bijective de  $\mathbb{R}^n$ . Alors :

- a)  $T$  échange les parties boréliennes entre elles, et les parties négligeables entre elles.

b) Pour toute fonction  $f \geq 0$  borélienne on a

$$\int f(y) dy = |\det T| \int f(Tx) dx \leq +\infty$$

c) Pour toute fonction numérique (réelle ou complexe) borélienne, la condition que  $f$  soit intégrable est identique à la condition que  $f \circ T$  soit intégrable et dans ce cas on a la formule du changement (linéaire) de variables

$$\int f(y) dy = |\det T| \int f(Tx) dx$$

**Preuve.** Il suffit de vérifier b), car c) s'obtient en remplaçant  $f$  par  $|f|$ , et en décomposant  $f$  en  $f = f^+ - f^-$  dans le cas réel. Or b) est vraie pour  $f = 1_B$ , avec  $B$  borélien, car en écrivant  $B = T(A)$ , avec  $A = T^{-1}(B)$ , on a, avec (2.3.2) :

$$\int 1_B(y) dy = \lambda_n B = \lambda_n [T(A)] = |\det T| \lambda_n(A)$$

et  $1_B(Tx) = 1_A(x)$ . On passe ensuite au cas des fonctions étagées puis au cas général par Beppo Lévi et le théorème d'approximation (1.4.6).  $\square$

**Remarque 1.** On interprète souvent le changement de variables  $y = Tx$  en disant que l'élément de volume  $dx$  se transforme en  $dy = |\det T| dx$ .

**Remarque 2.** On peut aussi traduire la formule du changement de variables en termes de mesure image. En effet la mesure image  $T(\lambda_n)$  vaut  $\frac{1}{|\det T|} \lambda_n$  et la mesure image  $T^{-1}(\lambda_n)$  vaut  $|\det T| \lambda_n$ .

**Remarque 3.** On prendra garde que lorsque  $T$  n'est plus supposée bijective, alors l'image directe  $T(B)$  d'un borélien peut ne pas être borélienne. On aura toutefois  $\lambda_n^*(T(B)) = 0$  puisque  $T(B)$  est nécessairement contenue dans un hyperplan.

### (2.3.7) Exemples

**Exemple 1.-** Reprendre (1.8.16) et (1.8.17) et revenir à l'intégration des fonctions sphériques sur  $\mathbb{R}^n$  en utilisant (2.3.3).

**Exemple 2 (Fonctions B et  $\Gamma$  d'Euler).-** On retrouve facilement la formule  $B(x, y) = \frac{\Gamma(x) \Gamma(y)}{\Gamma(x+y)}$ , en utilisant le théorème de Tonelli et la formule du changement de variables. En effet,

$$\begin{aligned}\Gamma(x) \Gamma(y) &= \int_0^{\infty} \int_0^{\infty} s^{x-1} t^{y-1} e^{-(s+t)} ds dt \\ &= \int \int_{0 \leq u \leq v} u^{x-1} (v-u)^{y-1} e^{-v} du dv\end{aligned}$$

avec  $s = u$ ,  $t = v - u$ ,  $ds dt = du dv$ . Puis

$$\begin{aligned}\Gamma(x) \Gamma(y) &= \int_0^{\infty} e^{-v} dv \left[ \int_0^v u^{x-1} (v-u)^{y-1} du \right] \\ &= \int_0^{\infty} e^{-v} dv \left[ v^{x+y-1} \int_0^1 w^{x-1} (1-w)^{y-1} dw \right]\end{aligned}$$

avec  $u = v w$ ,  $v$  fixé. D'où  $\Gamma(x) \Gamma(y) = B(x, y) \Gamma(x + y)$ .

**Exemple 3 (Volume du simplexe unité).** - Étant donné  $n$  vecteurs  $x_1, x_2, \dots, x_n$  de  $\mathbb{R}^n$  on appelle simplexe construit sur ces vecteurs l'ensemble  $S(x_1, \dots, x_n)$  des vecteurs  $\sum \lambda_k x_k$  tels que  $\lambda_k \geq 0$  et  $\sum \lambda_k \leq 1$ . On appelle simple unité de  $\mathbb{R}^n$  le simplexe  $S_n$  construit sur la base canonique  $(e_1, \dots, e_n)$ , soit encore

$$S_n = \{ x = (\xi_k); \xi_k \geq 0 \text{ et } \sum \xi_k \leq 1 \}$$

Pour calculer le volume  $\lambda_n(S_n) = W_n$ , on peut facilement opérer par récurrence sur  $n$ . On peut aussi chercher à évaluer des intégrales du type

$$I = \int_0^{\infty} \dots \int_0^{\infty} f(t_1 + \dots + t_n) dt_1 \dots dt_n$$

en introduisant la mesure image  $\nu = s(\lambda_n)$ , où  $s$  est la fonction  $s(x) = \xi_1 + \dots + \xi_n$  si  $x = (\xi_k)$ . On détermine alors  $\nu$  par sa fonction de répartition  $F(t) = \nu[0, t[$ , soit

$$\begin{aligned}F(t) &= \lambda_n \{ x = (\xi_k); \xi_k \geq 0 \text{ et } \sum \xi_k < t \} \\ &= t^n W_n\end{aligned}$$

puisque l'homothétie de rapport  $t$  transforme précisément le simplexe  $S_n$  en l'ensemble  $s^{-1}[0, t[$ , de même mesure que  $s^{-1}[0, t[$ . Alors  $\nu$  est la mesure sur  $[0, +\infty)$ , de densité

$$n t^{n-1} W_n, \text{ d'où } I = n W_n \int_0^{\infty} f(t) t^{n-1} dt.$$

En choisissant  $f(t) = e^{-t}$ , on obtient d'un côté  $I = 1$  et de l'autre  $I = n W_n \Gamma(n) = n! W_n$ . D'où :

- a) Le volume  $W_n$  du simplexe unité  $S_n$  est  $W_n = \frac{1}{n!}$ .
- b) Le volume du simplexe construit sur les vecteurs  $x_1, \dots, x_n$  de  $\mathbb{R}^n$  est égal à  $\frac{1}{n!} | \det(x_1, \dots, x_n) |$ .
- c) Pour toute fonction borélienne  $f \geq 0$  sur  $[0, +\infty)$  on a

$$\int_0^\infty \dots \int_0^\infty f(t_1 + \dots + t_n) dt_1 \dots dt_n = \frac{1}{(n-1)!} \int_0^\infty t^{n-1} f(t) dt$$

(2.3.8) **Exercice (Volume et déterminant de Gram)**

On fixe  $n$  vecteurs  $x_1, \dots, x_n$  dans l'espace  $\mathbb{R}^n$ , formant un système libre, et pour  $1 \leq p \leq n$  on introduit le déterminant de Gram

$$G(x_1, \dots, x_p) = \det [ \langle x_i, x_j \rangle ]_{1 \leq i, j \leq p}$$

- a) Prouver que le volume  $V_n(P)$  du paralléloèdre  $P$  construit sur  $x_1, \dots, x_n$  est donné par

$$V_n(P)^2 = G(x_1, \dots, x_n).$$

- b) On introduit la projection  $y_n$  de  $x_n$  sur l'espace vectoriel engendré par  $x_1, x_2, \dots, x_{n-1}$ , et soit  $z_n = x_n - y_n$ . Démontrer que

$$V_n(P)^2 = \|z_n\|^2 G(x_1, \dots, x_{n-1})$$

et en déduire que le volume  $V_n(P)$  est égal au produit du volume  $(n-1)$ -dimensionnel de l'une de ses bases par la hauteur correspondante (analogie avec le parallélogramme).

Déduire de là que  $V_n(P)$  est majoré par le produit des longueurs des vecteurs  $x_k$  et que l'on n'a

$$V_n(P) = \|x_1\| \|x_2\| \dots \|x_n\|$$

que si le système  $(x_k)$  est orthogonal.

- c) Soit  $M = [a_{ij}]$  une matrice carrée d'ordre  $n$ . Établir l'inégalité de Hadamard

$$| \det M | \leq \prod_{k=1}^n [ a_{k1}^2 + \dots + a_{kn}^2 ]^{1/2}$$

**La formule générale du changement de variables.** On fixe maintenant un ouvert  $U$  de  $\mathbb{R}^n$  et une application  $H : U \rightarrow \mathbb{R}^n$ , supposée être un difféomorphisme, ce qui signifie en détail :

- l'image  $H(U) = V$  est un ouvert de  $\mathbb{R}^n$ ,
- $H$  est un homéomorphisme de  $U$  sur  $V$ ,

- $H$  est de classe  $C^1$ , c'est-à-dire différentiable en tout point  $x \in U$ , l'application dérivée  $x \rightarrow DH_x$  de  $U$  dans  $L(\mathbb{R}^n)$  étant continue,
- enfin  $DH_x$  est, pour tout  $x \in U$ , une application linéaire inversible de  $\mathbb{R}^n$  sur lui-même, ce qui signifie encore que le jacobien  $J(x) = \det DH_x$  ne s'annule pas sur  $U$ .

Dans ces conditions on sait que l'application réciproque  $K : V \rightarrow U$  admet en tout point  $y \in V$ ,  $y = H(x)$ , une différentielle  $DK_y = (DH_x)^{-1}$ , qui est fonction continue de  $y$ . Ainsi  $K = H^{-1}$  est aussi un difféomorphisme tel que son jacobien  $J_K(y)$  est donné par  $J_K(y) = \frac{1}{J(x)}$  pour  $y = H(x)$ .

Alors  $H$ , étant un homéomorphisme, échange les parties boréliennes de  $U$  et  $V$ , qui sont d'ailleurs les parties boréliennes de  $\mathbb{R}^n$  contenues respectivement dans  $U$  et dans  $V$ . Le théorème qui suit va permettre de calculer la mesure de l'image  $H(A)$  pour tout borélien  $A$  de  $U$  selon une formule qui doit généraliser celle de (2.3.2). Mais de fait, en généralisant à la fois (2.3.2) et (2.3.6), on a :

(2.3.9) **Théorème (du changement de variables)**

Soit  $H : U \rightarrow V$  un difféomorphisme entre deux ouverts  $U$  et  $V$  de  $\mathbb{R}^n$ . On désigne par  $J(x) = \det DH_x$  le déterminant jacobien de  $H$  au point  $x \in U$ . On a :

- a) Pour tout borélien  $A$  de  $U$ , l'image  $H(A)$  est un borélien de  $V$  et

$$\lambda_n [H(A)] = \int_A |J(x)| dx$$

- b) Pour toute fonction  $f \geq 0$  borélienne sur  $V$  on a

$$\int_V f(y) dy = \int_U f [H(x)] |J(x)| dx$$

et plus généralement, pour tout borélien  $B$  de  $V$ ,

$$\int_B f(y) dy = \int_{H^{-1}(B)} f [H(x)] |J(x)| dx$$

- c) Pour toute fonction numérique (réelle ou complexe) borélienne sur  $V$ , la condition que  $f$  soit intégrable sur  $V$  est identique à la condition que la fonction  $|J(x)| f [H(x)]$  soit intégrable sur  $U$ , et dans ce cas on a la formule du changement de variables

$$\int_V f(y) dy = \int_U f [H(x)] |J(x)| dx$$

exprimant que dans la transformation  $y = H(x)$  l'élément de volume  $dy$  se calcule selon  $dy = |J(x)| dx$ .

Ce théorème n'est pas trivial et la preuve en est longue. Avant de l'aborder, remarquons que si  $H = T$  est linéaire, alors  $DH_x = T$  ne dépend pas de  $x$  et  $|J(x)| = |\det T|$ , ce qui montre bien en quoi on généralise (2.3.6).

Réciproquement, en posant  $T_x = DH_x$  pour simplifier, on voit que  $T_x$  est l'application linéaire tangente associée à  $H$  au point  $x$ , permettant, localement autour du point  $x$ , d'approcher l'application  $H$ . En ce sens l'intervention du facteur  $|J(x)| = |\det T_x|$  est très naturellement liée à (2.3.2).

Plaçons sur  $\mathbb{R}^n$ , et ceci uniquement pour faciliter la preuve, la norme  $\|x\| = \text{Sup } |\xi_k|$ , de manière que les boules soient en fait des pavés cubiques. L'espace  $L(\mathbb{R}^n)$  se trouve alors lui aussi normé avec

$$\|T\|_L = \text{Sup } \{ \|Tx\|, \|x\| \leq 1 \},$$

et l'on a  $\|Tx\| \leq \|T\|_L \|x\|$ ,  $\|T_1 T_2\|_L \leq \|T_1\|_L \|T_2\|_L$  et  $\|I\|_L = 1$ .

**Preuve du théorème.** Reprenons l'anneau  $\mathfrak{R}$  de (2.3.1) et fixons un pavé  $P = \prod [a_k, b_k] \in \mathfrak{R}$ , tel que  $\bar{P} = \prod [a_k, b_k] \subset U$ .

a) Par continuité uniforme des applications  $x \rightarrow T_x, x \rightarrow T_x^{-1}$  et  $x \rightarrow J(x)$  sur  $\bar{P}$ , on a :

$$(*) \quad \begin{cases} \forall \varepsilon > 0, \exists \delta > 0 \quad x, y \in P \text{ et } \|x - y\| \leq \delta \Rightarrow \begin{cases} \|T_x - T_y\|_L \leq \varepsilon \\ |J(x) - J(y)| \leq \varepsilon \end{cases} \\ \exists M < +\infty \quad \|T_x^{-1}\|_L \leq M \text{ pour tout } x \in P \end{cases}$$

Fixons  $\varepsilon > 0$ , et décomposons  $P$  en somme finie disjointe de petits pavés  $Q$ , tous de même arête  $2s \leq 2\delta$ , et considérons l'un d'eux, de centre  $a$ . Pour  $x \in Q$ , on a  $\|T_x - T_a\|_L \leq \varepsilon$ , donc aussi  $\|T_a^{-1} T_x - I\|_L \leq \varepsilon \|T_a^{-1}\|_L \leq M \varepsilon$ , donc encore  $\|T_a^{-1} T_x\|_L \leq 1 + M \varepsilon$ . Considérons

l'application  $G = T_a^{-1} H$ , telle que  $D G_x = T_a^{-1} T_x$ , de sorte que  $\|D G_x\|_L \leq 1 + M \varepsilon$  pour  $x \in Q$ . En appliquant la formule classique des accroissements finis, valable puisque  $Q$  est convexe, on obtient, pour  $x \in Q$

$$\|G(x) - G(a)\| \leq (1 + M \varepsilon) \|x - a\| \leq (1 + M \varepsilon) s$$

de sorte que  $G(Q)$  est contenu dans la boule de centre  $G(a)$  et de rayon  $(1 + M \varepsilon) s$ , c'est-à-dire dans le pavé de centre  $G(a)$  et d'arête  $2(1 + M \varepsilon) s$ , par le choix de la norme  $\|\cdot\|$ . Ainsi

$$\lambda_n [G(Q)] \leq (1 + M \varepsilon)^n (2s)^n = (1 + M \varepsilon)^n \lambda_n (Q).$$

Mais  $G(Q) = T_a^{-1} H(Q)$ , donc avec (2.3.2) on obtient

$$\lambda_n [G(Q)] = |J(a)|^{-1} \lambda_n [H(Q)]$$

puis  $\lambda_n [H(Q)] \leq (1 + M \varepsilon)^n |J(a)| \lambda_n (Q)$  .

Or sur  $Q$  on a  $|J(a)| \leq \varepsilon + |J(x)|$ , et en intégrant sur  $Q$  on obtient

$$|J(a)| \lambda_n (Q) \leq \varepsilon \lambda_n (Q) + \int_Q |J(x)| dx$$

ce qui, en résumé, fournit l'inégalité

$$\lambda_n [H(Q)] \leq (1 + M \varepsilon)^n \left[ \varepsilon \lambda_n (Q) + \int_Q |J(x)| dx \right]$$

Par sommation sur les petits cubes  $Q$  on arrive alors à

$$\lambda_n [H(P)] \leq (1 + M \varepsilon)^n \left[ \varepsilon \lambda_n (P) + \int_P |J(x)| dx \right]$$

où l'on peut faire tendre  $\varepsilon$  vers 0, ce qui donne

$$(**) \quad \lambda_n [H(P)] \leq \int_P |J(x)| dx$$

b) Posons alors  $\mu = H^{-1}(\lambda_n)$  et  $\nu = |J(x)| dx$ , de manière que (\*\*) signifie  $\mu P \leq \nu P$ , et fixons un borélien  $A$  de  $U$ . En recouvrant  $A$  par une réunion dénombrable de pavés  $P_n \subset U$ ,  $A \subset \bigcup P_n$ , on a

$$\mu A \leq \sum \mu P_n \leq \sum \nu P_n$$

et en faisant varier ces recouvrements, on obtient

$$\mu A \leq \nu^* A = \nu A$$

car  $A$  est borélien, donc mesurable pour  $\nu$ . On a donc obtenu

$$(***) \quad \lambda_n [H(A)] \leq \int_A |J(x)| dx$$

c) En considérant maintenant un borélien  $B$  de  $V$ , et en posant  $A = H^{-1}(B)$ , on peut récrire (\*\*\*) sous la forme

$$\lambda_n (B) \leq \int_{H^{-1}(B)} |J(x)| dx = \int 1_B [H(x)] |J(x)| dx$$

inégalité qui permet, grâce aux fonctions étagées et au théorème d'approximation (1.4.6), d'obtenir

$$\int_V f(y) dy \leq \int_U f [H(x)] |J(x)| dx$$

pour  $f \geq 0$  borélienne sur  $V$ .

En posant alors  $g(x) = f[H(x)] |J(x)|$ , et en échangeant les rôles de  $U$  et  $V$ , et ceux de  $H$  et  $K = H^{-1}$ , on a aussi

$$\int_U g(x) dx \leq \int_V g[K(y)] |J_K(y)| dy$$

Or  $g[K(y)] |J_K(y)| = f(y) |J(Ky)| |J_K(y)| = f(y)$  puisque  $J_K(y) = \frac{1}{J(x)}$  lorsque  $x = Ky$ .

On arrive ainsi enfin à l'égalité

$$\int_U g(x) dx = \int_V f(y) dy$$

qui donne la partie b) du théorème, d'où l'on tire aisément a) avec  $f = 1_B$  et  $B = H(A)$ , et aussi c) par le raisonnement habituel considérant  $|f|$  et la décomposition  $f = f^+ - f^-$ . La preuve est ainsi complète.  $\square$

La formule fondamentale étant ainsi obtenue portons un effort tout particulier dans le domaine des applications.

*Passage en coordonnées polaires.* Dans le plan  $(r, \theta)$ , soit  $U$  l'ouvert défini par  $r > 0$  et  $|\theta| < \pi$ . Dans le plan  $(x, y)$  soit  $V$  l'ouvert dont le complémentaire est la demi-droite réelle négative  $\{x \leq 0, y = 0\}$ . Alors  $U$  et  $V$  sont en correspondance bijective par le difféomorphisme  $H: U \rightarrow V$  qui définit le passage en coordonnées polaires

$$H(r, \theta) = (x, y) \text{ avec } x = r \cos \theta, \quad y = r \sin \theta$$

Le jacobien  $J(r, \theta)$  de  $H$  est donné par

$$J = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r$$

et ainsi la formule du changement de variables s'écrit, à la fois sur le plan  $\mathbb{R}^2$  et sur le premier quadrant

$$\int \int_{\mathbb{R}^2} f(x, y) dx dy = \int_0^{\infty} \int_{-\pi}^{\pi} f(r \cos \theta, r \sin \theta) r dr d\theta$$

$$\int_0^{\infty} \int_0^{\frac{\pi}{2}} f(x, y) dx dy = \int_0^{\infty} \int_0^{\frac{\pi}{2}} f(r \cos \theta, r \sin \theta) r dr d\theta$$

quand  $f$  vérifie bien entendu les bonnes hypothèses. On résume cela sous la forme :

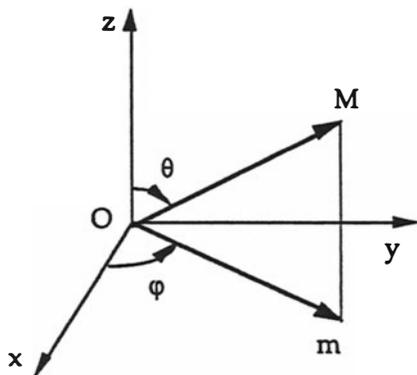
|| L'élément d'aire dans  $\mathbb{R}^2$  est  $dx dy$  en coordonnées cartésiennes et  $r dr d\theta$  en coordonnées polaires.

*Exemple.* Soit  $I = \int_0^{\infty} \exp\left(-\frac{x^2}{2}\right) dx$  l'intégrale de Gauss. Alors

$$I^2 = \int_0^{\infty} \int_0^{\infty} \exp\left[-\frac{1}{2}(x^2 + y^2)\right] dx dy = \int_0^{\infty} \int_0^{\frac{\pi}{2}} \exp\left(-\frac{r^2}{2}\right) r dr d\theta = \frac{\pi}{2}$$

d'où  $I = \sqrt{\frac{\pi}{2}}$  et  $\int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2}\right) dx = 2I = \sqrt{2\pi}$ .

*Passage en coordonnées sphériques.* En coordonnées sphériques on repère un point  $(x, y, z)$  de  $\mathbb{R}^3$  par son rayon polaire  $r = (x^2 + y^2 + z^2)^{1/2} = OM$ , sa colatitude  $\theta = (\overrightarrow{Oz}, \overrightarrow{OM})$ , et sa longitude  $\varphi = (\overrightarrow{Ox}, \overrightarrow{Om})$ , où  $m$  est la projection de  $M$  sur le plan  $xOy$ .



En termes rigoureux, l'ouvert  $U$  dans l'espace  $(r, \theta, \varphi)$  est défini par  $r > 0$ ,  $0 < \theta < \pi$ ,  $|\varphi| < \pi$ , et l'ouvert  $V$  dans l'espace  $(x, y, z)$  est le complémentaire du demi-hyperplan fermé  $\{x \leq 0, y = 0\}$ .

Le difféomorphisme  $H$  qui définit le passage en coordonnées sphériques s'écrit :

$$H(r, \theta, \varphi) = (x, y, z) \quad \text{avec} \quad \begin{cases} x = r \sin \theta \cos \varphi \\ y = r \sin \theta \sin \varphi \\ z = r \cos \theta \end{cases}$$

D'où le jacobien  $J(r, \theta, \varphi)$

$$J = \begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} & \frac{\partial y}{\partial \varphi} \\ \frac{\partial z}{\partial r} & \frac{\partial z}{\partial \theta} & \frac{\partial z}{\partial \varphi} \end{vmatrix} = \begin{vmatrix} \sin \theta \cos \varphi & r \cos \theta \cos \varphi & -r \sin \theta \sin \varphi \\ \sin \theta \sin \varphi & r \cos \theta \sin \varphi & r \sin \theta \cos \varphi \\ \cos \theta & -r \sin \theta & 0 \end{vmatrix}$$

soit  $J = r^2 \sin \theta$ . Ainsi la formule du changement de variables en coordonnées sphériques s'écrit, pour l'espace  $\mathbb{R}^3$  et le premier octant :

$$\iiint_{\mathbb{R}^3} f(x, y, z) dx dy dz = \int_0^{\infty} \int_0^{\pi} \int_{-\pi}^{\pi} f(r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta) r^2 \sin \theta dr d\theta d\varphi$$

$$\int_0^{\infty} \int_0^{\infty} \int_0^{\infty} f(x, y, z) \, dx \, dy \, dz =$$

$$\int_0^{\infty} \int_0^{\frac{\pi}{2}} \int_0^{\frac{\pi}{2}} f(r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta) r^2 \sin \theta \, dr \, d\theta \, d\varphi$$

soit en résumant :

|| L'élément de volume dans  $\mathbb{R}^3$  est  $dx \, dy \, dz$  en coordonnées cartésiennes et  $r^2 \sin \theta \, dr \, d\theta \, d\varphi$  en coordonnées sphériques.

**Exemple 1.** Reprendre l'exercice (1.12.13) sur la loi du  $\chi^2$  à 3 degrés de liberté.

**Exemple 2.** Volume de la boule  $x^2 + y^2 + z^2 \leq 1$ . C'est le volume classique de la sphère; on a

$$V = \int_0^1 dr \int_0^{\pi} d\theta \int_{-\pi}^{\pi} r^2 \sin \theta \, d\varphi = 4\pi \int_0^1 r^2 \, dr = \frac{4}{3}\pi$$

**Exemple 3.** Potentiel newtonien créé par une boule. On sait que le potentiel créé par un point  $M$ , affecté d'une masse  $m$ , électrique ou gravitationnelle, en un point  $A$  est  $\frac{m}{AM}$ . Il suit de là que si un corps  $S$  admet une densité continue  $\rho(x, y, z)$ , le potentiel créé par  $S$  en un point  $A$  quelconque de l'espace (extérieur ou intérieur à  $S$ ) est

$$V_S(A) = \iiint_S \frac{\rho(x, y, z)}{AM} \, dx \, dy \, dz$$

Cherchons à déterminer la fonction  $V_S$  lorsque  $S$  est la boule  $x^2 + y^2 + z^2 \leq R^2$ , et lorsque la densité  $\rho$  ne dépend que de la distance  $r = OM$ , densité que nous écrirons  $\rho(r)$  pour simplifier. Le problème a évidemment une symétrie sphérique de sorte que le calcul de  $V(A)$  pour  $A = (0, 0, a)$  suffit. Alors  $AM = (r^2 + a^2 - 2ar \cos \theta)^{1/2}$ , donc, avec  $a > 0$  :

$$\begin{aligned} V(A) &= \iiint_S \frac{\rho(r)}{AM} r^2 \sin \theta \, dr \, d\theta \, d\varphi \\ &= 2\pi \int_0^R r^2 \rho(r) \, dr \int_0^{\pi} \frac{\sin \theta \, d\theta}{(r^2 + a^2 - 2ar \cos \theta)^{1/2}} \end{aligned}$$

$$\begin{aligned}
 &= \frac{2\pi}{a} \int_0^R r \rho(r) \left[ (r^2 + a^2 - 2ar \cos \theta)^{1/2} \right]_0^\pi dr \\
 &= \frac{2\pi}{a} \int_0^R r \rho(r) [(r+a) - |a-r|] dr
 \end{aligned}$$

- Ainsi si  $A$  est extérieur à  $S$ , alors  $a > R$ , donc

$$V(A) = \frac{4\pi}{a} \int_0^R r^2 \rho(r) dr = \frac{M}{a}$$

où  $M$  est la masse totale de  $S$ . On retrouve le fait bien connu que le potentiel en  $A$  est le même que celui créé par le centre de  $S$  affecté de la masse totale  $M$  de  $S$ .

- Si  $A$  est intérieur à  $S$ , alors  $0 < a < R$  et

$$V(A) = V(a) = \frac{4\pi}{a} \left[ \int_0^a r^2 \rho(r) dr + \int_a^R ar \rho(r) dr \right]$$

d'où  $V'(a) = -\frac{4\pi}{a^2} \int_0^a r^2 \rho(r) dr$  et  $(a^2 V')' = -4\pi a^2 \rho(a)$ , soit encore

$V'' + \frac{2}{a} V' = -4\pi \rho$ . On retrouve la loi de Poisson  $\Delta V = -4\pi \rho$ , une fois rappelée l'expression du laplacien pour une fonction  $V$  sur  $\mathbb{R}^3$  ne dépendant que de la distance à l'origine.

- Si la densité  $\rho$  est constante on a donc

$$V(a) = \frac{M}{a} \text{ si } a \geq R \text{ et } V(a) = \frac{M}{2R} \left[ 3 - \frac{a^2}{R^2} \right] \text{ si } a \leq R$$

où  $M$  est toujours la masse totale de  $S$ .

**Exemple 4.** Aire des surfaces de  $\mathbb{R}^3$ . Il n'est pas dans notre propos de définir les volumes  $n$ -dimensionnels des variétés de dimension  $n$  plongés dans un espace  $\mathbb{R}^N$ . Toutefois le cas  $n = 2$ ,  $N = 3$ , des surfaces ordinaires plongées dans  $\mathbb{R}^3$  est intéressant pour les applications.

Appelons morceau de surface  $S$  la donnée d'un ouvert  $U$  de  $\mathbb{R}^2$  (espace des paramètres  $(u, v)$ ) et d'une application injective  $(u, v) \rightarrow M(u, v)$  de  $U$  dans  $\mathbb{R}^3$  ayant les propriétés suivantes :

$$(*) \quad \begin{cases} \text{a- } M \text{ est de classe } C^1 \text{ sur l'ouvert } U \\ \text{b- Pour tout } (u, v) \in U \text{ les vecteurs } \frac{\partial M}{\partial u} \text{ et } \frac{\partial M}{\partial v} \text{ forment un} \\ \quad \text{système libre} \end{cases}$$

On sait alors que  $S$  admet en chaque point  $M$  un plan tangent, engendré par les vecteurs  $\frac{\partial M}{\partial u}$  et  $\frac{\partial M}{\partial v}$ .

La partie, notée encore  $S$ , image de  $U$  par l'application  $M$  est le morceau de surface étudié, de sorte que le couple  $(S, M)$  définit un élément de surface paramétrée. Il est clair que tout difféomorphisme d'un autre ouvert  $V$  de  $\mathbb{R}^2$  sur  $U$  réalise une nouvelle paramétrisation de  $S$  vérifiant encore les conditions (\*).

Posant, ce qui est classique,

$$E = \left\| \frac{\partial M}{\partial u} \right\|^2, \quad F = \left\langle \frac{\partial M}{\partial u}, \frac{\partial M}{\partial v} \right\rangle, \quad G = \left\| \frac{\partial M}{\partial v} \right\|^2, \quad H = (EG - F^2)^{1/2} = \left\| \frac{\partial M}{\partial u} \wedge \frac{\partial M}{\partial v} \right\|,$$

où les normes sont euclidiennes, on obtient la normale unitaire à  $S$  sous la forme

$$v = \frac{1}{H} \left( \frac{\partial M}{\partial u} \wedge \frac{\partial M}{\partial v} \right).$$

A tout élément d'aire  $du \, dv$  dans l'espace des paramètres correspond, en première approximation, l'élément d'aire défini dans le plan tangent par le parallélogramme associé aux accroissements  $\frac{\partial M}{\partial u} du$  et  $\frac{\partial M}{\partial v} dv$ , que l'on peut calculer en le remplaçant par l'élément de volume dans  $\mathbb{R}^3$  construit sur les vecteurs  $\frac{\partial M}{\partial u} du$ ,  $\frac{\partial M}{\partial v} dv$  et  $v$ , et dont la valeur est  $H \, du \, dv$ . Ce raisonnement heuristique invite alors, étant donné une partie borélienne  $B$  de  $S$ , à définir la mesure de  $B$  dans  $S$  par la formule

$$m(B) = \iint_{M^{-1}(B)} H \, du \, dv$$

On constate alors sans difficulté que  $m(B)$  ne change pas lorsqu'on pratique un changement de paramètres de classe  $C^1$ , car si  $u = \varphi(s, t)$ ,  $v = \psi(s, t)$ , alors le jacobien

$$J(s, t) \text{ n'est autre que } J = \begin{pmatrix} \frac{\partial \varphi}{\partial s} & \frac{\partial \varphi}{\partial t} \\ \frac{\partial \psi}{\partial s} & \frac{\partial \psi}{\partial t} \end{pmatrix},$$

et précisément  $\frac{\partial M}{\partial s} \wedge \frac{\partial M}{\partial t} = J \cdot \left( \frac{\partial M}{\partial u} \wedge \frac{\partial M}{\partial v} \right)$ , d'où  $\left\| \frac{\partial M}{\partial s} \wedge \frac{\partial M}{\partial t} \right\| = |J| H$  et

$$\left\| \frac{\partial M}{\partial s} \wedge \frac{\partial M}{\partial t} \right\| ds dt = H |J| ds dt = H \, du \, dv.$$

En particulier lorsqu'on choisit pour  $S$  la sphère de rayon  $R$ , paramétrée en coordonnées sphériques

$$x = R \sin \theta \cos \varphi$$

$$y = R \sin \theta \sin \varphi \quad 0 < \theta < \pi, \quad |\varphi| < \pi$$

$$z = R \cos \theta$$

on obtient  $dm = R^2 \sin \theta d\theta d\varphi$ , ce qui redonne  $4\pi R^2$  comme aire de la sphère.

## 2.4 Indépendance et mesure produit

Revenons aux questions d'indépendance introduites en 1.12. Pour le cas d'une famille finie de  $p$  variables aléatoires réelles  $X_1, X_2, \dots, X_p$  sur un espace  $(\Omega, \Sigma, P)$ , on peut énoncer un critère extrêmement utile en pratique.

### (2.4.1) Théorème

Soit  $X_1, X_2, \dots, X_p$ ,  $p$  variables aléatoires réelles et  $X = (X_1, X_2, \dots, X_p)$  le vecteur qu'elles constituent dans  $\mathbb{R}^p$ . Une condition nécessaire et suffisante pour que  $X_1, X_2, \dots, X_p$  soient (globalement) indépendantes est que la loi  $\mu$  de  $X$  soit le produit des lois  $\mu_i$  des  $X_i$  :

$$\mu = \mu_1 \otimes \mu_2 \otimes \dots \otimes \mu_p$$

**Preuve.** Elle résulte de la définition de la mesure produit et de la définition de l'indépendance, dès que l'on a remarqué les égalités :

$$P\{X_1 \in B_1, \dots, X_p \in B_p\} = P\{X \in B_1 \times B_2 \times \dots \times B_p\} = \mu\{B_1 \times B_2 \times \dots \times B_p\}$$

$$\prod_{k=1}^p P\{X_k \in B_k\} = \prod_{k=1}^p \mu_k(B_k)$$

pour tout choix des boréliens  $B_k$  de  $\mathbb{R}$ . □

Nous verrons plus loin des applications importantes de ce théorème lorsque nous introduirons la convolution. Pour l'instant cherchons plutôt les applications du côté des mesures gaussiennes.

**Vecteurs gaussiens et lois gaussiennes (centrées).** Commençons par introduire  $p$  variables indépendantes  $X_1, X_2, \dots, X_p$ , suivant chacune la loi LG  $(0, 1)$ . D'après le théorème, le vecteur  $X = (X_1, \dots, X_p)$  a pour loi la probabilité  $\gamma$  de densité

$$g_\gamma(x) = \frac{1}{(2\pi)^{p/2}} \exp \left[ -\frac{1}{2} (x_1^2 + \dots + x_p^2) \right]$$

soit encore, en introduisant la norme euclidienne

$$g_\gamma(x) = \frac{1}{(2\pi)^{p/2}} \exp \left[ -\frac{1}{2} \|x\|^2 \right]$$

La matrice de covariance, introduite en 1.12 dans le cas général, est ici donnée par  $c_{ij} = E(X_i X_j) = E(X_i) E(X_j) = 0$  si  $i \neq j$  et  $c_{kk} = E(X_k^2) = 1$ ; c'est donc la matrice unité de  $\mathbb{R}^p$ , de sorte qu'on pourrait noter  $\gamma$  sous la forme LG  $(0, I_p)$ .

Considérons maintenant un sous-espace  $E \subset \mathbb{R}^p$ . En fixant une base *orthonormale* dans  $E$ , on peut définir sur  $E$  une mesure de Borel  $\lambda_E$ . Et d'après (2.3.3) cette mesure ne dépend pas de la base choisie, puisqu'on passe d'une base orthonormale à une autre par une isométrie de  $E$ . On peut donc aussi définir sur  $E$  une probabilité gaussienne canonique  $\gamma_E$  en choisissant pour  $\gamma_E$  la probabilité de densité

$$g_E(s) = \frac{1}{(2\pi)^{m/2}} \exp \left[ -\frac{1}{2} \|s\|^2 \right] \quad \begin{array}{l} m = \dim E \\ s \in E \end{array}$$

On a alors :

(2.4.2) *Proposition*

Soit  $E$  et  $F$  deux sous-espaces orthogonaux de  $\mathbb{R}^p$ . On a  $\gamma_{E \oplus F} = \gamma_E \otimes \gamma_F$  et en particulier pour  $F = E^\perp$  on a

$$\gamma = \gamma_E \otimes \gamma_{E^\perp}$$

*Preuve.* Puisque  $\lambda_{E \oplus F} = \lambda_E \otimes \lambda_F$ , il suffit de constater que la densité  $g_{E \oplus F}(s, t)$  n'est autre que  $g_E(s) \cdot g_F(t)$ , comme conséquence du théorème de Pythagore

$$\|(s, t)\|^2 = \|s + t\|^2 = \|s\|^2 + \|t\|^2 \quad s \in E, t \in F$$

et d'utiliser (2.2.10). □

Passons maintenant à la définition générale d'une mesure gaussienne sur  $\mathbb{R}^p$ , en nous restreignant au cas des mesures centrées.

(2.4.3) *Définition*

On dit qu'une probabilité  $\mu$  est gaussienne sur  $\mathbb{R}^p$  lorsqu'il existe une application linéaire  $M \in L(\mathbb{R}^p)$  telle que  $\mu$  soit la mesure image  $\mu = M(\gamma)$ .

Dans ce cas, la matrice de covariance, ou l'opérateur de covariance, est donnée par

$$\begin{aligned} \langle \Gamma x, y \rangle &= \int \langle x, t \rangle \langle y, t \rangle d\mu(t) = \int \langle x, Ms \rangle \langle y, Ms \rangle d\gamma(s) \\ &= \int \langle M^*x, s \rangle \langle M^*y, s \rangle d\gamma(s) = \langle M^*x, M^*y \rangle \\ &= \langle MM^*x, y \rangle \end{aligned}$$

soit  $\Gamma = MM^*$ . Et la question qui se pose est de savoir comment la mesure  $\mu$  dépend effectivement de l'opérateur  $M$ , ou de l'opérateur  $\Gamma$ . En fait,  $\mu$  est complètement

déterminée par  $\Gamma$ , ce que nous pouvons démontrer élémentairement dans le cas où  $\Gamma$  est un isomorphisme, donc aussi  $M$  puisque  $\det \Gamma = (\det M)^2$ .

(2.4.4) *Proposition*

Soit  $M$  un isomorphisme de  $\mathbb{R}^p$  et  $\Gamma = MM^*$ . Alors la mesure gaussienne  $\mu = M(\gamma)$  est définie par la densité

$$g_\mu(x) = \frac{1}{(2\pi)^{p/2}} \frac{1}{(\det \Gamma)^{1/2}} \exp \left[ -\frac{1}{2} \langle \Gamma^{-1}x, x \rangle \right]$$

*Preuve.* Soit  $f : \mathbb{R}^p \rightarrow [0, +\infty)$  une fonction borélienne positive.

Alors  $\int f(x) d\mu(x) = \int f(Mt) d\gamma(t)$ , et le changement de variables  $x = Mt$ , donc  $dx = |\det M| dt$ , donne

$$\begin{aligned} \int f(x) d\mu(x) &= \frac{1}{(2\pi)^{p/2}} \frac{1}{|\det M|} \int f(x) \exp \left[ -\frac{1}{2} \|M^{-1}x\|^2 \right] dx \\ &= \int f(x) g_\mu(x) dx \end{aligned}$$

puisque  $|\det M| = (\det \Gamma)^{1/2}$  et  $\|M^{-1}x\|^2 = \langle (M^{-1})^* M^{-1}x, x \rangle = \langle \Gamma^{-1}x, x \rangle$ . On a donc le résultat en choisissant  $f = 1_B$ , avec  $B$  borélien de  $\mathbb{R}^p$ .  $\square$

(2.4.5) *Corollaire*

La mesure gaussienne normale  $\gamma$  est invariante par toute isométrie de  $\mathbb{R}^p$ .

*Preuve.* Car si  $M = U$  est une isométrie alors  $\Gamma = UU^* = I$  et  $g_\mu = g_\gamma$ .  $\square$

Le fait que l'opérateur de covariance  $\Gamma$  détermine la mesure  $\mu$  permet de donner une réciproque à la proposition (1.12.24) sur la covariance des variables indépendantes, réciproque tout à fait spéciale au cas gaussien.

(2.4.6) *Théorème*

Soit  $X = (X_1, X_2, \dots, X_p)$  un vecteur gaussien, c'est-à-dire dont la loi  $\mu$  est une mesure gaussienne (centrée) sur  $\mathbb{R}^p$ . Pour que les variables  $X_1, X_2, \dots, X_p$  soient (globalement) indépendantes, il faut et il suffit que la matrice de covariance  $\Gamma$  soit diagonale, autrement dit que  $\text{cov}(X_i, X_j) = 0$  pour tous  $i \neq j$ .

*Preuve.* La condition est évidemment nécessaire. Réciproquement, soit  $\Gamma = [c_{ij}]$  avec  $c_{ij} = 0$  si  $i \neq j$  et  $c_{kk} = E(X_k^2) = \sigma_k^2 \geq 0$ .

Supposons d'abord  $\sigma_k > 0$  pour tout  $k$ , de sorte que  $\Gamma$  est inversible, et que  $\mu$  est définie par la densité  $g_\mu$  de (2.4.4). Or

$$g_{\mu}(x) = \frac{1}{(2\pi)^{p/2}} \frac{1}{\sigma_1 \sigma_2 \dots \sigma_p} \exp \left[ -\frac{1}{2} \sum_{k=1}^p \frac{x_k^2}{\sigma_k^2} \right]$$

de sorte que  $g_{\mu} = g_1 \otimes g_2 \otimes \dots \otimes g_p$  avec

$$g_k(t) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma_k} \exp \left[ -\frac{t^2}{2\sigma_k^2} \right], \quad t \in \mathbb{R}$$

Il est alors immédiat que la loi de la variable  $X_k$  est définie par la densité  $g_k$ , soit  $d\mu_k = g_k(t) dt$ , et ainsi  $\mu = \mu_1 \otimes \dots \otimes \mu_p$ , ce qui donne l'indépendance.

Maintenant si  $\sigma_k \geq 0$ , examinons le cas où  $\sigma_k^2 = E(X_k^2) = 0$  pour un indice  $k$  fixé. On a évidemment  $X_k = 0$  p.s. Et puisque la fonction nulle est indépendante de toute tribu contenue dans  $\Sigma$ , on voit que la question reste ramenée à l'ensemble des indices  $k$  tels que  $\sigma_k > 0$ , d'où le résultat général.  $\square$

(2.4.7) *Exercice*

- a) Établir que toutes les mesures gaussiennes (centrées) sur  $\mathbb{R}$  sont les lois LG  $(0, \sigma)$ ,  $\sigma > 0$ , augmentées de la mesure de Dirac  $\delta_0$ , correspondant à  $\sigma = 0$ .
- b) Soit  $X = (X_1, X_2, \dots, X_p)$  un vecteur gaussien (centrée) sur  $\mathbb{R}^p$ , de loi  $\mu = M(\gamma)$ , de covariance  $\Gamma = MM^*$ . Pour chaque  $a \in \mathbb{R}^p$ , soit

$$Z_a = \langle a, X \rangle = \sum_{k=1}^p a_k X_k. \text{ On pose}$$

$$\sigma_a^2 = E(Z_a^2) = \langle \Gamma a, a \rangle = \|M^* a\|^2$$

On suppose  $\sigma_a > 0$ . Établir qu'il existe une isométrie  $U$  de  $\mathbb{R}^p$  telle que  $U e_1 = \frac{1}{\sigma_a} M^* a$  et en déduire que la variable  $Z_a$  est gaussienne sur  $\mathbb{R}$ , de loi LG  $(0, \sigma_a)$ . Qu'en est-il lorsque  $\sigma_a = 0$  ?

On a ainsi prouvé que si  $(X_1, \dots, X_p)$  est gaussien toutes ses marges

$$Z_a = \sum_{k=1}^p a_k X_k, \text{ et en particulier les } X_k, \text{ sont des variables gaussiennes.}$$

(2.4.8) *Exercice*

Soit  $X$  une variable aléatoire réelle suivant une loi LG  $(0, 1)$ . On fixe  $a > 0$  et on considère la variable  $Y$  définie par les conditions :

$$Y = X \quad \text{si} \quad |X| \leq a$$

$$Y = -X \quad \text{si} \quad |X| > a$$

- a) Établir que  $Y$  suit la même loi que  $X$ .

- b) Le vecteur  $Z = (X, Y)$  est-il gaussien ?
- c) Les variables  $X$  et  $Y$  sont-elles indépendantes ?
- d) Quelle est la loi de la variable  $\frac{X + Y}{2}$  ?
- e) Montrer qu'il existe une valeur (unique) de  $a$  telle que  
 $E(XY) = \text{cov}(X, Y) = 0$ .

En déduire que le théorème (2.4.6) n'est plus vrai si l'on suppose seulement que les variables  $X_k$  sont gaussiennes sans imposer le caractère gaussien du vecteur  $(X_1, \dots, X_p)$ .



## CHAPITRE 3 : CONVOLUTION ET TRANSFORMATION DE FOURIER

### 3.1 Les espaces $L^1$ et $L^2$

On revient à la donnée d'un espace mesuré  $(\Omega, \Sigma, \mu)$ , la mesure  $\mu$  étant supposée  $\sigma$ -finie. Rappelons qu'on a désigné, en (1.8.7), par  $\mathfrak{F}^1 = \mathfrak{F}^1(\mu) = \mathfrak{F}^1(\Omega, \Sigma, \mu)$  l'espace vectoriel des fonctions (mesurables) intégrables par rapport à  $\mu$ . On supposera qu'il s'agit de fonctions réelles ou complexes.

Pour toute  $f \in \mathfrak{F}^1$  on introduit le nombre

$$N_1(f) = \int |f| \, d\mu < +\infty$$

et on vérifie immédiatement que l'application  $N_1(\cdot) : \mathfrak{F}^1 \rightarrow [0, +\infty)$  est une semi-norme, c'est-à-dire que l'on a

$$N_1(f) \geq 0 ; N_1(\lambda f) = |\lambda| N_1(f) ; N_1(f + g) \leq N_1(f) + N_1(g)$$

pour  $f, g \in \mathfrak{F}^1$  quelconques. De plus  $N_1(f) = 0$  équivaut à  $f = 0$   $\mu$ pp, c'est-à-dire à  $f \in \mathcal{N}$ , où  $\mathcal{N}$  est le sous-espace des fonctions (intégrables) négligeables.

L'espace  $\mathfrak{F}^2 = \mathfrak{F}^2(\Omega, \Sigma, \mu)$ . A côté de  $\mathfrak{F}^1$  on introduit aussi l'ensemble des fonctions mesurables  $f$  telles que  $|f|^2 \in \mathfrak{F}^1$ , noté  $\mathfrak{F}^2$ . On dit que  $\mathfrak{F}^2$  est l'espace des fonctions (mesurables) de carré intégrable.

On remarquera que la condition  $|f|^2 \in \mathfrak{F}^1$  n'est pas suffisante, à elle seule, pour que  $f \in \mathfrak{F}^2$ , en l'absence de mesurabilité. Pour plus de précision, on peut voir en exercice que l'on a en fait l'équivalence

$$f \in \mathfrak{F}^2 \Leftrightarrow f \cdot |f| \in \mathfrak{F}^1.$$

Pour prouver que  $\mathfrak{F}^2$  est un espace vectoriel on introduit le nombre

$$N_2(f) = \left( \int |f|^2 \, d\mu \right)^{1/2} < +\infty$$

pour toute  $f \in \mathfrak{F}^2$ . On a alors :

#### (3.1.1) Proposition (Inégalité de Cauchy-Schwarz)

Pour toutes  $f, g \in \mathfrak{F}^2$ , la fonction  $fg$  est élément de  $\mathfrak{F}^1$  et

$$N_1(fg) \leq N_2(f) N_2(g)$$

**Preuve.** La formule de la moyenne (1.8.12), appliquée à  $fg$ , permet de se ramener à supposer  $f \geq 0$  et  $g \geq 0$ . La condition  $2fg \leq f^2 + g^2$  permet de voir que  $fg \in \mathfrak{F}^1$ , et la considération du trinôme positif

$$T(\lambda) = \int (f + \lambda g)^2 d\mu = N_2(f)^2 + 2\lambda N_1(fg) + \lambda^2 N_2(g)^2,$$

dont le discriminant  $\Delta$  est négatif ou nul, donne le résultat.  $\square$

On en déduit aisément :

(3.1.2) **Proposition**

L'ensemble  $\mathfrak{F}^2$  est un espace vectoriel sur lequel l'application  $N_2(\cdot)$  est une semi-norme, telle que  $N_2(f) = 0$  ssi  $f \in \mathcal{N}$ .

**Preuve.** Pour  $f, g \in \mathfrak{F}^2$ , on a déjà  $fg \in \mathfrak{F}^1$ , donc  $|f+g|^2 \in \mathfrak{F}^1$  et ainsi  $f+g \in \mathfrak{F}^2$ . De plus

$$\begin{aligned} N_2^2(f+g) &= N_2^2(f) + 2 \operatorname{Re} \int f \bar{g} d\mu + N_2^2(g) \\ &\leq N_2^2(f) + 2 N_2(f) N_2(g) + N_2^2(g) = [N_2(f) + N_2(g)]^2 \end{aligned}$$

d'où  $N_2(f+g) \leq N_2(f) + N_2(g)$ . Le reste est évident.  $\square$

(3.1.3) **Exercice**

- On suppose  $\mu$  finie. Montrer que  $\mathfrak{F}^2 \subset \mathfrak{F}^1$ . Réciproque.
- On suppose que  $\mu = P$  est une probabilité. Montrer que  $N_1(f) \leq N_2(f)$  pour toute  $f \in \mathfrak{F}^2$ .

**Les espaces  $L^1$  et  $L^2$ .** - Les espaces  $\mathfrak{F}^1$  et  $\mathfrak{F}^2$  ne sont pas normés puisque  $N_1$  et  $N_2$  sont seulement des semi-normes. Par ailleurs, les conditions  $N_1(f-g) = 0$ , ou  $N_2(f-g) = 0$ , sont équivalentes à  $f-g \in \mathcal{N}$ , et impliquent  $N_1(f) = N_1(g)$ , ou  $N_2(f) = N_2(g)$ . On peut donc introduire :

(3.1.4) **Définition**

On désigne par  $L^1$  (resp.  $L^2$ ) l'ensemble des classes d'équivalence de fonctions de  $\mathfrak{F}^1$  (resp.  $\mathfrak{F}^2$ ) modulo l'égalité  $\mu$ pp. C'est aussi l'espace vectoriel quotient  $\mathfrak{F}^1/\mathcal{N}$  (resp.  $\mathfrak{F}^2/\mathcal{N}$ ).

On peut alors transporter la semi-norme  $N_1$  sur l'espace  $L^1$ . Désignons en effet (provisoirement) par  $\dot{f} = f + \mathcal{N}$  la classe d'équivalence de  $f$ . En posant  $\|\dot{f}\|_1 = N_1(f)$  si  $f \in \mathfrak{F}^1$  et  $\|\dot{f}\|_2 = N_2(f)$  si  $f \in \mathfrak{F}^2$ , on obtient une définition cohérente, et il est immédiat qu'alors  $\|\cdot\|_1$  est une norme sur  $L^1$  et  $\|\cdot\|_2$  une norme sur  $L^2$ .

Les espaces  $L^1$  et  $L^2$  étant ainsi normés, la première question à poser est celle de leur complétude. Pour cela on a besoin d'un lemme général assez commode.

(3.1.5) *Lemme*

Soit  $(E, \| \cdot \|)$  un espace normé. Pour que  $E$  soit complet, autrement dit un espace de Banach, il faut et il suffit que pour toute suite  $(x_n)$  de  $E$  telle que  $\sum \|x_n\| < +\infty$ , la série  $\sum x_n$  soit convergente dans  $E$ , ce qui revient à dire que toute série normalement convergente est convergente dans  $E$ .

**Preuve.** La condition est nécessaire car  $\| \sum_p^q x_n \| \leq \sum_p^q \| x_n \|$ . Réciproquement, supposons la vérifiée, et soit  $(y_n)$  une suite de Cauchy dans  $E$ . Pour montrer qu'elle est convergente dans  $E$ , il suffit de construire une sous-suite  $(y_{n_k})$  qui soit convergente. Or, il est facile, grâce à la condition de Cauchy, de construire la suite des entiers  $n_k$ , strictement croissante, de manière que  $\| y_{n_{k+1}} - y_{n_k} \| \leq 2^{-k}$ . Posons alors  $x_k = y_{n_{k+1}} - y_{n_k}$ , de sorte que  $\sum \| x_k \| < +\infty$ , et qu'ainsi la série  $\sum x_k$  est convergente dans  $E$ , ce qui signifie que les sommes partielles

$$\sum_{j=1}^{k-1} x_j = y_{n_k} - y_{n_1}$$

convergent vers un élément  $z \in E$ . Alors  $y_{n_k} \rightarrow y = y_{n_1} + z$  et tout est dit. □

A partir de là on obtient :

(3.1.6) *Théorème (Fischer-Riesz)*

L'espace  $L^1$ , normé par  $\| \cdot \|_1$ , est un espace de Banach.

**Preuve.** Fixons une suite  $(f_n)$  de  $\mathcal{L}^1$  telle que  $\sum N_1(f_n) < +\infty$ . On peut utiliser (1.8.10) pour voir qu'il existe une fonction  $F \in \mathcal{L}^1$  telle que l'on ait  $\sum f_n = F$   $\mu$ pp. Il reste à vérifier que la classe  $\dot{F}$  est, dans  $L^1$ , la somme de la série  $\sum \dot{f}_n$ . Or cela résulte de

$$\begin{aligned} \| \dot{F} - \sum_1^n \dot{f}_k \|_1 &= N_1 \left( F - \sum_1^n f_k \right) = N_1 \left( \sum_{n+1}^\infty f_k \right) \\ &\leq \int \left( \sum_{n+1}^\infty |f_k| \right) d\mu = \sum_{n+1}^\infty N_1(f_k) = \varepsilon_n \downarrow 0. \end{aligned} \quad \square$$

*L'espace de Hilbert  $L^2$ .* - Supposons les fonctions à valeurs complexes. Il est alors immédiat que la norme  $\| \cdot \|_2$  de  $L^2$  provient du produit scalaire

$$(\dot{f} | \dot{g}) = \int f \bar{g} d\mu$$

puisque l'intégrale  $\int f \bar{g} \, d\mu$  ne dépend pas du choix de  $f$  et  $g$  dans leur classe.

Ainsi  $L^2$  apparaît comme un espace préhilbertien, et si l'on démontre qu'il est complet, on aura en fait un espace de Hilbert.

(3.1.7) **Théorème (Fischer-Riesz)**

┆ L'espace  $L^2$ , normé par  $\| \cdot \|_2$ , est un espace de Hilbert.

**Preuve.** Appliquons (3.1.5) en considérant une suite  $(f_n)$  de  $\mathfrak{F}^2$  telle que  $\sum N_2(f_n) < +\infty$ .

Posons  $g = \sum |f_n| \leq +\infty$ , ce qui détermine une fonction  $g \in \bar{\mathfrak{F}}_+(\Sigma)$ . Avec Beppo Lévi on a

$$\begin{aligned} \int g^2 \, d\mu &= \lim_{n \rightarrow \infty} \int \left( \sum_1^n |f_k| \right)^2 \, d\mu = \lim_{n \rightarrow \infty} \left[ N_2 \left( \sum_1^n |f_k| \right) \right]^2 \\ &\leq \lim_{n \rightarrow \infty} \left[ \sum_1^n N_2(f_k) \right]^2 \leq \left[ \sum_1^\infty N_2(f_k) \right]^2 < +\infty \end{aligned}$$

Il suit de là que  $g$  est  $\mu$ pp finie, d'où l'on déduit l'existence d'une fonction  $h \in \mathfrak{F}^2$  telle que  $h = g \, \mu$ pp, et d'une fonction mesurable  $F$  telle que  $F = \sum f_n \, \mu$ pp. Comme de plus on a

$|F| \leq h \, \mu$ pp, on obtient  $F \in \mathfrak{F}^2$ , et il ne reste plus qu'à voir que la série  $\sum f_n$  converge vers  $F$  dans l'espace  $L^2$ . Or on a presque partout

$$\left| F - \sum_1^n f_k \right| = \left| \sum_{n+1}^\infty f_k \right| \leq h$$

ce qui permet l'application du théorème de convergence dominée à la suite

$$\left| F - \sum_1^n f_k \right|^2, \text{ et donne la condition } N_2 \left( F - \sum_1^n f_k \right) \rightarrow 0. \quad \square$$

**Remarque.** On rencontre ici, dès 1907, la première justification de l'intérêt apporté par le nouvel outil qu'est l'intégrale de Lebesgue, qui contrairement à l'intégrale de Riemann, amène à la considération d'espaces complets. C'est ainsi que l'espace  $L^2$  a fourni le premier exemple "fonctionnel" d'espace de Hilbert, prolongeant le cas de l'espace  $\mathfrak{L}^2$  des suites de carré sommable. Il est ainsi directement à l'origine de la "géométrisation de l'analyse", de la création des espaces fonctionnels, de l'apparition de la notion de norme et des problèmes de convergence associés.

**Convention.** Pour simplifier l'écriture nous désignerons désormais par la même notation

$f$  les éléments de  $\mathfrak{F}^1$  ou  $\mathfrak{F}^2$  et ceux de  $L^1$  ou  $L^2$ , et nous écrirons  $\|f\|_1$  plutôt que  $\|f\|_1$ . Cela revient en pratique à identifier deux fonctions (mesurables)  $\mu$ pp égales, et à parler de  $L^1$  et  $L^2$  comme espaces de fonctions. C'est évidemment abusif mais l'usage en est bien établi.

**Théorèmes de densité.**- Il n'est pas toujours commode de raisonner avec des fonctions intégrables et la tentation (ou la nécessité) se fait souvent sentir de ne travailler qu'avec de "bons" sous-espaces de  $\mathfrak{L}^1$  ou  $L^1$ . On entend d'ailleurs par "bons" principalement des sous-espaces partout denses. Et ils sont de deux sortes : construits à partir des fonctions étagées dans le cadre général abstrait d'un espace  $(\Omega, \Sigma, \mu)$ , ou à partir des fonctions continues dans le cas d'une mesure borélienne sur  $\mathbb{R}^n$  ou sur un espace métrique localement compact dénombrable à l'infini.

Étant donnée une fonction étagée  $g$  sur la tribu  $\Sigma$ , écrite selon

$$g = \sum \lambda_k 1_{A_k} \quad \text{avec } \lambda_k \neq 0, i \neq j \Rightarrow A_i \cap A_j = \emptyset$$

on voit que  $|g| = \sum |\lambda_k| 1_{A_k}$  et  $|g|^2 = \sum |\lambda_k|^2 1_{A_k}$ , d'où résulte que  $g \in \mathfrak{L}^1$  équivaut à  $g \in \mathfrak{L}^2$ , ou encore aux conditions  $\mu A_k < +\infty$  pour tout  $k$ .

On désignera donc par  $\mathfrak{E}^1 = \mathfrak{E}^2 = \mathfrak{E} \cap \mathfrak{L}^1 = \mathfrak{E} \cap \mathfrak{L}^2$  l'espace des fonctions étagées intégrables (ou de carré intégrable). Alors :

(3.1.8) **Théorème**

┆ L'espace  $\mathfrak{E}^1 = \mathfrak{E}^2$  est partout dense dans chacun des espaces  $L^1$  ou  $L^2$ .

**Preuve.** Il suffit, pour le cas  $L^1$ , d'approcher chaque fonction  $f \geq 0$  de  $\mathfrak{L}^1$  par une fonction étagée  $g$  au sens de la norme  $\|\cdot\|_1$ . Or il existe une suite  $g_n \in \mathfrak{E}$  telle que  $g_n \uparrow f$ , d'où  $\int (f - g_n) d\mu \downarrow 0$  par monotonie décroissante puisque les intégrales sont finies. Mais alors  $\|f - g_n\|_1 \rightarrow 0$ .  $\square$

On peut même diminuer l'espace  $\mathfrak{E}^1$  en le remplaçant par l'espace  $\mathfrak{E}^1(\mathfrak{R})$  des fonctions étagées sur n'importe quel anneau  $\mathfrak{R}$  qui engendre la tribu  $\Sigma$ , pourvu que  $\mu$  soit  $\sigma$ -finie sur  $\mathfrak{R}$ .

(3.1.9) **Corollaire**

┆ Soit  $\mathfrak{R}$  un anneau engendrant la tribu  $\Sigma$ , tel que la mesure  $\mu$  soit  $\sigma$ -finie sur  $\mathfrak{R}$ . Alors l'espace  $\mathfrak{E}^1(\mathfrak{R})$  des fonctions intégrables étagées sur  $\mathfrak{R}$  (combinaisons linéaires finies d'éléments  $1_B$  avec  $B \in \mathfrak{R}$  et  $\mu B < +\infty$ ) est dense dans chacun des espaces  $L^1$  ou  $L^2$ .

**Preuve.** D'après (3.1.8) il suffit d'approcher toute fonction  $1_A$  avec  $A \in \Sigma$  et  $\mu A < +\infty$ . Or la théorie de la mesure extérieure assure l'existence, pour tout  $\varepsilon > 0$ , d'une suite  $(C_n)$  choisie dans l'anneau  $\mathfrak{R}$ , telle que  $A \subset \bigcup C_n$  et  $\sum \mu C_n \leq \mu A + \frac{\varepsilon}{2}$ . Choisissons  $N$  tel

que  $\sum_{N+1}^{\infty} \mu C_n \leq \frac{\varepsilon}{2}$ , et posons  $C = \bigcup_{n \leq N} C_n \in \Sigma$  et  $B = \bigcup_{n \leq N} C_n \in \mathfrak{R}$ . Alors on a  $A \subset C$ ,  $B \subset C$ ,  $\mu(C \setminus A) \leq \frac{\varepsilon}{2}$  et  $\mu(C \setminus B) \leq \frac{\varepsilon}{2}$ , d'où  $\mu(A \Delta B) \leq \varepsilon$ . Mais

$$\|1_A - 1_B\|_1 = \|1_A - 1_B\|_2^2 = \mu(A \Delta B) \leq \varepsilon$$

ce qui suffit. □

Il est maintenant facile de tirer de (3.1.9) des résultats relatifs à l'espace  $\mathbb{R}^n$  ou plus généralement au cas d'un espace  $T$  métrique localement compact dénombrable à l'infini, déjà rencontré au sujet du théorème de Riesz-Alexandroff.

(3.1.10) *Théorème*

Soit  $\mu$  une mesure borélienne (localement finie) sur un espace  $T$  métrique localement compact et dénombrable à l'infini. Alors l'espace  $\mathcal{K}(T)$  des fonctions continues à support compact est dense dans chacun des espaces  $L^1$  ou  $L^2$ .

*Preuve.* Soit  $B$  un borélien tel que  $\mu B < +\infty$ . Par le théorème de régularité (1.11.1) il existe, pour tout  $\varepsilon > 0$ , un compact  $K \subset B$  tel que  $\mu(B \setminus K) \leq \varepsilon$ . D'après la preuve de (1.11.1), il existe un ouvert  $U \supset K$  tel que  $\mu(U \setminus K) \leq \varepsilon$ , et il existe aussi une fonction  $g \in \mathcal{K}(T)$  telle que  $1_K \leq g \leq 1_U$ .

$$\begin{aligned} \text{Alors } |1_B - g| &\leq |1_B - 1_K| + |1_K - g| \\ &\leq |1_B - 1_K| + |1_U - 1_K| \\ &\leq 1_{B \setminus K} + 1_{U \setminus K} \end{aligned}$$

d'où  $\|1_B - g\|_1 \leq 2\varepsilon$  et  $\|1_B - g\|_2 \leq 2\sqrt{\varepsilon}$ . □

Le cas de l'espace  $\mathbb{R}^n$  est encore plus particulier, puisqu'on a même

(3.1.11) *Théorème*

Soit  $\mu$  une mesure borélienne sur  $\mathbb{R}^n$  (localement finie). l'espace vectoriel engendré par les fonctions  $g = g_1 \otimes g_2 \otimes \dots \otimes g_n$ , où chaque  $g_k$  est une fonction indéfiniment dérivable et à support compact sur  $\mathbb{R}$ , est dense dans  $L^1$  ou  $L^2$ .

*Preuve.* Avec (3.1.9) appliqué à l'anneau  $\mathcal{R}$  formé des réunions finies (disjointes) de pavés  $P = \prod [a_k, b_k[$ , il suffit d'approcher une fonction  $1_P$ . Or fixons  $\varepsilon > 0$  et choisissons, ce qui est possible,  $a'_k$  et  $b'_k$  tels que  $a'_k < a_k$  et  $a_k < b'_k < b_k$  de façon que le pavé compact  $K = \prod [a'_k, b'_k] \subset P$  et le pavé ouvert  $U = \prod ]a'_k, b'_k[ \supset P$  soient tels que  $\mu(U \setminus K) \leq \varepsilon$ . Il est alors aisé, avec l'exercice (3.1.12) qui suit, de construire, pour chaque  $k$ , une fonction  $g_k$ , de classe  $C^\infty$ , telle que  $0 \leq g_k \leq 1$ ,  $g_k \equiv 1$  sur  $[a_k, b'_k]$  et  $g_k \equiv 0$  en dehors de  $]a'_k, b'_k[$ . La fonction  $g = g_1 \otimes \dots \otimes g_n$ , définie par  $g(x) = g_1(x_1) g_2(x_2) \dots g_n(x_n)$  est alors telle que  $1_K \leq g \leq 1_U$ , d'où  $|g - 1_P| \leq 1_{U \setminus K}$  et  $\|g - 1_P\|_1 \leq \varepsilon$ ,  $\|g - 1_P\|_2 \leq \sqrt{\varepsilon}$ . □

(3.1.12) *Exercice*

1) Montrer que la fonction  $\varphi$ , définie sur  $\mathbb{R}$  par

$$\varphi(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \exp\left(-\frac{1}{x}\right) & \text{si } x > 0 \end{cases}$$

est indéfiniment dérivable sur  $\mathbb{R}$ .

2) En déduire que la fonction  $\psi(x) = \varphi(x) \varphi(1-x)$  est indéfiniment dérivable sur  $\mathbb{R}$  et telle que  $\psi(x) > 0$  pour  $x \in ]0, 1[$  et  $\psi(x) = 0$  pour  $x \notin ]0, 1[$ .

3) On pose  $A = \int_0^1 \psi(t) dt$ . Établir que la fonction

$$g(x) = \frac{1}{A} \int_{-\infty}^x \psi(t) dt$$

est indéfiniment dérivable sur  $\mathbb{R}$  et telle que  $0 \leq g(x) \leq 1$ ,  $g(x) = 0$  si  $x \leq 0$  et  $g(x) = 1$  si  $x \geq 1$ .

## 3.2 Produit de convolution

Dans ce paragraphe, de même que dans le suivant consacré à la transformation de Fourier, l'espace  $L^1$  (ou  $L^2$ ), écrit sans autre précision, se rapporte à la mesure de Borel  $\lambda_n$  sur  $\mathbb{R}^n$ , notée aussi  $d\lambda_n(x) = dx$  pour simplifier. On rappelle que pour toute  $f \in L^1$  on a les formules suivantes, pour  $a \in \mathbb{R}^n$ ,

$$\int f(x+a) dx = \int f(a-x) dx = \int f(x) dx$$

traduisant l'invariance de  $\lambda_n$  par les translations et la symétrie par rapport à l'origine.

Commençons par le produit de convolution de deux fonctions.

(3.2.1) *Définition*

Soient  $f$  et  $g$  deux éléments de  $L^1$ . On appelle produit de convolution de  $f$  et  $g$ , noté  $f \star g$ , l'élément  $h$  de  $L^1$  suffisamment bien défini par l'égalité presque partout.

$$h(x) = \int f(x-t) g(t) dt$$

Fixons en effet  $f \in \mathfrak{L}^1$  et  $g \in \mathfrak{L}^1$ . En appliquant le théorème de Fubini à la fonction  $F(x, t) = f(x - t) g(t)$ , définie sur  $\mathbb{R}^n \times \mathbb{R}^n$ , on voit que la fonction  $t \rightarrow f(x - t) g(t)$  est intégrable sur  $\mathbb{R}^n$  pour presque tout  $x$ . Autrement dit la fonction  $x \rightarrow \int f(x - t) g(t) dt$  est presque partout définie et toute version  $h$ , c'est-à-dire toute fonction mesurable presque partout égale, est élément de  $\mathfrak{L}^1$ , donc définit un élément de  $L^1$ , noté encore  $h$ . Il est alors clair que si on modifie  $f$  et  $g$  presque partout en  $f_1$  et  $g_1$  éléments de  $\mathfrak{L}^1$ , on obtient une version  $h_1$  telle que  $h_1 = h$  presque partout. Ainsi l'élément  $h \in L^1$  ne dépend que des éléments  $f \in L^1$  et  $g \in L^1$ .

Les propriétés élémentaires de ce produit de convolution sont regroupées dans les deux énoncés suivants :

(3.2.2) *Théorème*

- a) Sur l'espace  $L^1$  le produit de convolution est commutatif, associatif et distributif par rapport à l'addition.
- b) En ajoutant aux opérations vectorielles de  $L^1$  le produit de convolution  $\star$ , on obtient une algèbre de Banach appelée algèbre convolutive  $L^1$ . En particulier on a

$$\|f \star g\|_1 \leq \|f\|_1 \|g\|_1$$

*Preuve.* Par Fubini on a déjà

$$\begin{aligned} \|f \star g\|_1 &\leq \iint |f(x - t) g(t)| dt dx \\ &\leq \int |g(t)| \left[ \int |f(x - t)| dx \right] dt = \|f\|_1 \|g\|_1 \end{aligned}$$

Par ailleurs la commutativité  $f \star g = g \star f$  est évidente par le changement de variable  $s = x - t$ . Pour prouver l'associativité, fixons une troisième fonction  $h \in L^1$ , qui n'a rien à voir avec  $f \star g$ . Par Fubini, on a

$$\begin{aligned} [(f \star g) \star h](x) &= \iint f(x - t - s) g(s) h(t) ds dt \\ &= \iint f(x - s - t) g(t) h(s) ds dt \\ &= [(f \star h) \star g](x) \end{aligned}$$

par symétrie  $(s, t) \rightarrow (t, s)$ . D'où, avec la commutativité,

$$\begin{aligned} (f \star g) \star h &= (f \star h) \star g = (h \star f) \star g = (h \star g) \star f \\ &= f \star (g \star h) \end{aligned}$$

ce qui fournit l'associativité. □

(3.2.3) *Proposition*

Soit  $f, g \in L^1$ . Pour toute fonction  $\varphi$ , Borel-mesurable et bornée sur  $\mathbb{R}^n$ , on a

$$\int (f \star g)(x) \varphi(x) dx = \iint \varphi(s+t) f(s) g(t) ds dt$$

**Preuve.** Avec Fubini on a, successivement

$$\begin{aligned} \int (f \star g)(x) \varphi(x) dx &= \iint f(x-t) g(t) \varphi(x) dt dx \\ &= \int g(t) \left[ \int f(x-t) \varphi(x) dx \right] dt \end{aligned}$$

et le changement de variable  $s = x - t$  donne, à  $t$  fixé,

$$\begin{aligned} \int (f \star g)(x) \varphi(x) dx &= \int g(t) \left[ \int f(s) \varphi(s+t) ds \right] dt \\ &= \iint \varphi(s+t) f(s) g(t) ds dt \quad \square \end{aligned}$$

**Produit de convolution de deux mesures bornées.** Considérons sur l'espace  $\mathbb{R}^n$  deux mesures boréliennes bornées  $\mu$  et  $\nu$  et désignons par  $s$  l'application (somme)  $(x, y) \rightarrow x + y$  de  $\mathbb{R}^n \times \mathbb{R}^n$  dans  $\mathbb{R}^n$ . On peut alors introduire la mesure bornée  $\mu \otimes \nu$  sur  $\mathbb{R}^n \times \mathbb{R}^n$ , puis son image  $\lambda = s(\mu \otimes \nu)$  par l'application continue  $s$ . Avec la formule d'intégration par rapport à une mesure image on voit que l'on a

$$\int \varphi d\lambda = \int \varphi(u+v) d\mu(u) d\nu(v)$$

pour toute fonction  $\varphi$  Borel-mesurable et bornée sur  $\mathbb{R}^n$ . La proposition (3.2.3) montre ainsi que si  $\mu$  admet la densité  $f$  par rapport à la mesure  $dx$ , et si  $\nu$  admet la densité  $g$  alors la mesure  $\lambda$  admet la densité  $h = f \star g$ . Dans le cas général on notera donc, sans incohérence,  $\lambda = \mu \star \nu$  en disant que  $\lambda$  est la mesure (bornée) produit de convolution de  $\mu$  et  $\nu$ . En résumé on a :

(3.2.4) *Théorème*

a) Pour toute partie borélienne  $A$  de  $\mathbb{R}^n$ , on a

$$\begin{aligned} (\mu \star \nu)(A) &= (\mu \otimes \nu)(s^{-1}(A)) \\ &= \int \mu(A-t) d\nu(t) = \int \nu(A-t) d\mu(t) \end{aligned}$$

b) Pour toute fonction  $\varphi$  Borel-mesurable et bornée sur  $\mathbb{R}^n$  on a

$$\int \varphi d(\mu \star \nu) = \iint \varphi(u+v) d\mu(u) d\nu(v)$$

c) Lorsque  $\mu$  et  $\nu$  sont des probabilités, il en est de même de  $\mu \star \nu$ .

d) Enfin l'opération de convolution est commutative et associative sur l'ensemble  $M^1 = M^1(\mathbb{R}^n)$  des mesures boréliennes bornées.

**Exemple.** Soit  $\delta_a$  la mesure de Dirac au point  $a \in \mathbb{R}^n$ . Vérifier que la mesure  $\mu \star \delta_a$  n'est autre que la translatée  $\tau_a \mu$ , image de  $\mu$  dans la translation  $\tau_a : x \rightarrow x + a$ , définie par  $(\tau_a \mu)(A) = \mu(A - a)$ . En particulier  $\delta_a \star \delta_b = \delta_{a+b}$ . En particulier aussi, lorsque  $a = 0$ ,  $\mu \star \delta_0 = \mu$ . D'où

(3.2.5) **Proposition**

La mesure de Dirac  $\delta_0$  à l'origine est élément neutre pour le produit de convolution sur l'ensemble  $M^1$  des mesures bornées.

**Remarque.** Lorsque la mesure  $\nu$  admet une densité  $g$ , (qui est alors nécessairement élément de  $L^1$  et positive ou nulle) on peut facilement montrer que la mesure  $\mu \star \nu$

admet la densité  $h$ , définie presque partout par  $h(x) = \int g(x - t) d\mu(t)$ .

On note quelquefois cette fonction  $h$  sous la forme  $h = \mu \star g$ .

Plus généralement on peut définir  $h = \mu \star g$  par la formule précédente pour tout choix de  $g \in L^1$  ( $g$  n'est plus nécessairement positive). On a alors,  $g \in L^1$  et  $\|h\|_1 \leq \mu(\mathbb{R}^n) \|g\|_1$ .

**Remarque.** Il peut toutefois arriver que le produit de convolution  $\mu \star \nu$  de deux mesures admette une densité  $h$  par rapport à la mesure de Borel  $dx$  sans que  $\mu$  et  $\nu$  soient densitables. Un exemple simple (?) est fourni en choisissant pour  $\mu$  la probabilité sur  $\mathbb{R}^2$  égale à la masse +1 uniformément répartie sur le cercle  $x^2 + y^2 = 1$ . Elle est défini par l'égalité

$$\int \varphi d\mu = \frac{1}{2\pi} \int_0^{2\pi} \varphi(\cos \theta, \sin \theta) d\theta$$

pour toute fonction  $\varphi$  Borel-mesurable et bornée. On peut alors vérifier, avec quelques calculs, que la mesure  $\mu \star \mu$  admet la densité

$$h(x, y) = \begin{cases} \frac{1}{\pi^2} \frac{1}{r\sqrt{4-r^2}} & \text{si } r = \sqrt{x^2 + y^2} < 2 \\ 0 & \text{si } r \geq 2 \end{cases}$$

Et bien entendu  $\mu$  n'est pas densitable puisqu'elle est concentrée sur le cercle  $x^2 + y^2 = 1$ , qui est de mesure nulle pour la mesure de Borel sur  $\mathbb{R}^2$ .

**Somme de vecteurs aléatoires indépendants.** Fixons un espace probabilisé  $(\Omega, \Sigma, P)$  et deux vecteurs aléatoires  $X : \Omega \rightarrow \mathbb{R}^n$  et  $Y : \Omega \rightarrow \mathbb{R}^n$ , à valeurs dans le même espace  $\mathbb{R}^n$ . Alors le vecteur somme  $Z = X + Y$  est aussi à valeurs dans l'espace  $\mathbb{R}^n$ . On peut donc introduire les trois probabilités  $\mu, \nu, \lambda$  sur  $\mathbb{R}^n$  respectivement lois de  $X, Y$  et  $Z$ . La question est de savoir s'il existe une relation simple entre ces trois probabilités. Dans le cas général il n'en est rien. Mais dans le cas où  $X$  et  $Y$  sont des vecteurs aléatoires indépendants on sait que la loi du vecteur aléatoire  $(X, Y)$ , à valeurs dans  $\mathbb{R}^n \times \mathbb{R}^n$ , n'est autre que la probabilité  $\mu \otimes \nu$ . Alors la loi du vecteur aléatoire  $Z = X + Y$  est

évidemment l'image par l'application somme  $s : (x, y) \rightarrow x + y$  de la probabilité  $\mu \otimes \nu$ , autrement dit la probabilité  $\mu \star \nu$ .

D'où le résultat :

(3.2.6) *Théorème*

Soient  $X$  et  $Y$  deux vecteurs aléatoires indépendants, à valeurs dans l'espace  $\mathbb{R}^n$ , de lois respectives  $\mu$  et  $\nu$ . Alors le vecteur aléatoire somme  $Z = X + Y$  a pour loi la probabilité  $\lambda = \mu \star \nu$ .

(3.2.7) *Corollaire*

Soient  $X_1, X_2, \dots, X_p$  des vecteurs aléatoires indépendants (globalement), à valeurs dans l'espace  $\mathbb{R}^n$ , de lois respectives  $\mu_1, \mu_2, \dots, \mu_p$ . Alors le vecteur aléatoire somme  $Z = X_1 + X_2 + \dots + X_p$  a pour loi la probabilité

$$\lambda = \mu_1 \star \mu_2 \star \dots \star \mu_p.$$

(3.2.8) *Exemples*

**Exemple 1.** Soit  $X$  et  $Y$  deux variables aléatoires indépendantes suivant chacune une loi de Poisson de paramètres respectifs  $\alpha$  et  $\beta$ . Alors la variable  $Z = X + Y$  suit une loi de Poisson de paramètres  $\gamma = \alpha + \beta$ .

**Exemple 2.** Soit  $X$  et  $Y$  deux variables aléatoires indépendantes suivant chacune une loi binomiale d'ordres respectifs  $m$  et  $n$  et de même paramètre  $p$ . Alors la variable  $Z = X + Y$  suit une loi binomiale  $(m+n, p)$ .

En particulier, soit  $X_1, X_2, \dots, X_n$ ,  $n$  variables indépendantes suivant toutes une loi de Bernoulli de même paramètre  $p$ . Alors la variable  $Z = X_1 + X_2 + \dots + X_n$ , suit une loi binomiale  $(n, p)$ .

**Exemple 3.** Soit  $X$  et  $Y$  deux variables indépendantes suivant des lois de Laplace-Gauss de paramètres respectifs  $(m, \sigma^2)$  et  $(m', \sigma'^2)$ . Alors la variable  $Z = X + Y$  suit une loi de Laplace-Gauss de paramètre  $(m+m', \sigma^2 + \sigma'^2)$ . On pourra se ramener à (2.4.7) en supposant d'abord  $m = m' = 0$ , et en remarquant que le vecteur  $(X, Y)$  est gaussien; alors la variable  $Z$  est gaussienne et on calculera sa variance.

**Exemple 4.** Soit  $X$  et  $Y$  deux variables indépendantes suivant des lois du  $\chi^2$  respectives à  $m$  et  $n$  degrés de liberté. On voit alors (sans calculs) que la variable  $Z = X + Y$  suit une loi du  $\chi^2$  à  $(m+n)$  degrés de liberté. On retrouvera ce résultat en effectuant le produit de convolution  $f_m \star f_n$  des densités correspondantes. Comparer à (2.2.16a).

### 3.3 Transformation de Fourier et fonction caractéristique

Les notations restent celles de 3.2 relativement à l'espace  $L^1 = L^1(\mathbb{R}^n, \lambda_n)$ . Le produit scalaire classique sur  $\mathbb{R}^n$  est noté  $\langle x, y \rangle = \sum_{k=1}^n x_k y_k$  ; il est associé à la norme euclidienne  $\|x\| = \langle x, x \rangle^{1/2}$ .

#### (3.3.1) Définition

- a) On appelle transformée de Fourier d'une fonction  $f \in L^1$  la fonction  $\mathcal{F}f$  définie sur  $\mathbb{R}^n$  par

$$(\mathcal{F}f)(x) = \int e^{i\langle x, t \rangle} f(t) dt$$

- b) On appelle transformée de Fourier d'une mesure bornée  $\mu \in L^1$  la fonction  $\mathcal{F}\mu$  définie sur  $\mathbb{R}^n$  par

$$(\mathcal{F}\mu)(x) = \int e^{i\langle x, t \rangle} d\mu(t)$$

- c) Lorsque  $\mu$  est une probabilité on dit que  $\mathcal{F}\mu$  est la *fonction caractéristique* (ou la f.c.) de  $\mu$ .

Il est clair que  $\mathcal{F}f$  et  $\mathcal{F}\mu$  sont des fonctions bornées, continues, d'après le théorème (1.10.1) et telles que  $\|\mathcal{F}f\| \leq \|f\|_1$  et  $\|\mathcal{F}\mu\| \leq (\mathcal{F}\mu)(0) = \mu(\mathbb{R}^n)$ . De plus, on a les propriétés de linéarité évidentes suivantes :

$$\mathcal{F}(\alpha f + \beta g) = \alpha \mathcal{F}f + \beta \mathcal{F}g \quad f, g \in L^1 ; \alpha, \beta \in \mathbb{C}$$

$$\mathcal{F}(\alpha \mu + \beta \nu) = \alpha \mathcal{F}\mu + \beta \mathcal{F}\nu \quad \mu, \nu \in M^1 ; \alpha, \beta \geq 0$$

Par ailleurs, si  $f$  est réelle et paire, alors  $\mathcal{F}f$  est une fonction réelle et paire donnée par

$$(\mathcal{F}f)(x) = \int f(t) \cos \langle x, t \rangle dt.$$

Si  $f$  est réelle et impaire alors  $(\mathcal{F}f)(x) = i \int f(t) \sin \langle x, t \rangle dt$  et  $\mathcal{F}f$  est imaginaire pure et impaire. De même si  $\mu$  est une mesure symétrique, c'est-à-dire invariante par la symétrie  $t \rightarrow -t$ , alors  $\mathcal{F}\mu$  est réelle et paire et  $(\mathcal{F}\mu)(x) = \int \cos \langle x, t \rangle d\mu(t)$ .

De plus pour toute  $\mu \in M^1$ , on a  $(\mathcal{F}\mu)(-x) = \overline{(\mathcal{F}\mu)(x)}$ .

De plus pour toute  $\mu \in M^1$ , on a  $(\mathcal{F}\mu)(-x) = \overline{(\mathcal{F}\mu)(x)}$ .

Donnons maintenant l'un des premiers résultats fondamentaux.

#### (3.3.2) Théorème (Riemann-Lebesgue)

Pour toute  $f \in L^1$  la fonction  $\mathcal{F}f$  est élément de l'espace  $C_0(\mathbb{R}^n)$ .

**Preuve.** Soit  $C^\infty(\mathbb{R}^n)$  l'espace des fonctions continues et bornées sur  $\mathbb{R}^n$ , muni de sa norme uniforme. Déjà la transformation de Fourier  $\mathcal{F}$  opère continûment de  $L^1$  dans  $C^\infty(\mathbb{R}^n)$ , puisque  $\|\mathcal{F}f\| \leq \|f\|_1$ . Il suffit donc de prouver que l'on a  $\mathcal{F}f \in C_0(\mathbb{R}^n)$  pour les fonctions  $f$  d'un sous-espace dense de  $L^1$ , ou encore pour les fonctions  $f$  d'un ensemble total dans  $L^1$ . Avec les résultats (3.1.9) et (3.1.11) on peut donc se ramener au cas où  $f = 1_P$  est la fonction indicatrice d'un pavé  $P = \prod [a_k, b_k[$ . Mais alors si  $g = \mathcal{F}f$ , on voit que

$$g(x) = g_1(x_1) g_2(x_2) \dots g_n(x_n), \text{ avec } g_k(t) = \int_{a_k}^{b_k} e^{ist} ds.$$

Or on a  $\|g_k\| \leq (b_k - a_k)$  et  $|g_k(t)| \leq \frac{2}{|t|}$  pour  $t \neq 0$ . Et puisqu'il existe toujours un indice  $k$  tel que  $|x_k| \geq \frac{\|x\|}{\sqrt{n}}$ , on voit facilement que l'on a

$$|(\mathcal{F}f)(x)| \leq A^{n-1} \frac{2\sqrt{n}}{\|x\|}, \quad x \neq 0$$

avec  $A = \max_k (b_k - a_k)$ , d'où  $\mathcal{F}f \in C_0(\mathbb{R}^n)$ .  $\square$

Pour une mesure bornée  $\mu$  on n'a pas en général  $\mathcal{F}\mu \in C_0(\mathbb{R}^n)$  comme on peut voir avec  $\mu = \delta_0$  puisque  $\mathcal{F}\delta_0 = 1$ , ou avec  $\mu = \delta_a$  puisque  $(\mathcal{F}\delta_a)(x) = e^{i\langle x, a \rangle}$ . Toutefois :

### (3.3.3) Théorème

Pour toute  $\mu \in M^1(\mathbb{R}^n)$  la fonction  $\mathcal{F}\mu$  est bornée et uniformément continue sur  $\mathbb{R}^n$ .

**Preuve.** On part de l'égalité  $|e^{ia} - e^{ib}|^2 = 4 \sin^2\left(\frac{b-a}{2}\right)$  pour  $a, b$  réels, égalité qui fournit l'inégalité

$$|(\mathcal{F}\mu)(x) - (\mathcal{F}\mu)(y)| \leq 2 \int |\sin \langle \frac{x-y}{2}, t \rangle| d\mu(t)$$

Il suffit alors de voir que la fonction  $x \rightarrow \int |\sin \frac{\langle x, t \rangle}{2}| d\mu(t)$  est continue, ce qui ramène à (1.10.1), et nulle à l'origine.  $\square$

Donnons maintenant quelques évidences utiles.

### (3.3.4) Proposition

On fixe la fonction  $f \in L^1$  et la mesure bornée  $\mu$ .

a) Pour tout  $R > 0$ , la fonction  $f\left(\frac{t}{R}\right)$  a pour transformée de Fourier la fonction  $R^n (\mathcal{F}f)(Rx)$ .

- b) Pour tout  $a \in \mathbb{R}^n$  la fonction translatée  $\tau_a f$  (définie par  $t \rightarrow f(t - a)$ ) et la mesure translatée  $\tau_a \mu$  ont respectivement pour transformées de Fourier les fonctions  $e^{i\langle x, a \rangle} \mathcal{F}f$  et  $e^{i\langle x, a \rangle} \mathcal{F}\mu$ .
- c) Pour tout  $a \in \mathbb{R}^n$  la fonction  $e^{-i\langle a, t \rangle} f(t)$  a pour transformée de Fourier la fonction  $\tau_a (\mathcal{F}f)(x) = (\mathcal{F}f)(x - a)$ .

(3.3.5) **Proposition**

On fixe les entiers  $m$  et  $n$ . Alors sur l'espace  $\mathbb{R}^{m+n} = \mathbb{R}^m \times \mathbb{R}^n$ , on a :

- a)  $\mathcal{F}(f \otimes g) = \mathcal{F}f \otimes \mathcal{F}g$  pour  $f \in L^1(\mathbb{R}^m)$  et  $g \in L^1(\mathbb{R}^n)$
- b)  $\mathcal{F}(\mu \otimes \nu) = \mathcal{F}\mu \otimes \mathcal{F}\nu$  pour  $\mu \in M^1(\mathbb{R}^m)$  et  $\nu \in M^1(\mathbb{R}^n)$

**Liaison avec le produit de convolution.** On aborde là à la deuxième propriété importante, qui donne toute sa richesse à la transformation de Fourier.

(3.3.6) **Théorème**

- a) Pour  $f, g \in L^1$  on a  $\mathcal{F}(f \star g) = \mathcal{F}f \cdot \mathcal{F}g$ . En conséquence la transformation de Fourier  $\mathcal{F}$  est un homomorphisme continu de l'algèbre de Banach convolutive  $L^1(\mathbb{R}^n)$  dans l'algèbre de Banach multiplicative  $C_0(\mathbb{R}^n)$ .
- b) Pour  $\mu, \nu \in M^1$  on a  $\mathcal{F}(\mu \star \nu) = \mathcal{F}\mu \cdot \mathcal{F}\nu$ .
- c) Enfin pour  $g \in L^1$  et  $\mu \in M^1$  on a  $\mathcal{F}(\mu \star g) = \mathcal{F}\mu \cdot \mathcal{F}g$ .

**Preuve.** Il suffit d'utiliser (3.2.3) pour a), (3.2.4) pour b), et la remarque qui suit (3.2.5) pour c), en choisissant pour  $\varphi$  la fonction  $\varphi(t) = e^{i\langle x, t \rangle}$ ,  $x$  fixé.  $\square$

**Remarque.** L'égalité  $\mathcal{F}\delta_0 = 1$  signifie donc, avec (3.3.6) que la transformation de Fourier transporte l'élément unité convolutif en l'élément unité multiplicatif. Par ailleurs le sens profond du théorème est que beaucoup de calculs dans l'algèbre convolutive  $L^1(\mathbb{R}^n)$  pourront être grandement simplifiés par passage à l'image de Fourier puisqu'ils se transformeront en calculs multiplicatifs plus habituels. Nous en verrons plus loin des exemples lorsqu'on aura montré qu'en plus des propriétés (3.3.6) la transformation de Fourier possède la propriété d'injectivité.

**Fonction caractéristique d'un vecteur aléatoire.** Étant donné un vecteur aléatoire  $X$  à valeurs dans  $\mathbb{R}^n$ , dont la loi est notée  $\mu_X$ , on convient d'appeler fonction caractéristique de  $X$  la fonction caractéristique de  $\mu_X$ , que l'on note alors sous la forme  $\varphi_X = \mathcal{F}\mu_X$ . On obtient là une fonction complexe définie sur l'espace  $\mathbb{R}^n$  par

$$\varphi_X(x) = E[\exp(i\langle x, X(\bullet) \rangle)]$$

On peut maintenant vérifier élémentairement les propriétés suivantes :

- a) Pour tout  $a \in \mathbb{R}^n$  on a  $\varphi_{X+a}(x) = e^{i\langle x, a \rangle} \varphi_X(x)$
- b) Plus généralement, si  $X$  et  $Y$  sont deux vecteurs aléatoires *indépendants* à valeurs dans  $\mathbb{R}^n$  on a  $\varphi_{X+Y} = \varphi_X \cdot \varphi_Y$ .
- c) Pour toute application linéaire  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$  le vecteur aléatoire  $Y = AX$ , à valeurs dans  $\mathbb{R}^n$ , a pour fonction caractéristique
- $$\varphi_Y(y) = \varphi_{AX}(y) = \varphi_X(A^*y)$$
- où  $A^*$  est l'application transposée de  $A$ .
- d) Si  $X$  et  $Y$  sont deux vecteurs aléatoires *indépendants* à valeurs respectivement dans  $\mathbb{R}^m$  et  $\mathbb{R}^n$  et si  $Z$  est le vecteur aléatoire  $(X, Y)$ , à valeurs dans  $\mathbb{R}^{m+n}$  on a  $\varphi_Z = \varphi_X \otimes \varphi_Y$ .

Examinons maintenant quelques exemples concrets.

*Loi de Poisson.* La distribution de Poisson sur  $\mathbb{R}$  (ou sur  $\mathbb{N}$ ), de paramètre  $\lambda$ , soit

$P_\lambda = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \delta_k$ , admet pour fonction caractéristique la fonction

$$\mathcal{F} P_\lambda(x) = \exp[\lambda(e^{ix} - 1)]$$

*Loi binomiale.* Pour l'ordre  $n$  et le paramètre  $p$ ,  $0 < p < 1$ , elle est explicitée par

$\mu = \sum_{k=0}^n C_n^k p^k q^{n-k} \delta_k$ , avec  $q = 1 - p$ , et sa fonction caractéristique vaut

$$\mathcal{F} \mu(x) = (p e^{ix} + q)^n = [1 - p(1 - e^{ix})]^n$$

Pour  $n = 1$ , on obtient la loi de Bernoulli  $\beta = q \delta_0 + p \delta_1$ , et la formule convolutive  $\mu = \beta^n$  redonne la formule multiplicative  $\mathcal{F} \mu = (\mathcal{F} \beta)^n$ , avec  $F\beta(x) = p e^{ix} + q$ .

*Loi d'Euler.* Elle est définie dans le cas général par les paramètres  $\alpha > 0$  et  $\beta > 0$  associés à la densité

$$f_{\alpha,\beta}(t) = \begin{cases} 0 & \text{si } t \leq 0 \\ \frac{1}{\beta^\alpha \Gamma(\alpha)} t^{\alpha-1} \exp\left(-\frac{t}{\beta}\right) & \text{si } t > 0 \end{cases}$$

La fonction caractéristique  $\varphi_{\alpha,\beta} = \mathcal{F} f_{\alpha,\beta}$  est donc donnée par

$$\begin{aligned}\varphi_{\alpha,\beta}(x) &= \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_0^\infty t^{\alpha-1} \exp\left(-\frac{t}{\beta}\right) e^{ixt} dt \\ &= \frac{1}{\Gamma(\alpha)} \int_0^\infty u^{\alpha-1} e^{-u(1-i\beta x)} du \\ &= \varphi_{\alpha,1}(\beta x)\end{aligned}$$

Le calcul de  $\varphi_{\alpha,1} = \varphi_\alpha$ , proposé en (2.2.16) par la méthode de l'équation différentielle, donne donc

(3.3.7) **Proposition**

a) La loi générale d'Euler admet pour fonction caractéristique la fonction

$$\varphi_{\alpha,\beta}(x) = (1 + \beta^2 x^2)^{-\frac{\alpha}{2}} e^{i\alpha \operatorname{Arc} \operatorname{tg}(\beta x)}$$

b) En particulier, la loi du  $\chi^2$  à  $n$  degrés de liberté, correspondant à  $\beta = 2$  et  $\alpha = \frac{n}{2}$ , admet la fonction caractéristique

$$\varphi_n(x) = (1 + 4x^2)^{-\frac{n}{2}} e^{i\frac{n}{2} \operatorname{Arc} \operatorname{tg}(2x)}$$

L'exemple 4 de (3.2.8) explique pourquoi on a  $\varphi_{m+n} = \varphi_m \cdot \varphi_n$ . Plus généralement, on peut vérifier que  $f_{\alpha,\beta} \star f_{\alpha',\beta} = f_{\alpha+\alpha',\beta}$  comme conséquence de la formule  $B(\alpha, \alpha') = \frac{\Gamma(\alpha) \Gamma(\alpha')}{\Gamma(\alpha + \alpha')}$ , et retrouver ainsi que  $\varphi_{\alpha+\alpha',\beta} = \varphi_{\alpha,\beta} \cdot \varphi_{\alpha',\beta}$ .

**Loi de Cauchy.** Elle est définie, pour le paramètre  $a > 0$ , par la densité

$$f_a(t) = \frac{1}{\pi} \frac{a}{a^2 + t^2}$$

sur  $\mathbb{R}$  et sa fonction caractéristique,  $\varphi_a$  est donc donnée par

$$\varphi_a(x) = \frac{2a}{\pi} \int_0^\infty \frac{\cos(tx)}{a^2 + t^2} dt = \frac{2}{\pi} \int_0^\infty \frac{\cos ux}{1 + u^2} du$$

Le calcul proposé en (1.10.4), Ex. 2, fournit donc  $\varphi_a(x) = \exp[-a|x|]$ . Ici encore on pourra vérifier directement que  $f_a \star f_b = f_{a+b}$ , d'où  $\varphi_{a+b} = \varphi_a \cdot \varphi_b$ .

**Lois gaussiennes sur  $\mathbb{R}^n$ .** Il s'agit là d'un exemple extrêmement important, à la fois pour la théorie et pour la pratique, que nous détaillerons à peu près complètement. Commençons par la loi normale sur  $\mathbb{R}$ .

(3.3.8) **Proposition**

Soit  $\gamma = \text{LG}(0, 1)$ , de densité  $\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)$  sur  $\mathbb{R}$ .

Alors  $(\mathcal{F}\gamma)(x) = \exp\left(-\frac{x^2}{2}\right)$ .

*Preuve.* On renvoie à (1.10.4), Ex. 1. □

(3.3.9) **Corollaire 1**

La loi gaussienne normale sur  $\mathbb{R}^n$ , de densité

$$g(t) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{\|t\|^2}{2}\right),$$

admet pour fonction caractéristique, la fonction

$$G(x) = \exp\left[-\frac{\|x\|^2}{2}\right].$$

*Preuve.* Avec (3.3.5) et la proposition. □

(3.3.10) **Corollaire 2**

Soit  $\Gamma$  une matrice réelle  $n \times n$ , symétrique et de type défini positif. Suivant (2.4.4) on introduit la mesure gaussienne centrée  $\mu$  admettant  $\Gamma$  comme matrice de covariance, de densité

$$g(t) = \frac{1}{(2\pi)^{n/2}} \frac{1}{(\det \Gamma)^{1/2}} \exp\left[-\frac{1}{2} \langle \Gamma^{-1}t, t \rangle\right]$$

sur  $\mathbb{R}^n$ . Alors la fonction caractéristique  $\varphi = \mathcal{F}\mu$  est donnée par

$$\varphi(x) = \exp\left[-\frac{1}{2} \langle \Gamma x, x \rangle\right]$$

soit  $\varphi(x) = \exp\left[-\frac{1}{2} V(x)\right]$ , où  $V(x)$  est la forme quadratique de covariance de  $\mu$ .

*Preuve.* Soit  $\gamma$  la mesure gaussienne normale sur  $\mathbb{R}^n$ . On sait que  $\mu = M(\gamma)$ , avec  $\Gamma = MM^*$ , d'où suit que  $\varphi(x) = \varphi_\mu(x) = \varphi_\gamma(M^*x)$  avec  $\langle Mt, x \rangle = \langle t, M^*x \rangle$ . Mais

$$\varphi_\gamma(M^*x) = \exp\left[-\frac{1}{2} \|M^*x\|^2\right] = \exp\left[-\frac{1}{2} \langle \Gamma x, x \rangle\right].$$
 □

**Loi de Cauchy sur  $\mathbb{R}^n$ .** On peut généraliser naturellement la loi de Cauchy normale sur  $\mathbb{R}$ , c'est-à-dire de paramètre  $a = 1$ , au cas de la dimension  $n$ . Traitons ici la question sous la forme d'exercices.

(3.3.11) **Exercice**

Vérifier que la loi de Cauchy normale sur  $\mathbb{R}$  est la loi du quotient  $Y = \frac{X}{Z}$  de deux variables  $X$  et  $Z$  indépendantes suivant toutes deux la loi  $\gamma = LG(0, 1)$ .

(3.3.12) **Exercice**

Soit  $X_1, X_2, \dots, X_n, Z$  ( $n+1$ ) variables indépendantes et de même loi  $\gamma = LG(0, 1)$ . On introduit le vecteur  $X = (X_1, \dots, X_n)$  et le vecteur  $Y = \frac{X}{Z}$ , ainsi que la loi  $\nu$  de  $Y$ , appelée loi de Cauchy normale sur  $\mathbb{R}^n$ .

a) Prouver, avec Fubini, qu'il existe une constante  $K_n$  telle que, pour toute fonction  $\varphi$ , borélienne et bornée sur  $\mathbb{R}^n$ , on ait :

$$\int \varphi(y) d\nu(y) = K_n \int_{\mathbb{R}^n} \frac{\varphi(t)}{[1 + \|t\|^2]^{\frac{n+1}{2}}} dt$$

b) Établir que

$$K_n = \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} |u|^n d\gamma(u) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{n+1}{2}}}$$

et retrouver, en passant, le résultat de (1.10.10).

c) En déduire que la loi  $\nu$  admet sur  $\mathbb{R}^n$  la densité

$$g(t) = \frac{\pi^{\frac{n+1}{2}}}{\Gamma\left(\frac{n+1}{2}\right)} \frac{1}{[1 + \|t\|^2]^{\frac{n+1}{2}}}$$

et retrouver (3.3.11) pour  $n = 1$ .

(3.3.13) **Exercice**

Les notations restant les mêmes, on veut déterminer la fonction caractéristique de la loi  $\nu$  sur  $\mathbb{R}^n$ , notée  $\psi$ .

a) Établir que pour  $x \in \mathbb{R}^n$  on a

$$\psi(x) = \int_{\mathbb{R}} \exp\left[-\frac{1}{2} \frac{\|x\|^2}{t^2}\right] d\gamma(t) = K(\|x\|)$$

où l'on a posé

$$\begin{aligned} K(a) &= \int_{\mathbb{R}} \exp\left[-\frac{1}{2} \frac{a^2}{t^2}\right] d\gamma(t) \\ &= \frac{2}{\sqrt{2\pi}} \int_0^{\infty} \exp\left[-\frac{1}{2} \left(t^2 + \frac{a^2}{t^2}\right)\right] dt \end{aligned}$$

pour  $a \geq 0$ .

b) En déduire, avec les résultats de (1.10.4, Ex. 2), l'égalité

$$\psi(x) = (\mathcal{F}\nu)(x) = \exp[-\|x\|]$$

*Transformation de Fourier et différentiabilité.* Donnons quelques résultats élémentaires, décrits en dimension  $n = 1$  pour simplifier, mais facilement adaptables en dimension quelconque.

(3.3.14) *Proposition*

a) Soit  $f \in \mathcal{L}^1(\mathbb{R})$  une fonction admettant une dérivée continue  $f'$  telle que  $f' \in \mathcal{L}^1(\mathbb{R})$ . Alors

$$(\mathcal{F}f)(x) = \frac{i}{x} (\mathcal{F}f')(x)$$

b) Soit  $f \in \mathcal{L}^1(\mathbb{R})$  une fonction telle que  $\int |t f(t)| dt < +\infty$ . Alors la fonction  $\mathcal{F}f$  est continûment dérivable et

$$(\mathcal{F}f)'(x) = i \int_{-\infty}^{+\infty} t e^{itx} f(t) dt$$

c) Pour toute  $\mu \in M^1(\mathbb{R})$  admettant un moment d'ordre 1,  $\mathcal{F}\mu$  est de classe  $C^1$  et

$$(\mathcal{F}\mu)'(x) = i \int t e^{itx} d\mu(t)$$

**Preuve.** Les assertions b) et c) sont évidentes. Quant à a), on commence par montrer que la condition  $f' \in \mathcal{L}^1$  implique l'existence de limites  $\ell$  et  $\ell'$  de  $f(t)$  quand  $t \rightarrow +\infty$  ou

$t \rightarrow -\infty$ . Mais la condition  $f \in \mathfrak{L}^1$  entraîne les égalités  $\mathcal{L} = \mathcal{L}' = 0$ . Il ne suffit plus ensuite qu'à faire une intégration par parties.  $\square$

(3.3.15) *Corollaire*

a) Soit  $f \in \mathfrak{L}^1(\mathbb{R})$  une fonction admettant jusqu'à l'ordre  $p \geq 2$  des dérivées continues  $f^{(k)} \in \mathfrak{L}^1(\mathbb{R})$ . Alors

$$(\mathcal{F} f)(x) = \left(\frac{i}{x}\right)^p (\mathcal{F} f^{(p)})(x)$$

En particulier, on a  $\mathcal{F} f \in \mathfrak{L}^1(\mathbb{R})$ .

b) Soit  $f \in \mathfrak{L}^1(\mathbb{R})$  une fonction telle que  $\int |t|^p |f(t)| dt < +\infty$  pour un entier  $p \geq 2$ . Alors la fonction  $\mathcal{F} f$  est de classe  $C^p$  sur  $\mathbb{R}$  et

$$(\mathcal{F} f)^{(p)}(x) = i^p \int_{-\infty}^{+\infty} t^p e^{itx} f(t) dt$$

c) Soit  $\mu \in M^1(\mathbb{R})$  une mesure admettant un moment d'ordre  $p \geq 2$ . Alors  $\mathcal{F} \mu$  est de classe  $C^p$  et

$$(\mathcal{F} \mu)^{(p)}(x) = i^p \int t^p e^{itx} d\mu(t)$$

*Remarque.* Le corollaire précédent montre en gros que plus  $f$  a de bonnes propriétés de dérivabilité et plus sa transformée de Fourier  $\mathcal{F} f$  a un bon comportement à l'infini. Réciproquement le comportement de  $f$  à l'infini, qui conditionne la finitude des intégrales  $\int |t|^p |f(t)| dt$ , influe sur les propriétés de dérivabilité de  $f$ .

*L'espace de Schwartz*  $\mathcal{S} = \mathcal{S}(\mathbb{R}^n)$ . Supposons tout d'abord  $n = 1$  et introduisons l'espace  $\mathcal{S} = \mathcal{S}(\mathbb{R})$  des fonctions indéfiniment dérivables  $f$  telles que les fonctions  $x^k f^{(p)}(x)$  soient bornées pour tous entiers  $k \geq 0$  et  $p \geq 0$ . Il revient au même de supposer que  $\lim_{|x| \rightarrow \infty} x^k f^{(p)}(x) = 0$  pour tous  $k \geq 0$  et  $p \geq 0$ . On dit encore que  $\mathcal{S}$  est l'espace des fonctions  $C^\infty$  à décroissance rapide. Alors (3.3.14) signifie que  $\mathcal{S}$  est invariant par la transformation de Fourier et que l'opérateur de dérivation  $D: \mathcal{S} \rightarrow \mathcal{S}$ , et l'opérateur  $X: \mathcal{S} \rightarrow \mathcal{S}$  de multiplication par la variable  $x$ , sont reliés entre eux par la transformation de Fourier selon les formules

$$X \mathcal{F} = i \mathcal{F} D ; \quad D \mathcal{F} = i \mathcal{F} X$$

Pour le cas  $n \geq 2$ , il faut remplacer l'opérateur  $X$  par chacun des opérateurs  $X_1, \dots, X_n$  de multiplication par les variables respectives  $x_1, \dots, x_n$  et l'opérateur  $D$  par les opérateurs de dérivation partielle  $D_k = \frac{\partial}{\partial x_k}$ .

En notation condensée on introduit alors la notion de multi-indices  $\alpha = (\alpha_1, \dots, \alpha_n)$ , ainsi que les opérateurs associés  $X^\alpha = X_1^{\alpha_1} \dots X_n^{\alpha_n}$  et  $D^\alpha = D_1^{\alpha_1} \dots D_n^{\alpha_n}$ .

L'espace  $\mathcal{S} = \mathcal{S}(\mathbb{R}^n)$  se définit donc par la condition que pour tous multi-indices  $\alpha, \beta$  la fonction  $X^\alpha D^\beta f$  soit bornée sur  $\mathbb{R}^n$ , ou bien encore soit élément de l'espace  $C_0(\mathbb{R}^n)$ . Et il est alors facile de vérifier que la transformation de Fourier  $\mathcal{F}$  laisse  $\mathcal{S}$  invariant, et que les opérateurs  $X^\alpha$  et  $D^\alpha$  sont liés à  $\mathcal{F}$  selon les formules

$$X^\alpha \mathcal{F} = i^{|\alpha|} \mathcal{F} D^\alpha ; \quad D^\alpha \mathcal{F} = i^{|\alpha|} \mathcal{F} X^\alpha$$

avec  $|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n$ .

L'intérêt de l'espace  $\mathcal{S}$  est donc concentré dans ces formules une fois vu qu'on a affaire à un espace suffisamment grand. Or c'est bien le cas puisque :

(3.3.16) *Exercice*

Démontrer que l'espace  $\mathcal{S} = \mathcal{S}(\mathbb{R}^n)$  est partout dense dans chacun des espaces  $L^1, L^2$  et  $C_0(\mathbb{R}^n)$ . On pourra introduire l'espace  $\mathcal{D} = \mathcal{D}(\mathbb{R}^n)$  des fonctions  $C^\infty$  à support compact et s'inspirer de (3.1.11).

(3.3.17) *Exercice*

Démontrer que pour  $f, g \in \mathcal{S} = \mathcal{S}(\mathbb{R}^n)$  on a  $f \star g \in \mathcal{S}$  et  $fg \in \mathcal{S}$ , autrement dit que  $\mathcal{S}$  est à la fois algèbre convolutive et algèbre multiplicative, la transformation de Fourier faisant passer du produit de convolution au produit multiplicatif ordinaire.

*La formule d'inversion pour les fonctions.* Il nous faut maintenant résoudre deux problèmes : le premier consiste à prouver que l'opérateur  $\mathcal{F} : L^1 \rightarrow C_0(\mathbb{R}^n)$  est injectif; le second consiste à déterminer d'une certaine façon, l'opérateur inverse  $\mathcal{F}^{-1}$ , défini sur l'image  $\mathcal{F} L^1$ . C'est ce qu'on appelle la *synthèse spectrale*.

Pour cela on introduit, parallèlement à  $\mathcal{F}$ , l'opérateur  $\bar{\mathcal{F}}$  défini sur  $L^1$  par

$$(\bar{\mathcal{F}} f)(x) = \int e^{-i\langle x, t \rangle} f(t) dt$$

en remarquant que si l'on note  $\overset{v}{f}$  la fonction symétrique de  $f$  définie par  $\overset{v}{f}(t) = f(-t)$ , alors  $\bar{\mathcal{F}} f = \mathcal{F} \overset{v}{f} = (\mathcal{F} f)^\vee$ . On voit donc que l'opérateur  $\bar{\mathcal{F}}$  a les mêmes propriétés générales que l'opérateur  $\mathcal{F}$ .

Lorsque  $f = g$  est la fonction de Gauss introduite en (3.3.9), densité de la loi gaussienne normale sur  $\mathbb{R}^n$ , on a

$$\mathcal{F} g = G = \exp\left[-\frac{\|x\|^2}{2}\right] = (2\pi)^{\frac{n}{2}} g$$

de sorte que  $\overline{\mathcal{F}} G = \mathcal{F} G = (2\pi)^n g$  et ainsi  $\overline{\mathcal{F}} \mathcal{F} g = (2\pi)^n g$ .

On va voir qu'en fait ce résultat est presque général. Pour cela on introduit, pour tout  $\sigma > 0$ , la fonction

$$g_\sigma(t) = \frac{1}{(2\pi)^{\frac{n}{2}}} \frac{1}{\sigma^n} \exp\left[-\frac{\|t\|^2}{2\sigma^2}\right]$$

densité de la loi gaussienne sur  $\mathbb{R}^n$  dont la matrice de covariance est  $\Gamma_\sigma = \sigma^2 I$ , et dont la transformée de Fourier  $G_\sigma = \mathcal{F} g_\sigma$  est donnée par  $G_\sigma(x) = \exp\left[-\frac{\sigma^2 \|x\|^2}{2}\right]$ , d'après (3.3.4) ou (3.3.10).

Alors pour toute  $f \in L^1$ , on a  $g_\sigma \star f \in L^1$  et  $G_\sigma \mathcal{F} f = \mathcal{F}(g_\sigma \star f) \in \mathcal{F} L^1$ , donc  $G_\sigma \mathcal{F} f \in L^1 \cap \mathcal{F} L^1$ . Et l'on a

(3.3.18) *Lemme*

Pour toute  $f \in L^1$ , on a l'égalité

$$\overline{\mathcal{F}} (G_\sigma \mathcal{F} f) = (2\pi)^n g_\sigma \star f$$

*Preuve.* On part de l'égalité

$$\overline{\mathcal{F}} (G_\sigma \mathcal{F} f)(x) = \int G_\sigma(t) e^{-i\langle x, t \rangle} (\mathcal{F} f)(t) dt$$

et l'on remplace  $\mathcal{F} f(t)$  par sa valeur intégrale, en remarquant que la fonction  $(f \otimes G_\sigma)(s, t) = f(s) G_\sigma(t)$  est élément de l'espace  $L^1(\mathbb{R}^n \times \mathbb{R}^n)$  ce qui permet d'appliquer Fubini. D'où

$$\begin{aligned} \overline{\mathcal{F}} (G_\sigma \mathcal{F} f)(x) &= \iint f(s) G_\sigma(t) e^{-i\langle x-s, t \rangle} ds dt \\ &= \int f(s) (\overline{\mathcal{F}} G_\sigma)(x-s) ds = [(\overline{\mathcal{F}} G_\sigma) \star f](x) \end{aligned}$$

et on termine avec l'égalité  $\overline{\mathcal{F}} G_\sigma = \mathcal{F} G_\sigma = (2\pi)^n g_\sigma$  obtenue à partir de l'égalité  $G_\sigma = \mathcal{F}(g_\sigma)$  en changeant  $\sigma$  en  $\frac{1}{\sigma}$ . □

L'idée qui suit consiste à étudier ce qui se passe lorsque  $\sigma \downarrow 0$ . On remarquera déjà que  $G_\sigma \uparrow 1$ , ce qui permet de penser qu'en un certain sens on doit avoir  $\gamma_\sigma \rightarrow \delta$ , où  $\gamma_\sigma$  est la mesure gaussienne de densité  $g_\sigma$ . On arrive ainsi à la notion, que l'on pourrait développer, d'approximation de l'unité convolutive  $\delta$  par des éléments de  $L^1$ , puisque bien entendu l'unité convolutive "exacte"  $\delta$  n'est pas élément de  $L^1$ . Alors :

(3.3.19) *Théorème (Synthèse spectrale)*

Tout élément  $f \in L^1$  est complètement déterminé par sa transformée de Fourier  $\mathcal{F}f$  selon

$$f = \lim_{\sigma \downarrow 0} \frac{1}{(2\pi)^n} \bar{\mathcal{F}} (G_\sigma \mathcal{F}f)$$

où la limite est prise dans l'espace  $L^1$ , c'est-à-dire selon la norme  $\|\cdot\|_1$  de  $L^1$ .

*Preuve.* Avec le lemme il suffit de prouver que  $g_\sigma \star f \rightarrow f$  dans  $L^1$ . Or

$$(g_\sigma \star f)(x) - f(x) = \int [f(x-t) - f(x)] g_\sigma(t) dt$$

puisque  $\int g_\sigma(t) dt = 1$ , et ainsi, avec Fubini-Tonelli

$$\begin{aligned} \|g_\sigma \star f - f\|_1 &\leq \int \left[ \int |f(x-t) - f(x)| dx \right] g_\sigma(t) dt \\ &\leq \int \omega_1(f, t) g_\sigma(t) dt \end{aligned}$$

où l'on a posé  $\omega_1(f, t) = \int |f(x-t) - f(x)| dx$ , expression qui définit le module de continuité  $\omega_1(f, \cdot)$  de  $f$  dans  $L^1$ , tel déjà que  $\omega_1(f, t) \leq 2 \|f\|_1$ . En admettant provisoirement que  $\omega_1(f, t) \rightarrow 0$  quand  $t \rightarrow 0$  dans  $\mathbb{R}^n$ , on voit que pour tout  $\varepsilon > 0$ , il existe  $\delta > 0$  tel que  $\omega_1(f, t) \leq \varepsilon$  pour  $\|t\| \leq \delta$ , d'où

$$\int_{\|t\| \leq \delta} \omega_1(f, t) g_\sigma(t) dt \leq \varepsilon \int g_\sigma(t) dt = \varepsilon$$

$$\int_{\|t\| > \delta} \omega_1(f, t) g_\sigma(t) dt \leq 2 \|f\|_1 \int_{\|t\| > \delta} d\gamma_\sigma(t)$$

$$\leq 2 \|f\|_1 \int_{\|u\| > \frac{\delta}{\sigma}} d\gamma(u)$$

par le changement de variable  $u = \frac{t}{\sigma}$ . En résumé on a

$$\|g_\sigma \star f - f\|_1 \leq \varepsilon + 2 \|f\|_1 \gamma \left[ B \left( 0, \frac{\delta}{\sigma} \right)^c \right]$$

et quand  $\sigma \downarrow 0$  on voit que  $B \left( 0, \frac{\delta}{\sigma} \right)^c \downarrow \emptyset$ , de sorte que

$$\limsup_{\sigma \downarrow 0} \|g_\sigma \star f - f\|_1 \leq \varepsilon .$$

De fait on a donc  $\limsup_{\sigma \downarrow 0} \|g_\sigma \star f - f\|_1 = 0$ , ce qui termine la preuve.  $\square$

Avant de revenir à la preuve du fait que  $\omega_1(f, t) \rightarrow 0$  quand  $t \rightarrow 0$  dans  $\mathbb{R}^n$ , tirons les principales conséquences du théorème.

(3.3.20) *Corollaire 1*

Sur l'espace  $L^1$  la transformation de Fourier  $\mathcal{F}$  est un opérateur injectif. Autrement dit deux fonctions intégrable  $f$  et  $g$  telles que  $\mathcal{F}f = \mathcal{F}g$  sont nécessairement presque partout égales.

(3.3.21) *Corollaire 2 (Formule d'inversion)*

Soit  $f \in L^1$  une fonction intégrable telle que son image de Fourier  $\mathcal{F}f$  soit elle aussi intégrable. Alors  $f$  est presque partout égale à la fonction continue

$$\frac{1}{(2\pi)^n} \overline{\mathcal{F} \mathcal{F} f}.$$

En particulier si  $f$  est continue et telle que  $f \in L^1$  et  $\mathcal{F}f \in L^1$  alors  $f(x) = \frac{1}{(2\pi)^n} (\overline{\mathcal{F} \mathcal{F} f})(x)$  en tout point  $x \in \mathbb{R}^n$ .

*Preuve.* L'hypothèse  $\mathcal{F}f \in L^1$  assure que  $G_\sigma \mathcal{F}f \rightarrow \mathcal{F}f$  dans  $L^1$  car

$$\|G_\sigma \mathcal{F}f - \mathcal{F}f\|_1 \leq \int [1 - G_\sigma(x)] |\mathcal{F}f(x)| dx$$

et il suffit d'appliquer le théorème de convergence dominée de Lebesgue puisque  $G_\sigma(x) \uparrow 1$  quand  $\sigma \downarrow 0$ . Par le théorème de Riemann-Lebesgue on voit donc que  $(2\pi)^n g_\sigma \star f = \overline{\mathcal{F}}(G_\sigma \mathcal{F}f)$  tend vers  $\overline{\mathcal{F}} \mathcal{F}f$  dans l'espace  $C_0(\mathbb{R}^n)$ , donc en norme uniforme, tandis que  $(2\pi)^n g_\sigma \star f \rightarrow (2\pi)^n f$  dans l'espace  $L^1$  d'après (3.3.19). Par intégration sur une partie borélienne et bornée  $A$  de  $\mathbb{R}^n$ , on en déduit sans difficulté que

$$\int_A (2\pi)^n f(x) dx = \int_A (\overline{\mathcal{F} \mathcal{F} f})(x) dx,$$

d'où le résultat.  $\square$

**Modules de continuité dans  $L^1$  et  $L^2$ .** Rappelons que toute fonction  $f \in C_0(\mathbb{R}^n)$  est bornée et uniformément continue. Si pour tout  $a \in \mathbb{R}^n$  on désigne par  $\tau_a$  l'opérateur de translation par  $a$ , qui opère sur les fonctions selon

$$(\tau_a f)(t) = f(t - a)$$

on voit que la continuité uniforme de  $f$  sur  $\mathbb{R}^n$  se traduit par le fait que le module de continuité

$$\omega(f, a) = \|\tau_a f - f\| = \sup_t |f(t - a) - f(t)|$$

tend vers zéro quand  $a \rightarrow 0$  dans  $\mathbb{R}^n$ .

Lorsque  $f$  est élément de  $L^1$  et  $L^2$  respectivement, alors aussi la translátée  $\tau_a f$ . Ainsi l'opérateur  $\tau_a$  définit une isométrie (surjective) de  $L^1$  dans lui-même, ou de  $L^2$  dans lui-même, puisque

$$\|\tau_a f\|_1 = \int |f(t-a)| dt = \int |f(t)| dt = \|f\|_1$$

pour  $f \in L^1$ . Il est donc naturel d'introduire, comme pour le cas de  $C_0(\mathbb{R}^n)$ , les modules de continuité

$$\omega_1(f, a) = \|\tau_a f - f\|_1 \quad \text{si } f \in L^1$$

$$\omega_2(f, a) = \|\tau_a f - f\|_2 \quad \text{si } f \in L^2$$

Et alors :

(3.3.22) *Proposition*

- a) Pour toute  $f \in L^1$  on a  $\omega_1(f, a) \rightarrow 0$  quand  $a \rightarrow 0$  dans  $\mathbb{R}^n$ .
- b) Pour toute  $f \in L^2$  on a  $\omega_2(f, a) \rightarrow 0$  quand  $a \rightarrow 0$  dans  $\mathbb{R}^n$ .

*Preuve.* Prouvons b) et pour cela fixons  $f \in L^2$ . Pour tout  $\varepsilon > 0$ , il existe d'après (3.1.10) une fonction  $g \in \mathcal{K}(\mathbb{R}^n)$  telle que  $\|f - g\|_2 \leq \varepsilon$ . Comme on a  $\|\tau_a f - f\|_2 \leq 2\|f\|_2$ , on en déduit  $\|\tau_a f - f\|_2 \leq 2\varepsilon + \|\tau_a g - g\|_2$ , de sorte que la question est ramenée au cas  $g \in \mathcal{K}(\mathbb{R}^n)$ . Or il existe  $r > 0$  tel que la fonction  $g$  ait son support contenu dans la boule euclidienne  $B_r = B(0, r)$  de rayon  $r$ . Alors pour  $a \in \mathbb{R}^n$ , tel que  $\|a\| \leq 1$ , on a  $\text{supp}(\tau_a g) \subset B_{r+1}$  et ainsi

$$\|\tau_a f - f\|_2^2 = \int_{B_{r+1}} |g(t-a) - g(t)|^2 dt \leq \omega^2(g, a) \lambda_n[B_{r+1}]$$

On termine en remarquant que  $\omega(g, a) \rightarrow 0$  quand  $a \rightarrow 0$  d'après la continuité uniforme de  $g$ .  $\square$

*L'espace de Wiener  $\mathcal{W}(\mathbb{R}^n)$ .* Certaines difficultés dans l'étude de la transformation de Fourier proviennent du fait que  $\mathcal{F}$  opère de l'espace  $L^1$  dans un *autre* espace  $C_0(\mathbb{R}^n)$ , et on peut alors imaginer qu'il est intéressant d'introduire certains sous-espaces  $E \subset L^1 \cap C_0$ , invariants par l'opérateur  $\mathcal{F}$ . C'est évidemment le cas pour l'espace de Schwartz  $\mathcal{S} = \mathcal{S}(\mathbb{R}^n)$ . Mais ce n'est pas le cas pour l'espace  $\mathcal{K}(\mathbb{R}^n)$  des fonctions continues à support compact, ni même pour l'espace  $\mathcal{D}(\mathbb{R}^n)$  des fonctions  $C^\infty$  à support compact. Par exemple, il est facile de vérifier, pour  $n = 1$ , que si  $f \in \mathcal{K}(\mathbb{R})$  alors  $\mathcal{F}f$  est en fait une fonction analytique, de sorte que  $\mathcal{K} \cap \mathcal{F}\mathcal{K} = 0$

Alors si  $E \subset L^1 \cap C_0$  est tel que  $\mathcal{F}E \subset E$  on aura  $\mathcal{F}f \in L^1$  pour toute  $f \in E$  et ainsi, en accord avec la seconde partie de (3.3.21), on voit apparaître le plus grand espace  $E$  comme étant défini par la condition  $f \in L^1$  et  $\mathcal{F}f \in L^1$ . Mais la formule de réciprocity (3.3.21) montre aussi que les conditions précédentes équivalent à la seule condition  $f \in L^1 \cap \mathcal{F}L^1$ . D'où la définition :

(3.3.23) *Définition*

On appelle espace de Wiener l'espace  $\mathcal{W} = L^1 \cap \mathcal{F}L^1$ . C'est aussi l'espace des fonctions continues  $f$  telles que  $f \in L^1$  et  $\mathcal{F}f \in L^1$ . C'est encore le plus grand sous-espace de  $L^1 \cap C_0$  invariant par transformation de Fourier.

On munit généralement  $\mathcal{W}$  de la norme

$$N(f) = \|f\|_1 + \|\mathcal{F}f\|_1$$

qui en fait un espace de Banach.

Avec ce choix il est clair que  $\mathcal{F}$  et  $\overline{\mathcal{F}}$  opèrent tous deux de  $\mathcal{W}$  dans  $\mathcal{W}$ . Mais en fait on a :

(3.3.24) *Théorème*

La transformée de Fourier  $\mathcal{F}$  est un isomorphisme de l'espace  $\mathcal{W}$  sur lui-même, admettant l'opérateur  $(2\pi)^{-n} \overline{\mathcal{F}}$  comme isomorphisme réciproque. De plus pour  $f, g \in \mathcal{W}$  on a  $f \star g \in \mathcal{W}$  et  $fg \in \mathcal{W}$ , de sorte que  $\mathcal{W}$  est en fait à la fois une algèbre convolutive et une algèbre multiplicative.

**Preuve.** La première partie n'est autre que (3.3.21). Si maintenant  $f$  et  $g$  sont éléments de  $\mathcal{W}$ , alors  $f \star g \in L^1$  et  $\mathcal{F}(f \star g) = \mathcal{F}f \cdot \mathcal{F}g$  est élément de  $L^1$  car  $\mathcal{F}f \in L^1$  et  $\mathcal{F}g \in C_0$ . Donc  $\mathcal{W}$  est stable par convolution, donc aussi par multiplication puisque l'isomorphisme  $\mathcal{F}$  transforme précisément l'une en l'autre.  $\square$

Ainsi, si l'espace  $\mathcal{K} = \mathcal{K}(\mathbb{R}^n)$  est bien adapté à la théorie des espaces  $L^1$  et  $L^2$ , surtout d'ailleurs à cause des résultats de densité, l'espace de Wiener  $\mathcal{W}$  est bien adapté à la transformation de Fourier. Et ceci sans même perdre le bénéfice de la densité puisque :

(3.3.25) *Proposition*

L'espace de Wiener  $\mathcal{W}$  est dense dans chacun des espaces  $L^1$ ,  $L^2$  et  $C_0(\mathbb{R}^n)$ .

**Preuve.** On sait avec (3.1.10) que  $\mathcal{K}$  est dense dans  $L^1$  et  $L^2$ , et aussi évidemment dans  $C_0$ . Par ailleurs pour toute  $f \in L^1$  la fonction continue  $g_\sigma \star f$  est élément de  $\mathcal{W}$  puisque  $\mathcal{F}(g_\sigma \star f) = G_\sigma \mathcal{F}f \in L^1$ . Il suffit donc de prouver que pour toute  $f \in \mathcal{K}$  on a  $g_\sigma \star f \rightarrow f$  dans chacun des espaces  $L^1$ ,  $L^2$  et  $C_0$ . Pour  $L^1$ , c'est (3.3.19), avec même  $f \in L^1$ . Pour  $C_0$ , on a, en introduisant le module de continuité  $\omega(f, t)$ ,

$$\|g_\sigma \star f - f\| \leq \int \omega(f, t) g_\sigma(t) dt$$

et la preuve de (3.3.19) montre que  $\int \omega(f, t) g_\sigma(t) dt \rightarrow 0$  quand  $\sigma \downarrow 0$ . Enfin pour  $L^2$  il suffit de voir, avec toujours  $f \in \mathcal{K}$ , que l'on a

$$\|g_\sigma \star f - f\|_2^2 \leq \|g_\sigma \star f - f\| \|g_\sigma \star f - f\|_1$$

pour conclure.  $\square$

(3.3.26) *Exercice.* (Les espaces  $\mathcal{F}L^1$  et  $C_0$ )

- a) Montrer que  $\mathcal{F}L^1$  est dense dans  $C_0(\mathbb{R}^n)$ .  
 b) Soit  $f \in L^1(\mathbb{R})$  une fonction impaire. Établir l'égalité pour  $a, b > 0$  :

$$\int_b^a \frac{(\mathcal{F}f)(x)}{x} dx = 2i \int_0^\infty f(t) \left[ \int_{at}^{bt} \frac{\sin u}{u} du \right] dt$$

et en déduire que

$$\int_0^\infty \frac{(\mathcal{F}f)(x)}{x} dx = i\pi \int_0^\infty f(t) dt$$

où l'intégrale, au premier membre, est une intégrale de Riemann impropre.

- c) Soit  $g \in C_0(\mathbb{R})$  la fonction impaire définie par

$$g(x) = \frac{1}{1 + |\ln x|} \quad \text{pour } x \geq 0.$$

Montrer que  $g \notin \mathcal{F}L^1$ .

(3.3.27) *Exercice*

On fixe la fonction  $f \in L^1 \cap C_0(\mathbb{R}^n)$  supposée telle que  $\mathcal{F}f \geq 0$ .

- a) Démontrer que  $\int G_\sigma(x) (\mathcal{F}f)(x) dx = (2\pi)^n \int g_\sigma(t) f(t) dt$ .  
 b) En déduire que  $f$  est élément de l'espace de Wiener  $\mathcal{W}$ .

(3.3.28) *Exercice*

Sur l'espace  $\mathcal{W} = \mathcal{W}(\mathbb{R}^n)$  on considère l'opérateur

$$U = \frac{1}{(2\pi)^{\frac{n}{2}}} \mathcal{F}.$$

Établir que  $U^2 f = f$  pour toute  $f \in \mathcal{W}$  et en déduire que  $U^4 = I$ , opérateur identité.

*La formule d'inversion pour les mesures bornées.* On revient aux espaces  $L^1(\mathbb{R}^n)$  et  $M^1(\mathbb{R}^n)$  et, de même que pour l'espace  $L^1$ , on veut montrer que la transformation de Fourier est injective sur  $M^1 = M^1(\mathbb{R}^n)$ , cône positif des mesures bornées. On peut même se poser la question de l'injectivité sur l'ensemble  $L^1 \cup M^1$ .

(3.3.29) *Lemme*

On fixe  $\mu, \nu \in M^1$  et  $f \in L^1$ . Alors

$$\text{a) } \int (\mathcal{F}\nu) d\mu = \int (\mathcal{F}\mu) d\nu$$

$$\text{b) } \int (\mathcal{F}f) d\mu = \int (\mathcal{F}\mu)(x) f(x) dx$$

*Preuve.* Avec Fubini. □

(3.3.30) *Proposition*

Pour toute  $\mu \in M^1$  et toute  $g \in \mathcal{W}$ , on a

$$\int g d\mu = \frac{1}{(2\pi)^n} \int (\overline{\mathcal{F}g})(x) (\mathcal{F}\mu)(x) dx$$

*Preuve.* On applique le lemme à la fonction  $f = (2\pi)^{-n} \overline{\mathcal{F}g}$ , compte tenu que  $g = \mathcal{F}f$ . □

La proposition signifie que la transposée de Fourier  $\mathcal{F}\mu$  détermine complètement l'intégrale par rapport à  $\mu$  sur l'espace de Wiener  $\mathcal{W}$ . Comme  $\mathcal{W}$  est partout dense dans  $C_0(\mathbb{R}^n)$  d'après (3.5.25), on a ainsi un moyen de déterminer l'intégrale par rapport à  $\mu$  sur  $C_0(\mathbb{R}^n)$ . D'où le résultat :

(3.3.31) *Théorème. (Injectivité)*

Soit  $\mu$  et  $\nu$  deux mesures bornées sur  $\mathbb{R}^n$ . Pour que  $\mu = \nu$  il faut et il suffit que  $\mathcal{F}\mu = \mathcal{F}\nu$ .

*Preuve.* L'égalité de  $\mu$  et  $\nu$  sur  $C_0(\mathbb{R}^n)$  garantit que  $\mu = \nu$  d'après (1.11.1), partie unicité du théorème de Riesz-Alexandroff. □

En introduisant les images  $\mu_a$  de  $\mu$  par chacune des applications linéaires  $x \rightarrow \langle a, x \rangle$ , où  $a \in \mathbb{R}^n$ , on voit que  $(\mathcal{F}\mu)(a) = \mathcal{F}\mu_a(1)$ , d'où

(3.3.32) *Corollaire 1*

Pour que deux mesures bornées  $\mu$  et  $\nu$  sur  $\mathbb{R}^n$  soient égales, il faut et il suffit que leurs marges  $\mu_a$  et  $\nu_a$  soient les mêmes pour tout  $a \in \mathbb{R}^n$ .

On remarquera que si  $(e_1, e_2, \dots, e_n)$  est la base canonique de  $\mathbb{R}^n$ , la condition que les marges  $\mu_k = \mu_{e_k}$  soient les mêmes pour  $\mu$  et  $\nu$  ne suffit pas à assurer l'égalité  $\mu = \nu$ .

(3.3.33) *Corollaire 2*

Soit  $X$  et  $Y$  deux vecteurs aléatoires à valeurs dans le même espace  $\mathbb{R}^n$ . Les assertions suivantes sont équivalentes :

- Les vecteurs  $X$  et  $Y$  ont même loi.
- Ils ont même fonction caractéristique.
- Pour tout  $a \in \mathbb{R}^n$ , les variables aléatoires  $\langle a, X \rangle$  et  $\langle a, Y \rangle$  ont même loi.

*Remarque.* La preuve du théorème général d'injectivité donnée ici est assez indirecte, mais c'est la plus rapide en dimension  $n$  quelconque. On prendra garde que la formule (3.3.10) ne peut s'étendre au cas des fonctions boréliennes bornées  $g$ . En effet si c'était le cas, on aurait nécessairement  $\mu A = 0$  pour tout borélien  $A$  négligeable pour la mesure de Borel (puisqu'alors  $\overline{\mathcal{F}} 1_A = 0$ ), ce qui est faux en général.

Le cas particulier  $n = 1$  peut toutefois se traiter directement sans passer par l'espace de Wiener  $\mathcal{W}$ . En effet

(3.3.34) *Proposition*

Soit  $\varphi = \mathcal{F}\mu$  avec  $\mu \in M^1(\mathbb{R})$ . Pour  $a < b$  on a les formules, dites de Perron-Stieltjes,

$$a) \mu(\{a\}) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T e^{-iax} \varphi(x) dx$$

$$b) \frac{1}{2} \mu(\{a\} + \frac{1}{2} \mu(\{b\} + \mu] a, b[ = \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-iax} - e^{-ibx}}{ix} \varphi(x) dx$$

*Preuve.* Il suffit de prouver a) pour  $a = 0$  et de remplacer ensuite  $\mu$  par sa translatée  $\tau_{-a} \mu$ . Or, avec Fubini, on a

$$\frac{1}{2\pi} \int_{-T}^T \varphi(x) dx = \int J_T(t) d\mu(t)$$

$$\text{où } J_T(t) = \frac{1}{2\pi} \int_{-T}^T e^{itx} dx = \begin{cases} 1 & \text{si } t = 0 \\ \frac{\sin tT}{tT} & \text{si } t \neq 0 \end{cases}$$

et il suffit d'appliquer le théorème de convergence dominée puisque  $|J_T(t)| \leq 1$  et  $J_T(t) \rightarrow 1_{\{0\}}(t)$  lorsque  $T \rightarrow \infty$ .

Pour prouver b) on procède de même avec

$$\frac{1}{2\pi} \int_{-T}^T \frac{e^{-iax} - e^{-ibx}}{ix} \varphi(x) dx = \int K_T(t) d\mu(t)$$

où 
$$K_T(t) = \frac{1}{\pi} \int_0^T \frac{\sin x(b-t) - \sin x(a-t)}{x} dx = L_T(b-t) - L_T(a-t)$$

avec  $L_T(\alpha) = \frac{1}{\pi} \int_0^T \frac{\sin \alpha x}{x} dx$ . Or  $L_T(0) = 0$  et  $L_T$  est une fonction impaire. De plus pour  $\alpha > 0$  on a

$$L_T(\alpha) = \frac{1}{\pi} \int_0^{\alpha T} \frac{\sin u}{u} du = H(\alpha T)$$

avec  $H(x) = \frac{1}{\pi} \int_0^x \frac{\sin u}{u} du$ .

Or on sait que  $H$  est continue et telle que  $H(x) \rightarrow \frac{1}{2}$  quand  $x \rightarrow +\infty$ . D'où résulte que  $H$  est bornée, donc  $|K_T(t)| \leq 2 \|H\|$ , ce qui donne la condition de convergence dominée. Et puisque  $\lim_{T \rightarrow \infty} H(\alpha T)$  vaut respectivement  $-\frac{1}{2}$ ,  $0$  ou  $\frac{1}{2}$  suivant que  $\alpha < 0$ ,  $\alpha = 0$  ou  $\alpha > 0$ , on voit que

$$\lim_{T \rightarrow \infty} K_T(t) = \begin{cases} 0 & \text{si } t \notin [a, b] \\ \frac{1}{2} & \text{si } t = a \text{ ou } t = b \\ 1 & \text{si } t \in ]a, b[ \end{cases}$$

ce qui donne b). □

On remarquera que la proposition garantit l'injectivité de la transformation de Fourier, car si  $\mathcal{F}\mu = \mathcal{F}\nu$ , alors  $\mu[a, b[ = \nu[a, b[$  pour tous  $a, b$ , donc  $\mu = \nu$  par le raisonnement habituel.

(3.3.35) *Exercice*

Soit  $\mu \in M^1(\mathbb{R}^n)$  et  $\varphi = \mathcal{F}\mu$ . Pour chaque  $T > 0$  soit  $P_T$  le cube  $[-T, +T]^n$ .

a) Prouver que  $\mu(\{0\}) = \lim_{T \rightarrow \infty} \frac{1}{(2T)^n} \int_{P_T} \varphi(x) dx$

b) En remplaçant  $\mu$  par la mesure  $\nu = \mu \star \overset{\vee}{\mu}$ , où  $\overset{\vee}{\mu}$  est la symétrisée de  $\mu$ , définie par  $\overset{\vee}{\mu}(A) = \mu(-A)$ , obtenir la relation

$$\sum_{a \in \mathbb{R}^n} \mu(\{a\})^2 = \lim_{T \rightarrow \infty} \frac{1}{(2T)^n} \int_{P_T} |\varphi(x)|^2 dx$$

(On remarquera que la somme  $\Sigma$  est dénombrable).

- c) Montrer que si l'on a  $\varphi \in L^1$ ,  $\varphi \in L^2$  ou  $\varphi \in C_0(\mathbb{R}^n)$ , alors  $\mu$  est une mesure diffuse sur  $\mathbb{R}^n$ .

(3.3.36) **Exercice. (Injectivité dans le cas mixte)**

- a) Soit  $\mu \in M^1(\mathbb{R}^n)$  et  $f \in L^1$ . On suppose  $\mathcal{F}\mu = \mathcal{F}f$ . Montrer alors que l'on a  $f \geq 0$  presque partout et que  $\mu$  est la mesure de densité  $f$ , autrement dit  $d\mu(x) = f(x) dx$ . On prouvera d'abord que  $f$  est réelle et on introduira les mesures  $\nu^+$  et  $\nu^-$ , de densités respectives  $f^+$  et  $f^-$ .
- b) On suppose  $\mathcal{F}\mu = \varphi \in L^1$ . Montrer que  $\mu$  admet la densité  $f = (2\pi)^{-n} \overline{\mathcal{F}\varphi}$ .
- c) On voit donc que chacune des conditions  $\mathcal{F}\mu \in L^1$  ou  $\mathcal{F}\mu \in \mathcal{F}L^1$  implique que  $\mu$  est Borel-densifiable. On va voir qu'il n'en est rien avec la condition  $\mathcal{F}\mu \in C_0(\mathbb{R}^n)$ . Pour cela revenons au cas  $n = 2$  et à la probabilité  $\mu$  sur  $\mathbb{R}^2$  égale à la mesure uniforme sur le cercle unité  $x^2 + y^2 = 1$ , définie par

$$\iint f(x, y) d\mu(x, y) = \frac{1}{2\pi} \int_0^{2\pi} f(\cos \theta, \sin \theta) d\theta.$$

Montrer alors que

$$(\mathcal{F}\mu)(x, y) = \frac{1}{\pi} \int_0^\pi e^{ir \cos \theta} d\theta = \frac{1}{\pi} \int_{-1}^1 \frac{e^{irt}}{\sqrt{1-t^2}} dt$$

avec  $r = \sqrt{x^2 + y^2}$ , et prouver, avec le théorème de Riemann-Lebesgue, que  $\mathcal{F}\mu \in C_0(\mathbb{R}^n)$ . Conclure en remarquant que  $\mu$  n'est pas Borel-densifiable sur  $\mathbb{R}^2$ .

### 3.4 Transformation de Fourier et convolution sur $L^2$

On va maintenant pouvoir définir la transformation de Fourier sur l'espace  $L^2 = L^2(\mathbb{R}^n)$  grâce aux deux propriétés suivantes déjà démontrées :

- a)  $\mathcal{F}$  opère de l'espace de Wiener  $\mathcal{W}$  dans lui-même.  
 b) L'espace  $\mathcal{W}$  est partout dense dans  $L^2$

Rappelons auparavant que l'espace  $L^2$ , formé ici de fonctions réelles ou complexes, est un espace de Hilbert lorsqu'on le munit du produit scalaire  $(f | g) = \int f(x) \overline{g(x)} dx$ .

Désignons maintenant par  $\mathcal{W}_2$  l'espace de Wiener  $\mathcal{W}$  muni de la norme  $\|\cdot\|_2$  de  $L^2$ . Cela étant, on a pour  $f, g \in L^1$  quelconques, par Fubini

$$\begin{aligned} \int (\mathcal{F}f) \bar{g} dx &= \iint e^{i\langle x, t \rangle} f(t) \bar{g}(x) dx dt \\ &= \int \overline{f(\overline{\mathcal{F}}g)} dt \end{aligned}$$

En particulierisant  $f, g$  dans  $\mathcal{W}$ , et en tenant compte de (3.3.24) on a

(3.4.1) **Théorème**

Sur l'espace préhilbertien  $\mathcal{W}_2$  les opérateurs

$$U = (2\pi)^{-\frac{n}{2}} \mathcal{F} \quad \text{et} \quad \bar{U} = (2\pi)^{-\frac{n}{2}} \bar{\mathcal{F}}$$

sont deux isométries réciproques, adjointes l'une de l'autre.

*Preuve.* L'égalité précédente signifie  $(\mathcal{F}f | g) = (f | \bar{\mathcal{F}}g)$  pour  $f, g \in \mathcal{W}_2$ , donc aussi

$(Uf | g) = (f | \bar{U}g)$ . Mais si  $g = \mathcal{F}f$  alors  $\bar{\mathcal{F}}g = \bar{\mathcal{F}}\mathcal{F}f = (2\pi)^n f$ , d'où  $\|\mathcal{F}f\|_2^2 = (2\pi)^n \|f\|_2^2$  et

ainsi  $\|Uf\|_2^2 = \|f\|_2^2$ . □

(3.4.2) **Définition**

On appelle transformation de Fourier sur  $L^2$  l'opérateur, noté encore  $\mathcal{F} : L^2 \rightarrow L^2$ , qui prolonge canoniquement la transformation de Fourier sur

$\mathcal{W}_2$ . On définit de même l'opérateur  $\bar{\mathcal{F}} : L^2 \rightarrow L^2$ .

Par des raisonnements de continuité évidents, on a

(3.4.3) **Théorème (Plancherel)**

Sur l'espace  $L^2 = L^2(\mathbb{R}^n)$  les opérateurs  $U = (2\pi)^{-\frac{n}{2}} \mathcal{F}$  et  $\bar{U} = (2\pi)^{-\frac{n}{2}} \bar{\mathcal{F}}$  sont deux isométries réciproques, adjointes l'une de l'autre. Autrement dit, on a pour  $f, g \in L^2$

a)  $\mathcal{F}\bar{\mathcal{F}}f = \bar{\mathcal{F}}\mathcal{F}f = (2\pi)^n f$

b)  $(\mathcal{F}f | g) = (f | \bar{\mathcal{F}}g)$

c)  $(\mathcal{F}f | \mathcal{F}g) = (2\pi)^n (f | g)$

d)  $\|\mathcal{F}f\|_2^2 = (2\pi)^n \|f\|_2^2$

*Remarque.* Il convient de bien voir que lorsque  $f$  décrit l'espace  $L^2$  la fonction  $\mathcal{F}f$  n'a aucune propriété particulière de continuité sur  $\mathbb{R}^n$  puisqu'elle décrit, elle aussi, l'espace  $L^2$  tout entier. Par ailleurs il se pose un problème de notation lorsque  $f \in L^1 \cap L^2$ , car il existe a priori deux images de Fourier pour  $f$  : celle notée  $\varphi_1 \in C_0(\mathbb{R}^n)$  image de  $f \in L^1$ , et celle notée  $\varphi_2 \in L^2$ , image de  $f \in L^2$ . Or pour toute fonction  $h \in \mathcal{W}$  on a

$$\int \varphi_1 \bar{h} dx = \int f \overline{\overline{\mathcal{F}h}} dx = (f | \overline{\mathcal{F}h}) = (\varphi_2 | h)$$

où la première égalité s'obtient avec Fubini et la dernière avec (3.4.3). Fixons une fonction  $g \in \mathcal{W}$  qui ne s'annule pas sur  $\mathbb{R}^n$ , par exemple la fonction de Gauss. On a alors  $\bar{g}h \in \mathcal{W}$ ,  $\varphi_1 g \in C_0 \cap L^1$  donc  $\varphi_1 g \in L^2$  et  $\varphi_2 g \in L^2$ , de sorte qu'en remplaçant  $h$  par  $\bar{g}h$  on obtient  $((\varphi_1 - \varphi_2)g | h) = 0$ , ce qui prouve que la fonction  $(\varphi_1 - \varphi_2)g \in L^2$  est orthogonale à  $\mathcal{W}_2$ . Elle est donc nulle dans  $L^2$  puisque  $\mathcal{W}_2$  est dense dans  $L^2$ , et ainsi  $(\varphi_1 - \varphi_2)g = 0$  presque partout, d'où  $\varphi_1 = \varphi_2$  presque partout. En résumé, on vérifie la cohérence de la notation  $\mathcal{F}f$  sous la forme :

(3.4.4) *Proposition*

Pour toute  $f \in L^1 \cap L^2$  l'image de Fourier  $\mathcal{F}f \in L^2$  est exactement la classe de la fonction  $\mathcal{F}f \in C_0(\mathbb{R}^n)$  définie par (3.3.1).

**Calcul de  $\mathcal{F}f$  pour  $f \in L^2$ .** La définition de  $\mathcal{F}f \in L^2$  étant donnée par passage à la limite, le calcul de  $\mathcal{F}f$  se fait par approximation. Donnons toutefois deux cas particuliers intéressants.

(3.4.5) *Proposition*

Soit  $f \in L^1$  tel que  $\overline{\mathcal{F}f} \in L^2$ . Alors la fonction  $h = \overline{\mathcal{F}f}$  est telle que  $\mathcal{F}h = (2\pi)^n f$  dans  $L^2$ .

*Preuve.* Reprenons les fonctions gaussiennes  $g_\sigma$  et  $G_\sigma$  de (3.3.18). On sait que  $g_\sigma \star f \rightarrow f$  dans  $L^1$  quand  $\sigma \downarrow 0$  et  $\overline{\mathcal{F}}(g_\sigma \star f) = G_\sigma \overline{\mathcal{F}f}$ . Or  $G_\sigma \overline{\mathcal{F}f} \rightarrow \overline{\mathcal{F}f}$  dans  $L^2$ , comme on voit avec le théorème de convergence dominée de Lebesgue. Donc  $g_\sigma \star f = (2\pi)^{-n} \mathcal{F}(G_\sigma \overline{\mathcal{F}f})$  tend vers  $(2\pi)^{-n} \mathcal{F} \overline{\mathcal{F}f} = (2\pi)^{-n} \mathcal{F}h$  dans l'espace  $L^2$ . Ainsi  $g_\sigma \star f$  a pour limite  $f$  dans

l'espace  $L^1$  et  $(2\pi)^{-n} \mathcal{F}h$  dans l'espace  $L^2$ , d'où l'on déduit  $\int_A f dx = (2\pi)^{-n} \int_A \mathcal{F}h dx$  pour tout borélien borné  $A$ , et ainsi  $f = (2\pi)^{-n} \mathcal{F}h$  presque partout, d'où l'égalité  $\mathcal{F}h = (2\pi)^n f$  dans  $L^2$ .  $\square$

*Exemple 1.* On fixe  $a > 0$  et soit  $f_a(t) = Y(t) e^{-at}$ , où  $Y$  est la fonction de Heaviside  $Y = 1_{]0,+\infty[}$ . Alors  $f_a \in L^1$  et  $(\overline{\mathcal{F}} f_a)(x) = \frac{1}{a + ix}$ , de sorte que  $\overline{\mathcal{F}} f_a \in L^2$ . Ainsi la fonction  $\frac{1}{a + it}$  appartient à  $L^2$  sans appartenir à  $L^1$  et

$$\mathcal{F} \left( \frac{1}{a + it} \right) = 2\pi Y(x) e^{-ax}$$

n'est pas continue. En séparant partie paire et partie impaire, on obtient

$$\mathcal{F} \left( \frac{a}{a^2 + t^2} \right) = \pi e^{-a|x|}$$

$$\mathcal{F} \left( \frac{it}{a^2 + t^2} \right) = -\pi \varepsilon(x) e^{-a|x|} \quad \text{avec } \varepsilon(x) = \text{sgn}(x).$$

*Exemple 2.* Soit  $f_a = 1_{[-a,a]}$ . Alors  $\overline{\mathcal{F}} f_a(x) = 2 \frac{\sin ax}{x}$  et ainsi  $\overline{\mathcal{F}} f_a \in L^2$ , sans être élément de  $L^1$ . Alors la fonction

$$\mathcal{F} \left( \frac{\sin at}{t} \right) = \pi 1_{[-a,a]}$$

n'est pas non plus continue.

Le second cas particulier est basé sur le fait que pour tout compact  $K$  de  $\mathbb{R}^n$  la fonction  $1_K$  est élément de  $L^2$ , donc  $1_K f \in L^1$  pour toute  $f \in L^2$ . Soit alors  $(K_p)$  une suite croissante de compacts telle que  $\bigcup K_p = \mathbb{R}^n$ , de sorte que l'on a  $1_{K_p} \uparrow 1$  et par Beppo Lévi  $1_{K_p} f \rightarrow f$  dans  $L^2$ . Alors  $\mathcal{F} f$  est la limite de  $\mathcal{F}(1_{K_p})$  et par (3.4.4) on a :

(3.4.6) **Proposition**

Soit  $(K_p)$  une suite croissante de compacts, de réunion  $\mathbb{R}^n$  (par exemple  $K_p = B(0, p)$  ou  $K_p = [-p, p]^n$ ). Alors pour toute  $f \in L^2$  l'image de Fourier  $\mathcal{F} f \in L^2$  est la limite dans  $L^2$  de la suite des fonctions, éléments de  $C_0(\mathbb{R}^n)$ ,

$$x \rightarrow \int_{K_p} e^{i\langle x, t \rangle} f(t) dt$$

**Produit de convolution dans  $L^2$ .** Pour  $f, g \in L^2$  la fonction  $t \rightarrow f(x - t) g(t)$  est évidemment intégrable pour tout  $x \in \mathbb{R}^n$ . On est donc en droit de poser

$$(f \star g)(x) = \int f(x - t) g(t) dt$$

et par Cauchy-Schwarz on a  $|(f \star g)(x)| \leq \|f\|_2 \|g\|_2$ .

Par ailleurs si  $h = f \star g$  alors

$$|h(x) - h(y)| \leq \|g\|_2 \left[ \int |f(x-t) - f(y-t)|^2 dt \right]^{1/2}$$

et l'on reconnaît dans  $[ \ ]^{1/2}$  le module de continuité  $\omega_2(f, x-y)$  d'où

$$|h(x) - h(y)| \leq \|g\|_2 \omega_2(f, x-y)$$

ce qu'on peut traduire encore par l'inégalité

$$\omega(h, \delta) \leq \|g\|_2 \omega_2(f, \delta) .$$

En particulier  $h = f \star g$  est uniformément continue et bornée et  $\|f \star g\| \leq \|f\|_2 \|g\|_2$ .  
Rajoutons à cela que pour  $f, g \in \mathcal{K}(\mathbb{R}^n)$ , on a aussi  $h = f \star g \in \mathcal{K}(\mathbb{R}^n)$ , car si  $f$  a pour support compact  $A$  et  $g$  pour support compact  $B$  alors  $\text{supp}(f \star g) \subset A + B$ , et  $A + B$  est compact. Et ainsi, par densité de  $\mathcal{K}(\mathbb{R}^n)$  dans  $L^2$ , on voit que  $f \star g \in C_0(\mathbb{R}^n)$ .

D'où :

(3.4.7) *Proposition*

Le produit de convolution

$$(f \star g)(x) = \int f(x-t) g(t) dt$$

définit une application bilinéaire continue de  $L^2 \times L^2$  dans l'espace  $C_0(\mathbb{R}^n)$ , qui est symétrique.

Remarquons encore que si le produit  $\star$  est bien commutatif, la question de l'associativité ne se pose pas puisqu'on ne peut pas définir un produit de convolution du type  $C_0 \star L^2$ .

De même la condition  $f \star g \in C_0(\mathbb{R}^n)$  ne permet pas en général de calculer  $\mathcal{F}(f \star g)$ , de sorte que la formule de (3.3.6) ne peut se prolonger. Par contre la condition  $f \in L^2$  et  $g \in L^2$  implique  $fg \in L^1$ , de sorte qu'on peut calculer  $\mathcal{F}(fg)$ . Et alors :

(3.4.8) *Proposition*

Soit  $f, g \in L^2$  quelconques.

a) On a l'égalité dans  $C_0(\mathbb{R}^n)$

$$\mathcal{F}(fg) = (2\pi)^{-n} \mathcal{F}f \star \mathcal{F}g$$

b) En particulier on a aussi  $f \star g \in \mathcal{F}L^1 \subset C_0(\mathbb{R}^n)$ .

*Preuve.* La formule  $\mathcal{F}(f \star g) = \mathcal{F}f \cdot \mathcal{F}g$ , valable pour  $f, g \in \mathcal{W}_2$  est aussi équivalente, compte tenu de la formule de réciprocity et du caractère bijectif de  $\mathcal{F} : \mathcal{W}_2 \rightarrow \mathcal{W}_2$ , à la formule  $\mathcal{F}(fg) = (2\pi)^{-n} \mathcal{F}f \star \mathcal{F}g$ . Or l'application  $(f, g) \rightarrow \mathcal{F}(fg)$  et l'application  $(f, g) \rightarrow (2\pi)^{-n} \mathcal{F}f \star \mathcal{F}g$  sont toutes deux définies sur  $L^2$  et continues sur  $L^2$ . Leur coïncidence sur  $\mathcal{W}_2 \times \mathcal{W}_2$  assure donc leur égalité, d'où a).

Par le théorème de Plancherel on peut aussi écrire a) sous la forme

$$f \star g = (2\pi)^{-n} \mathcal{F} \left[ \overline{\mathcal{F} f} \bullet \overline{\mathcal{F} g} \right]$$

ce qui donne  $f \star g \in \mathcal{F}L^1$ , en précisant donc un peu (3.4.7). □

Terminons par quelques exercices.

(3.4.9) *Exercice*

- a) Utiliser l'exemple 2 après (3.4.5) pour déterminer la transformée de Fourier de la fonction  $\frac{\sin^2 t}{t^2}$ .

En déduire la valeur des intégrales  $\int_0^\infty \frac{\sin^2 t}{t^2} dt$  et  $\int_0^\infty \frac{\sin^4 t}{t^4} dt$ .

- b) En utilisant la transformée de Fourier de la fonction  $e^{-a|t|}$ , avec  $a > 0$ , obtenir l'égalité, pour  $a > 0$ :

$$\int_0^\infty e^{-at} \frac{\sin t}{t} dt = \text{Arc tg } \frac{1}{a}$$

Que peut-on en déduire lorsque  $a \downarrow 0$  ?

- c) Évaluer de même l'intégrale

$$\int_0^\infty e^{-at} \frac{\sin^2 t}{t^2} dt \quad \text{pour } a \geq 0.$$

- d) Retrouver les résultats de (1.10.4. Ex 3), à savoir les égalités

$$\frac{1}{\pi} \int_0^\infty \frac{\sin t}{t} \frac{a dt}{a^2 + t^2} = \frac{1}{2a} (1 - e^{-a}), \quad a > 0$$

$$\frac{1}{\pi} \int_0^\infty \frac{\sin^2 t}{t^2} \frac{a dt}{a^2 + t^2} = \frac{1}{2a^2} (2a - 1 + e^{-a}), \quad a > 0$$

(3.4.10) *Exercice*

Utiliser (3.4.8) pour déterminer l'image de Fourier de la fonction

$$\frac{ab}{\pi^2} \frac{1}{(a^2 + t^2)(b^2 + t^2)} \quad a, b > 0$$

Examiner le cas  $a = b$ .

(3.4.11) *Exercice*

Soit  $f \in L^2(\mathbb{R})$  une fonction telle que  $f(t) = 0$  si  $|t| > \frac{1}{2}$  et  $\int |f(t)|^2 dt = 1$ .

- Montrer que  $\varphi = \mathcal{F}f$  est élément de l'espace  $L^2 \cap C_0(\mathbb{R})$ .
- Soit  $\mu$  la mesure de densité  $\frac{1}{2\pi} |\varphi(x)|^2$  par rapport à la mesure de Borel sur  $\mathbb{R}$ . Montrer que les fonctions  $e^{inx}$ ,  $n \in \mathbb{Z}$ , forment un système orthonormal dans l'espace de Hilbert  $L^2(\mu)$ . On introduira les translatées  $f_n = \tau_n f$ .

(3.4.12) *Exercice*

- Montrer que l'algèbre convolutive  $L^1 = L^1(\mathbb{R}^n)$  n'a pas d'élément unité.
- Montrer que l'équation  $f \star f = 0$  n'a pas d'autre solution que la solution  $f = 0$ , dans chacun des espaces  $L^1$  et  $L^2$ .
- Montrer que l'équation  $f \star f = f$  n'a pas d'autre solution que la solution  $f = 0$  dans l'espace  $L^1$ .
- Peut-on facilement décrire toutes les solutions de l'équation  $f \star f = f$  dans l'espace  $L^2$ ? Donner des exemples concrets pour  $n = 1$ .
- Résoudre l'équation  $\mu \star \mu = \mu$  dans l'ensemble  $M^1(\mathbb{R}^n)$ .

(3.4.13) *Exercice*

Pour toute partie  $A \subset \mathbb{R}^n$  on désigne par  $A - A$  l'ensemble des points  $a - b$ , où  $a \in A$  et  $b \in A$ . On suppose que  $A$  est une partie intégrable de  $\mathbb{R}^n$ .

- Montrer que tout point  $x \in \mathbb{R}^n$  tel que  $(1_A \star \overset{\vee}{1}_A)(x) > 0$  appartient à l'ensemble  $A - A$ .
- En déduire que si la mesure de Borel de  $A$  est strictement positive, alors  $A - A$  est un voisinage de zéro dans  $\mathbb{R}^n$  (théorème de Steinhaus).