Yves Coudène

L3

La mesure du hasard

Théorie de la mesure, intégration et probabilités

Cours et exercices corrigés



ellipses

Références sciences

La mesure du hasard

Théorie de la mesure, intégration et probabilités

Cours et exercices corrigés

Yves Coudène

L3

ellipses

Collection Références sciences

dirigée par Paul de Laboulaye
paul.delaboulaye@editions-ellipses.fr

Retrouvez tous les livres de la collection et des extraits sur www.editions-ellipses.fr

ISBN 9782340-118966
Dépôt légal : octobre 2026
©Ellipses Édition Marketing S.A.
8/10 rue la Quintinie 75015 Paris

DANGER LE PHOTOCOPILLAGE TUE LE LIVRE

Le Code de la propriété intellectuelle et artistique n'autorisant, aux termes des alinéas 2 et 3 de l'article L. 122-5, d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale, ou partielle, faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause, est illicite » (alinéa 1er de l'article L. 122-4).
Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

www.editions-ellipses.fr

Introduction

Il revient à Blaise Pascal le mérite d'avoir compris que les mathématiques pouvaient s'appliquer à l'étude du hasard et d'avoir proposé de développer une discipline *entièrement nouvelle* qu'il nomme « aleae geometria », la géométrie du hasard. Il faut cependant attendre le début du vingtième siècle pour que les progrès de l'analyse mathématique puissent fournir des fondements rigoureux à ce que nous appelons maintenant la théorie moderne des probabilités.

Dans son traité sur le triangle arithmétique écrit en 1654 [Pa], Blaise Pascal étudie en détail les coefficients binomiaux afin de résoudre des problèmes de combinatoire relatifs aux jeux de hasard [Der]. Il fonde ainsi la théorie des probabilités discrètes. L'idée des probabilités continues apparaît un siècle plus tard dans les travaux de De Moivre et Laplace, sous une forme intuitive reliant probabilités et aires sous les courbes. On leur doit une découverte majeure qui est aujourd'hui utilisée dans tous les domaines des sciences humaines et exactes. De toutes les lois mathématiques dont la science moderne fait usage, la loi de De Moivre-Laplace est certainement la plus célèbre. Elle est passée dans le langage courant sous le nom de loi normale, elle est aussi dénommée loi gaussienne en l'honneur de Carl Friedrich Gauss. Conceptuellement simple et mathématiquement profonde, elle est devenue le symbole de l'universalité des phénomènes aléatoires. Son étude va requérir tout le génie des analystes du dix-neuvième siècle mais les méthodes du calcul différentiel et intégral disponibles alors échouent à lui donner toute sa généralité.

Il faut attendre la naissance de la théorie des ensembles à la fin du dix-neuvième siècle et les travaux de Borel et Lebesgue [Leb] en théorie de la mesure puis en intégration pour que les bases de la théorie moderne des probabilités puissent être posées par Kolmogorov dans un ouvrage majeur écrit en 1933 [Ko]. C'est cette théorie qui va être présentée ici.

Le passage de la deuxième à la troisième année de Licence marque une rupture dans la formation du mathématicien, qui fait parallèle à la révolution qui s'est opérée au sein des mathématiques à la fin du dix-neuvième siècle. Du point de vue de

la mesure, l'analyse classique s'intéresse à des ensembles réguliers. Des coniques d'Apollonius aux courbes algébriques de Fermat, du cône d'Archimède aux surfaces de Riemann, ce sont des objets lisses qu'on s'efforce de mesurer en leur attribuant une longueur, une aire ou un volume. Le calcul différentiel et intégral, reposant sur l'intégrale de Riemann, est l'outil qui permet de mener à bien les calculs. Il s'avère que les ensembles que l'on cherche à mesurer en théorie des probabilités sont nettement moins réguliers que ceux qui appartiennent à la géométrie chère à Pascal. Quant aux limites qui font surgir des lois continues de phénomènes discrets, elles nécessitent des méthodes nouvelles qui amènent à repenser la notion même de convergence.

Le problème central qui motive la théorie de l'intégration peut se résumer ainsi : sous quelles conditions peut-on intervertir limite et intégrale ? Dans le cadre riemannien, la convergence simple d'une suite de fonctions intégrables ne garantit ni l'intégrabilité de la limite, ni la convergence des intégrales. Cette faiblesse théorique rend les passages à la limite compliqués, périlleux et difficiles à justifier. C'est au début du vingtième siècle qu'Henri Lebesgue révolutionne l'analyse en proposant d'intégrer des fonctions non plus en découpant leur intervalle de définition (l'axe des abscisses), mais en découpant l'ensemble des valeurs d'arrivée (l'axe des ordonnées). Ce changement de perspective, apparemment anodin, permet de définir une notion d'intégrale pour des fonctions beaucoup plus complexes que des fonctions continues. Il en découle une théorie de l'intégration robuste et souple, que nous allons exposer dans la première partie de ce livre.

Parallèlement à ces développements analytiques, le calcul des probabilités a longtemps souffert d'un manque de formalisme rigoureux, oscillant entre combinatoire et intuition physique. L'apport fondamental d'Andreï Kolmogorov en 1933 a été de comprendre que la théorie des probabilités n'est autre que la théorie de la mesure et de l'intégration, appliquée à des espaces de mesure totale égale à 1. Ainsi, ce cours n'est pas la juxtaposition de deux disciplines disjointes, mais l'exposition d'une théorie unifiée.

– Une variable aléatoire n'est rien d'autre qu'une fonction mesurable.

– L'espérance mathématique est exactement l'intégrale de Lebesgue.

– Un événement est un ensemble mesurable.

Cette identification met à la disposition de la théorie des probabilités les méthodes de la théorie de l'intégrale de Lebesgue donnant ainsi à la fois rigueur et généralité aux lois des grands nombres et au théorème central limite. C'est ce que nous allons exposer dans la seconde partie de ce livre.

Genèse de l'ouvrage

La première partie de l'ouvrage repose sur des notes d'un cours d'intégration qui a été enseigné en troisième année de licence à l'université de Rennes et de Brest. Doté d'un volume horaire de vingt-quatre heures distribuées sur un semestre, il traite l'intégralité de la théorie de la mesure et de l'intégrale de Lebesgue intervenant dans la suite du cursus, à la fois en analyse fonctionnelle et en probabilité.

Des compléments plus poussés prépareront également le lecteur à des cours spécialisés dans le domaine des systèmes dynamiques, de la théorie ergodique, des processus stochastiques ou encore de la géométrie fractale, où l'on doit disposer de méthodes fonctionnant dans des espaces métriques généraux et non plus juste dans $\mathbf{R}^d$ ou dans des espaces localement compact. Parmi ces compléments figurent certains résultats publiés pour la première fois en français dans un ouvrage universitaire. C'est le cas par exemple du théorème de Stone-Weierstrass mesurable figurant à la fin du sixième chapitre.

La seconde partie de ce livre repose sur des notes de cours enseignées en troisième année de licence à l'université Sorbonne université de 2020 à 2026. Là encore, le volume horaire était de 24h sur un semestre et supposait d'avoir suivi un cours d'intégration tel celui exposé dans la première partie. On s'est concentré sur les théorèmes de convergence classiques, essentiellement dans le cadre indépendant : loi faible et forte des grands nombres, théorème central limite, théorème des trois séries, loi du 0-1 de Kolmogorov.

Étant donné le volume horaire dévoué à la théorie des probabilités dans les maquettes actuelles de licence, il est préférable d'aborder les notions d'espérance conditionnelle et de chaînes de Markov au niveau du master, où elles pourront être étudiées en détail. Aussi n'avons-nous pas proposé de compléments dans ce cours sur ces sujets, qu'on pourra trouver exposés dans le livre de Y. Lacroix et L. Mazliak publié aux éditions Ellipses [LM], ou dans [Ne1], [LG] et [Dur].

À qui s'adresse ce livre

Ce cours se situe au niveau de la troisième année de licence et couvre la partie du programme concernant l'intégration et les probabilités. Il sera utile aussi bien aux étudiants poursuivant des études universitaires qu'aux élèves en première année de grandes écoles persévérant dans le domaine des mathématiques.

Pour la première partie, il suffit d'avoir suivi un cours portant sur l'intégrale de Riemann, tel qu'il est enseigné dans les cycles universitaires en première et seconde année ou en classes préparatoires. Pour la seconde partie, un minimum de familiarité avec les probabilités discrètes est conseillé, sous peine d'être découragé par la tour-

nure abstraite que prend la théorie moderne des probabilités. Même si cette théorie repose sur une théorie de l'intégration différente de celle de Riemann, les motivations restent les mêmes et ce texte poursuit le travail effectué dans les cours des années antérieures. Il est important d'être familier avec les concepts de base des probabilités et avec les exemples classiques, tels qu'ils sont exposés par exemple dans mon ouvrage intitulé *La géométrie du hasard* et publié aux éditions Ellipses [Co].

En rédigeant cet ouvrage, j'ai aussi pensé aux étudiants qui préparent le concours de l'agrégation de mathématiques, voir page 310. Certains passages de ce livre reprennent des thèmes que j'ai eu l'occasion de présenter aux agrégatifs à l'époque où j'intervenais dans la préparation de Rennes et de Brest. On m'a parfois reproché de proposer des développements à l'oral du concours pour lesquels il n'y avait pas de référence dans nos bibliothèques universitaires, voilà qui est en partie réparé.

Plan de l'ouvrage

Si besoin est, on peut commencer la lecture par l'annexe consacrée à des rappels de théorie des ensembles. Elle comporte une section sur la dénombrabilité, notion sur laquelle repose toute la théorie.

La première partie de ce livre est dédiée à l'intégrale de Lebesgue. Elle commence par des préliminaires de théorie de la mesure et se poursuit par la définition de l'intégrale. On a cherché à en donner une définition concrète en terme de sommes dont la convergence est évidente, avant de revenir à une approche plus usuelle par le biais des fonctions étagées. La mesure de Lebesgue est construite au chapitre 4, suivant l'approche de Carathéodory. Il est important de détailler cette construction, quitte à aller ensuite un peu plus vite sur la construction des mesures produits qui n'apporte rien de plus sur le plan conceptuel mais dont les détails sont assez long à exposer. Les espaces L^p jouent un rôle important en analyse fonctionnelle, ils sont étudiés au chapitre 6. Le plus célèbre est l'espace de Hilbert L^2 qui constitue un cadre bien adapté à la théorie de Fourier. Celle-ci est développée d'abord dans le cadre des séries puis ensuite sous sa forme intégrale. Le but final est d'exposer la formule d'inversion de Fourier, elle sera utilisée dans la seconde partie pour démontrer le théorème central limite.

Les notations utilisées en théorie de l'intégration et en probabilités ne sont pas les mêmes. Une annexe contient les énoncés d'un certain nombre de théorèmes démontrés dans la première partie mais cette fois-ci exprimés en termes probabilistes, pour aider le lecteur à faire la transition.

La seconde partie est consacrée à l'étude des théorèmes limites pour les suites de variables aléatoires indépendantes. Le texte est organisé de façon à parvenir assez

rapidement à la preuve de la loi forte des grands nombres, au chapitre 12, qui est faite pour des variables de carrés intégrables. Le cas intégrable est traité plus tard, dans le chapitre concernant les séries aléatoires, comme corollaire des théorèmes de convergence pour ces séries. Le second objectif est le théorème central limite, atteint au chapitre 14. Il faut pour cela étudier en détail les différents types de convergence et les relations qui s'établissent entre eux. On étudie d'abord le cas unidimensionnel avant de passer à l'étude des vecteurs gaussiens et à la version multidimensionnelle du théorème central limite. Le dernier chapitre est consacré à l'étude des séries de variables aléatoires indépendantes et à la loi du 0-1 de Kolmogorov. On a fait en sorte que cette seconde partie puisse être lue indépendamment de la première grâce à des rappels d'intégration en fin d'ouvrage.

Chaque chapitre se termine par une section d'exercices. Ce sont des énoncés de niveaux variés qui ont été traités en séance de travaux dirigés dans les cours afférents. Ils sont présentés par ordre de difficulté. Au début de la section consacrée aux exercices du chapitre se trouve un renvoi à la troisième partie du livre, vers des indications aidant à leurs résolution. Il ne faut pas hésiter à les consulter. Chaque indication renvoie ensuite à la correction de l'exercice correspondant, présentée en détail un peu plus loin dans la troisième partie, où sont rassemblés à la fois les énoncés et leurs solutions. Les exercices situés dans l'annexe consacrée à la théorie des ensembles sont corrigés dans cette annexe.

Des compléments figurent à la fin de certains chapitres. Ils peuvent être abordés lors d'une seconde lecture et approfondissent le corps du texte. Les résultats qui sont traités dans ces compléments s'avèrent utiles dans certains cours de master spécialisés. ou pour la préparation à certains concours.

Je me suis parfois écarté de la terminologie usuelle. En intégration, j'ai préféré parler de théorème de convergence croissante plutôt que de théorème de convergence monotone, car il n'y a pas de théorème de convergence décroissante pour les fonctions positives. Pour être précis, c'est le théorème de convergence dominée qui joue le rôle de théorème de convergence décroissante pour ces fonctions. En probabilités, j'ai préféré parler de variables à densité plutôt que de variables continues pour éviter toute confusion, car une variable dite continue n'est pas une fonction continue. Il s'agit d'une variable dont la loi est absolument continue relativement à la mesure de Lebesgue. Elle admet une densité, laquelle n'est pas continue en général. Enfin, pour les noms des probabilistes russes, j'ai adopté la translittération scientifique avec Markov et Kolmogorov plutôt que la translittération phonétique plus classique avec les graphies françaises Markoff et Kolmogoroff.

Nouveautés

En sus d'un corpus classique destiné à préparer au mieux l'étudiant à l'analyse et aux probabilités de niveau master, ce livre expose un certain nombre de résultats qui mériteraient d'être mieux connus mais qui sont rarement mentionnés dans la littérature. Ils sont susceptibles d'intéresser aussi bien les étudiants qui débutent leur apprentissage de la théorie, les lecteurs souhaitant consolider leurs connaissances en vue d'un concours ou encore les mathématiciens professionnels qui seront sans doute surpris par certains énoncés.

– Il n'est pas nécessaire que la convergence soit croissante dans le théorème de convergence croissante (corollaire 3, page 34). C'est aisé à démontrer et possède quelques applications intéressantes.

– Le théorème de Stone-Weierstrass mesurable (théorème 28, page 83) est une généralisation naturelle aux espaces L^p du fameux théorème de densité dans le cadre continu. Il s'applique à des algèbres de fonctions qui ne sont pas nécessairement continues.

– Il est bien connu que la boule unité d'un espace de Hilbert est compacte pour la topologie faible. Elle est aussi séquentiellement compacte (théorème 37, page 106), un fait qui n'est pas toujours mentionné dans les livres d'analyse fonctionnelle et qui n'est pas évident car la boule unité n'est pas nécessairement métrisable relativement à cette topologie faible.

– Les ouvrages traitant la transformée de Fourier dans le cadre L^2 sont parfois pauvres en exemples. Nous donnons une table de transformées de Fourier (annexe C, page 331) qui contient une liste raisonnable de fonctions L^2 dont la transformée est explicite. On pourra consulter [Ob] pour plus d'exemples.

– Il est possible de calculer la série des inverses des carrés des entiers grâce à la synthèse de Fourier. La réciproque est moins connue, il est possible de montrer que les exponentielles complexes forment une base hilbertienne de $L^2(\mathbf{R}/\mathbf{Z})$ grâce à cette série (théorème 32, page 95), dont la valeur peut bien sûr être calculée sans faire appel à la théorie de Fourier (exercice 5.4 page 69, corrigé page 249).

– Le théorème de Dirichlet ponctuel pour les séries de Fourier a un analogue pour les transformées de Fourier (théorème 41, page 119) qui est tout aussi utile et n'est pas plus difficile à démontrer.

– Nous faisons le lien entre la formule des résidus et la transformée de Fourier (complément page 120). La formule des résidus est traitée dans le cours d'analyse complexe de licence ou de maîtrise et ce lien n'est en général pas mentionné.

– On peut définir la notion d'espace de Sobolev H^1 sans faire appel à la transformée de Fourier. C'est la classe des fonctions dont le taux de variation converge au sens L^2, cette caractérisation apparaît naturellement lorsqu'on essaie de calculer la transformée de la dérivée d'une fonction L^2 (proposition 44, page 115).

– La synthèse de Fourier affirme que les sommes partielles symétriques de la série de Fourier d'une fonction L^2 convergent vers la fonction en norme L^2. Il existe un théorème analogue pour la transformée de Fourier faisant intervenir les intégrales partielles (corollaire 12, page 113).

– La convergence faible des mesures au sens des analystes coïncide avec la convergence étroite des probabilistes lorsqu'on considère des mesures de probabilité (théorème 48, page 174). Ce calcul est souvent omis dans les livres de probabilités.

– La démonstration de la loi forte des grands nombres page 211 qui termine le chapitre 16 est courte et astucieuse. Elle est due à Jacques Neveu et repose sur la loi du 0-1 de Kolmogorov.

– Il n'est pas aisé de rendre compte de simulations numériques sur le papier. On s'y est tout de même essayé, le lecteur pourra méditer sur de longues suites de chiffres illustrant la loi des grands nombres (fin du chapitre 12) et sur quelques graphes de fréquences cherchant à converger vers la loi normale (fin du chapitre 14).

– Enfin nous donnons une preuve relativement courte de la formule de changement de variables pour les intégrales multiples (théorème 21 page 66). C'est une formule importante qu'on voit dès la seconde année de licence de mathématiques, mais les preuves qui figurent dans la littérature sont souvent longues et parfois illisibles.

Remerciements

J'aimerais remercier les professeurs qui m'ont appris la théorie de l'intégration et les probabilités, Patrick Gérard, Jean-Pierre Kahane à Orsay et les collègues avec qui j'ai eu le plaisir de les enseigner, Jean-François Le Gall à l'Ecole Normale Supérieure, Stéphane Leborgne à Rennes, Jean-Marc Derrien et Benoit Saussol à Brest, Michèle Thieullen, Irina Kourkova et Camille Tardif à Jussieu et bien d'autres amis qui je l'espère trouveront dans ce livre trace des conversations que nous avons eu autour de la nécessité et du hasard. Je remercie également Sabine Chanthara, secrétaire à l'université de Brest avec qui j'ai eu le plaisir de travailler et qui a contribué à la frappe du manuscrit d'intégration à une époque où la direction du département de mathématique de Brest occupait une grande partie de mon temps. Enfin je remercie Valentin et Philémon Coudène pour leur relecture attentive du manuscrit.

Table des matières

Notations **1**

I Théorie de la mesure et intégrale de Lebesgue **3**

1 Théorie de la mesure **5**
1 Tribus 5
2 Mesures 7
3 Propriétés des mesures 9
4 Fonctions mesurables 11
5 Complément : mesure et topologie 15
6 Complément : un ensemble non borélien 16
7 Exercices 17

2 L'intégrale de Lebesgue **19**
1 Intégration des fonctions positives 19
2 Monotonie 22
3 Linéarité 24
4 Exercices 28

3 Propriétés de l'intégrale **31**
1 Densité des fonctions étagées 31
2 Limites inférieures et supérieures 33
3 Lien avec l'intégrale de Riemann 36
4 Intégrale des fonctions à valeurs complexes 37
5 Lien entre dérivée et intégrale 42
6 Complément : fonctions de dérivée intégrable 43
7 Exercices 44

4 La mesure de Lebesgue **47**
1 Généralisation de la notion de longueur 47
2 Construction de Carathéodory 49
3 Mesure de Lebesgue sur **R** 51
4 Tribus complètes 52
5 Unicité et approximation 52
6 Complément : lemme de la classe monotone 54
7 Complément : critère de Carathéodory 55
8 Complément : théorème de représentation de Riesz 56
9 Exercices 57

5 Intégrales multiples **59**
1 Intégrales dépendant d'un paramètre 59
2 Mesure produit 61
3 Théorème de Fubini 63
4 Intégration par parties 65
5 Formule du changement de variables 66
6 Exercices 69

6 Espaces L^p **71**
1 La norme L^p 71
2 Inégalités L^p et complétude 72
3 Densité dans les espaces L^p 76
4 Convolution 78
5 Complément : densité L^p, le cas métrique 82
6 Complément : théorème de Stone-Weierstrass mesurable 83
7 Exercices 86

7 Espaces de Hilbert **89**
1 Analyse hilbertienne 89
2 Séries de Fourier, point de vue L^2 94
3 Exemples de bases hilbertiennes 98
4 Séries de Fourier, convergence ponctuelle 99
5 Complément : le cas C^1 par morceaux 103
6 Complément : dualité et compacité faible 104
7 Exercices 106

8 Transformée de Fourier **109**
1 Transformée de Fourier dans L^2 109
2 Formule d'inversion de Fourier 112
3 Propriétés de la transformée 115
4 Complément : inversion de Fourier ponctuelle 119
5 Complément : transformée des fractions rationnelles 120
6 Exercices 122

II Théorie moderne des probabilités **125**

9 Formalisme de Kolmogorov **127**
1 Le cas discret 127
2 Le cas continu 129
3 Le cas des espaces produits 129
4 Complément : critère de Kolmogorov 131
5 Exercices 134

10 Variables aléatoires **135**
1 Définition d'une variable aléatoire 135
2 Espérance et variance 136
3 Inégalités 137
4 Loi d'une variable aléatoire 139
5 Loi d'un multiplet de variables aléatoires 142
6 Exercices 143

11 Indépendance **145**
1 Indépendance d'événements et de variables aléatoires 145
2 Lemme de Borel-Cantelli 146
3 Loi d'un multiplet de variables indépendantes 148
4 Exercices 151

12 Loi des grands nombres **153**
1 Loi faible des grands nombres 154
2 Loi forte des grands nombres 155
3 Illustration numérique 161
4 Exercices 166

13 Convergence de suites aléatoires **169**
1 Les différents types de convergence 169
2 Fonction caractéristique et transformée de Fourier 172
3 Convergence en loi et fonction caractéristique 174
4 Autres propriétés de la convergence en loi 178
5 Exercices 181

14 Théorème central limite **183**
1 Fonction caractéristique de la loi normale 183
2 Théorème central limite 184
3 Illustration numérique 187
4 Exercices 191

15 Vecteurs aléatoires **193**
1 Variables aléatoires à valeurs vectorielles 193
2 Définition des vecteurs gaussiens 195
3 Loi des vecteurs gaussiens 197
4 Théorèmes limites pour les vecteurs aléatoires 199
5 Exercices 201

16 Séries de variables aléatoires indépendantes **203**
1 Loi du 0-1 de Kolmogorov 203
2 Convergence des séries aléatoires 205
3 Retour sur la loi des grands nombres 209
4 Exercices 212

III Exercices corrigés **213**

17 Indications pour les exercices **215**

18 Correction des exercices **225**
1 Théorie de la mesure 225
2 L'intégrale de Lebesgue 230
3 Propriétés de l'intégrale 236
4 La mesure de Lebesgue 241
5 Intégrales multiples 246
6 Espaces L^p 253
7 Espaces de Hilbert 260

8 Transformée de Fourier 265
9 Formalisme de Kolmogorov 272
10 Variables aléatoires 277
11 Indépendance 282
12 Loi des grands nombres 286
13 Convergence de suites de variables aléatoires 290
14 Théorème central limite 296
15 Vecteurs aléatoires 302
16 Séries de variables aléatoires indépendantes 306
17 Un mot pour les agrégatifs 310

A Rappels de théorie des ensembles **313**
1 Opérations ensemblistes 314
2 Relations ensemblistes 315
3 Image directe et réciproque 317
4 Dénombrabilité 318
5 Exercices 319
6 Solution des exercices de théorie des ensembles 321

B Résumé d'intégration pour les probabilités **325**
1 Théorèmes de convergence 325
2 Intégrales dépendant d'un paramètre 327
3 Intégrales multiples 327
4 Espaces L^p 328
5 Unicité des mesures 329

C Tables **331**
1 Tables de transformées de Fourier 331
2 Formulaire de théorie de Fourier 333
3 Lois classiques 334

Bibliographie **335**

Index **337**

Notations

Les ensembles des nombres entiers, entiers relatifs, réels et complexes sont notés respectivement $\mathbf{N}$, $\mathbf{Z}$, $\mathbf{R}$, $\mathbf{C}$.

$\mathbf{1}_A$ ..... fonction indicatrice de l'ensemble A
A^c ..... complémentaire de l'ensemble A
$A \triangle B$ ..... différence symétrique des ensembles A et B
$B(x, r)$ ..... boule ouverte de centre x de rayon r
C^∞ ..... ensemble des fonctions indéfiniment différentiables
C_c ..... ensemble des fonctions continues à support compact
$Cov(X, Y)$ ..... covariance de X et Y
δ_ω ..... mesure de Dirac au point ω
$E(X)$ ..... espérance de X
F_X ..... fonction de répartition
λ ..... mesure de Lebesgue
L^p ..... espace des classes de fonctions L^p
$\overline{\lim}$, $\underline{\lim}$ ..... limite supérieure et inférieure
μ ..... mesure
$\mathbf{N}^*$ ..... nombres entiers non nuls
Ω ..... ensemble de résultats
$\circ$ ..... composition
$\emptyset$ ..... ensemble vide
P ..... mesure de probabilité
P_X ..... loi de la variable aléatoire X
$P_{(X,Y)}$ ..... loi du couple (X, Y)
$\mu \otimes \nu$, $P \otimes Q$ ..... produit des mesures μ et ν ou P et Q
$p.p.$ ..... presque partout
$p.s.$ ..... presque sûrement
S_n ..... somme de X_1 à X_n
$\sigma(X)$, $V(X)$ ..... écart-type, variance de X
Σ ..... matrice de covariance
${}^tA$ ..... transposée de la matrice A
$\mathcal{T}$ ..... tribu
$\mathcal{T}_1 \otimes \mathcal{T}_2$ ..... produit des tribus $\mathcal{T}_1$ et $\mathcal{T}_2$
$\coprod$ ..... union disjointe
$\langle f, g \rangle$, $\langle X, Y \rangle$ ..... produit scalaire dans L^2
$\|f\|_p$, $\|X\|_p$ ..... norme L^p

Première partie

Théorie de la mesure et intégrale de Lebesgue

Chapitre 1

Théorie de la mesure

Ce chapitre a deux objectifs : d'une part introduire le concept abstrait d'espace mesuré qui va servir de base à la théorie de l'intégrale de Lebesgue, d'autre part, montrer comment les espaces topologiques usuels comme $\mathbf{R}^d$ possèdent une structure naturelle d'espace mesuré qui permet de généraliser les concepts de longueur, d'aire et de volume.

La théorie de la mesure repose sur le formalisme de la théorie des ensembles. Nous allons utiliser les opérations ensemblistes et faire usage du concept de dénombrabilité. Des rappels figurent dans l'annexe A avec quelques exercices pour vérifier que ces notions sont bien comprises.

1 Tribus

Définition 1 *Soit X un ensemble. Un ensemble $\mathcal{T}$ de parties de X est appelé une tribu ou σ-algèbre s'il vérifie les propriétés suivantes :*

– $\emptyset \in \mathcal{T}$,

– *pour tout* $A \in \mathcal{T}$, $A^c \in \mathcal{T}$,

– *pour toute suite* $(A_i)_{i\in\mathbf{N}}$ *d'éléments de* $\mathcal{T}$, $\bigcup_{i\in\mathbf{N}} A_i \in \mathcal{T}$.

On a alors $X \in \mathcal{T}$ car X est le complémentaire de l'ensemble vide et $\bigcap_{i\in\mathbf{N}} A_i \in \mathcal{T}$ en vertu de la relation

$$\bigcap_{i\in\mathbf{N}} A_i = \Big(\bigcup_{i\in\mathbf{N}} A_i^c\Big)^c.$$

La plus grande tribu de parties de X est l'ensemble de toutes les parties de X, noté $\mathcal{P}(X)$. La plus petite ne contient que deux éléments, l'ensemble vide et X lui-même.

Proposition 1 *Soit X un ensemble et $\{\mathcal{T}_i\}_{i\in I}$ une famille de tribus de parties de X. Alors $\bigcap_{i\in I} \mathcal{T}_i$ est une tribu.*

Démonstration
Pour tout $i \in I$, nous avons $\emptyset \in \mathcal{T}_i$ car $\mathcal{T}_i$ est une tribu. Par conséquent $\emptyset \in \cap\mathcal{T}_i$. De même, si A appartient à $\mathcal{T}_i$ pour tout i, il en va de même de A^c et A^c est dans l'intersection des $\mathcal{T}_i$.

Soit $(A_j)_{j\in\mathbf{N}}$ une suite d'éléments appartenant à $\cap\mathcal{T}_i$. Les A_j appartiennent à tous les $\mathcal{T}_i$, il en va de même de l'union $\cup A_j$ car les $\mathcal{T}_i$ sont des tribus, nous avons donc $\cup A_j \subset \cap\mathcal{T}_i$. L'intersection des tribus $\mathcal{T}_i$ est bien invariante par union dénombrable et passage au complémentaire.

Définition 2 *La tribu engendrée par un ensemble quelconque de parties de X est égale à l'intersection de toutes les tribus qui contiennent ces parties.*

Soit X un espace topologique. La tribu des boréliens *de X est la tribu engendrée par les sous-ensembles ouverts de X. On la note $\mathcal{B}(X)$.*

La tribu engendrée par un ensemble de parties est donc la plus petite tribu contenant ces parties. Notons que tous les ouverts de X sont boréliens, tout comme les intersections dénombrables d'ouverts de X, ainsi que les fermés, car leur complémentaire est ouvert.

Proposition 2 *La tribu des boréliens de $\mathbf{R}$ est engendrée par les intervalles ouverts. Pour tout $d \in \mathbf{N}^*$, la tribu des boréliens de $\mathbf{R}^d$ est engendrée par les boules ouvertes.*

Démonstration
Il suffit de vérifier que tout ouvert U de $\mathbf{R}^d$ peut s'écrire comme union dénombrable de boules ouvertes. Montrons que U est union des boules centrées sur les points de coordonnées rationnelles et de rayon rationnel contenues dans U :

$$U = \bigcup_{\substack{(x,r)\in\mathbf{Q}^d\times\mathbf{Q}_+^* \\ tels\ que\ B(x,r)\subset U}} B(x,r)$$

Comme U est ouvert, nous savons qu'il existe une boule ouverte $B(y,s)$, $s > 0$, contenue dans U. Soit $r \in \mathbf{Q}_+^*$ inférieur à $s/2$ et x un point à coordonnées rationnelles dans $B(y,r)$. Nous avons alors

$$y \in B(x,r) \subset B(y,s) \subset U.$$

Tout point $y \in U$ est contenu dans une boule $B(x, r)$ incluse dans U avec (x, r) dans $\mathbf{Q}^d \times \mathbf{Q}_+^*$. La proposition est démontrée.

La preuve précédente fonctionne dans tout espace métrique qui admet un sous-ensemble dénombrable dense (ici $\mathbf{Q}^d$). Un tel espace métrique est dit *séparable*.

2 Mesures

Définition 3 *Soit X un ensemble et $\mathcal{T}$ une tribu de parties de X. Une mesure μ définie sur $\mathcal{T}$ est une fonction $\mu : \mathcal{T} \to [0, +\infty]$ satisfaisant les propriétés suivantes :*

$-\ \mu(\emptyset) = 0,$

$-\ \mu\Big(\coprod_{i\in\mathbf{N}} A_i\Big) = \sum_{i\in\mathbf{N}} \mu(A_i).$

Cette dernière propriété est appelée σ-additivité. Elle implique l'additivité finie : pour tout $n \in \mathbf{N}$ et $A_0, \ldots, A_n$ dans $\mathcal{T}$ disjoints deux à deux,

$$\mu\Big(\coprod_{i=0}^{n} A_i\Big) = \sum_{i=0}^{n} \mu(A_i).$$

De fait, il suffit de poser $A_i = \emptyset$ pour $i > n$. On a alors

$$\coprod_{i\in\mathbf{N}} A_i = \coprod_{i=0}^{n} A_i, \quad \sum_{0}^{\infty} \mu(A_i) = \sum_{0}^{n} \mu(A_i).$$

Exemples
Soit X un ensemble, $x_0 \in X$. La *mesure de Dirac* au point x_0 est la mesure définie sur la tribu $\mathcal{P}(X)$ par

$$\delta_{x_0}(A) = \begin{cases} 1 & \text{si } x_0 \in A \\ 0 & \text{sinon.} \end{cases}$$

pour tout $A \subset X$.

Un autre exemple est donné par la *mesure de comptage*, qui associe à un sous-ensemble quelconque $A \subset X$ son nombre d'élément si A est fini et $+\infty$ sinon.

La proposition suivante est due à Henri Lebesgue (1901).

Théorème 1 *Il existe une unique mesure sur la tribu des boréliens de $\mathbf{R}$ qui associe à tout intervalle sa longueur. Elle est notée λ ou $\lambda_{\mathbf{R}}$, on l'appelle la* mesure de Lebesgue *de la droite réelle.*

En d'autres termes, il existe une unique mesure $\lambda : \mathcal{B}(\mathbf{R}) \to [0, +\infty]$ *satisfaisant pour tout* $a, b \in \mathbf{R}$ *tels que* $a \leq b$*,*

$$\lambda([a, b]) = b - a.$$

Ce théorème montre qu'il existe une mesure qui étend la notion de longueur à l'ensemble des boréliens de $\mathbf{R}$. La démonstration sera faite au chapitre 4.

Il existe également pour tout $d \in \mathbf{N}^*$ une mesure sur $\mathbf{R}^d$ qui étend les notions usuelles d'aire et de volume d-dimensionnel. On l'appelle encore mesure de Lebesgue sur $\mathbf{R}^d$.

Proposition 3 *Il existe une unique mesure* $\lambda_{\mathbf{R}^d} : \mathcal{B}(\mathbf{R}^d) \to [0, +\infty]$ *satisfaisant*

$$\lambda_{\mathbf{R}^d}\Big(\prod_{i=1}^{d} [a_i, b_i]\Big) = \prod_{i=1}^{d} (b_i - a_i)$$

pour tout $a_1, \ldots, a_d$ *et* $b_1, \ldots, b_d$ *dans* $\mathbf{R}$ *tels que* $a_1 \leq b_1$*, …,* $a_d \leq b_d$*.*

La mesure de Lebesgue des singletons est égale à 0 : $\lambda(\{x\}) = \lambda(\prod [x_i, x_i]) = 0$ pour tout $x = (x_1, \ldots, x_d) \in \mathbf{R}^d$. On dit que cette mesure est *non-atomique*. Pour une telle mesure, tous les ensembles dénombrables sont de mesure nulle, car ils sont union dénombrable disjointe de singletons.

Les ensembles de mesure nulle jouent un rôle important en théorie de la mesure, ce qui amène à la définition suivante.

Définition 4 *On dit qu'un ensemble est négligeable s'il est de mesure nulle.*

Nous pouvons maintenant introduire la notion d'espace mesuré.

Définition 5 *Un* espace mesuré $(X, \mathcal{T}, \mu)$ *est la donnée*

– d'un ensemble X*,*

– d'une tribu $\mathcal{T}$ *de parties de* X*,*

– d'une mesure μ *définie sur cette tribu.*

La mesure μ est dite *finie* si $\mu(X) < \infty$. On parle de *mesure de probabilité* ou d'*espace probabilisé* quand $\mu(X) = 1$. On parle de *mesure borélienne* lorsqu'elle est définie sur la tribu des boréliens d'un espace topologique.

3 Propriétés des mesures

Soit $(X, \mathcal{T}, \mu)$ un espace mesuré. Voici quelques propriétés satisfaites par μ.

Monotonie : *soit $A, B \in \mathcal{T}$. Si $A \subset B$, alors $\mu(A) \leq \mu(B)$.*

Pour démontrer cette propriété, on note que B peut s'écrire $B = A \coprod (B \setminus A)$ et on applique la propriété d'additivité :

$$\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A).$$

Remarquons que si de plus $\mu(A)$ est fini, nous avons aussi

$$\mu(B \setminus A) = \mu(B) - \mu(A).$$

Attention, cette formule n'a plus de sens lorsque $\mu(A)$ vaut $+\infty$ car l'expression à droite de l'égalité vaut $\infty - \infty$. Il n'est possible de soustraire une quantité aux deux termes d'une égalité que si cette quantité est finie.

Sous-additivité : *soit $(A_i)_{i \in \mathbf{N}}$ une suite d'éléments de $\mathcal{T}$. Alors*

$$\mu\Big(\bigcup_{i \in \mathbf{N}} A_i\Big) \leq \sum_{i=0}^{\infty} \mu(A_i).$$

Montrons cette inégalité en posant $A'_0 = A_0$ et $A'_k = A_k \setminus \Big(\bigcup_{i=0}^{k-1} A_i\Big)$. On a alors

$$\bigcup_{i \in \mathbf{N}} A_i = \coprod_{i \in \mathbf{N}} A'_i, \qquad A'_i \subset A_i,$$

si bien que

$$\mu\Big(\bigcup_{i \in \mathbf{N}} A_i\Big) = \mu\Big(\coprod_{i \in \mathbf{N}} A'_i\Big) = \sum_{i=0}^{\infty} \mu(A'_i) \leq \sum_{i=0}^{\infty} \mu(A_i).$$

Comme application immédiate, nous en déduisons qu'une union dénombrable d'ensembles négligeables est négligeable.

Convergence croissante : *soit $(A_i)_{i \in \mathbf{N}}$ une suite d'éléments de $\mathcal{T}$ croissante pour l'inclusion : $A_i \subset A_{i+1}$ pour tout i. Alors*

$$\mu\Big(\bigcup_{i \in \mathbf{N}} A_i\Big) = \lim_{i \to \infty} \mu(A_i).$$

Pour montrer cette convergence, on procède comme plus haut en faisant usage des ensembles A'_i.

$$\mu\Big(\bigcup_{i\in\mathbf{N}} A_i\Big) = \sum_{i=0}^{\infty} \mu(A'_i) = \lim_{N\to\infty} \sum_{i=0}^{N} \mu(A'_i) = \lim_{N\to\infty} \mu\Big(\coprod_{i=0}^{N} A'_i\Big) = \mu(A_N),$$

cette dernière égalité découlant de la croissance des A_i : $\coprod_{i=0}^{N} A'_i = \bigcup_{i=0}^{N} A_i = A_N$.

Convergence décroissante : *soit $(A_i)_{i\in\mathbf{N}}$ une suite d'éléments de $\mathcal{T}$ décroissante pour l'inclusion : $A_{i+1} \subset A_i$ pour tout i. On suppose de plus que A_0 est de mesure finie : $\mu(A_0) < \infty$. Alors*

$$\mu\Big(\bigcap_{i\in\mathbf{N}} A_i\Big) = \lim_{i\to\infty} \mu(A_i).$$

On se ramène à la convergence croissante en posant $B_i = A_0 \setminus A_i$. La suite B_i est croissante et comme A_0 est de mesure finie,

$$\mu(B_i) = \mu(A_0) - \mu(A_i)$$

et cette suite converge vers $\mu(\cup_{i\in\mathbf{N}} B_i)$. L'union des B_i vaut $A_0 \setminus \bigcap_{i\in\mathbf{N}} A_i$, la limite de la suite vaut donc $\mu(A_0) - \mu(\bigcap_{i\in\mathbf{N}} A_i)$ et le résultat est démontré.

Illustrons les propriétés précédentes par un résultat de convergence sur les séries.

Définition 6 *On dit qu'une propriété portant sur les points d'un espace mesuré est vraie* presque partout *si l'ensemble des points pour lesquels cette propriété est satisfaite est de complémentaire négligeable.*

Proposition 4 *Soit $(X, \mathcal{T}, \mu)$ un espace mesuré et $(A_k)_{k\in\mathbf{N}}$ une suite d'éléments de $\mathcal{T}$ satisfaisant*

$$\sum_{k=0}^{\infty} \mu(A_k) < \infty.$$

Alors presque tout $x \in X$ n'appartient qu'à un nombre fini de A_k.

Démonstration

Il s'agit de montrer que l'ensemble des points qui appartiennent à une infinité de A_k est de mesure nulle. Un point $x \in X$ appartient à une infinité de A_k si on peut

trouver pour tout entier n un entier $k \geq n$ tel que x appartienne à A_k. Traduisons cette propriété en termes ensemblistes.

$$\{x \in X \mid \forall\, n \in \mathbf{N},\ \exists\, k \geq n \text{ tel que } \ x \in A_k\} = \bigcap_{n\in\mathbf{N}} \Big(\bigcup_{k\geq n} A_k \Big).$$

Par sous-additivité,

$$\mu\Big(\bigcup_{k\geq n} A_k \Big) \ \leq \ \sum_{k=n}^{\infty} \mu(A_k) \ \longrightarrow_{n\to\infty} \ 0.$$

La série tend vers 0 quand n tend vers l'infini car c'est le reste d'une série convergente. Posons $B_n = \cup_{k\geq n} A_k$. Les ensembles B_n forment une suite de parties de X décroissante pour l'inclusion. Tous ces ensembles sont de mesure finie. Par convergence décroissante, nous avons donc

$$\mu\Big(\bigcap_{n\in\mathbf{N}} B_n \Big) = \lim_{n\to\infty} \mu(B_n) = 0.$$

La proposition est démontrée.

La proposition précédente illustre une méthode générale. Pour démontrer qu'une propriété est vraie presque partout, on traduit cette propriété en termes ensemblistes et on estime la mesure des ensembles obtenus à l'aide des inégalités vues plus haut.

Corollaire 1 *Soit $(X, \mathcal{T}, \mu)$ un espace mesuré, (c_k) une suite de nombres complexes et (A_k) une suite d'éléments de $\mathcal{T}$ satisfaisant*

$$\sum_{k=0}^{\infty} \mu(A_k) < \infty.$$

Alors la série $\displaystyle\sum_{k\in\mathbf{N}} c_k \mathbf{1}_{A_k}(x)$ converge pour presque tout $x \in X$.

De fait, pour presque tout $x \in X$, nous avons vu qu'il n'y a qu'un nombre fini de termes de cette série qui sont non nuls.

4 Fonctions mesurables

Nous définissons une classe de fonctions qui généralise celle des fonctions continues. Ces fonctions sont celles que nous allons pouvoir intégrer relativement à une mesure.

Définition 7 *Soit X et Y deux ensembles, $\mathcal{T}$ une tribu de parties de X, $\mathcal{U}$ une tribu de parties de Y. Une fonction $f : X \to Y$ est dite* mesurable *relativement à ces tribus si l'image réciproque de tout élément de $\mathcal{U}$ est dans $\mathcal{T}$:*

$$\forall\, B \in \mathcal{U}, \quad f^{-1}(B) \in \mathcal{T}.$$

Remarquons que la composée de deux fonctions mesurables f, g est mesurable en vertu de la relation $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$. Pour démontrer la mesurabilité, on peut se contenter de regarder les images réciproques d'une famille de parties qui engendre la tribu $\mathcal{U}$.

Lemme 1 *Soit X, Y deux ensembles, $\mathcal{T}$ et $\mathcal{U}$ des tribus de parties de X et Y. On se donne une famille C de parties de Y qui engendre $\mathcal{U}$ et $f : X \to Y$. Alors f est mesurable si et seulement si*

$$\forall\, C \in C, \quad f^{-1}(C) \in \mathcal{T}.$$

Démonstration
Posons $\mathcal{U}' = \{B \in \mathcal{U} \mid f^{-1}(B) \in \mathcal{T}\}$, montrons que ce sous-ensemble de $\mathcal{U}$ est une tribu.

– $\emptyset \in \mathcal{U}'$ car $f^{-1}(\emptyset) = \emptyset \in \mathcal{T}$.

– Soit $B \in \mathcal{U}'$. Nous avons $f^{-1}(B^c) = f^{-1}(B)^c \in \mathcal{T}$, donc $B^c \in \mathcal{U}'$.

– Soit $B_i \in \mathcal{U}'$, $i \in \mathbf{N}$. $f^{-1}(\bigcup_{i\in\mathbf{N}} B_i) = \bigcup_{i\in\mathbf{N}} f^{-1}(B_i) \in \mathcal{T}$. Donc $\bigcup_{i\in\mathbf{N}} B_i \in \mathcal{U}$.

Par hypothèse, C est incluse dans la tribu $\mathcal{U}'$. Comme $\mathcal{U}$ est la plus petite tribu contenant les éléments de C, elle est incluse dans $\mathcal{U}'$ si bien que $\mathcal{U} = \mathcal{U}'$.

La preuve précédente illustre une méthode générale : pour démontrer qu'une propriété vraie pour un ensemble de parties s'étend à la tribu engendrée par cet ensemble, il suffit de montrer que l'ensemble des parties satisfaisant cette propriété est une tribu.

Lorsque X et Y sont des espaces topologiques, on peut considérer la tribu des boréliens sur chacun de ces espaces. Rappelons que c'est la tribu engendrée par les sous-ensembles ouverts de l'espace.

Définition 8 *Soit X et Y deux espaces topologiques, une fonction $f : X \to Y$ est dite* borélienne *si elle est mesurable relativement aux tribus boréliennes de X et de Y.*

D'après le lemme précédent, pour montrer qu'une fonction est borélienne, il suffit de vérifier que les images réciproques des ouverts de Y sont des ensembles boréliens de X. Comme l'image réciproque d'un ouvert par une fonction continue est un ouvert, on en déduit la proposition suivante.

Proposition 5 *Soit X et Y deux espaces topologiques. Toute fonction continue définie de X dans Y est borélienne.*

On s'intéresse maintenant aux fonctions à valeurs réelles ou complexes définies sur un espace mesuré quelconque.

Définition 9 *Soit f une fonction définie sur X à valeurs dans $[0,\infty]$, $\mathbf{C}$ ou $\mathbf{R}^d$. Une telle fonction est dite* $\mathcal{T}$-mesurable*, ou juste mesurable s'il n'y a pas d'ambiguité concernant la tribu considérée sur X, si les images réciproques des ensembles boréliens de l'espace d'arrivée sont dans $\mathcal{T}$.*

Nous avons vu au premier chapitre que la tribu des boréliens de $\mathbf{R}^d$ est engendrée par les boules ouvertes. On en déduit qu'une fonction à valeurs dans $\mathbf{R}^d$ est mesurable si et seulement si les images réciproques des boules ouvertes sont dans $\mathcal{T}$. Pour une fonction à valeurs dans $\mathbf{R}$, il suffit de vérifier que les images réciproques des intervalles sont dans $\mathcal{T}$. On peut même se restreindre aux images réciproques des intervalles de la forme $[a,\infty[$ (ou $]a,\infty[$) car tous les intervalles de $\mathbf{R}$ s'expriment à partir de ceux-là de la façons suivante.

$$]a,b[= \bigcup_{n\in\mathbf{N}^*} \left[a+\frac{1}{n},\infty\right[\ \setminus \ [b,\infty[, \qquad]a,b[=]a,\infty[\setminus \bigcap_{n\in\mathbf{N}^*} \left]b-\frac{1}{n},\infty\right[.$$

La notion de mesurabilité est très souple, elle est préservée par la plupart des opérations usuelles sur les fonctions. En particulier, une limite simple de fonctions mesurables est mesurable. Cette propriété, qui n'est pas vraie pour les fonctions continues, permettra de passer aisément à la limite dans les intégrales.

Proposition 6 *On se place sur un espace mesuré $(X,\mathcal{T},\mu)$.*
Soit $f,g : X \to \mathbf{R}$ des fonctions mesurables et $\lambda \in \mathbf{R}$. Alors

$$\lambda f + g, \quad fg, \quad min(f,g), \quad max(f,g) \quad \textit{sont mesurables.}$$

Soit (f_n) une suite de fonctions mesurables définies de X dans $[-\infty,\infty]$. Alors

$$\inf_{n\in\mathbf{N}} f_n \textit{ et } \sup_{n\in\mathbf{N}} f_n \quad \textit{sont mesurables.}$$

L'ensemble des points où la suite (f_n) converge est dans $\mathcal{T}$:

$$\{x \in X \mid (f_n(x))_{n\in\mathbf{N}} \text{ est une suite convergente}\} \in \mathcal{T}.$$

Si la suite (f_n) converge simplement, sa limite est mesurable.

Démonstration

– Commençons par montrer que la fonction $x \mapsto (f(x), g(x))$ définie de X dans $\mathbf{R}^2$ est mesurable si f et g sont mesurables. Pour cela, montrons que l'image réciproque des boules ouvertes de $\mathbf{R}^2$ pour la norme $\|(y_1, y_2)\|_\infty = \max(|y_1|, |y_2|)$ sont mesurables. Ces boules sont de la forme

$$B((y_1, y_2), r) =]y_1 - r, y_1 + r[\times]y_2 - r, y_2 + r[$$

avec $y_1, y_2 \in \mathbf{R}$ et $r \in \mathbf{R}_+$, ce qui donne

$$(f, g)^{-1}\Big(B((y_1, y_2), r)\Big) = f^{-1}(\,]y_1 - r, y_1 + r[\,) \cap g^{-1}(\,]y_2 - r, y_2 + r[\,) \in \mathcal{T}.$$

Les fonctions $(x, y) \mapsto \lambda x + y$, $(x, y) \mapsto xy$, $(x, y) \mapsto \max(x, y)$ définies de $\mathbf{R}^2$ dans $\mathbf{R}$ sont continues donc mesurables. On obtient encore des fonctions mesurables en les composant avec $x \mapsto (f(x), g(x))$.

– Posons $F(x) = \inf\{f_n(x) \mid n \in \mathbf{N}\} \in [-\infty, \infty]$. Soit $a \in \mathbf{R}$, remarquons que $F(x)$ est supérieur ou égal à a si et seulement si il en va de même pour tous les $f_n(x)$.

$$F^{-1}([a, \infty[) = \bigcap_{n\in\mathbf{N}} f_n^{-1}([a, \infty[) \in \mathcal{T}$$

On en déduit que F est mesurable. On procède de manière similaire pour $\sup\ f_n$.

– La suite $(f_n(x))$ converge si et seulement si elle est de Cauchy, c'est-à-dire si

$$\forall\, K \in \mathbf{N}^*,\ \exists\, N \in \mathbf{N},\ \forall\, m, n \geq N, \quad |f_m(x) - f_n(x)| < 1/K,$$

$$\begin{aligned}\{x \mid (f_n(x)) \text{ converge}\} &= \bigcap_{K\in\mathbf{N}^*} \bigcup_{N\in\mathbf{N}} \bigcap_{m,n\geq N} \{x \in X \mid |f_m(x) - f_n(x)| < 1/K\} \\ &= \bigcap_{K\in\mathbf{N}^*} \bigcup_{N\in\mathbf{N}} \bigcap_{m,n\geq N} \Psi_{m,n}^{-1}(\,]-1/K, 1/K[\,) \in \mathcal{T}\end{aligned}$$

car $\Psi_{m,n} = f_m - f_n$ est mesurable.

– Enfin, si la suite (f_n) converge vers f,

$$f^{-1}(]a,\infty[) = \{x \in X \mid \forall\, N \in \mathbf{N},\ \exists\, n \geq N,\ f_n(x) > a\}$$
$$= \bigcap_{N\in\mathbf{N}} \bigcup_{n\geq N} f_n^{-1}(\,]a,\infty[\,) \in \mathcal{T}.$$

La proposition est démontrée.

5 Complément : mesure et topologie

Même si les fonctions boréliennes ne sont pas continues en général, on va montrer qu'elles possèdent une certaine régularité lorsque l'espace sous-jacent est complet.

Définition 10 *Soit X un espace topologique et μ une mesure définie sur la tribu des boréliens de X. La mesure μ est dite* intérieurement régulière *si pour tout ensemble borélien $A \subset X$,*

$$\sup\{\mu(K) \mid K \text{ compact},\ K \subset A\} = \mu(A)$$

Théorème 2 (Oxtoby-Ulam) *Toute mesure borélienne finie, définie sur un espace métrique séparable complet, est intérieurement régulière.*

Démonstration
Considérons une suite $(x_i)_{i\in\mathbf{N}}$ dense dans X et $n \in \mathbf{N}^*$. La famille de boules fermées $\{\overline{B}(x_i, 1/n)\}_{i\in\mathbf{N}}$ recouvre X si bien qu'il existe $N \in \mathbf{N}$ dépendant de n tel que

$$\mu\Big(X \setminus \bigcup_{i=0}^{N} \overline{B}(x_i, 1/n)\Big) < \frac{\varepsilon}{2^n}.$$

Soit $K' = \bigcap_{n\in\mathbf{N}^*} \bigcup_{i=0}^{N} \overline{B}(x_i, 1/n)$. Cet ensemble est fermé, précompact donc compact.

On a de plus l'inégalité : $\mu(K'^c) < \varepsilon$. Soit $A \subset X$ un ensemble borélien. Par régularité extérieure, on peut trouver $F \subset A$ fermé tel que $\mu(A \setminus F) < \varepsilon$. Posons $K = K' \cap F \subset A$. Cet ensemble est compact et $\mu(A \setminus K) < 2\varepsilon$.

Théorème 3 (Lusin) *Soit X un espace topologique séparé, μ une mesure finie intérieurement régulière, Y un espace métrique séparable et $f : X \to Y$ une fonction borélienne. Alors pour tout ensemble borélien $A \subset X$, pour tout $\varepsilon > 0$, on peut trouver un compact $K \subset A$ tel que $\mu(A \setminus K) < \varepsilon$ et $f_{|K}$ est continue.*

Démonstration
Fixons $n \in \mathbf{N}^*$. Soit E_i une famille dénombrable disjointe d'ensembles boréliens de diamètre plus petit que $2/n$ qui recouvre Y. On peut construire les E_i par récurrence à partir d'une suite $\{y_i\}_{i\in\mathbf{N}^*}$ dense dans Y en posant $E_i = B(y_i, 1/n) \setminus \cap_{j<i}\ E_j$.

Par régularité, pour tout i, on peut trouver un borélien A_i' et un compact K_i tel que

$$K_i \subset A_i' \subset A \cap f^{-1}(E_i), \quad \mu(A \cap f^{-1}(E_i) \setminus K_i) < \varepsilon/2^{i+n}.$$

On a $\mu(A \setminus \cup K_i) < \frac{\varepsilon}{2^n}$ si bien qu'il existe $N \in \mathbf{N}$, dépendant de n, tel que

$$\mu\Big(A \setminus \bigcup_{i=0}^{N} K_i\Big) < \frac{\varepsilon}{2^n}\,.$$

Définissons une fonction f_n sur $K_n' = \cup_0^N K_i$ en posant $f(x) = y_i$ pour tout $x \in K_i$. Les compacts K_i sont disjoints, la fonction f_n est donc continue. Elle vérifie de plus $d(f_n(x), f(x)) < 1/n$ sur K_n'. La suite f_n converge vers f uniformément sur $\cap_n K_n'$, ce qui montre que f est continue sur ce compact. Ceci termine la preuve.

6 Complément : un ensemble non borélien

Soit $\alpha \in \mathbf{R}$. On considère la translation $t_\alpha(x) = x + \alpha$ et on note le translaté d'un ensemble $E \subset \mathbf{R}$ par $t_\alpha(E) = \alpha + E$. Remarquons que la mesure de Lebesgue est invariante par translation : pour tout ensemble borélien $A \subset \mathbf{R}$, $\lambda(A) = \lambda(\alpha + A)$. De fait, on vérifie sans difficulté que la fonction $A \mapsto \lambda(\alpha + A)$ est une mesure qui associe à un intervalle sa longueur. Cette propriété caractérise la mesure λ uniquement.

La proposition suivante, due à Waclaw Sierpiński (1882-1969), est de nature purement ensembliste. Elle va permettre d'exhiber un sous-ensemble de $\mathbf{R}$ non borélien.

Proposition 7 *Il existe des ensembles $E_i \subset [0, 1[$, $i \in \mathbf{Z}$, disjoints deux à deux et des réels $a_i \in [0, 2[$ tels que*

$$[0, 1[= \bigsqcup_{i\in\mathbf{Z}} E_i, \quad [0, 2[= \bigsqcup_{i\in\mathbf{Z}} a_i + E_i.$$

En conséquence, il ne peut pas y avoir de mesure μ satisfaisant $0 < \mu([0, 1[\,) < \infty$, invariante par translation et définie sur la tribu engendrée par les E_i et leurs translatés

car pour une telle mesure,

$$\mu([0,1[) = \sum_{i\in\mathbf{Z}} \mu(E_i) = \sum_{i\in\mathbf{Z}} \mu(a_i + E_i) = \mu([0,2[\,) = 2\mu([0,1[\,).$$

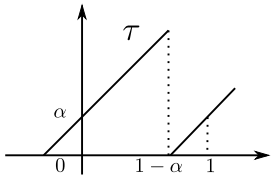
En particulier, on voit en prenant pour μ la mesure de Lebesgue λ qu'au moins un des ensembles E_i n'est pas borélien.

Démonstration
Considérons un réel irrationnel α et posons $\tau(x) = x + \alpha \bmod 1$.

L'application itérée n-ième de τ est une bijection de $[0,1[$ de la forme

$$\begin{aligned} \tau^n(x) &= x + 1 - \alpha_n \quad \text{sur } [0,\alpha_n[\\ &= x - \alpha_n \qquad \text{sur } [\alpha_n, 1[\end{aligned}$$

avec $\alpha_n = 1 + \text{partie entière}(n\alpha) - n\alpha$.



Notons que pour $x \in [0,1[$, les termes de la suite $\{\tau^n(x)\}_{n\in\mathbf{Z}}$ sont tous distincts et l'ensemble de ces suites constitue une partition de $[0,1[$. Choisissons un point dans chaque élément de cette partition afin de construire un nouvel ensemble A qui vérifie

$$[0,1[= \bigsqcup_{x\in A} \{\tau^n(x)\}_{n\in\mathbf{Z}} = \bigsqcup_{n\in\mathbf{Z}} \tau^n(A) = \bigsqcup_{k\in\mathbf{Z}} \tau^{2k}(A) \cup \bigsqcup_{k\in\mathbf{Z}} \tau^{2k+1}(A).$$

Il suffit de translater $\tau^{2k}(A)$ par τ^{-k} de manière à l'envoyer sur $\tau^k(A)$ puis de translater $\tau^{2k+1}(A)$ par $t_1 \circ \tau^{-k-1}$ pour l'envoyer sur $t_1(\tau^k(A))$ afin de reconstituer les intervalles $[0,1[$ et $[1,2[$. Rappelons que $t_1(x) = x + 1$. On peut tout expliciter comme suit :

$$\begin{aligned} E_{4i} &= \tau^{2i}(A) \cap [0,\alpha_{-i}[\,, & E_{4i+2} &= \tau^{2i+1}(A) \cap [0,\alpha_{-i-1}[\,, \\ E_{4i+1} &= \tau^{2i}(A) \cap [\alpha_{-i},1[\,, & E_{4i+3} &= \tau^{2i+1}(A) \cap [\alpha_{-i-1},1[\,, \end{aligned}$$

$$a_{4i} = 1 - \alpha_{-i}, \quad a_{4i+1} = -\alpha_{-i}, \quad a_{4i+2} = 2 - \alpha_{-i-1}, \quad a_{4i+3} = 1 - \alpha_{-i-1}.$$

7 Exercices

Les indications pour les exercices qui suivent se trouvent page 215.

Exercice 1.1
On considère l'ensemble des parties A de $\mathbf{R}$ qui satisfont la propriété suivante :

$$x \in A \quad \text{implique} \quad -x \in A.$$

Montrer que cet ensemble de parties est une tribu. Parmi les ensembles suivants, quels sont ceux qui appartiennent à cette tribu ?

$$\mathbf{N},\ \mathbf{Z},\ [0,1],\ \mathbf{Q},\ \{x \in \mathbf{R} \mid x^6 - 9x^2 + 1 = 0\},\ \emptyset,\ ]-2,2[\ ,\ \{-\sqrt{2}/2, 1/\sqrt{2}\}$$

Exercice 1.2
Soit X un ensemble, $f : X \to \mathbf{R}$ une application. Montrer que l'ensemble des parties de X de la forme $f^{-1}(B)$, avec B partie quelconque de $\mathbf{R}$, est une tribu. Quelle est la tribu obtenue lorsqu'on considère l'application $f : \mathbf{R} \to \mathbf{R}$ donnée par $f(x) = x^2$? Faire le lien avec l'exercice précédent.

Exercice 1.3
Soit $(X, \mathcal{T}, \mu)$ un espace mesuré. Montrer que pour tout A, $B \in \mathcal{T}$,

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B).$$

Exercice 1.4
– Soit A et B deux sous-ensembles boréliens de $[0,1]$ tels que $\lambda(A) > \frac{1}{2}$ et $\lambda(B) > \frac{1}{2}$. Montrer que $A \cap B$ est non vide.

– On considère trois sous-ensembles boréliens A, B et C tous de mesure de Lebesgue strictement supérieure à $2/3$. Montrer que $A \cap B \cap C$ est non vide.

– Généraliser à un nombre fini quelconque de sous-ensembles boréliens de $[0,1]$.

– Généraliser à un ensemble mesuré quelconque $(X, \mathcal{T}, \mu)$ tel que $\mu(X) < \infty$.

Exercice 1.5
On note $\mathbf{Q}$ l'ensemble des nombres rationnels. Montrer que $\lambda(\mathbf{Q}) = 0$.

Exercice 1.6
Soit $(X, \mathcal{T}, \mu)$ un espace mesuré tel que $\mu(X) < \infty$ et $f : X \to [0, \infty[$ une fonction mesurable. Montrer que pour tout $\varepsilon > 0$, il existe $A \subset X$ mesurable tel que $\mu(A^c) < \varepsilon$ et f est bornée sur A.

Exercice 1.7
Soit I un intervalle de $\mathbf{R}$ et $f : I \to \mathbf{R}$ une fonction dérivable. Montrer que f et f' sont des fonctions boréliennes. Donner un exemple de fonction f dérivable dont la dérivée n'est pas continue.