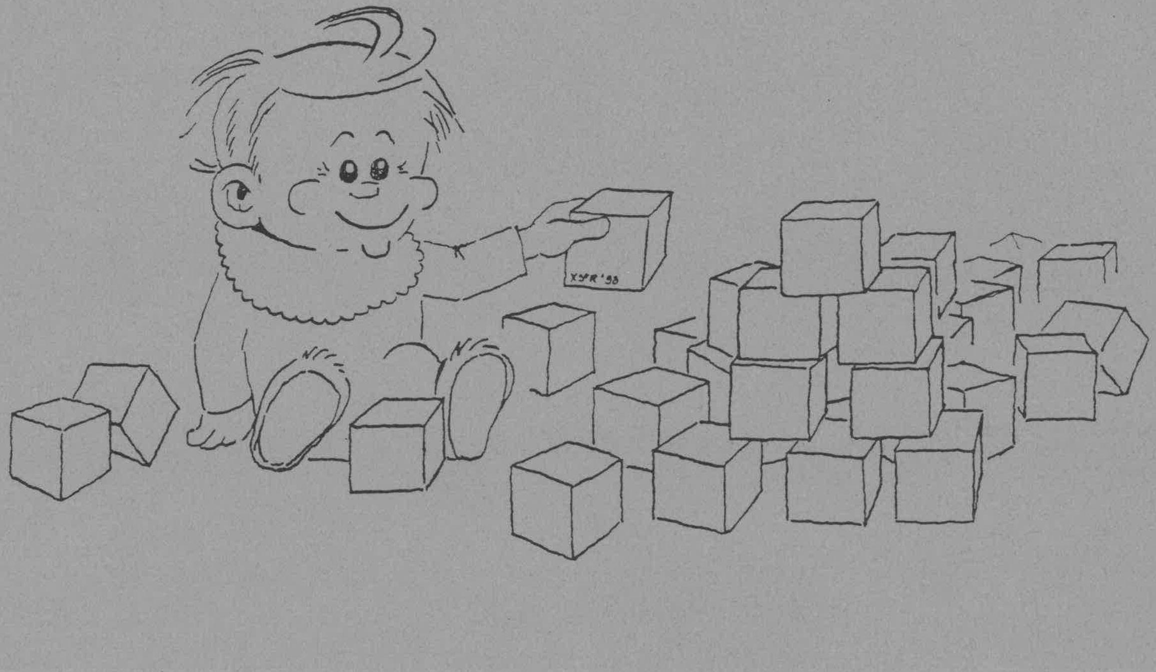
Xavier Saint Raymond

Intégrales simples et multiples

ou

L'intégrale de Lebesgue dans $\mathbb{R}^n$

Agrémenté de quelques dessins de l'auteur.



IREM des Pays de Loire

Édition du 1er novembre 1998

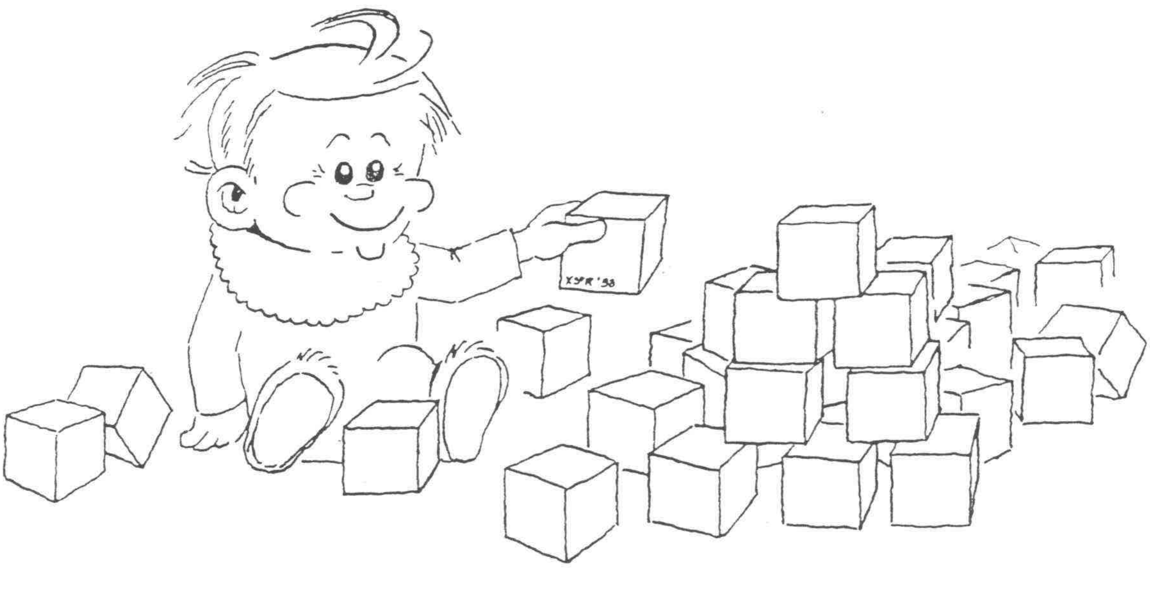
Xavier Saint Raymond

Intégrales simples et multiples

ou

L'intégrale de Lebesgue dans $\mathbb{R}^n$

Agrémenté de quelques dessins de l'auteur.



IREM des Pays de Loire

Édition du 1er novembre 1998

Avertissement

Ce petit polycopié présente l'essentiel de la théorie de l'intégration de Lebesgue dans le cadre (restreint) de la mesure de Lebesgue sur $\mathbb{R}^n$. L'intégrale y est construite directement à partir des fonctions en escalier sans passer ni par les tribus ni par la mesure de Lebesgue qui n'est définie qu'après coup[(1)]. Si l'on compare cette construction avec celle, plus traditionnelle, par les tribus et mesures, on la trouvera sans doute plus compliquée techniquement, notamment à cause du lemme d'Égorov du paragraphe 2, mais elle présente aussi des avantages : grâce à la signification très géométrique des intégrales des fonctions en escalier et à l'absence de notions intermédiaires, notre présentation est sans doute plus motivante pour les étudiants. Dans ce format, ce cours vise en réalité deux publics bien différents : les étudiants en licence de mathématiques, et les candidats au concours de l'agrégation de mathématiques.

À l'intention des étudiants de licence, nous avons distingué typographiquement deux niveaux de lecture et avons ajouté en appendice un paragraphe de *préliminaires indispensables*. Dans ce dernier paragraphe se trouvent inclus un peu de logique mathématique et théorie des ensembles généralement méconnues de nos étudiants, quelques notions topologiques sur les fonctions de plusieurs variables réelles qui vont peut-être un peu au-delà de ce qui est effectivement enseigné en premier cycle, et des conseils de rédaction pour les devoirs. Quant aux deux niveaux typographiques, ils ont été conçus pour éviter en première lecture les démonstrations trop techniques qui risquent de faire perdre de vue les idées essentielles, et qui sont signalées dans le texte par des caractères plus petits.

Les candidats à l'agrégation trouveront naturellement dans ce polycopié un exposé assez complet de la théorie de l'intégration dans $\mathbb{R}^n$. Ils y trouveront en outre un certain nombre de compléments utiles : la notion de famille sommable de nombres complexes (paragraphe 2), la dérivabilité presque partout des fonctions monotones (paragraphe 6), un théorème fondamental de l'analyse dans le cadre des fonctions lipschitziennes et mesurables bornées (paragraphe 6), la transformation de Fourier et la convolution des fonctions intégrables (paragraphe 7), les fonctions gaussiennes et la formule d'inversion de Fourier (paragraphe 8), la transformation de Fourier des fonctions de carré intégrable (paragraphe 9), et un petit exposé succinct mais substantiel sur les séries de Fourier (paragraphe 10). Enfin, ils pourront élaborer eux-mêmes d'autres compléments à l'aide des nombreux exercices complétant chacun des dix paragraphes du cours.

L'auteur espère que ce polycopié pourra ainsi rendre service aux étudiants concernés, et tient à remercier l'IREM des Pays de Loire qui en a assuré l'impression et la diffusion.

Nantes, le 1er novembre 1998,
X. Saint Raymond.

---

(1) Ce cours s'inspire largement d'un cours de J. Dixmier à la Sorbonne, *L'intégrale de Lebesgue*, qui propose la même démarche. Merci par ailleurs à P. Gérard qui a bien voulu me prêter ses notes sur le lemme d'Égorov.

Introduction

On rencontre souvent en mathématiques des expressions du genre

$$m_1 f(1) + \ldots + m_N f(N)$$

qui représentent des "moyennes" des nombres $f(k)$ affectés des poids $m_k > 0$. Un exemple type de cette situation est fourni par la définition du barycentre d'un système de N points p_k de masses m_k, qui est le point g défini par la formule

$$m\,\overrightarrow{og} = m_1\overrightarrow{op_1} + \ldots + m_N\,\overrightarrow{op_N}\,,$$

où $m = m_1 + \ldots + m_N$ est la masse totale du système. Si maintenant au lieu de dépendre d'une variable discrète $k = 1, \ldots, N$, la quantité f dépendait d'une variable continue $x \in \mathbb{R}$, il nous faudrait remplacer la somme discrète $m_1 f(1) + \ldots + m_N f(N)$ par une "somme continue", c'est-à-dire par l'intégrale

$$\int f(x)\,dm(x)$$

où la notation $dm(x)$ indique que l'on attribue peut-être plus de poids à certains domaines décrits par la variable x. Lorsque le poids est uniformément réparti, on notera plus simplement $\int f(x)\,dx$ cette intégrale. Enfin, si la quantité f dépend de n variables réelles $(x_1, \ldots, x_n) = x \in \mathbb{R}^n$, l'analogue des sommes précédentes sera l'intégrale multiple

$$\int \cdots \int f(x_1, \ldots, x_n)\,dx_1 \ldots dx_n$$

que nous noterons (à nouveau) simplement $\int f(x)\,dx$.

Objectifs du cours

Tout au long de ce cours, nous poursuivrons un triple but :

(a) Définir les intégrales des fonctions de n variables réelles. Cette définition procède en deux temps : on définit d'abord l'intégrale d'une fonction en escalier comme une somme $m_1 f(x^1) + \ldots + m_N f(x^N)$, puis on étend cette intégrale aux fonctions plus générales à l'aide d'un procédé de passage à la limite que nous tâcherons de détailler. Bien entendu, les fonctions d'une seule variable réelle rentrent dans ce cadre comme un cas particulier.

(b) Décrire les principales propriétés de l'intégrale. Ces propriétés peuvent être établies rigoureusement à partir du moment où l'on dispose d'une définition précise de l'intégrale. En plus des propriétés classiques (linéarité, continuité et positivité), nous démontrerons les résultats suivants :
Théorème 1. *Si une suite de fonctions* (f_p) *converge vers* f, *alors la suite numérique* $\left(\int f_p(x)\,dx\right)$ *tend vers* $\int f(x)\,dx$;
Théorème 2. *Si* $f(x,y)$ *est de classe* C^1 *par rapport aux variables* $y = (y_1,\ldots,y_m)$, *alors la fonction* $g(y) = \int f(x,y)\,dx$ *aussi, et on a* $(\partial g/\partial y_j)(y) = \int(\partial f/\partial y_j)(x,y)\,dx$;
Théorème 3. *Les fonctions* $g(y) = \int f(x,y)\,dx$ *et* $h(x) = \int f(x,y)\,dy$ *vérifient l'identité* $\int g(y)\,dy = \int f(z)\,dz = \int h(x)\,dx$.

(c) Fournir des critères d'intégrabilité et des méthodes de calcul des intégrales. Nous nous attacherons en particulier à justifier les méthodes classiques de calcul : par recherche de primitive et intégration par parties pour les intégrales simples, par intégration itérée et changement de variables pour les intégrales multiples.

Il serait dans le prolongement naturel de ce cours de généraliser les résultats précédents à un cadre abstrait, car les arguments développés pour étudier l'intégrale des fonctions de n variables réelles s'adaptent bien au calcul d'intégrales de fonctions dépendant de variables décrivant un ensemble (abstrait) dont on sait mesurer tout une famille de parties dites *mesurables*. Mais pour limiter la taille de ce petit polycopié, nous n'aborderons pas ici cette généralisation, et nous nous contentons de renvoyer l'étudiant à l'un des très nombreux ouvrages d'intégration disponibles dans le commerce, par exemple à l'ouvrage récent de J. Gapaillard, *Intégration pour la licence.*

Remarque fondamentale

Les théorèmes cités ci-dessus sont incomplets : on n'y a indiqué que ce qu'il ne faut pas retenir, à savoir la conclusion. Ce qu'il sera important de retenir, ce seront les hypothèses, que nous détaillerons pendant le cours. Les conclusions de ces trois théorèmes peuvent aussi s'écrire

$$\lim_{p\to\infty}\int f_p(x)\,dx = \int \lim_{p\to\infty} f_p(x)\,dx\,,$$
$$\frac{\partial}{\partial y_j}\int f(x,y)\,dx = \int \frac{\partial f}{\partial y_j}(x,y)\,dx\,,$$
$$\text{Intég}_y\int f(x,y)\,dx = \int \text{Intég}_y\, f(x,y)\,dx$$

en notant Intég_y l'opération consistant à calculer l'intégrale par rapport aux variables y, et compte tenu de ce que cette opération, comme l'opération de dérivation d'ailleurs, comporte un passage à la limite, ces propriétés sont en fait des propriétés d'interversion d'intégrales et de limites. Cela nous ramène donc à l'un des problèmes fondamentaux de l'analyse : sous quelles hypothèses a-t-on le droit d'intervertir deux limites ?

On trouve facilement des exemples montrant qu'une telle interversion n'est pas toujours possible. Ainsi considérons le cas de la suite de fonctions d'une variable réelle $(f_p)_{p\in\mathbb{Z}}$ définie par $f_p(x) = 2^p f_0(2^p x)$, où f_0 est la fonction s'annulant en dehors de l'intervalle $[0, 2]$ et vérifiant

$$f_0(x) = x \quad \text{pour } x \in [0, 1], \qquad f_0(x) = 2 - x \quad \text{pour } x \in [1, 2].$$

En effectuant le changement de variable $y = 2^p x$, on voit que $\int f_p(x)\,dx = \int f_0(y)\,dy = 1$ pour tout $p \in \mathbb{Z}$, et donc que cette suite numérique (constante) tend vers 1 lorsque p tend vers $\pm\infty$. Et pourtant, lorsque p tend vers $+\infty$, il y a convergence simple de cette suite de fonctions vers la fonction nulle, ce qui montre qu'on ne peut passer à la limite dans l'intégrale ; ce qui empêche de passer à la limite est ici un phénomène de *concentration de masse*. Et de la même façon, on voit facilement que la suite (f_p) converge uniformément vers 0 lorsque p tend vers $-\infty$, ce qui ne permet pas davantage de passer à la limite à cause cette fois d'un phénomène de *fuite de masse à l'infini*. Dans nos théorèmes d'interversion d'intégrales et de limites, nous devrons donc poser des hypothèses qui empêchent ces deux phénomènes de se produire.

Pourquoi passer à la limite dans les intégrales ?

Ce problème intervient naturellement, par exemple, dans la situation suivante.

Théorème (de Cauchy-Lipschitz). *Soit $F : \mathbb{R}^2 \to \mathbb{R}$ une fonction continue vérifiant $|F(x,y) - F(x,z)| \le C\,|y - z|$ pour tous x, y et $z \in \mathbb{R}$, la constante C étant indépendante de x, y et z. Alors pour tous x_0 et $y_0 \in \mathbb{R}$, il existe une unique fonction $f : \mathbb{R} \to \mathbb{R}$ de classe C^1 telle que*

$$f'(x) = F(x, f(x)) \qquad \textit{et} \qquad f(x_0) = y_0\,.$$

En intégrant l'équation différentielle proposée entre les bornes x_0 et x, on voit que la solution f doit vérifier

$$f(x) = y_0 + \int_{x_0}^{x} F(t, f(t))\,dt\,,$$

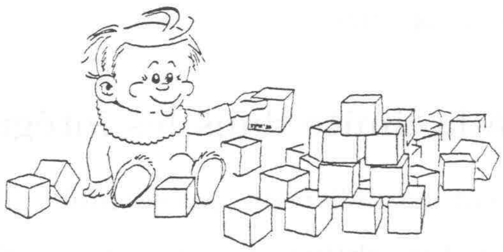
et que réciproquement, toute fonction continue f solution de cette équation intégrale est une solution de classe C^1 du problème posé. On résout cette équation intégrale par approximations successives en posant $f_0(x) = y_0$ puis

$$f_{p+1}(x) = y_0 + \int_{x_0}^{x} F(t, f_p(t))\,dt\,.$$

On montre ensuite[1] que la suite (f_p) converge vers une fonction continue f, ce qui permet d'affirmer que

$$f(x) = y_0 + \lim_{p\to\infty} \int_{x_0}^{x} F(t, f_p(t))\, dt \quad \text{et} \quad \int_{x_0}^{x} F(t, f(t))\, dt = \int_{x_0}^{x} \lim_{p\to\infty} F(t, f_p(t))\, dt\,.$$

On voit donc que pour conclure, il suffit de disposer d'un théorème d'interversion de limite et d'intégrale. ∎



[1] On montre en effet par récurrence que pour tout $t\in[x_0,x]$ et tout $p>0$, $|f_p(t)-f_{p-1}(t)|\leq \sup_{[x_0,x]} |F(\,.\,,y_0)|\,(t-x_0)^p/p!$, ce qui entraîne la convergence uniforme sur tout intervalle compact de la suite (f_p). C'est aussi cette propriété qui permet d'intervertir la limite et l'intégrale comme nous le verrons au corollaire 4.5.

1. Fonctions en escalier

À quoi servent les intégrales ? L'une des premières réponses qui viennent à l'esprit, c'est qu'elles servent à calculer des volumes, des aires et des longueurs, et plus généralement à *mesurer*. Il est donc naturel de commencer par poser la question : qu'est-ce que mesurer ? Il semble que du point de vue pratique, cela consiste à compter combien de fois on peut ranger un objet, choisi préalablement comme unité, dans l'objet à mesurer. Dans cette perspective, il convient d'étudier d'abord les objets qui se rangent le plus commodément.

Définition 1.1. *On appelle* pavé *de* $\mathbb{R}^n$ *tout ensemble* $P \subset \mathbb{R}^n$ *de la forme*

$$P = [a_1, b_1] \times \ldots \times [a_n, b_n] = \{\, (x_1, \ldots, x_n) \in \mathbb{R}^n \,;\, a_j \leq x_j \leq b_j \,\},$$

et on appelle mesure *de ce pavé le nombre réel positif* $|P| = \prod_{j=1}^{n} (b_j - a_j)$.

Soulignons que les pavés considérés dans cette définition sont compacts : ce sera toujours sous-entendu lorsque nous parlerons de pavés. Nous utiliserons fréquemment les propriétés élémentaires suivantes des pavés, pour l'énoncé desquelles il est commode de convenir que $\varnothing$ est aussi un pavé.

Lemme 1.2. *Soit* P *un pavé de* $\mathbb{R}^n$. *Alors :*
(a) *Pour tout* $x_0 \in \mathbb{R}^n$, *le translaté de* P *de vecteur* x_0, *c'est-à-dire l'ensemble* $P + x_0 = \{\, x + x_0 \,;\, x \in P \,\}$, *est un pavé de même mesure que* P.
(b) *L'intersection de* P *avec un autre pavé est encore un pavé.*
(c) *Les pavés de* $\mathbb{R}^{n+m} \simeq \mathbb{R}^n \times \mathbb{R}^m$ *sont les produits cartésiens d'un pavé* P *de* $\mathbb{R}^n$ *par un pavé* Q *de* $\mathbb{R}^m$, *avec la relation* $|P \times Q| = |P|\,|Q|$.

Démonstration. En écrivant $P = [a_1, b_1] \times \ldots \times [a_n, b_n]$ et $x_0 = (c_1, \ldots, c_n)$, le pavé translaté vaut $P + x_0 = [a_1 + c_1, b_1 + c_1] \times \ldots \times [a_n + c_n, b_n + c_n]$, et c'est donc un pavé de même mesure que P, d'où (a). L'intersection de deux intervalles compacts est clairement un intervalle compact (sauf si cette intersection est vide), et la propriété (b) résulte alors de l'identité $\big(\prod_{j=1}^{n} A_j\big) \cap \big(\prod_{j=1}^{n} B_j\big) = \prod_{j=1}^{n} (A_j \cap B_j)$. Enfin un pavé de $\mathbb{R}^{n+m}$ est de la forme $[a_1, b_1] \times \ldots \times [a_{n+m}, b_{n+m}] = P \times Q$ en notant $P = [a_1, b_1] \times \ldots \times [a_n, b_n]$ et $Q = [a_{n+1}, b_{n+1}] \times \ldots \times [a_{n+m}, b_{n+m}]$, et la relation entre les mesures de la propriété (c) provient alors de l'associativité de la multiplication. ∎

Lorsque nous étudierons des familles de pavés empiétant les uns sur les autres, nous pourrons utiliser le résultat suivant qui sert à démêler les enchevêtrements.

Lemme 1.3. *Si $P_1, \ldots, P_N$ sont des pavés de $\mathbb{R}^n$, alors il existe un ensemble fini d'indices I, et pour $i \in I$ des points $x^i \in \mathbb{R}^n$ et des réels positifs m_i tels que pour tout indice $k \leq N$, $|P_k| = \sum_{i;x^i \in P_k} m_i$.*

Démonstration. En notant $P_k = [a_{k,1}, b_{k,1}] \times \ldots \times [a_{k,n}, b_{k,n}]$ pour $k = 1, \ldots, N$, on range pour tout j fixé les nombres $a_{k,j}$ et $b_{k,j}$ comme une suite croissante $c_{1,j} \leq c_{2,j} \leq \ldots \leq c_{2N,j}$, puis on appelle x^i et m_i respectivement le point central et la mesure du pavé

$$Q_i = [c_{i_1-1,1}, c_{i_1,1}] \times \ldots \times [c_{i_n-1,n}, c_{i_n,n}]$$

pour tout $i = (i_1, \ldots, i_n) \in I = \{2, \ldots, 2N\}^n$. Pour tout indice k fixé, on a $x^i \in P_k$ si et seulement si l'indice i vérifie $a_{k,j} \leq c_{i_j-1,j} \leq c_{i_j,j} \leq b_{k,j}$ pour tout $j \leq n$, et alors la somme de toutes les mesures m_i pour ces indices redonne bien la mesure de P_k par factorisation. ■

Dans la suite, nous noterons $\mathbb{1}_A$ la fonction *indicatrice*[1] d'une partie A de $\mathbb{R}^n$, c'est-à-dire la fonction qui vaut $\mathbb{1}_A(x) = 1$ si $x \in A$ et $\mathbb{1}_A(x) = 0$ sinon.

Comme nous l'avons rappelé en tête de ce paragraphe, les intégrales servent d'abord à mesurer. Ainsi, ce que nous attendons de l'intégrale d'une fonction positive $f : \mathbb{R}^n \to \mathbb{R}_+$, c'est qu'elle mesure l'espace de $\mathbb{R}^{n+1}$ compris entre l'hyperplan $y = 0$ et le graphe $y = f(x)$. Dans le cas d'une fonction $f(x) = c\,\mathbb{1}_P(x)$ où c est un réel positif et P un pavé de $\mathbb{R}^n$, cet espace est un pavé de $\mathbb{R}^{n+1}$ de mesure égale à $c\,|P|$. Dans le cas d'une somme (finie) de fonctions $c_k\,\mathbb{1}_{P_k}$ pour des pavés disjoints P_k, on est amené à considérer une réunion finie de pavés disjoints de $\mathbb{R}^{n+1}$, dont la mesure totale sera naturellement prise égale à la somme des mesures de ces différents pavés, soit $\sum_{k=1}^N c_k\,|P_k|$. Plus généralement, dans le cas de coefficients complexes et de pavés empiétant les uns sur les autres, on prendra la définition suivante.

Proposition et définition 1.4. *On appelle* fonction en escalier *dans $\mathbb{R}^n$ toute combinaison linéaire (à coefficients complexes) de fonctions indicatrices de pavés. Si $f = \sum_{k=1}^N c_k\,\mathbb{1}_{P_k}$ est une fonction en escalier, le nombre complexe $\sum_{k=1}^N c_k\,|P_k|$ ne dépend que de la fonction f, mais pas de l'écriture utilisée, et ce nombre est appelé* intégrale *de la fonction f. On le note $\int f(x)\,dx$.*

Démonstration. Il nous faut vérifier que si une fonction en escalier f admet deux écritures

$$f = \sum_{k=1}^N c_k\,\mathbb{1}_{P_k} = \sum_{\ell=1}^M d_\ell\,\mathbb{1}_{Q_\ell},$$

alors on a l'égalité $\sum_{k=1}^N c_k\,|P_k| = \sum_{\ell=1}^M d_\ell\,|Q_\ell|$. Pour cela, on utilise les points x^i et les mesures m_i du lemme 1.3 associés à la collection des pavés P_k et Q_ℓ, qui

(1) Cette fonction est aussi appelée fonction *caractéristique* de A.

permettent d'écrire

$$\sum_{k=1}^{N} c_k\,|P_k| = \sum_{k=1}^{N} c_k\Big(\sum_{i;x^i\in P_k} m_i\Big) = \sum_{i\in I} m_i\Big(\sum_{k;x^i\in P_k} c_k\Big) = \sum_{i\in I} m_i\, f(x^i) =$$

$$= \sum_{i\in I} m_i\Big(\sum_{\ell;x^i\in Q_\ell} d_\ell\Big) = \sum_{\ell=1}^{M} d_\ell\Big(\sum_{i;x^i\in Q_\ell} m_i\Big) = \sum_{\ell=1}^{M} d_\ell\,|Q_\ell|\,,$$

ce qui prouve l'égalité voulue. ∎

Les premières propriétés des fonctions en escalier sont données dans l'énoncé suivant.

Proposition 1.5. *Soit $f : \mathbb{R}^n \to \mathbb{C}$ une fonction en escalier. Alors :*
(a) *Pour tout $x_0 \in \mathbb{R}^n$, la fonction translatée $x \mapsto f(x - x_0)$ est en escalier.*
(b) *Pour tout $c \in \mathbb{C}$ et toute fonction $g : \mathbb{R}^n \to \mathbb{C}$ en escalier, les fonctions $\overline{f}$, $f + c\,g$ et fg sont en escalier.*
(c) *Si g est une fonction en escalier dans $\mathbb{R}^m$, la fonction $h : (x, y) \mapsto f(x)\,g(y)$ est en escalier dans $\mathbb{R}^n \times \mathbb{R}^m$.*
(d) *Si $g : \mathbb{C} \to \mathbb{C}$ vérifie $g(0) = 0$, alors la fonction $g \circ f$ est en escalier.*
(e) *La fonction $|f|$ est en escalier.*
(f) *Si $f_1, \ldots, f_N$ sont des fonctions en escalier à valeurs réelles, alors il en est de même des fonctions $\sup(f_1, \ldots, f_N)$ et $\inf(f_1, \ldots, f_N)$.*
(g) *Si $P_1, \ldots, P_N$ sont des pavés de $\mathbb{R}^n$ et $A = \cup_{k=1}^{N} P_k$, alors $\mathbb{1}_A$ est en escalier.*

Démonstration. Dans toute cette démonstration, on écrira $f = \sum_{k=1}^{N} c_k\,\mathbb{1}_{P_k}$.

Les propriétés (a), (b) et (c) sont simples à obtenir. Pour la fonction translatée cela résulte de ce que $\mathbb{1}_P(x - x_0) = \mathbb{1}_{P+x_0}(x)$ est la fonction indicatrice du pavé $P + x_0$, pour la fonction conjuguée cela résulte de la formule $\overline{f} = \sum_{k=1}^{N} \overline{c_k}\,\mathbb{1}_{P_k}$, pour les combinaisons linéaires cela résulte de la définition, tandis que pour les produits cela résulte de l'identité $\mathbb{1}_A\,\mathbb{1}_B = \mathbb{1}_{A\cap B}$ et de ce que l'intersection de deux pavés est encore un pavé. Enfin, si $g = \sum_{\ell=1}^{M} d_\ell\,\mathbb{1}_{Q_\ell}$ pour des pavés Q_ℓ de $\mathbb{R}^m$, la fonction $h(x, y) = f(x)\,g(y)$ s'écrit donc

$$h(x, y) = \sum_{k=1}^{N}\sum_{\ell=1}^{M} c_k\, d_\ell\,\mathbb{1}_{P_k}(x)\mathbb{1}_{Q_\ell}(y) = \sum_{k,\ell} e_{k,\ell}\,\mathbb{1}_{R_{k,\ell}}(x, y)$$

où les $e_{k,\ell} = c_k\, d_\ell$ sont des nombres complexes et les $R_{k,\ell} = P_k \times Q_\ell$ sont des pavés de $\mathbb{R}^n \times \mathbb{R}^m$.

Pour établir la propriété (d), on note $K = \{1, 2, \ldots, N\}$. Pour tout $I \subset K$, on considère les ensembles $P_I = \cap_{k\in I} P_k$ qui sont des pavés (éventuellement vides), puis on définit des nombres complexes c_I récursivement suivant le nombre d'éléments de I en posant $c_\varnothing = 0$, puis $c_I = g\big(\sum_{k\in I} c_k\big) - \sum_{J\subset I,\, J\neq I} c_J$ lorsque

I n'est pas vide. Alors la fonction $g \circ f$ est égale à la fonction en escalier $h = \sum_{I \subset K} c_I \, \mathbb{1}_{P_I}$ puisque pour tout $x \in \mathbb{R}^n$ on a $f(x) = \sum_{k \in I(x)} c_k$ en notant $I(x) = \{ k \in K \,;\, x \in P_k \}$, ce qui implique que $g \circ f(x) = g\big(\sum_{k \in I(x)} c_k\big)$, d'où $h(x) = 0 = g(0) = g \circ f(x)$ si $I(x) = \emptyset$, et

$$h(x) = \sum_{I \subset K} c_I \, \mathbb{1}_{P_I}(x) = \sum_{I \subset I(x)} c_I = c_{I(x)} + \sum_{J \subset I(x),\, J \neq I(x)} c_J =$$
$$= g\Big(\sum_{k \in I(x)} c_k\Big) = g \circ f(x)$$

si $I(x) \neq \emptyset$, par définition de $c_{I(x)}$.

On obtient alors la propriété (e) en remarquant que $|f| = g \circ f$ si on note g la fonction $g(z) = |z|$ qui vérifie $g(0) = 0$. Dans le cas de deux fonctions en escalier f_1 et f_2, la propriété (f) découle des propriétés (b) et (e) puisque

$$\sup(f_1, f_2) = \frac{f_1 + f_2}{2} + \frac{|f_1 - f_2|}{2} \qquad \text{et} \qquad \inf(f_1, f_2) = \frac{f_1 + f_2}{2} - \frac{|f_1 - f_2|}{2},$$

et le cas d'un nombre fini de fonctions en escalier s'en déduit par récurrence. Enfin, la propriété (g) résulte de (f) puisque $\mathbb{1}_A = \sup(\mathbb{1}_{P_1}, \dots, \mathbb{1}_{P_N})$. ∎

Enfin, les propriétés des intégrales des fonctions en escalier sont résumées dans l'énoncé suivant.

Proposition 1.6. *Soit $f : \mathbb{R}^n \to \mathbb{C}$ une fonction en escalier. Alors :*
(a) Invariance *: pour tout $x_0 \in \mathbb{R}^n$ fixé, $\int f(x - x_0)\, dx = \int f(x)\, dx$.*
(b) Linéarité *: pour toute constante $c \in \mathbb{C}$ et toute fonction $g : \mathbb{R}^n \to \mathbb{C}$ en escalier, $\int \overline{f}(x)\, dx = \overline{\int f(x)\, dx}$ et $\int (f + c\, g)(x)\, dx = \int f(x)\, dx + c \int g(x)\, dx$.*
(c) Tensorialité *: si $g : \mathbb{R}^m \to \mathbb{C}$ est en escalier, la fonction en escalier (sur $\mathbb{R}^{n+m}$) $h : (x, y) \mapsto f(x)\, g(y)$ vérifie $\int h(z)\, dz = \big(\int f(x)\, dx\big)\big(\int g(y)\, dy\big)$.*
(d) Continuité *: $\big|\int f(x)\, dx\big| \leq \int |f(x)|\, dx$.*
(e) Positivité *: si f est à valeurs réelles positives, alors $\int f(x)\, dx \geq 0$.*
(f) Croissance *: si $f \leq g$ sont en escalier dans $\mathbb{R}^n$ et à valeurs réelles, alors $\int f(x)\, dx \leq \int g(x)\, dx$.*

Démonstration. En reprenant les notations de la démonstration de la proposition précédente, on a $f(x - x_0) = \sum_{k=1}^N c_k \, \mathbb{1}_{P_k + x_0}(x)$, ce qui prouve la propriété d'invariance (par translation) puisque les $P_k + x_0$ sont des pavés de mêmes mesures que les P_k. La propriété de linéarité, qui implique que l'intégrale d'une fonction à valeurs réelles est un nombre réel, découle directement des définitions, et pour la propriété de tensorialité, il suffit d'utiliser que les pavés $P_k \times Q_\ell$ de $\mathbb{R}^n \times \mathbb{R}^m$ ont pour mesure $|P_k \times Q_\ell| = |P_k|\,|Q_\ell|$, ce qui permet de factoriser la somme des $c_k\, d_\ell\, |P_k \times Q_\ell|$.

En ce qui concerne la continuité, on peut écrire $|f| = \sum_{\ell=1}^M d_\ell \, \mathbb{1}_{Q_\ell}$ puisque cette fonction est en escalier. Nous utilisons alors les points x^i et les mesures m_i

du lemme 1.3, mais associés à la famille de pavés $P_1, \ldots, P_N, Q_1, \ldots, Q_M$, qui permettent d'écrire

$$\left|\int f(x)\,dx\right| = \left|\sum_{k=1}^{N} c_k\,|P_k|\right| = \left|\sum_{i\in I} m_i \Big(\sum_{k;x^i\in P_k} c_k\Big)\right| = \left|\sum_{i\in I} m_i\, f(x^i)\right| \leq$$

$$\leq \sum_{i\in I} m_i\,|f(x^i)| = \sum_{i\in I} m_i \Big(\sum_{\ell;x^i\in Q_\ell} d_\ell\Big) = \sum_{\ell=1}^{M} d_\ell\,|Q_\ell| = \int |f(x)|\,dx\,,$$

ce qui prouve l'inégalité cherchée.

Enfin, la propriété de positivité résulte de la propriété de continuité, tandis que la propriété de croissance s'obtient en combinant linéarité et positivité. ∎

Pour la démonstration du lemme d'Égorov au paragraphe 2, nous aurons besoin des compléments géométriques suivants.

Définition 1.7. *On appelle* domaine pavable de $\mathbb{R}^n$ *toute partie $D \subset \mathbb{R}^n$ telle que la fonction $\mathbb{1}_D$ soit en escalier.*

On peut remarquer que toute réunion finie de domaines pavables est un domaine pavable puisque si $D = D_1 \cup \ldots \cup D_N$, alors la fonction $\mathbb{1}_D = \sup(\mathbb{1}_{D_1}, \ldots, \mathbb{1}_{D_N})$ est en escalier d'après la proposition 1.5 (f).

Lemme 1.8.
(a) *Pour toute fonction f en escalier et toute partie $A \subset \mathbb{C} \setminus \{0\}$, l'image réciproque par f de A est un domaine pavable.*
(b) *Tout domaine pavable D est contenu dans une réunion finie $P_1 \cup \ldots \cup P_N$ de pavés vérifiant $\sum_{k=1}^{N} |P_k| = \int \mathbb{1}_D(x)\,dx$.*
(c) *Pour tout domaine pavable D, il existe des pavés $Q_1, \ldots, Q_N$ de mesures nulles tels que $D \cup Q_1 \cup \ldots \cup Q_N$ soit compact.*

Démonstration. Pour démontrer la propriété (a), considérons une fonction f en escalier, $c \neq 0$ un nombre complexe, $D = f^{-1}(c)$ et P un pavé tel que $f = 0$ en dehors de P. Alors $\varepsilon = \inf_{x\in P,\, x\notin D} |f(x) - c\,\mathbb{1}_P(x)| > 0$ puisque $f - c\,\mathbb{1}_P$ est une fonction en escalier qui ne s'annule que sur D et en dehors de P. On a donc $\mathbb{1}_D = \sup\big(\mathbb{1}_P - \varepsilon^{-1}\,|f - c\,\mathbb{1}_P|\ ,\ 0\ \big)$, et c'est une fonction en escalier d'après la proposition 1.5. Comme f ne prend qu'un nombre fini de valeurs, l'image réciproque par f d'une partie $A \subset \mathbb{C} \setminus \{0\}$ est une réunion finie de domaines de la forme $f^{-1}(c)$ du type ci-dessus, et c'est donc un domaine pavable.

Pour démontrer la propriété (b), on écrit $\mathbb{1}_D = \sum_{\ell=1}^{M} c_\ell\,\mathbb{1}_{Q_\ell}$, et on reprend la démonstration du lemme 1.3 qui fournit un ensemble fini d'indices I et des pavés P_i pour $i \in I$ adaptés à la famille des Q_ℓ et dont on note x^i les centres et m_i les mesures. En posant $J = \{\, i \in I\,;\ x^i \in D\,\}$, on voit que par construction des P_i, on a $D \subset \cup_{i\in J} P_i$ et $\int \mathbb{1}_D(x)\,dx = \sum_{i\in I} m_i\,\mathbb{1}_D(x^i) = \sum_{i\in J} |P_i|$.

Pour démontrer la propriété (c), on écrit $\mathbb{1}_D = \sum_{k=1}^{M} c_k\,\mathbb{1}_{P_k}$, puis on note $Q_1, \ldots, Q_N$ la famille de pavés de mesures nulles dont la réunion forme les bords des pavés P_k. En dehors de la réunion de ces pavés Q_ℓ, la fonction $\sum_{k=1}^{M} c_k\,\mathbb{1}_{P_k}$ est localement constante, ce qui prouve que la frontière de D est contenue dans cette réunion. On peut donc écrire que $D \cup Q_1 \cup \ldots \cup Q_N = \overline{D} \cup Q_1 \cup \ldots \cup Q_N$, et cette dernière réunion est bien compacte puisque $\overline{D}$ est compact comme fermé borné et que les Q_ℓ sont aussi compacts. ∎

Exercices

1.1. Soient A et B deux parties d'un ensemble E. On note $A \Delta B = (A \setminus B) \cup (B \setminus A)$ la *différence symétrique* de A et B.

(a) Montrer que l'on a $\mathbb{1}_{A\cap B} = \mathbb{1}_A \mathbb{1}_B = \inf(\mathbb{1}_A, \mathbb{1}_B)$, $\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \mathbb{1}_B = \sup(\mathbb{1}_A, \mathbb{1}_B)$, $\mathbb{1}_{E\setminus A} = 1 - \mathbb{1}_A$ et $\mathbb{1}_{A\Delta B} = \mathbb{1}_A + \mathbb{1}_B - 2\,\mathbb{1}_A \mathbb{1}_B$.

(b) En déduire que $\mathscr{P}(E)$ muni du produit $(A, B) \mapsto A\Delta B$ est un groupe abélien.

1.2. Soit $f : \mathbb{R}^n \to \mathbb{C}$ une fonction en escalier. Montrer que cette fonction ne prend qu'un nombre fini de valeurs.

1.3. On considère l'espace vectoriel $\mathbb{R}^m$ muni d'une norme $u \mapsto \|u\|$. Une fonction $f : \mathbb{R}^n \to \mathbb{R}^m$ est dite *en escalier* s'il existe des pavés $P_k \subset \mathbb{R}^n$ et des vecteurs $u_k \in \mathbb{R}^m$ tels que $f = \sum_{k=1}^N \mathbb{1}_{P_k} u_k$.

(a) Soient $f : \mathbb{R}^n \to \mathbb{R}^m$ une fonction et $f_1 : \mathbb{R}^n \to \mathbb{R}, \ldots, f_m : \mathbb{R}^n \to \mathbb{R}$ ses composantes dans la base canonique de $\mathbb{R}^m$. Montrer que f est en escalier si et seulement si chaque f_j est en escalier.

(b) Montrer que si $f : \mathbb{R}^n \to \mathbb{R}^m$ est en escalier, le vecteur $\int f(x)\,dx = \sum_{k=1}^N |P_k|\, u_k$ ne dépend que de f, mais pas de l'écriture $f = \sum_{k=1}^N \mathbb{1}_{P_k} u_k$ utilisée, et que les composantes de $\int f(x)\,dx$ dans la base canonique de $\mathbb{R}^m$ valent $\int f_1(x)\,dx, \ldots, \int f_m(x)\,dx$.

(c) Montrer que si $f : \mathbb{R}^n \to \mathbb{R}^m$ est en escalier, alors $\|f\| : \mathbb{R}^n \to \mathbb{R}$ est aussi en escalier, et que $\|\int f(x)\,dx\| \le \int \|f(x)\|\,dx$.

1.4. Soit $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{C}$ une fonction en escalier. On notera x les variables dans $\mathbb{R}^n$, y les variables dans $\mathbb{R}^m$ et $z = (x, y)$ les variables dans $\mathbb{R}^n \times \mathbb{R}^m$.

(a) Montrer que pour tout $y \in \mathbb{R}^m$ (fixé), la fonction $x \mapsto f(x, y)$ est en escalier dans $\mathbb{R}^n$, puis que la formule $g(y) = \int f(x, y)\,dx$ définit une fonction $g : \mathbb{R}^m \to \mathbb{C}$ en escalier.

(b) Montrer que pour tout $x \in \mathbb{R}^n$ (fixé), la fonction $y \mapsto f(x, y)$ est en escalier dans $\mathbb{R}^m$, puis que la formule $h(x) = \int f(x, y)\,dy$ définit une fonction $h : \mathbb{R}^n \to \mathbb{C}$ en escalier.

(c) Montrer que $\int g(y)\,dy = \int f(z)\,dz = \int h(x)\,dx$.

1.5. Soit $\phi : \mathbb{R}^n \to \mathbb{R}^n$ la fonction affine définie par $\phi(y) = x_0 + A\,y$ où $x_0 \in \mathbb{R}^n$ et A est une matrice $n \times n$ diagonale dont les éléments diagonaux sont des réels $a_j > 0$. On notera $d = \prod_{j=1}^n a_j > 0$ le déterminant de A.

(a) Montrer que pour tout pavé $P \subset \mathbb{R}^n$, l'ensemble $\phi(P)$ est un pavé dont on exprimera la mesure en fonction de celle de P et du nombre d.

(b) Montrer que pour toute fonction $f : \mathbb{R}^n \to \mathbb{C}$ en escalier la fonction $f \circ \phi$ est en escalier, puis exprimer l'intégrale de $f \circ \phi$ en fonction de celle de f et du nombre d.

1.6. On rappelle que l'on appelle *domaine pavable* toute partie $D \subset \mathbb{R}^n$ telle que $\mathbb{1}_D$ soit en escalier (voir définition 1.7 ci-dessus), et on considère une fonction $f : \mathbb{R}^n \to \mathbb{C}$ en escalier.

(a) Montrer que tout produit cartésien d'intervalles bornés est un domaine pavable, et que l'intersection ou la réunion de deux domaines pavables est un domaine pavable.

(b) Montrer que la fonction f s'écrit de façon unique sous la forme $f = \sum_{k=1}^N c_k\, \mathbb{1}_{D_k}$ où les c_k sont non nuls et deux à deux distincts, et les D_k sont des domaines pavables deux à deux disjoints.

2. Convergence en moyenne et convergence presque partout

La définition de l'intégrale d'une fonction en escalier conduit à considérer la notion suivante.

Définition 2.1. *Soit* $(f_p)_{p\in\mathbb{N}}$ *une suite de fonctions en escalier sur* $\mathbb{R}^n$. *On dit que* (f_p) *est* de Cauchy *(sous-entendu : pour la* convergence en moyenne*) si pour tout* $\varepsilon > 0$, *il existe* $N \in \mathbb{N}$ *tel que*

$$\int |f_p(x) - f_q(x)|\, dx \leq \varepsilon \qquad \text{pour tous } p,\ q \geq N\,.$$

Remarque. Si (f_p) est une suite de Cauchy de fonctions en escalier, alors la suite de nombres complexes $\left(\int f_p(x)\, dx\right)$ a une limite pour $p \to \infty$. En effet, cette suite est de Cauchy puisque par linéarité et continuité de l'intégrale on a

$$\left| \int f_p(x)\, dx - \int f_q(x)\, dx \right| \leq \int |f_p(x) - f_q(x)|\, dx\,.$$

Nous donnerons plus loin un résultat essentiel sur les suites de Cauchy de fonctions en escalier : c'est le lemme d'Égorov, qui affirme qu'elles possèdent toujours des sous-suites qui convergent *presque partout* en un sens que nous allons maintenant définir. Cette définition s'appuie sur la notion de *dénombrabilité*, qui est essentielle en théorie de l'intégration.

Définition 2.2. *Soit* E *un ensemble.*
(a) *On dit que* E *est* au plus dénombrable *s'il existe une application surjective de* $\mathbb{N}$ *sur* E.
(b) *Si* $(m_q)_{q\in E}$ *est une famille de réels* ≥ 0, *on note* $\sum_{q\in E} m_q$ *la borne supérieure sur toutes les parties finies* $F \subset E$ *des sommes* $\sum_{q\in F} m_q$. *Si on autorise les* m_q *à prendre la valeur* ∞, *on conviendra en outre que* $\sum_{q\in E} m_q = \infty$ *dès que l'un des* m_q *au moins prend cette valeur.*

Intuitivement, un ensemble au plus dénombrable est un ensemble dont on peut numéroter les éléments sans en oublier aucun. Les ensembles au plus dénombrables possédant une infinité d'éléments seront dits *dénombrables* tout court. On peut démontrer que si E est (infini) dénombrable, alors il existe aussi une application *bijective* de $\mathbb{N}$ sur E, et c'est pourquoi les éléments d'un ensemble dénombrable seront souvent notés comme une suite $(q_p)_{p\in\mathbb{N}}$. Par exemple, les entiers naturels pairs $\{0, 2, 4, 6, \dots\}$ forment un ensemble dénombrable car l'application $p \mapsto 2p$ est bijective.

La notion de somme d'un nombre infini de termes positifs définie au (b) généralise la notion de somme finie. Il ne faut pas la confondre avec la notion de série, où le nombre de termes à sommer est dénombrable et où l'on indique dans quel ordre on doit effectuer la sommation (voir cependant la proposition suivante). Cette définition n'a un sens raisonnable que parce que tous les réels à sommer sont positifs. Dans la suite, on conviendra que $0 \times \infty = 0$.

Proposition 2.3.
(a) *Toute réunion au plus dénombrable d'ensembles au plus dénombrables est au plus dénombrable.*
(b) *Les sommes de réels positifs de la définition 2.2 sont commutatives et associatives, et on a* $c\sum_{q\in E} m_q = \sum_{q\in E} c\,m_q$ *pour tout* $c \geq 0$.
(c) *Si* $(m_q)_{q\in E}$ *est une famille de réels positifs dont la somme est finie, alors l'ensemble des indices* q *pour lesquels* $m_q > 0$ *est au plus dénombrable.*
(d) *Si* $(m_q)_{q\in\mathbb{N}}$ *est une suite de réels positifs, alors* $\sum_{q=0}^{\infty} m_q = \sum_{q\in\mathbb{N}} m_q$, *que ce nombre soit fini ou infini.*

Démonstration. En répétant éventuellement plusieurs fois le même ensemble, la famille dont on considère la réunion en (a) peut être notée comme une suite $(E_p)_{p\in\mathbb{N}}$, et de la même façon, on notera $(s_{p,q})_{q\in\mathbb{N}}$ les éléments de E_p. Pour tout $r \in \mathbb{N}$, l'ensemble $F_r = \{ s_{p,q} \,;\, p+q = r \}$ possède au plus $r+1$ éléments, et on peut donc numéroter successivement l'élément de F_0, puis ceux de F_1, de F_2 et ainsi de suite. En procédant de la sorte, tout élément $s_{p,q}$ recevra un numéro inférieur à $(p+q+1)(p+q+2)/2$, ce qui prouve qu'on numérote ainsi tous les éléments de la réunion sans exception (surjectivité).

Les sommes de la définition 2.2 sont clairement commutatives puisqu'aucun ordre n'intervient dans la définition (les sommes finies sont commutatives), et la propriété de multiplication par un réel positif s'obtient immédiatement par passage à la limite dans les sommes finies. Si E est réunion d'une famille $(E_i)_{i\in I}$ d'ensembles deux à deux disjoints, et si F est une partie finie de E, il existe une partie finie $J \subset I$ telle que $F \subset \cup_{i\in J} E_i$, et par associativité des sommes finies, on peut écrire

$$\sum_{q\in F} m_q = \sum_{i\in J} \Big(\sum_{q\in F\cap E_i} m_q \Big) \leq \sum_{i\in J} \Big(\sum_{q\in E_i} m_q \Big) \leq \sum_{i\in I} \Big(\sum_{q\in E_i} m_q \Big),$$

les deux dernières inégalités résultant de la définition 2.2. De plus, si on avait une inégalité stricte $S = \sum_{q\in E} m_q < \sum_{i\in I} \big(\sum_{q\in E_i} m_q \big)$, alors il existerait par définition une partie finie $J \subset I$ telle que $\sum_{i\in J} \big(\sum_{q\in E_i} m_q \big) > S$, puis pour chaque $i \in J$, il existerait un nombre $S_i < \sum_{q\in E_i} m_q$ et une partie finie $F_i \subset E_i$ tels que $\sum_{i\in J} S_i = S$ et que $\sum_{q\in F_i} m_q > S_i$. En notant $F = \cup_{i\in J} F_i$, on aurait donc

$$\sum_{q\in F} m_q = \sum_{i\in J} \Big(\sum_{q\in F_i} m_q \Big) > \sum_{i\in J} S_i = S = \sum_{q\in E} m_q,$$

ce qui contredirait la définition de cette dernière somme puisque $F \subset E$ est finie. On a donc prouvé la propriété d'associativité $\sum_{q\in E} m_q = \sum_{i\in I} \big(\sum_{q\in E_i} m_q \big)$.

Pour la propriété (c), il suffit de noter $S = \sum_{q\in E} m_q$, de remarquer que $E_p = \{ q \in E \,;\, m_q \geq 2^{-p} \}$ possède au plus $2^p S$ éléments, et d'en déduire que l'ensemble des indices $q \in E$ pour lesquels $m_q > 0$ est réunion dénombrable des ensembles finis E_p.

Enfin, en notant $\mathscr{F}$ l'ensemble des parties finies de $\mathbb{N}$, on a

$$\sum_{q=0}^{N} m_q = \sum_{q\in\{0,\dots,N\}} m_q \leq \sup_{F\in\mathscr{F}} \Big(\sum_{q\in F} m_q\Big) = \sum_{q\in\mathbb{N}} m_q$$

d'où $\sum_{q=0}^{\infty} m_q \leq \sum_{q\in\mathbb{N}} m_q$ en prenant la limite pour $N \to \infty$. Réciproquement, si $F \in \mathscr{F}$ et $N_F = \max_{q\in F} q$,

$$\sum_{q\in F} m_q \leq \sum_{q=0}^{N_F} m_q \leq \sum_{q=0}^{\infty} m_q\,,$$

d'où $\sum_{q\in\mathbb{N}} m_q \leq \sum_{q=0}^{\infty} m_q$ en prenant la borne supérieure sur toutes les parties finies $F \subset \mathbb{N}$. ∎

Application. D'après la propriété (a) de la proposition 2.3, l'ensemble des entiers relatifs $\mathbb{Z} = \mathbb{N} \cup (-\mathbb{N})$ est dénombrable. C'est aussi le cas de l'ensemble des rationnels $\mathbb{Q} = \cup_{q\in\mathbb{N}^*} \frac{1}{q}\mathbb{Z}$, et par récurrence, de l'ensemble $\mathbb{Q}^n = \cup_{q\in\mathbb{Q}} \big(\mathbb{Q}^{n-1} \times \{q\}\big)$ des n-uplets de rationnels.

Remarques. Les propriétés (b) et (d) de la proposition 2.3 prouvent l'affirmation suivante : si E est un ensemble dénombrable (auquel cas ses éléments peuvent être écrits comme une suite $(q_p)_{p\in\mathbb{N}}$), et si $(m_q)_{q\in E}$ est une famille de réels positifs, la somme de la série $\sum_{p=0}^{\infty} m_{q_p}$ ne dépend pas de la numérotation $p \mapsto q_p$, et vaut $\sum_{q\in E} m_q$. Enfin la propriété (c) montre que l'on peut toujours se ramener à cette situation lorsque $\sum_{q\in E} m_q < \infty$.

Lorsque les nombres à sommer ne sont plus positifs, on a recours à la notion suivante que nous n'utiliserons cependant pas dans ce cours d'intégration.

Proposition et définition 2.4. *Soient E un ensemble, et $\mathscr{F}$ l'ensemble de ses parties finies. Pour toute famille $(z_q)_{q\in E}$ de nombres complexes, les trois propriétés suivantes sont équivalentes :*
(a) *Il existe un nombre $S \in \mathbb{C}$ tel que : $\forall \varepsilon > 0,\ \exists F_\varepsilon \in \mathscr{F} : \forall F \in \mathscr{F},\ F \supset F_\varepsilon \Rightarrow \big|\sum_{q\in F} z_q - S\big| \leq \varepsilon$.*
(b) $\sup_{F\in\mathscr{F}} \big|\sum_{q\in F} z_q\big| < \infty$.
(c) $\sum_{q\in E} |z_q| < \infty$.
Lorsque ces propriétés sont vérifiées, on dit que la famille $(z_q)_{q\in E}$ est sommable, *et l'unique nombre S de la propriété* (a) *est appelé* somme *de la famille $(z_q)_{q\in E}$.*

Démonstration. Montrons d'abord que le nombre S de la propriété (a) est unique (s'il existe). En effet, si S et S' vérifient la propriété (a) et si F_ε et F'_ε sont des parties finies associées à un $\varepsilon > 0$, alors la partie finie $F = F_\varepsilon \cup F'_\varepsilon$ contient F_ε et F'_ε, si bien que $\big|\sum_{q\in F} z_q - S\big| \leq \varepsilon$ et $\big|\sum_{q\in F} z_q - S'\big| \leq \varepsilon$, d'où l'on tire $|S - S'| \leq 2\varepsilon$ par inégalité triangulaire. Ce résultat étant obtenu pour tout $\varepsilon > 0$, on en déduit donc que $S = S'$.

Si maintenant la famille $(z_q)_{q\in E}$ vérifie cette propriété (a), il existe une partie finie F_1 telle que pour toute $F \in \mathscr{F}$ contenant F_1, on ait $\big|\sum_{q\in F_1} z_q - S\big| \leq 1$. Alors $M = \sup_{F\subset F_1} \big|\sum_{q\in F} z_q\big| < \infty$ puisqu'il n'y a qu'un nombre fini de parties $F \subset F_1$, et d'autre part, pour toute $F \in \mathscr{F}$ disjointe de F_1, on peut écrire

$$\Big|\sum_{q\in F} z_q\Big| = \Big|\Big(\sum_{q\in F\cup F_1} z_q - S\Big) - \Big(\sum_{q\in F_1} z_q - S\Big)\Big| \leq \Big|\sum_{q\in F\cup F_1} z_q - S\Big| + \Big|\sum_{q\in F_1} z_q - S\Big| \leq 2$$

par construction de F_1. Comme toute $F \in \mathscr{F}$ est réunion d'une partie contenue dans F_1 et d'une partie disjointe de F_1, il en résulte que l'on a montré que $|\sum_{q\in F} z_q| \leq M+2$ pour toute $F \in \mathscr{F}$, soit (b).

Ensuite, si la famille $(z_q)_{q\in E}$ vérifie la propriété (b), c'est-à-dire si $M = \sup_{F\in\mathscr{F}} |\sum_{q\in E} z_q| < \infty$, on écrit $E = E_r \cup E_i$ en notant $E_r = \{\, q \in E\,;\, |\mathscr{Re}\, z_q| \geq |\mathscr{Im}\, z_q|\,\}$ et $E_i = \{\, q \in E\,;\, |\mathscr{Im}\, z_q| \geq |\mathscr{Re}\, z_q|\,\}$, puis $E_r = E_+ \cup E_-$ en notant $E_+ = \{\, q \in E\,;\, \mathscr{Re}\, z_q \geq |\mathscr{Im}\, z_q|\,\}$ et $E_- = \{\, q \in E\,;\, \mathscr{Re}\, z_q \leq -|\mathscr{Im}\, z_q|\,\}$. Pour tout $q \in E_+$, on a $|z_q| \leq \sqrt{2}\,\mathscr{Re}\, z_q$, si bien que pour toute partie finie $F \subset E_+$, on a l'inégalité $\sum_{q\in F} |z_q| \leq \sqrt{2}\,\mathscr{Re}(\sum_{q\in F} z_q) \leq \sqrt{2}\,M$, d'où $\sum_{q\in E_+} |z_q| \leq \sqrt{2}\,M$ par passage à la borne supérieure. On montre de même que $\sum_{q\in E_-} |z_q| \leq \sqrt{2}\,M$ d'où $\sum_{q\in E_r} |z_q| \leq 2\sqrt{2}\,M$, puis encore que $\sum_{q\in E_i} |z_q| \leq 2\sqrt{2}\,M$, ce qui prouve la propriété (c).

Soit enfin une famille $(z_q)_{q\in E}$ vérifiant la propriété (c), c'est-à-dire que $M = \sum_{q\in E} |z_q| < \infty$. Alors pour tout entier $p > 0$, il existe une partie finie $F_p \in \mathscr{F}$ telle que $\sum_{q\in F_p} |z_q| \geq M - (1/p)$, et donc telle que $\sum_{q\in E\setminus F_p} |z_q| \leq (1/p)$ (propriété d'associativité). Alors la suite (S_p) définie par $S_p = \sum_{q\in F_p} z_q$ est de Cauchy puisque

$$|S_p - S_r| = \Big|\sum_{q\in F_p\setminus F_r} z_q - \sum_{q\in F_r\setminus F_p} z_q\Big| \leq \sum_{q\in E\setminus F_r} |z_q| + \sum_{q\in E\setminus F_p} |z_q| \leq \frac{1}{r} + \frac{1}{p}\,.$$

La suite (S_p) converge donc vers une limite $S \in \mathbb{C}$, et on obtient $|S_p - S| \leq (1/p)$ en passant à la limite pour $r \to \infty$ dans l'inégalité précédente. Alors pour tout $\varepsilon > 0$, il existe un entier $p \geq (2/\varepsilon)$, et si une partie finie F contient la partie F_p correspondante, on peut écrire

$$\Big|\sum_{q\in F} z_q - S\Big| \leq \Big|\sum_{q\in F} z_q - S_p\Big| + |S_p - S| \leq \sum_{q\in F\setminus F_p} |z_q| + \frac{1}{p} \leq \frac{2}{p} \leq \varepsilon\,,$$

ce qui prouve la propriété (a). ∎

Compléments. Les propriétés (c) des propositions 2.3 et 2.4 montrent que si $(z_q)_{q\in E}$ est sommable, l'ensemble des $q \in E$ tels que $z_q \neq 0$ est au plus dénombrable. Alors la somme S est celle de la série absolument convergente $\sum_{p=0}^{\infty} z_{q_p}$ quelle que soit la numérotation $p \mapsto q_p$ des indices q tels que $z_q \neq 0$. En particulier, cette somme est commutative et associative, ce que l'étudiant pourra démontrer à titre d'exercice.

Définition 2.5. *On dit qu'une partie A de $\mathbb{R}^n$ est* négligeable *si pour tout $\varepsilon > 0$, il existe une suite $(P_q)_{q\in\mathbb{N}}$ de pavés de $\mathbb{R}^n$ telle que $A \subset \bigcup_{q\in\mathbb{N}} P_q$ et $\sum_{q\in\mathbb{N}} |P_q| \leq \varepsilon$. On dit ensuite qu'une propriété dépendant de $x \in \mathbb{R}^n$ est vraie* presque partout *(en abrégé :* pp*) si l'ensemble des x pour lesquels elle est fausse est négligeable. En particulier, on dit que la suite de fonctions (f_p) converge presque partout vers la fonction f si l'ensemble des x tels que $f(x)$ n'est pas limite de la suite $(f_p(x))$ est négligeable.*

Exemples. Tout pavé de mesure nulle est négligeable puisqu'il est recouvert par lui-même. Tout ensemble dénombrable (par exemple $\mathbb{Q}^n \subset \mathbb{R}^n$) est négligeable puisqu'un point est un pavé de mesure nulle. L'hyperplan d'équation $x_1 = 0$ est négligeable puisqu'il est recouvert par la suite de pavés $(\{0\} \times [-p,p]^{n-1})_{p\in\mathbb{N}}$ qui sont tous de mesure nulle. Toute partie d'un ensemble négligeable est négligeable.

Il n'y a pas unicité de la limite presque partout d'une suite de fonctions (f_p) : si f en est une limite presque partout, la même suite converge aussi presque partout vers la fonction g si et seulement si $f = g$ presque partout, c'est-à-dire si l'ensemble des x tels que $f(x) \neq g(x)$ est négligeable. On peut montrer que l'égalité presque partout est une relation d'équivalence.

Proposition 2.6.
(a) *Tout translaté* $A + x_0$ *d'un ensemble négligeable* A *est négligeable.*
(b) *Toute réunion dénombrable d'ensembles négligeables est négligeable.*
(c) *Si* $A \subset \mathbb{R}^n$ *est négligeable,* $A \times \mathbb{R}^m \subset \mathbb{R}^n \times \mathbb{R}^m$ *est aussi négligeable.*
(d) *Un pavé est négligeable si et seulement s'il est de mesure nulle.*

Démonstration. Si A est négligeable et $\varepsilon > 0$, on peut trouver une suite $(P_q)_{q\in\mathbb{N}}$ de pavés recouvrant A et dont la somme des mesures est inférieure à ε. Alors la suite de pavés $(P_q + x_0)_{q\in\mathbb{N}}$ recouvre $A + x_0$, et comme cette construction peut être effectuée pour tout $\varepsilon > 0$, on en déduit que $A + x_0$ est négligeable, soit (a).

Soient $(A_p)_{p\in\mathbb{N}}$ une suite d'ensembles négligeables, A leur réunion, et $\varepsilon > 0$. Pour tout $p \in \mathbb{N}$, soit $(P_{p,q})_{q\in\mathbb{N}}$ une suite de pavés telle que $A_p \subset \bigcup_{q\in\mathbb{N}} P_{p,q}$ et $\sum_{q\in\mathbb{N}} |P_{p,q}| \leq \varepsilon/2^{p+1}$. La famille de pavés $(P_{p,q})_{(p,q)\in\mathbb{N}^2}$ est dénombrable et on a $A \subset \bigcup_{(p,q)\in\mathbb{N}^2} P_{p,q}$ et $\sum_{(p,q)\in\mathbb{N}^2} |P_{p,q}| \leq \sum_{p\in\mathbb{N}} (\varepsilon/2^{p+1}) = \varepsilon$, d'où (b).

Pour la propriété (c) on écrit que l'ensemble $A \times \mathbb{R}^m$ est la réunion dénombrable des ensembles $B_q = A \times [-q, q]^m$ pour $q \in \mathbb{N}$. D'autre part, si la suite de pavés $(P_p)_{p\in\mathbb{N}}$ recouvre A, alors la suite de pavés $Q_p = P_p \times [-q, q]^m$ recouvre B_q et vérifie $\sum_{p\in\mathbb{N}} |Q_p| = (2q)^m \sum_{p\in\mathbb{N}} |P_p|$. Il en résulte que chaque B_q est négligeable, puis que $A \times \mathbb{R}^m$ aussi.

Enfin, si un pavé P est de mesure nulle, il est clairement négligeable puisqu'il est recouvert par lui-même, et réciproquement, si P est négligeable, alors il est de mesure nulle puisque $|P| = \int \mathbb{1}_P(x)\,dx \leq \varepsilon$ pour tout $\varepsilon > 0$ d'après l'inégalité du lemme suivant. ∎

Lemme 2.7 (de Borel-Lebesgue). *Soient* A *une partie de* $\mathbb{R}^n$ *telle que* $\mathbb{1}_A$ *soit en escalier, et* $(P_q)_{q\in\mathbb{N}}$ *une suite de pavés telle que* $\overline{A} \subset \bigcup_{q\in\mathbb{N}} P_q$. *Alors* $\int \mathbb{1}_A(x)\,dx \leq \sum_{q\in\mathbb{N}} |P_q|$.

Démonstration. Si A est contenu dans une réunion finie $P_0 \cup \ldots \cup P_N$, on a $\mathbb{1}_A \leq \sum_{q=0}^N \mathbb{1}_{P_q}$ d'où $\int \mathbb{1}_A(x)\,dx \leq \sum_{q=0}^N |P_q|$ par croissance de l'intégrale. Dans le cas général où l'on a une suite infinie (P_q), on se ramène au cas précédent en utilisant la compacité de $\overline{A}$ de la façon suivante. Soit $\varepsilon > 0$ fixé. On voit facilement que chaque P_q est contenu dans l'intérieur d'un pavé Q_q vérifiant $|Q_q| \leq |P_q| + (\varepsilon/2^q)$ (il suffit d' "élargir un peu" les côtés de P_q). Comme $\overline{A} \subset \bigcup_{q\in\mathbb{N}} P_q$, le compact $\overline{A}$ est recouvert par les intérieurs des Q_q dont on peut extraire un sous-recouvrement fini, ce qui donne $A \subset Q_0 \cup \ldots \cup Q_N$. On a donc

$$\int \mathbb{1}_A(x)\,dx \leq \sum_{q=0}^{N} |Q_q| \leq \sum_{q\in\mathbb{N}} |P_q| + 2\,\varepsilon\,,$$

ce qui entraîne le résultat cherché puisque $\varepsilon > 0$ est arbitraire. ∎

Le théorème suivant est fondamental du point de vue théorique : toute la construction de l'intégrale des fonctions intégrables reposera sur ce théorème. Il faut absolument en comprendre l'énoncé, et les étudiants désireux de poursuivre en maîtrise de mathématiques s'efforceront d'en apprendre aussi la démonstration.

Théorème 2.8 (Lemme d'Égorov). *Soit (f_p) une suite de Cauchy de fonctions en escalier. Alors :*
(a) *Il existe une suite extraite $(f_{p_q})_{q\in\mathbb{N}}$ et une fonction f telles que $f_{p_q} \to f$ presque partout. Plus précisément, il existe une suite de pavés $(P_r)_{r\in\mathbb{N}}$ telle que $\sum_{r\in\mathbb{N}} |P_r| \leq 1$ et que pour tout $N \in \mathbb{N}$, $f_{p_q} \to f$ uniformément sur le complémentaire de l'ensemble $A_N = \cup_{r>N} P_r$.*
(b) *L'intégrale $\int |f_p(x)|\, dx$ tend vers 0 si et seulement s'il existe une suite extraite $(f_{p_q})_{q\in\mathbb{N}}$ qui converge vers 0 presque partout.*

Démonstration. La suite (f_p) étant de Cauchy, on peut trouver, pour tout entier q, un entier $p_q > p_{q-1}$ tel que l'intégrale de $|f_{p_q} - f_p|$ soit inférieure à 2^{-2q} pour $p \geq p_q$. De cette façon, on définit une sous-suite $(f_{p_q})_{q\in\mathbb{N}}$ vérifiant

$$\int |f_{p_q}(x) - f_{p_{q+1}}(x)|\, dx \leq 2^{-2q}.$$

Si on note $D_q = \{x \in \mathbb{R}^n \,;\, |f_{p_q}(x) - f_{p_{q+1}}(x)| > 2^{-q}\}$, on voit que $2^{-q}\,\mathbb{1}_{D_q} \leq |f_{p_q} - f_{p_{q+1}}|$, d'où la majoration

$$\int \mathbb{1}_{D_q}(x)\, dx \leq 2^q \int |f_{p_q}(x) - f_{p_{q+1}}(x)|\, dx \leq 2^{-q}.$$

D'après le lemme 1.8 parties (a) et (b), D_q est donc recouvert par des pavés $P_{N_{q-1}+1}, \ldots, P_{N_q}$ dont la somme des mesures ne dépasse pas 2^{-q}. La collection de tous ces pavés pour $q \geq 1$ vérifie donc $\sum_{r\in\mathbb{N}} |P_r| \leq 1$, et dans le complémentaire de $A_N = \cup_{r>N} P_r$, on a $|f_{p_q} - f_{p_{q+1}}| \leq 2^{-q}$ pour tout q tel que $N_q > N$, ce qui montre que la suite $(f_{p_q})_{q\in\mathbb{N}}$ y est uniformément de Cauchy. On notera f sa limite, qui est définie en tout point $x \notin A = \cap_{N\in\mathbb{N}} A_N$. Pour compléter la démonstration de la propriété (a), il suffit donc de montrer que A est négligeable, ce qui est clair puisqu'il est recouvert par chaque A_N, et que A_N est la réunion d'une suite de pavés dont la somme des mesures est inférieure au reste $\sum_{r>N} |P_r|$ (arbitrairement petit) de la série convergente $\sum_{r\in\mathbb{N}} |P_r|$.

Pour montrer que si $\left(\int |f_p(x)|\, dx\right)$ tend vers 0, alors (f_p) possède une sous-suite qui converge vers 0 presque partout, on répète l'argument ci-dessus : l'entier p_q est choisi tel que $\int |f_{p_q}(x)|\, dx \leq 2^{-2q}$, puis on détermine des pavés $P_{N_{q-1}+1}, \ldots, P_{N_q}$ dont la somme des mesures est inférieure à 2^{-q} et tels que $|f_{p_q}(x)| \leq 2^{-q}$ pour tout $x \notin \cup_{r=N_{q-1}+1}^{N_q} P_r$, et on conclut comme ci-dessus que (f_{p_q}) converge vers 0 presque partout.

Réciproquement, supposons qu'il existe une sous-suite $(f_{p_q})_{q\in\mathbb{N}}$ qui converge presque partout vers 0. En remarquant que

$$\left| \int |f_p(x)|\, dx - \int |f_q(x)|\, dx \right| \leq \int |f_p(x) - f_q(x)|\, dx,$$

on constate que la suite des intégrales $\left(\int |f_p(x)|\, dx\right)$ est de Cauchy dans $\mathbb{R}$, donc converge. Pour montrer que sa limite est nulle, il suffit de montrer qu'on peut en extraire une sous-suite qui converge vers 0, et nous prendrons la sous-suite qui correspond à la sous-suite (f_{p_q}) qui converge presque partout vers 0. Par ailleurs, quitte à extraire à nouveau une sous-suite, que l'on notera désormais $(g_q)_{q\in\mathbb{N}}$, on peut supposer qu'on a les conclusions de la partie (a) du théorème avec $f = 0$.

Fixons un $\varepsilon > 0$. Il existe alors p tel que pour tout $q \geq p$ on ait la majoration $\int |g_q(x) - g_p(x)|\, dx \leq \varepsilon/3$. Ce nombre p étant désormais fixé, on notera $C = \sup |g_p|$, $D = \{x \in \mathbb{R}^n \,;\, g_p(x) \neq 0\}$ et

$|D| = \int \mathbb{1}_D(x)\,dx$. D'après la partie (a) du théorème, il existe une suite de pavés $(P_r)_{r\in\mathbb{N}}$ telle que $\sum_{r\in\mathbb{N}} |P_r| \leq \varepsilon/(1+3\,C)$ et que la suite (g_q) converge uniformément vers 0 en dehors de $A = \cup_{r\in\mathbb{N}} P_r$. On se restreint alors désormais aux indices q assez grands pour que $|g_q(x)| \leq \varepsilon/(1+3\,|D|)$ si $x \notin A$.

Pour tout indice q ainsi fixé, on a $|g_q| = |g_q|(\mathbb{1}_{A_q} + \mathbb{1}_{B_q})$ en notant

$$A_q = \{\, x \in \mathbb{R}^n \,;\, |g_q(x)| > \varepsilon/(1+3\,|D|)\,\} \ (\subset A) \qquad \text{et} \qquad B_q = \{\, x \in \mathbb{R}^n \,;\, 0 < |g_q(x)| \leq \varepsilon/(1+3\,|D|)\,\}\,.$$

Comme $|g_q| \leq |g_q - g_p| + |g_p|$, que $\mathbb{1}_{A_q} + (1-\mathbb{1}_D)\mathbb{1}_{B_q} \leq 1$ et que $|g_p|(1-\mathbb{1}_D) = 0$, on peut donc écrire

$$\begin{aligned}\int |g_q(x)|\,dx &= \int |g_q(x)|\,\big(\mathbb{1}_{A_q}(x) + (1-\mathbb{1}_D(x))\mathbb{1}_{B_q}(x)\big)\,dx + \int |g_q(x)|\,\mathbb{1}_D(x)\,\mathbb{1}_{B_q}(x)\,dx \leq \\ &\leq \int |g_q(x) - g_p(x)|\,dx + \int |g_p(x)|\,\mathbb{1}_{A_q}(x)\,dx + \int |g_q(x)|\,\mathbb{1}_D(x)\,\mathbb{1}_{B_q}(x)\,dx \leq \\ &\leq \frac{\varepsilon}{3} + C\int \mathbb{1}_{A_q}(x)\,dx + \frac{\varepsilon}{1+3\,|D|}\int \mathbb{1}_D(x)\,dx\,.\end{aligned}$$

Le dernier terme écrit étant inférieur à $\varepsilon/3$, on pourra donc conclure que la suite des intégrales $\left(\int |g_q(x)|\,dx\right)$ converge vers 0 si on démontre que $\int \mathbb{1}_{A_q}(x)\,dx \leq \varepsilon/(1+3\,C)$. Or d'après la partie (c) du lemme 1.8, il existe des pavés $Q_1, \ldots, Q_N$ de mesures nulles tels que $A_q \cup Q_1 \cup \ldots \cup Q_N$ soit compact, ce qui entraîne que $\overline{A_q}$ est recouvert par la suite (P_r) à laquelle on adjoint les pavés $Q_1, \ldots, Q_N$. On peut alors conclure grâce au lemme de Borel-Lebesgue (lemme 2.7). ∎

Exercices

2.1. Montrer qu'un ensemble infini E est dénombrable si et seulement s'il existe une bijection de $\mathbb{N}$ sur E.

2.2. Montrer que si A et B sont dénombrables, alors $A \times B$ aussi, et en déduire que tout ouvert X de $\mathbb{R}^n$ est réunion d'une suite de boules (on considérera les boules $B(q, 1/p) \subset X$ avec $(q,p) \in \mathbb{Q}^n \times \mathbb{N}^*$).

2.3. Montrer qu'un ensemble A est négligeable si et seulement s'il existe une suite de pavés $(P_q)_{q\in\mathbb{N}}$ telle que la série $\sum_{q=0}^{\infty} |P_q|$ soit convergente et que $A \subset \cup_{q>N} P_q$ pour tout entier N.

2.4. Montrer que si A est négligeable, alors $\mathring{A} = \emptyset$ (on montrera que si $\mathring{A} \neq \emptyset$, alors A contient un petit pavé de mesure non nulle). En déduire qu'un ouvert non vide de $\mathbb{R}^n$ n'est jamais dénombrable.

2.5. Soit $f : \mathbb{R} \to \mathbb{R}^2$ une fonction *lipschitzienne*, c'est-à-dire vérifiant $|f(x) - f(y)| \leq C\,|x-y|$ pour tous x et $y \in \mathbb{R}$, la constante C étant indépendante de x et y. En coupant l'intervalle $[-p, p]$ en N morceaux, montrer que $f([-p,p])$ est négligeable pour tout entier $p \in \mathbb{N}$, puis en déduire que $f(\mathbb{R})$ est négligeable.

2.6. Soient $f : \mathbb{R}^{n-1} \to \mathbb{R}$ une fonction continue, et $G \subset \mathbb{R}^n$ son *graphe*, c'est-à-dire l'ensemble $G = \{\, x \in \mathbb{R}^n \,;\, x_n = f(x_1, \ldots, x_{n-1})\,\}$. Montrer que G est négligeable (on montrera que le graphe de $f_{|[-p,p]^{n-1}}$ est négligeable pour tout $p \in \mathbb{N}$ grâce à l'uniforme continuité).

En déduire que tout hyperplan de $\mathbb{R}^n$ est négligeable, ainsi que toute sphère de $\mathbb{R}^n$.

2.7. Montrer que la relation $f = g$ pp définit une relation d'équivalence sur l'ensemble des fonctions $f : \mathbb{R}^n \to \mathbb{C}$.

2.8. Soit (f_p) une suite de Cauchy de fonctions en escalier qui converge presque partout vers une fonction f. Montrer que les suites $(\mathscr{R}e\, f_p)$, $(\mathscr{I}m\, f_p)$, $(\overline{f_p})$ et $(|f_p|)$ sont des suites de Cauchy de fonctions en escalier qui convergent presque partout vers $\mathscr{R}e\, f$, $\mathscr{I}m\, f$, $\overline{f}$ et $|f|$.

2.9. Soient $\phi : \mathbb{R}^n \to \mathbb{R}^n$ la fonction affine considérée à l'exercice **1.5** de la page 10, et (f_p) une suite de fonctions en escalier dans $\mathbb{R}^n$.

(a) Montrer qu'une partie $B \subset \mathbb{R}^n$ est négligeable si et seulement si $\phi(B)$ est négligeable, et en déduire que la suite $(f_p \circ \phi)$ converge presque partout si et seulement si la suite (f_p) converge presque partout.

(b) Montrer que la suite $(f_p \circ \phi)$ est de Cauchy si et seulement si la suite (f_p) est de Cauchy.

2.10. On écrit tout entier p sous la forme $p = 2^q + k$ avec $0 \leq k < 2^q$, puis on note f_p la fonction indicatrice de l'intervalle $[2^{-q}k, 2^{-q}(k+1)]$, ce qui définit une suite de fonctions $f_p : \mathbb{R} \to \mathbb{R}$ en escalier. Montrer que cette suite est de Cauchy, mais qu'elle ne converge en aucun point. En extraire une sous-suite qui converge (presque) partout.

Eudoxe de Cnide
(vers 408 - 355 avant J.-C.)

Archimède
(287 - 212 avant J.-C.)

Les principaux créateurs de ce qui tint lieu de théorie de l'intégration dans l'antiquité grecque : la théorie des rapports de grandeurs et la méthode d'exhaustion.

N.B. En l'absence de document fiable datant de l'antiquité grecque, l'auteur a avantageusement remplacé le portrait de l'illustre Eudoxe par celui du non moins illustre Zéphyrin Brioché, peut-être plus connu encore sous le nom de "savant Cosinus".

3. Fonctions intégrables

Soient (f_p) et (g_p) deux suites de Cauchy de fonctions en escalier qui convergent presque partout vers la même fonction $f : \mathbb{R}^n \to \mathbb{C}$. Alors la suite de fonctions en escalier $(f_p - g_p)$ converge presque partout vers 0. De plus, on a par inégalité triangulaire

$$\int |(f_p - g_p)(x) - (f_q - g_q)(x)|\, dx \leq \int |f_p(x) - f_q(x)|\, dx + \int |g_p(x) - g_q(x)|\, dx\,,$$

ce qui prouve que la suite $(f_p - g_p)$ est de Cauchy. Il résulte alors du lemme d'Égorov partie (b) que l'intégrale de $|f_p - g_p|$ tend vers 0, et cela implique que les suites numériques $\big(\int f_p(x)\, dx\big)$ et $\big(\int g_p(x)\, dx\big)$ ont même limite puisque

$$\left| \int f_p(x)\, dx - \int g_p(x)\, dx \right| = \left| \int \big(f_p(x) - g_p(x)\big)\, dx \right| \leq \int |f_p(x) - g_p(x)|\, dx\,.$$

Cette propriété permet de poser les définitions suivantes.

Définition 3.1. *On dit qu'une fonction $f : \mathbb{R}^n \to \mathbb{C}$ est* intégrable (au sens de Lebesgue) *s'il existe une suite de Cauchy de fonctions en escalier (f_p) qui converge vers f presque partout, et on appelle* intégrale *d'une telle fonction f la limite, notée $\int f(x)\, dx$, de la suite de nombres complexes $\big(\int f_p(x)\, dx\big)$. Pour toute partie $A \subset \mathbb{R}^n$, on dit aussi qu'une fonction $f : \mathbb{R}^n \to \mathbb{C}$ est* intégrable sur A *si la fonction $\mathbb{1}_A f$ est intégrable, et l'intégrale de $\mathbb{1}_A f$ est alors notée $\int_A f(x)\, dx$.*

Cette définition montre que si f est intégrable et $g = f$ presque partout, alors g est intégrable et de même intégrale. Par exemple $\mathbb{1}_{\mathbb{Q}}$ est intégrable (sur $\mathbb{R}$) et d'intégrale nulle. Cette observation permet aussi de parler de l'intégrale d'une fonction qui n'est définie que presque partout. Par ailleurs, on voit facilement que si f est intégrable, alors elle est intégrable sur tout pavé P, puisque si (f_q) est une suite de Cauchy de fonctions en escalier qui converge presque partout vers f, alors $(\mathbb{1}_P\, f_q)$ est une suite de Cauchy de fonctions en escalier qui converge presque partout vers la fonction $\mathbb{1}_P\, f$.

Bien entendu, les fonctions en escalier sont intégrables et leurs intégrales sont celles qui ont été définies au paragraphe 1 (prendre la suite constante $f_p = f$ dans la définition 3.1). Un deuxième exemple fondamental de fonction intégrable est fourni par la proposition suivante.

Proposition 3.2. *Toute fonction continue sur un pavé est intégrable sur ce pavé.*

Démonstration. Soit f une fonction continue sur un pavé P. L'essentiel de cette démonstration consiste à construire une suite (f_q) de fonctions en escalier nulles en dehors du pavé P et convergeant uniformément vers $\mathbb{1}_P f$. Pour cela, quitte à traiter séparément les parties réelle et imaginaire de f, on peut supposer que f est à valeurs réelles.

Soit donc f une fonction réelle continue sur le pavé P, et $c = \inf_P f$. On note $\mathscr{P}_0 = \{P\}$, puis $\mathscr{P}_1$ l'ensemble des 2^n pavés obtenus en coupant chaque côté de P par son milieu, puis $\mathscr{P}_2$ l'ensemble de pavés obtenu en répétant cette opération sur chaque élément de $\mathscr{P}_1$, et ainsi de suite. Pour tout pavé Q appartenant à l'une de ces familles, on note c_Q la valeur prise par f au centre du pavé Q, et pour tout entier q fixé, on pose $f_q = c\,\mathbb{1}_P + \sup_{Q\in\mathscr{P}_q}\big((c_Q - c)\,\mathbb{1}_Q\big)$. Cette fonction en escalier vérifie $f_q(x) = 0$ pour tout $x \notin P$, et pour $x \in P$, on a $f_q(x) = c_Q$ pour un pavé $Q \in \mathscr{P}_q$ contenant x. En notant R le demi-diamètre de P (c'est-à-dire la distance de son centre à chacun de ses sommets), on en déduit la majoration $|f_q(x) - f(x)| \leq \sup_{|x-y|\leq 2^{-q}R} |f(x) - f(y)|$ pour tout $x \in P$, ce qui prouve la convergence uniforme de la suite (f_q) vers la fonction $\mathbb{1}_P f$ puisque f étant continue sur le compact P, elle y est uniformément continue.

La fonction $\mathbb{1}_P f$ étant partout limite de la suite (f_q) construite ci-dessus, il ne reste plus qu'à montrer que cette suite est de Cauchy. Or si pour $\varepsilon > 0$ on choisit N en sorte que $\sup|f_q - \mathbb{1}_P f| \leq \varepsilon/(2\,|P|)$ pour $q \geq N$, alors pour p et $q \geq N$, on a la majoration $|f_p - f_q| \leq |f_p - \mathbb{1}_P f| + |\mathbb{1}_P f - f_q| \leq (\varepsilon/|P|)\,\mathbb{1}_P$, d'où $\int |f_p(x) - f_q(x)|\,dx \leq \varepsilon$. La suite (f_q) est donc de Cauchy, et la fonction f est intégrable sur P. ∎

Corollaire et définition 3.3. *Si la fonction continue $f : \mathbb{R}^n \to \mathbb{C}$ s'annule en dehors d'un pavé P, elle est intégrable sur $\mathbb{R}^n$. Une telle fonction f est appelée fonction* continue à support compact.

Démonstration. Puisque f s'annule en dehors de P, on a $f = \mathbb{1}_P f$ et l'intégrabilité de f résulte alors de la proposition 3.2 puisque f est continue sur P, donc intégrable sur P. ∎

Les premières propriétés des intégrales proviennent de la proposition 1.6 (propriétés des intégrales des fonctions en escalier) par passage à la limite. Elles s'énoncent ainsi.

Proposition 3.4. *Soit $f : \mathbb{R}^n \to \mathbb{C}$ une fonction intégrable. Alors :*
(a) Invariance *: pour tout $x_0 \in \mathbb{R}^n$ fixé, la fonction $x \mapsto f(x - x_0)$ est intégrable, et $\int f(x - x_0)\,dx = \int f(x)\,dx$.*
(b) Linéarité *: pour toute constante $c \in \mathbb{C}$ et toute fonction $g : \mathbb{R}^n \to \mathbb{C}$ intégrable, les fonctions $\overline{f}$ et $f + c\,g$ sont intégrables, et on a les identités $\int \overline{f}(x)\,dx = \overline{\int f(x)\,dx}$ et $\int (f + c\,g)(x)\,dx = \int f(x)\,dx + c \int g(x)\,dx$.*

(c) Tensorialité : *si $g : \mathbb{R}^m \to \mathbb{C}$ est intégrable, la fonction $h : (x,y) \mapsto f(x)\, g(y)$ est intégrable et $\int h(z)\, dz = \big(\int f(x)\, dx\big)\big(\int g(y)\, dy\big)$.*
(d) Continuité : *la fonction $|f|$ est intégrable, et $\left|\int f(x)\, dx\right| \leq \int |f(x)|\, dx$.*
(e) Positivité : *si f est à valeurs réelles positives, alors $\int f(x)\, dx \geq 0$.*
(f) Croissance : *si $f \leq g$ sont intégrables sur $\mathbb{R}^n$ et à valeurs réelles, alors $\int f(x)\, dx \leq \int g(x)\, dx$.*

Démonstration. Toutes les identités sur les intégrales s'obtiennent par passage à la limite dans les identités correspondantes pour les fonctions en escalier, pourvu que l'on prouve l'intégrabilité des fonctions considérées. Cette intégrabilité est immédiate pour la fonction translatée et pour les combinaisons linéaires, mais un petit peu plus délicate pour les produits tensoriels.

Soient en effet (f_p) et (g_p) deux suites de Cauchy de fonctions en escalier convergeant respectivement vers f en tout point $x \notin A$, et vers g en tout point $y \notin B$ avec $A \subset \mathbb{R}^n$ et $B \subset \mathbb{R}^m$ négligeables. Alors le produit $h_p(x,y) = f_p(x)\, g_p(y)$ définit une suite (h_p) de fonctions en escalier qui converge vers h en tout point $(x,y) \notin (A \times \mathbb{R}^m) \cup (\mathbb{R}^n \times B)$, ensemble qui est négligeable d'après la proposition 2.5 (c). De plus, cette suite est de Cauchy puisque l'on peut écrire avec $z = (x,y)$

$$\begin{aligned}
\int |h_p(z) - h_q(z)|\, dz \leq\\
&\leq \int |f_p(x) - f_q(x)|\, |g_p(y)|\, dz + \int |f_q(x)|\, |g_p(y) - g_q(y)|\, dz \leq\\
&\leq \left(\int |f_p(x) - f_q(x)|\, dx\right) \left(\int |g_p(y)|\, dy\right) +\\
&\qquad\qquad + \left(\int |f_q(x)|\, dx\right) \left(\int |g_p(y) - g_q(y)|\, dy\right)
\end{aligned}$$

d'après la proposition 1.6 (c), et que les suites $\big(\int |f_q(x)|\, dx\big)$ et $\big(\int |g_p(y)|\, dy\big)$ sont bornées car convergentes. On a ainsi prouvé que le produit tensoriel $h(x,y) = f(x)\, g(y)$ est intégrable sur $\mathbb{R}^n \times \mathbb{R}^m$.

Pour obtenir la continuité, il faut observer que si la suite (f_p) est de Cauchy et converge presque partout vers f, alors la suite $(|f_p|)$ est de Cauchy puisque $\big||f_p| - |f_q|\big| \leq |f_p - f_q|$, et converge presque partout vers $|f|$. Pour la positivité, il suffit de remarquer que pour f positive, le nombre $\int f(x)\, dx = \int |f(x)|\, dx$ est positif d'après (d), et enfin la croissance résulte comme à la proposition 1.6, des propriétés de linéarité et de positivité. ∎

Applications. De la propriété de linéarité, on peut aussi déduire la propriété d'*additivité* suivante : si $A \cap B$ est négligeable et si f est intégrable sur A et sur B, alors f est intégrable sur $A \cup B$ et on a $\int_{A \cup B} f(x)\, dx = \int_A f(x)\, dx + \int_B f(x)\, dx$. En effet, lorsque $A \cap B$ est négligeable, on a $\mathbb{1}_{A \cup B} f = \mathbb{1}_A f + \mathbb{1}_B f$ presque partout. Dans le cas de la dimension $n = 1$, cette formule prouve la *relation de Chasles* en prenant $A = [a,b]$ et $B = [b,c]$.

Par ailleurs, avec les mêmes arguments qu'à la proposition 1.5 (f), il résulte de la proposition précédente que si $f_1, \dots, f_N$ sont des fonctions intégrables à valeurs réelles, les fonctions $\sup(f_1, \dots, f_N)$ et $\inf(f_1, \dots, f_N)$ sont aussi intégrables.

Remarque importante. S'il est vrai que toute combinaison linéaire de fonctions intégrables est intégrable, cela *n'est plus vrai* en général pour un *produit* de fonctions intégrables, et "encore moins" pour une fonction composée de fonctions intégrables. Les seuls produits intégrables dont nous disposions pour le moment sont les produits tensoriels de la propriété (c) ci-dessus, c'est-à-dire les produits dont chaque facteur dépend de variables différentes. Nous étudierons ultérieurement aux paragraphes 5 et 9 des produits dont les différents facteurs dépendent des mêmes variables, mais il faudra alors faire d'autres hypothèses pour qu'un tel produit soit intégrable.

Comme on a clairement $\int |c\, f(x)|\, dx = |c| \int |f(x)|\, dx$ et $\int |(f+g)(x)|\, dx \leq \int |f(x)|\, dx + \int |g(x)|\, dx$, on voit que l'application $f \mapsto \int |f(x)|\, dx$ est une semi-norme sur l'espace vectoriel des fonctions intégrables. Cette semi-norme n'est pas une norme, mais on peut caractériser les fonctions dont l'intégrale du module est nulle.

Proposition 3.5. *Soit $f : \mathbb{R}^n \to \mathbb{C}$ une fonction intégrable. Alors $\int |f(x)|\, dx = 0$ si et seulement si $f = 0$ presque partout.*

Démonstration. Soit (f_p) une suite de Cauchy de fonctions en escalier qui converge presque partout vers f. Comme $(|f_p|)$ est aussi de Cauchy et converge presque partout vers $|f|$, on a par définition

$$\lim_{p\to\infty} \int |f_p(x)|\, dx = \int |f(x)|\, dx = 0\,.$$

Il résulte alors du lemme d'Égorov partie (b) que l'on peut extraire une sous-suite $(f_{p_q})_{q\in\mathbb{N}}$ qui converge presque partout vers 0, d'où $f = 0$ presque partout. ∎

Définition 3.6. *On dira que la suite (f_p) de fonctions intégrables* converge en moyenne *vers la fonction intégrable f si la suite numérique $\left(\int |f_p(x) - f(x)|\, dx\right)$ tend vers 0.*

Comme dans le cas de la convergence presque partout introduite au paragraphe 2, il n'y a pas unicité de la limite en moyenne d'une suite (f_p) de fonctions intégrables. En effet, si (f_p) converge en moyenne vers f et si $f = g$ presque partout, alors

$$\int |f_p(x)-g(x)|\, dx \leq \int |f_p(x)-f(x)|\, dx + \int |f(x)-g(x)|\, dx = \int |f_p(x)-f(x)|\, dx\,,$$

ce qui prouve que (f_p) converge aussi vers g. Réciproquement, si une suite (f_p) converge en moyenne vers deux fonctions f et g, alors

$$\int |f(x) - g(x)|\, dx \leq \int |f(x) - f_p(x)|\, dx + \int |f_p(x) - g(x)|\, dx$$

pour tout p, ce qui montre que $\int |f(x) - g(x)|\, dx = 0$ en passant à la limite, d'où $f = g$ presque partout.

Par ailleurs, il est clair que si la suite (f_p) converge en moyenne vers f, alors la suite numérique $\big(\int f_p(x)\, dx\big)$ converge vers $\int f(x)\, dx$, ce qui montre l'intérêt de ce type de convergence.

Le résultat suivant indique comment on peut approcher, au sens de la convergence en moyenne, une fonction intégrable.

Proposition 3.7 (Densité des fonctions en escalier ou continues à support compact). *Soit $f : \mathbb{R}^n \to \mathbb{C}$ une fonction intégrable. Alors :*
(a) *Il existe une suite de fonctions en escalier qui converge en moyenne vers f.*
(b) *Il existe une suite de fonctions continues à support compact qui converge en moyenne vers f.*

Démonstration. Comme par définition de l'intégrabilité, il existe une suite de Cauchy de fonctions en escalier convergeant presque partout vers f, il suffit de montrer que cette suite converge aussi en moyenne vers f. Or pour tout $p \in \mathbb{N}$ fixé, la suite $(|f_p - f_q|)_{q\in\mathbb{N}}$ de fonctions en escalier est de Cauchy et converge presque partout vers $|f_p - f|$, d'où

$$\int |f_p(x) - f(x)|\, dx = \lim_{q\to\infty} \int |f_p(x) - f_q(x)|\, dx \,,$$

et on conclut que cette quantité peut être rendue arbitrairement petite pour p assez grand puisque (f_p) est de Cauchy. On a donc prouvé la propriété (a).

Pour établir la propriété (b), il suffit d'après le (a) de le faire pour une fonction f en escalier, et par combinaison linéaire, il suffit même de le faire pour la fonction indicatrice $f = \mathbb{1}_P$ d'un pavé $P = [a_1, b_1] \times \ldots \times [a_n, b_n]$. Pour tout intervalle compact $[a, b]$, notons $g_{a,b,q}$ la fonction définie par

$$g_{a,b,q}(t) = \begin{cases} 1 + 2^q(t-a) & \text{pour } t \in [a - 2^{-q}, a]\,, \\ 1 & \text{pour } t \in [a, b]\,, \\ 1 + 2^q(b-t) & \text{pour } t \in [b, b + 2^{-q}]\,, \\ 0 & \text{pour } t \notin [a - 2^{-q}, b + 2^{-q}]\,. \end{cases}$$

Comme $g_{a,b,q} \geq \mathbb{1}_{[a,b]}$, le produit tensoriel $f_q(x) = \prod_{j=1}^n g_{a_j,b_j,q}(x_j)$ est une fonction continue à support compact supérieure à $\mathbb{1}_P$, ce qui permet d'écrire

$$\int |f_q(x) - \mathbb{1}_P(x)|\, dx = \int f_q(x)\, dx - \int \mathbb{1}_P(x)\, dx = \prod_{j=1}^n (b_j - a_j + 2^{-q}) - \prod_{j=1}^n (b_j - a_j)\,,$$

et on a donc prouvé le résultat cherché puisque cette quantité tend vers 0 lorsque q tend vers l'infini. ∎

Le dernier résultat du paragraphe montre que la classe des fonctions intégrables est complète pour la convergence en moyenne.

Théorème 3.8 (de complétude, et d'Égorov). *Soit (f_p) une suite de fonctions intégrables qui est de Cauchy pour la convergence en moyenne, c'est-à-dire que pour tout $\varepsilon > 0$, il existe un $N \in \mathbb{N}$ tel que*

$$\int |f_p(x) - f_q(x)|\, dx \leq \varepsilon \qquad \text{pour tous } p \text{ et } q \geq N\,.$$

Alors la suite (f_p) converge en moyenne vers une fonction intégrable f, et il existe une sous-suite $(f_{p_q})_{q\in\mathbb{N}}$ qui converge presque partout vers f.

Démonstration. Chaque fonction f_p est limite presque partout d'une suite de Cauchy de fonctions en escalier, qui converge aussi en moyenne comme on vient de le voir dans la démonstration précédente. On peut donc trouver une fonction en escalier g_p telle que $\int |f_p(x) - g_p(x)|\, dx \leq 2^{-p}$, et en utilisant le lemme d'Égorov partie (a), on peut même choisir g_p en sorte que $|f_p(x) - g_p(x)| \leq 2^{-p}$ pour tout $x \notin A_p$, où A_p est la réunion d'une suite de pavés dont la somme des mesures reste inférieure à 2^{-p}. La suite de fonctions en escalier (g_p) vérifie

$$\begin{aligned}
&\int |g_p(x) - g_q(x)|\, dx \leq \\
&\qquad \leq \int |g_p(x) - f_p(x)|\, dx + \int |f_p(x) - f_q(x)|\, dx + \int |f_q(x) - g_q(x)|\, dx \leq \\
&\qquad \leq 2^{-p} + \int |f_p(x) - f_q(x)|\, dx + 2^{-q}\,,
\end{aligned}$$

et c'est donc une suite de Cauchy. D'après le lemme d'Égorov partie (a), elle possède une sous-suite $(g_{p_q})_{q\in\mathbb{N}}$ qui converge presque partout (et aussi en moyenne, toujours comme à la démonstration de la proposition 3.7) vers une fonction intégrable f.

Pour voir que la suite (f_p) converge en moyenne vers f, on écrit la majoration

$$\begin{aligned}
&\int |f_p(x) - f(x)|\, dx \leq \\
&\qquad \leq \int |f_p(x) - f_{p_q}(x)|\, dx + \int |f_{p_q}(x) - g_{p_q}(x)|\, dx + \int |g_{p_q}(x) - f(x)|\, dx
\end{aligned}$$

où chacun des termes peut être rendu arbitrairement petit en prenant p et q assez grands. Enfin, pour voir que la suite extraite (f_{p_q}) converge presque partout vers f, il suffit de montrer que la suite $(f_p - g_p)$ converge presque partout vers 0 puisqu'on sait déjà que la suite (g_{p_q}) converge presque partout vers f. Or pour $x \notin A_p$, $|f_p(x) - g_p(x)| \leq 2^{-p}$, et donc pour $x \notin B_N = \cup_{p>N} A_p$, la suite $\big(f_p(x) - g_p(x)\big)$ tend vers 0. Comme c'est vrai pour tout N, on a aussi cette propriété pour $x \notin B = \cap_{N\in\mathbb{N}} B_N$. Et comme B est contenu dans chaque B_N, qui est réunion d'une suite de pavés dont la somme des mesures reste inférieure à 2^{-N}, on voit que cet ensemble B est négligeable, ce qui achève la démonstration du théorème. ∎

Exercices

3.1. Soient $\phi : \mathbb{R}^n \to \mathbb{R}^n$ la fonction affine considérée aux exercices **1.5** et **2.9** des pages 10 et 17, et $f : \mathbb{R}^n \to \mathbb{C}$ une fonction intégrable. Montrer que la fonction $f \circ \phi$ est intégrable, et calculer son intégrale en fonction de celle de f.

3.2. Soit $a > 0$. Déterminer des suites de Cauchy de fonctions en escalier qui convergent presque partout respectivement vers les fonctions $x\,\mathbb{1}_{[0,a]}(x)$ et $x^2\,\mathbb{1}_{[0,a]}(x)$ définies sur $\mathbb{R}$, puis en déduire les valeurs des intégrales $\int_{[0,a]} x\,dx$ et $\int_{[0,a]} x^2\,dx$.

3.3. Soit P un pavé de $\mathbb{R}^n$. Montrer que le produit d'une fonction intégrable sur P par une fonction continue sur P est intégrable sur P.

3.4. On rappelle que pour toute $f : \mathbb{R}^n \to \mathbb{R}$, on appelle *intégrale supérieure* et *intégrale inférieure* de f les deux nombres

$$\int^* f(x)\,dx = \inf\Big\{ \textstyle\int h(x)\,dx\,;\ h \text{ en escalier vérifiant } h \geq f \Big\}$$

$$\text{et} \qquad \int_* f(x)\,dx = \sup\Big\{ \textstyle\int g(x)\,dx\,;\ g \text{ en escalier vérifiant } g \leq f \Big\}.$$

Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction *intégrable au sens de Riemann*, c'est-à-dire telle que $\int^* f(x)\,dx = \int_* f(x)\,dx \in \mathbb{R}$.

(a) Montrer que la fonction f est bornée et qu'elle s'annule en dehors d'un compact de $\mathbb{R}^n$.

(b) Montrer qu'il existe une suite croissante $(g_p)_{p\in\mathbb{N}}$ et une suite décroissante $(h_p)_{p\in\mathbb{N}}$ de fonctions en escalier vérifiant $g_p \leq f \leq h_p$ pour tout entier p, et $\lim_{p\to\infty} \int g_p(x)\,dx = \lim_{p\to\infty} \int h_p(x)\,dx$.

(c) Montrer que la suite $(g_p)_{p\in\mathbb{N}}$ est de Cauchy et qu'elle possède une sous-suite qui converge presque partout vers f. En déduire que f est intégrable (au sens de Lebesgue), et que son intégrale est égale à la valeur commune de ses intégrales supérieure et inférieure.

3.5. Avec les notations de l'exercice **1.3** de la page 10, on dit qu'une suite de fonctions en escalier $f_p : \mathbb{R}^n \to \mathbb{R}^m$ est *de Cauchy* si pour tout $\varepsilon > 0$, il existe $N \in \mathbb{N}$ tel que $\int \|f_p(x) - f_q(x)\|\,dx \leq \varepsilon$ pour p et $q \geq N$. On dit ensuite que la fonction $f : \mathbb{R}^n \to \mathbb{R}^m$ est *intégrable* si f est limite presque partout d'une suite de Cauchy de fonctions en escalier (f_p), et on définit alors son intégrale par la formule $\int f(x)\,dx = \lim_{p\to\infty} \int f_p(x)\,dx$.

Montrer que la fonction $f : \mathbb{R}^n \to \mathbb{R}^m$ est intégrable si et seulement si chacune de ses composantes $f_j : \mathbb{R}^n \to \mathbb{R}$ est intégrable, et que $\int f(x)\,dx$ est le vecteur de composantes $\int f_1(x)\,dx, \ldots, \int f_m(x)\,dx$.

3.6. Soit $\mathbf{v} = (v_1, \ldots, v_n)$ une famille de n vecteurs de $\mathbb{R}^n$. On admettra (voir exercice **4.1** de la page 30) que la fonction indicatrice du *parallélépipède*

$$P(\mathbf{v}) = \Big\{ \sum_{j=1}^{n} c_j v_j\,;\ 0 \leq c_j \leq 1 \Big\}$$

est intégrable. On notera $\Delta(\mathbf{v}) = \int \mathbb{1}_{P(\mathbf{v})}(x)\,dx$ son intégrale.

(a) Montrer que $\Delta(\mathbf{v})$ ne dépend pas de l'ordre des vecteurs $v_1, \ldots, v_n$.

(b) Montrer que $\Delta(cv_1, v_2, \ldots, v_n) = |c|\,\Delta(v_1, \ldots, v_n)$ pour tout $c \in \mathbb{R}$ (on commencera par le montrer pour $c \in \mathbb{Z}$, puis pour $c \in \mathbb{Q}$).

(c) Montrer que $\Delta(v_1 + cv_2, v_2, \ldots, v_n) = \Delta(v_1, \ldots, v_n)$ pour tout $c \in \mathbb{R}$ (on le démontrera d'abord pour $|c| \leq 1$ grâce à l'invariance par translation de l'intégrale).

(d) En notant $\sigma(\mathbf{v}) = \text{signe}(\text{dét}(\mathbf{v}))$, on pose $D(\mathbf{v}) = \sigma(\mathbf{v})\,\Delta(\mathbf{v})$. Montrer que cette fonction D est $\mathbb{R}$-linéaire de chacun de ses arguments et qu'elle est alternée. En déduire que $\Delta(\mathbf{v}) = |\text{dét}(\mathbf{v})|$ (on pourra utiliser un choix particulier de vecteurs v_j).

3.7. Soient f une fonction intégrable positive sur un pavé P, et g une fonction continue sur P à valeurs réelles. On admettra que le produit fg est intégrable (voir exercice **3.3** de la page 25). Montrer la formule suivante, appelée *formule de la moyenne* : il existe $x_0 \in P$ tel que

$$\int_P f(x)\, g(x)\, dx = g(x_0) \int_P f(x)\, dx\,.$$

3.8. Soient $f_1 : \mathbb{R} \to \mathbb{C}, \ldots, f_n : \mathbb{R} \to \mathbb{C}$ des fonctions intégrables sur $\mathbb{R}$. Montrer que la fonction $f : \mathbb{R}^n \to \mathbb{C}$ définie par $f(x) = \prod_{j=1}^n f_j(x_j)$ est intégrable sur $\mathbb{R}^n$, et que l'on a $\int f(x)\, dx = \prod_{j=1}^n \left(\int f_j(t)\, dt \right)$.

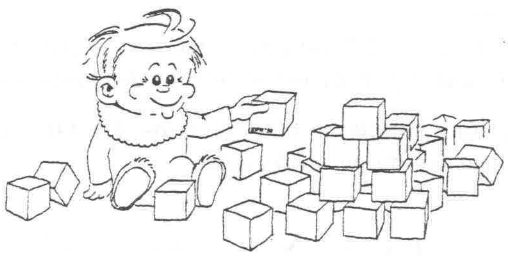
3.9. On définit $f : [-1,1] \to \mathbb{R}$ par la formule $f(x) = \sum_{k=1}^{\infty} \left(\sqrt{k} - \sqrt{k-1}\right) \mathbb{1}_{[-1/k,1/k]}(x)$.

(a) Montrer que cette série converge presque partout (partout sauf en 0).

(b) Montrer que f est intégrable (utiliser la définition de l'intégrabilité).

(c) Calculer f^2 et montrer que f^2 n'est pas intégrable (par l'absurde : montrer que son intégrale serait arbitrairement grande).

3.10. Montrer que si la suite $(f_p)_{p\in\mathbb{N}}$ de fonctions intégrables à valeurs complexes converge en moyenne vers la fonction f, alors les suites $(\mathscr{R}e\, f_p)$, $(\mathscr{I}m\, f_p)$, $(\overline{f_p})$ et $(|f_p|)$ convergent en moyenne vers $\mathscr{R}e\, f$, $\mathscr{I}m\, f$, $\overline{f}$ et $|f|$.



4. Théorèmes de convergence

Pour bien en comprendre la portée, l'étudiant aura intérêt à comparer les énoncés suivants aux phénomènes dont on a signalé dans l'introduction qu'ils empêchaient le passage à la limite dans les intégrales : phénomènes de concentration de masse ou de fuite de masse à l'infini.

Théorème 4.1 (de Beppo Levi, ou de convergence monotone). *Soit (f_p) une suite monotone de fonctions intégrables (à valeurs réelles). On suppose que la suite numérique $\left(\int f_p(x)\,dx\right)$ est bornée. Alors la suite (f_p) converge en moyenne et presque partout vers une fonction intégrable f.*

Démonstration. On peut par exemple supposer la suite (f_p) croissante, auquel cas la suite numérique $\left(\int f_p(x)\,dx\right)$ est croissante. Comme elle est aussi majorée, elle est convergente et donc de Cauchy. Comme la suite (f_p) est croissante, on a $\int |f_p(x) - f_q(x)|\,dx = \left|\int f_p(x)\,dx - \int f_q(x)\,dx\right|$ pour tous $p \geq q$, ce qui permet de conclure que la suite (f_p) est de Cauchy pour la convergence en moyenne. D'après le théorème de complétude, la suite (f_p) converge en moyenne vers une fonction intégrable f, et il existe une sous-suite $(f_{p_q})_{q\in\mathbb{N}}$ qui converge presque partout vers f. Comme pour chaque $x \in \mathbb{R}^n$ fixé la suite $(f_p(x))_{p\in\mathbb{N}}$ est croissante, elle converge dès qu'elle admet une sous-suite convergente. On en déduit que (f_p) converge vers f presque partout. ∎

Exemple. Soit K un compact de $\mathbb{R}^n$, c'est-à-dire un fermé borné. Alors sa fonction indicatrice $\mathbb{1}_K$ est intégrable.

En effet, K étant borné, il est contenu dans un grand pavé P. On reprend alors la construction effectuée pour démontrer la proposition 3.2 en notant $\mathscr{P}_0 = \{P\}$, puis $\mathscr{P}_1$, $\mathscr{P}_2$, ... les ensembles de pavés obtenus par division des éléments de l'ensemble précédent en 2^n pavés.

On note ensuite f_q la fonction indicatrice de la réunion de tous les éléments de $\mathscr{P}_q$ contenant au moins un point de K. Il s'agit d'une suite de fonctions en escalier donc intégrables, et il est facile de voir qu'elle est décroissante. Si $x \in K$, on a $f_q(x) = 1$ pour tout q, tandis que si $x \notin K$, ce point est le centre d'une boule qui ne rencontre pas K, et donc on a $f_q(x) = 0$ pour q assez grand. On voit ainsi que la suite décroissante (f_q) converge (partout) vers $\mathbb{1}_K$, et on obtient donc l'intégrabilité de $\mathbb{1}_K$ par le théorème puisque les intégrales des fonctions f_q sont toutes dans l'intervalle borné $[0, |P|]$.

Théorème 4.2 (de Lebesgue, ou de convergence dominée). *Soit* (f_p) *une suite de fonctions intégrables qui converge presque partout vers une fonction* f. *On suppose qu'il existe une fonction intégrable* F indépendante de p *telle que* $|f_p| \leq F$ *presque partout pour tout* p. *Alors la fonction* f *est intégrable et la suite* (f_p) *converge en moyenne vers* f.

Pour démontrer ce théorème, on commence par établir un lemme.

Lemme 4.3. *Soit* (f_p) *une suite de fonctions positives intégrables majorées presque partout par une fonction intégrable fixe* F. *Alors la fonction* $g(x) = \sup_p (f_p(x))$ *est intégrable.*

Démonstration. Pour tout $q \geq 1$, $g_q(x) = \sup_{1 \leq p \leq q}(f_p(x))$ est intégrable. Comme on a $g_q \leq F$ presque partout, la suite (g_q) est une suite croissante de fonctions intégrables, d'intégrales majorées par $\int F(x)\,dx$. Il résulte donc du théorème de convergence monotone que $g = \sup_p (f_p) = \lim_{q \to \infty} g_q$ est intégrable. ∎

Démonstration du théorème 4.2. Les fonctions $|f_p - f_q|$ sont majorées par la fonction fixe $2F$ qui est intégrable, et donc il résulte du lemme que la fonction $g_N(x) = \sup_{p,q \geq N} |f_p(x) - f_q(x)|$ est intégrable. Pour tous p et $q \geq N$, on a donc $|f_p - f_q| \leq g_N$, puis en intégrant et en prenant la borne supérieure sur tous les p et $q \geq N$, on trouve

$$\sup_{p,q \geq N} \int |f_p(x) - f_q(x)|\,dx \leq \int g_N(x)\,dx\,.$$

Par ailleurs, la suite $(g_N)_{N \in \mathbb{N}}$ est une suite décroissante de fonctions intégrables positives qui converge presque partout vers 0 (puisque $f_p(x)$ converge pour presque tout x), et il résulte du théorème de convergence monotone que la suite numérique $\int g_N(x)\,dx$ tend vers 0. Cela signifie que la suite (f_p) est de Cauchy pour la convergence en moyenne, et on déduit le résultat du théorème de complétude. ∎

Exemple. Soient $I \subset \mathbb{R}$ un intervalle ouvert de $\mathbb{R}$ et $f : I \to \mathbb{C}$ une fonction intégrable sur I. Étant donnés un complexe c et un point $a \in I$, on appelle *intégrale indéfinie* de f la fonction $F(x) = c + \int_{[a,x]} f(t)\,dt$ si $x \geq a$, $F(x) = c - \int_{[x,a]} f(t)\,dt$ si $x \leq a$. Nous allons montrer, à l'aide du théorème de convergence dominée, qu'une telle intégrale indéfinie est une fonction continue sur I et possédant des limites aux deux extrémités de I (ce qui implique que F est uniformément continue).

Soient $x_0 > a$ un point de I et (x_p) une suite de points de I tendant vers x_0. La fonction f étant intégrable sur I, elle est intégrable sur $[a, x_p]$. La suite de fonctions (f_p) définie par $f_p = \mathbb{1}_{[a,x_p]} f$ est donc une suite de fonctions intégrables, et elle converge vers $\mathbb{1}_{[a,x_0]} f$ en tout point, sauf peut-être en x_0. De plus, on a la majoration $|f_p| \leq \mathbb{1}_I |f|$, ce qui permet d'appliquer le théorème de convergence dominée puisque $|f|$ est intégrable sur I et indépendant de p. Il en résulte que la suite $(F(x_p))$ tend vers $F(x_0)$, et comme on a ce résultat pour toute suite (x_p) tendant vers x_0, on en déduit que F est continue en x_0. La continuité en un point $x_0 \leq a$ se montrerait par des arguments similaires, et l'existence d'une limite aux extrémités de l'intervalle aussi, en prenant une suite (x_p) de points de I tendant vers une de ces extrémités.

On prendra garde qu'une intégrale indéfinie n'est pas nécessairement dérivable lorsque l'on ne suppose pas f continue.

Corollaire 4.4 (d'intégration terme à terme). *Soit (u_p) une suite de fonctions intégrables. On suppose que la série numérique de terme général $\int |u_p(x)|\,dx$ est convergente. Alors la série de fonctions de terme général u_p converge presque partout vers une fonction intégrable f, et on a $\int f(x)\,dx = \sum_{p=0}^{\infty} \int u_p(x)\,dx$.*

Démonstration. La suite de fonctions (F_p) définie par $F_p(x) = \sum_{q=0}^{p} |u_q(x)|$ est une suite croissante de fonctions intégrables dont les intégrales $\int F_p(x)\,dx = \sum_{q=0}^{p} \int |u_q(x)|\,dx$ forment une suite numérique bornée par hypothèse. Il résulte donc du théorème de convergence monotone que cette suite converge presque partout vers une fonction F intégrable (et indépendante de p). Comme la convergence absolue d'une série numérique entraîne sa convergence, la suite de fonctions (f_p) définie par $f_p(x) = \sum_{q=0}^{p} u_q(x)$ converge elle aussi presque partout vers une fonction f, et la majoration $|f_p| \leq F_p \leq F$ presque partout permet d'appliquer le théorème de convergence dominée qui prouve que f est intégrable et que son intégrale est égale à la somme de la série numérique de terme général $\int u_p(x)\,dx$. ∎

Corollaire 4.5 (de convergence uniformément bornée sur un compact). *Soient K un compact de $\mathbb{R}^n$, et (f_p) une suite de fonctions intégrables sur K qui converge presque partout vers une fonction f. On suppose de plus que la suite est uniformément bornée, c'est-à-dire qu'il existe une constante M* indépendante de *p telle que $|f_p| \leq M$ presque partout sur K. Alors la fonction f est intégrable sur K et la suite $\int_K f_p(x)\,dx$ tend vers $\int_K f(x)\,dx$.*

Démonstration. Comme la fonction $\mathbb{1}_K$ est intégrable (voir plus haut), la fonction $F = M\,\mathbb{1}_K$ aussi, et il suffit donc d'appliquer le théorème de convergence dominée à la suite de fonctions $(\mathbb{1}_K f_p)$. ∎

Ce résultat permet en particulier de passer à la limite dans une intégrale $\int_K f_p(x)\,dx$ lorsque la suite (f_p) converge uniformément sur K. En effet, il suffit dans ce cas d'appliquer le corollaire à la suite $(f_p - f_q)_{p\in\mathbb{N}}$ pour q fixé, si ce nombre q a été choisi assez grand pour que $|f_p - f_q| \leq 1$ sur K lorsque $p \geq q$.

Remarque en guise de conclusion. Jusqu'à présent, nous avons discuté deux types de convergence : la convergence presque partout, et la convergence en moyenne. Ces notions sont équivalentes à deux restrictions près. Au théorème de complétude, nous avons vu que la convergence en moyenne entraînait la convergence presque partout *non pas de la suite elle-même, mais d'une sous-suite.* Réciproquement, le théorème de convergence dominée assure que la convergence presque partout implique la convergence en moyenne *pourvu que les termes de la suite soient dominés par une fonction intégrable fixe* (penser aux contre-exemples de l'introduction).

Exercices

4.1. Soit $B_n(1) = \{x \in \mathbb{R}^n \,;\, x_1^2 + \ldots + x_n^2 \leq 1\}$ la boule unité de $\mathbb{R}^n$. Montrer que la fonction indicatrice de $B_n(1)$ est intégrable. De même, montrer que la fonction indicatrice du parallélépipède $P(\mathbf{v})$ défini à l'exercice **3.6** de la page 25 est intégrable.

4.2. Soient $(c_p)_{p\in\mathbb{N}}$ une suite de réels positifs et $(f_p)_{p\in\mathbb{N}}$ une suite de fonctions intégrables positives.

(a) Montrer que la formule $\liminf_{p\to\infty} c_p = \lim_{p\to\infty}(\inf_{q\geq p} c_q)$ définit un nombre $\liminf_{p\to\infty} c_p \in [0, +\infty]$.

(b) On pose $g_p(x) = \inf_{q\geq p} f_q(x)$. Montrer que $(g_p)_{p\in\mathbb{N}}$ est une suite croissante de fonctions intégrables vérifiant $\int g_p(x)\,dx \leq \inf_{q\geq p} \int f_q(x)\,dx$.

(c) Montrer l'assertion suivante appelée *lemme de Fatou* : si $\liminf_{p\to\infty} \int f_p(x)\,dx < \infty$, alors la fonction $x \mapsto \liminf_{p\to\infty} f_p(x)$ est intégrable, et

$$\int \liminf_{p\to\infty} f_p(x)\,dx \leq \liminf_{p\to\infty} \int f_p(x)\,dx\,.$$

Que donne cet énoncé dans le cas de la suite définie par $f_p(x) = p\,\mathbb{1}_{[0,1/p]}(x)$?

4.3. Soient $(P_q)_{q\in\mathbb{N}}$ une suite de pavés de $\mathbb{R}^n$, $A_p = \cup_{q<p} P_q$, $A = \cup_{q\in\mathbb{N}} P_q$ et $f : \mathbb{R}^n \to \mathbb{C}$.

(a) Montrer que si f est intégrable sur A, alors f est intégrable sur tous les A_p avec l'égalité $\int_A f(x)\,dx = \lim_{p\to\infty} \int_{A_p} f(x)\,dx$.

(b) Montrer que f est intégrable sur A si et seulement si f est intégrable sur chaque A_p et la suite numérique $\left(\int_{A_p} |f(x)|\,dx\right)_{p\in\mathbb{N}}$ est bornée.

4.4. Soit $f : \mathbb{R}^n \to \mathbb{C}$ une fonction intégrable.

(a) Montrer que f est intégrable sur tout fermé F de $\mathbb{R}^n$: pour cela, on introduira la fonction continue positive $g(x) = \mathrm{dist}\,(x, F)$, et on considérera la suite $(f_p)_{p\in\mathbb{N}}$ définie par $f_p(x) = (1 + p\,g(x))^{-1} f(x)$.

(b) Montrer que f est intégrable sur tout ouvert X de $\mathbb{R}^n$.

4.5. Soient $f : \mathbb{R}^n \to \mathbb{C}$ une fonction intégrable, $x_0 \in \mathbb{R}^n$ et $f_\varepsilon(x) = f(x - \varepsilon\, x_0)$. Montrer que la fonction f_ε converge en moyenne vers f lorsque ε tend vers 0 (on commencera par le faire dans le cas d'une fonction f continue à support compact).

4.6. Soit (f_p) la suite de fonctions définies sur $\mathbb{R}$ par $f_p = \mathbb{1}_{[p-p^{-1},p]}$ pour $p \geq 1$. Montrer que cette suite converge en moyenne vers une fonction f que l'on précisera, et montrer qu'il n'existe pas de fonction F intégrable et indépendante de p telle que $|f_p| \leq F$ pour tout p.

4.7. Soit (f_p) une suite de fonctions intégrables sur $\mathbb{R}^n$ qui converge presque partout vers une fonction intégrable f.

(a) Montrer que si la suite (f_p) converge en moyenne vers f, alors $\int |f(x)|\,dx = \lim_{p\to\infty} \int |f_p(x)|\,dx$.

(b) Montrer à l'aide du lemme de Fatou (exercice **4.2** ci-dessus) que si $\int |f(x)|\,dx = \lim_{p\to\infty} \int |f_p(x)|\,dx$, alors $\int_A |f(x)|\,dx = \lim_{p\to\infty} \int_A |f_p(x)|\,dx$ pour toute partie $A \subset \mathbb{R}^n$ telle que f et les f_p soient intégrables sur A.

(c) Montrer que pour tout $\varepsilon > 0$, il existe un entier q et un ensemble $A \subset \mathbb{R}^n$ tels que $A \cup [-q, q]^n = \mathbb{R}^n$, $\int_A |f(x)|\,dx \leq \varepsilon/4$ et la suite (f_p) converge vers f uniformément sur $\mathbb{R}^n \setminus A$.

(d) En déduire la réciproque suivante de la question (a) : si $\int |f(x)|\,dx = \lim_{p\to\infty} \int |f_p(x)|\,dx$, alors la suite (f_p) converge en moyenne vers f.

5. Mesurabilité

La notion de mesurabilité étudiée dans ce paragraphe joue un rôle essentiel dans la généralisation de la théorie de l'intégration au cadre abstrait de la théorie de la mesure. Ici, nous l'utiliserons surtout pour le critère d'intégrabilité 5.2. À part ce critère, on pourra retenir que, sauf si on les construit exprès pour n'être pas mesurables, toutes les fonctions et toutes les parties de $\mathbb{R}^n$ sont mesurables[1].

Définition 5.1. *Une fonction $f : \mathbb{R}^n \to \mathbb{C}$ (resp. $f : \mathbb{R}^n \to [0, \infty] = \mathbb{R}_+ \cup \{\infty\}$) est dite* mesurable *(resp.* mesurable positive*) si c'est la limite presque partout d'une suite de fonctions en escalier. Pour toute fonction mesurable positive f, on définit son* intégrale $\int f(x)\,dx \in [0, \infty]$ *de la façon suivante : c'est $\int g(x)\,dx$ si f est égale presque partout à une fonction intégrable g, et c'est ∞ sinon. Enfin, une partie $A \subset \mathbb{R}^n$ est dite* mesurable *si $\mathbb{1}_A$ est mesurable, et la* mesure de A *est alors le nombre $|A| = \int \mathbb{1}_A(x)\,dx \in [0, \infty]$.*

Exemples. Toute fonction intégrable ou continue est mesurable. Pour les fonctions intégrables, cela résulte des définitions elles-mêmes. Pour une fonction continue f, on a montré que l'on pouvait trouver pour tout entier p une fonction en escalier f_p qui vérifie $|f_p(x) - f(x)| \leq 2^{-p}$ pour tout $x \in [-p, p]^n$, ce qui prouve que f est limite partout de la suite (f_p), et qu'elle est donc mesurable. Pour toute fonction f mesurable, l'intégrale $\int |f(x)|\,dx$ a un sens puisque la fonction $|f|$ est alors mesurable positive, et cette propriété va nous permettre d'énoncer le critère d'intégrabilité 5.2.

Les ensembles négligeables sont les ensembles mesurables de mesure nulle puisque d'après la proposition 3.5, $\mathbb{1}_A = 0$ presque partout si et seulement si $\int \mathbb{1}_A(x)\,dx = 0$. Tous les compacts sont mesurables et de mesure finie puisque leurs fonctions indicatrices sont intégrables. L'espace $\mathbb{R}^n$ tout entier est aussi mesurable puisque sa fonction indicatrice est continue ; on verra plus loin qu'il est de mesure infinie.

Comme on peut considérer en pratique que toutes les fonctions sont mesurables, le résultat suivant nous sera très utile.

Théorème 5.2 (Critère d'intégrabilité). *Pour qu'une fonction $f : \mathbb{R}^n \to \mathbb{C}$ soit intégrable, il faut et il suffit que f soit mesurable et que $\int |f(x)|\,dx < \infty$. En particulier, pour qu'une fonction mesurable f soit intégrable, il suffit que $|f|$ soit majoré par une fonction intégrable.*

---

[1] On peut même poser cela en axiome, à condition de renoncer à l'axiome du choix non-dénombrable, sans lequel d'ailleurs toute l'analyse classique reste valable.

Démonstration. Il est clair que toute fonction intégrable est mesurable et de module intégrable. Réciproquement, si $\int |f(x)|\,dx < \infty$, c'est que $|f|$ est intégrable. Plus généralement, s'il existe une fonction intégrable $g \geq |f|$ et si (f_p) est une suite de fonctions en escalier convergeant presque partout vers f, alors la suite (g_p) définie par

$$g_p(x) = \inf\Big(g(x), \sup\big(\mathscr{Re}\, f_p(x), -g(x)\big)\Big)$$

est une suite de fonctions intégrables vérifiant $|g_p| \leq g$ pour tout p, et convergeant presque partout vers $\mathscr{Re}\, f$. On a donc $\mathscr{Re}\, f$ intégrable par le théorème de convergence dominée, et comme on peut raisonner de la même manière pour la partie imaginaire, on a prouvé que $f = \mathscr{Re}\, f + i\,\mathscr{Im}\, f$ est intégrable. ∎

Exemples. Si A est une partie mesurable de $\mathbb{R}^n$ contenue dans un ensemble B et si f est intégrable sur B, alors f est intégrable sur A. En effet, d'après la proposition 5.3 (a) ci-dessous, $\mathbb{1}_A f$ est mesurable comme produit des deux fonctions mesurables $\mathbb{1}_A$ et $\mathbb{1}_B f$, et on conclut alors en remarquant que $|\mathbb{1}_A f|$ est majoré par la fonction intégrable $|\mathbb{1}_B f|$.

Le théorème 5.2 permet aussi de montrer que toute fonction mesurable bornée sur un compact est intégrable sur ce compact, puisque l'on a la majoration $|\mathbb{1}_K f| \leq (\sup_K |f|)\, \mathbb{1}_K$, et que la fonction indicatrice d'un compact est intégrable, et plus généralement que tout produit d'une fonction intégrable par une fonction mesurable bornée est intégrable.

Proposition 5.3. *Les fonctions mesurables possèdent les propriétés suivantes :*
(a) *Si f et g sont mesurables, $x_0 \in \mathbb{R}^n$ et $c \in \mathbb{C}$, les fonctions $x \mapsto f(x - x_0)$, $\overline{f}$, $f + c\,g$, fg et $|f|$ sont mesurables. Si de plus $f \neq 0$ presque partout, alors $1/f$ est mesurable.*
(b) *Si f est mesurable et $g : \mathbb{C} \to \mathbb{C}$ est continue, alors $g \circ f$ est mesurable.*
(c) *Si $(f_p)_{p\in\mathbb{N}}$ est une suite de fonctions mesurables convergeant presque partout vers une fonction f, alors f est mesurable.*
(d) *Si $(f_p)_{p\in\mathbb{N}}$ est une suite de fonctions mesurables à valeurs réelles, alors $\sup_{p\in\mathbb{N}}(f_p)$ et $\inf_{p\in\mathbb{N}}(f_p)$ sont mesurables.*

Démonstration. Pour obtenir la propriété (a), il suffit de remarquer que si f et g sont en escalier, les fonctions $x \mapsto f(x - x_0)$, $\overline{f}$, $f + c\,g$, fg et $|f|$ sont aussi en escalier ; si de plus une suite de fonctions en escalier (f_p) converge presque partout vers une fonction $f \neq 0$ presque partout, alors la suite (g_p) définie par $g_p(x) = 1/f_p(x)$ si $f_p(x) \neq 0$, $g_p(x) = 0$ sinon est une suite de fonctions en escalier qui converge presque partout vers $1/f$.

Pour la propriété (b), on remarque que si (f_p) est une suite de fonctions en escalier qui converge presque partout vers f, alors la suite $(g \circ f_p - g(0))$ est une suite de fonctions en escalier qui converge presque partout vers $g \circ f - g(0)$, ce qui prouve que $g \circ f$ est mesurable lorsque f l'est puisque la fonction constante $g(0)$ est mesurable.

Pour la propriété (c), on commence par construire une fonction intégrable h partout strictement positive en posant $h(x) = \sum_{q=1}^{\infty} 2^{-q} (2q)^{-n} \mathbb{1}_{[-q,q]^n}(x)$, l'intégrabilité de h provenant du théorème de convergence monotone. On pose alors

$$g_p(x) = h(x) \frac{f_p(x)}{h(x) + |f_p(x)|} \qquad \text{et} \qquad g(x) = h(x) \frac{f(x)}{h(x) + |f(x)|},$$

et comme h est strictement positive, on obtient facilement les majorations $|g_p| < h$ et $|g| < h$. Les fonctions g_p sont mesurables d'après la propriété (a) et majorées en module par la fonction intégrable h, et elles sont donc intégrables. Comme la suite (g_p) converge presque partout vers g, il résulte alors du théorème de convergence dominée que g est intégrable. Enfin, par définition de g, on peut écrire

$$|g| = h \frac{|f|}{h + |f|} \qquad \Rightarrow \qquad |f|\,(h - |g|) = h\,|g| \qquad \Rightarrow$$
$$\Rightarrow \qquad \frac{|f|}{h} = \frac{|g|}{h - |g|} \qquad \Rightarrow \qquad f = g\left(1 + \frac{|f|}{h}\right) = \frac{g\,h}{h - |g|},$$

ce qui permet de conclure que f est mesurable d'après (a).

Enfin pour la propriété (d), on commence par remarquer que les fonctions $\sup(f_1, f_2) = (f_1 + f_2 + |f_1 - f_2|)/2$ et $\inf(f_1, f_2) = (f_1 + f_2 - |f_1 - f_2|)/2$ sont mesurables lorsque f_1 et f_2 le sont, ce qui s'étend à un nombre fini de fonctions mesurables par récurrence. Dans le cas d'une suite (f_p), on peut écrire $\sup_{p\in\mathbb{N}} (f_p) = \lim_{q\to\infty} \sup_{p\leq q} (f_p)$ et $\inf_{p\in\mathbb{N}} (f_p) = \lim_{q\to\infty} \inf_{p\leq q} (f_p)$, ce qui permet de conclure d'après la propriété (c). ∎

Exemples. Si f est mesurable, alors $\mathscr{Re}\, f$ et $\mathscr{Im}\, f$ sont mesurables comme combinaisons linéaires de f et $\overline{f}$, et si $p > 0$, la fonction $|f|^p$ aussi comme composée $g \circ f$ par la fonction continue $g(z) = |z|^p$.

Si une fonction $F : \mathbb{R} \to \mathbb{C}$ est continue, et dérivable presque partout, alors sa dérivée est mesurable. En effet, c'est la limite presque partout de la suite (f_p) de fonctions continues (donc mesurables) définies par

$$f_p(x) = \frac{F\left(x + \frac{1}{p}\right) - F(x)}{\frac{1}{p}}.$$

Conventions pour les calculs avec l'infini. Lorsque l'on aborde les fonctions mesurables positives, on est amené à effectuer des comparaisons et des calculs avec l'infini. On convient donc que pour tout réel c, on a $c < \infty$ et $c + \infty = \infty + \infty = \infty$, que $c \times \infty = \infty \times \infty = \infty^c = \infty$ si $c > 0$, et que $0 \times \infty = 0$. Avec ces conventions, on peut énoncer les propriétés suivantes.

Proposition 5.4. *Soient f et g deux fonctions mesurables positives.*
(a) Croissance : *Si $f \leq g$, alors $\int f(x)\,dx \leq \int g(x)\,dx$, avec l'égalité si $f = g$ presque partout.*
(b) Linéarité restreinte : *Si $c \in [0, \infty]$, alors $f + c\,g$ est mesurable positive, et $\int \big(f(x) + c\,g(x)\big)\,dx = \int f(x)\,dx + c \int g(x)\,dx$.*
(c) Convergence croissante : *Si (f_p) est une suite croissante de fonctions mesurables positives convergeant (simplement) vers f, alors $\int f(x)\,dx = \lim_{p\to\infty} \int f_p(x)\,dx$.*

Démonstration. Commençons par démontrer la propriété (a). L'inégalité étant immédiate si $\int g(x)\,dx = \infty$, on peut supposer que g est égale presque partout à une fonction intégrable qu'on notera encore g. Il résulte alors du théorème 5.2 que f est aussi intégrable, et l'inégalité se réduit alors à la propriété de croissance des intégrales des fonctions intégrables. Quant au cas où $f = g$ presque partout, il résulte de ce qu'alors f est égale presque partout à une fonction intégrable si et seulement si g l'est aussi.

Pour démontrer la propriété de linéarité (b), nous le ferons séparément pour la somme et pour le produit par un scalaire. Si f et g sont (égales presque partout à des fonctions) intégrables, on a bien évidemment $\int (f+g)(x)\,dx = \int f(x)\,dx + \int g(x)\,dx$, tandis que si f (par exemple) n'est pas intégrable, on a $f \leq f + g$ puis $\infty = \int f(x)\,dx \leq \int (f+g)(x)\,dx$ d'après le (a), si bien que $\int (f+g)(x)\,dx = \infty = \int f(x)\,dx + \int g(x)\,dx$. Pour traiter les produits cf, il faut distinguer les cas suivant que c et $\int f(x)\,dx$ valent ∞, 0 ou un réel > 0 : si c et $\int f(x)\,dx < \infty$, c'est le cas (connu) des fonctions intégrables ; si c ou $\int f(x)\,dx = 0$, on a $cf = 0$ presque partout d'où $\int cf(x)\,dx = 0$; enfin si c et $\int f(x)\,dx > 0$ et que l'un d'eux est infini, on remarque que cf n'est pas intégrable d'où $\int cf(x)\,dx = \infty$.

Quant à la propriété (c), elle résulte du théorème de convergence monotone lorsque $\lim_{p\to\infty} \int f_p(x)\,dx < \infty$, et de la propriété (a) dans le cas contraire puisque $f_p \leq f$ pour tout p. ∎

Remarque. On prendra garde que la propriété (c) devient fausse si l'on remplace l'hypothèse de croissance de la suite par une hypothèse de décroissance.

Les propriétés des intégrales des fonctions mesurables positives nous conduisent enfin aux propriétés des ensembles mesurables et de leur mesure. Ainsi, la propriété (a) de la proposition précédente montre que si $A \subset B$ sont mesurables, alors $|A| \leq |B|$. On a aussi les résultats suivants.

Proposition 5.5.
(a) *L'espace $\mathbb{R}^n$ est mesurable et de mesure ∞.*
(b) *Le complémentaire d'un ensemble mesurable est mesurable.*
(c) *Si $(A_p)_{p\in\mathbb{N}}$ est une suite d'ensembles mesurables, alors $A = \cup_{p\in\mathbb{N}} A_p$ et $B = \cap_{p\in\mathbb{N}} A_p$ sont mesurables. De plus, on a $|A| \leq \sum_{p\in\mathbb{N}} |A_p|$, avec l'égalité si les A_p sont disjoints.*
(d) *Les ouverts et les fermés sont mesurables.*

Démonstration. On a déjà vu que $\mathbb{R}^n$ était mesurable, et comme $[-q,q]^n \subset \mathbb{R}^n$ et que $|[-q,q]^n| = (2q)^n$ pour tout entier q, on en déduit que $|\mathbb{R}^n| = \infty$. Si B est le complémentaire d'un ensemble mesurable A, on a $1\!\!1_B = 1 - 1\!\!1_A$ et donc B est mesurable. Pour la propriété (c), il suffit d'observer que

$$1\!\!1_A = \sup_{p\in\mathbb{N}} (1\!\!1_{A_p}) \qquad \text{et} \qquad 1\!\!1_B = \inf_{p\in\mathbb{N}} (1\!\!1_{A_p})$$

pour voir que A et B sont mesurables ; comme $1\!\!1_A \leq \sum_{p\in\mathbb{N}} 1\!\!1_{A_p}$ avec l'égalité si les A_p sont disjoints, on obtient les relations annoncées grâce aux différentes conclusions de la proposition 5.4. Quant à la propriété (d), elle résulte de ce que tout ouvert est le complémentaire d'un fermé, et que tout fermé F est la réunion de la suite de compacts $K_p = F \cap [-p,p]^n$. ∎

Exercices

5.1. Montrer que si $f : \mathbb{R}^n \to \mathbb{C}$ et $g : \mathbb{R}^m \to \mathbb{C}$ sont mesurables, alors la fonction $h : \mathbb{R}^{n+m} \to \mathbb{C}$ définie par $h(x,y) = f(x)\,g(y)$ est aussi mesurable.

5.2. Avec les notations des exercices **1.3** et **3.5** des pages 10 et 25, on considère une fonction $f : \mathbb{R}^n \to \mathbb{R}^m$ dont toutes les composantes sont mesurables.

(a) Montrer que la fonction $x \mapsto \|f(x)\|$ est mesurable positive.

(b) Montrer que f est intégrable si et seulement si $\int \|f(x)\|\,dx < \infty$, et qu'alors on a l'inégalité $\|\int f(x)\,dx\| \leq \int \|f(x)\|\,dx$.

5.3. Soit $[a,b]$ un intervalle compact de $\mathbb{R}$. On dit que la fonction $f : [a,b] \to \mathbb{C}$ est *réglée* si elle admet des limites à droite et à gauche en tout point de l'intervalle $]a,b[$, une limite à droite en a et une limite à gauche en b.

(a) Montrer que les fonctions en escalier, les fonctions continues et les fonctions monotones sont réglées.

(b) Montrer que toute fonction réglée est limite uniforme d'une suite de fonctions en escalier (pour tout $\varepsilon > 0$ fixé, on associera à chaque $x \in [a,b]$ un intervalle $]x-\delta(x), x+\delta(x)[$ dans lequel $f(y)$ est distant de moins de ε de la limite à droite (resp. à gauche) de f en x si y est situé à droite (resp. à gauche) de x).

(c) Montrer que toute fonction $f : \mathbb{R} \to \mathbb{C}$ admettant des limites à droite et à gauche en tout point est mesurable.

5.4. Soient $[a,b]$ un intervalle compact de $\mathbb{R}$ et $\phi : \mathbb{R} \to \mathbb{R}$ une fonction dérivable dont la dérivée ϕ' est croissante sur $[a,b]$ (une fonction ϕ possédant cette propriété est dite *convexe* sur $[a,b]$).

Soient $n \geq 1$ un entier, $m : \mathbb{R}^n \to \mathbb{R}_+$ une fonction positive intégrable d'intégrale égale à 1, et $f : \mathbb{R}^n \to [a,b]$ une fonction mesurable.

(a) Montrer que les fonctions $x \mapsto f(x)\,m(x)$ et $x \mapsto \big(\phi\circ f(x)\big)\,m(x)$ sont intégrables sur $\mathbb{R}^n$, et que le nombre $y_0 = \int f(x)\,m(x)\,dx$ appartient à l'intervalle $[a,b]$.

(b) Montrer que $\forall x \in \mathbb{R}^n$, $\phi(y_0) \leq \phi\circ f(x) - \phi'(y_0)\,\big(f(x)-y_0\big)$ (on pourra distinguer les x pour lesquels $f(x) \geq y_0$ et ceux pour lesquels $f(x) \leq y_0$).

(c) En déduire l'*inégalité de Jensen* : $\phi\big(\int f(x)\,m(x)\,dx\big) \leq \int \big(\phi\circ f(x)\big)\,m(x)\,dx$.

5.5. Soit $f_p = 1\!\!1_{[0,2^{-p}]\times\mathbb{R}} : \mathbb{R}^2 \to [0,1]$. Montrer que la suite (f_p) est une suite décroissante de fonctions mesurables positives dont on calculera les intégrales ainsi que l'intégrale de la limite. Comparer avec le théorème 5.4 (c) et commenter.

5.6. Soient $A_1, \ldots, A_N$ des parties mesurables disjointes de $\mathbb{R}^n$, $c_1, \ldots, c_N$ des nombres complexes non nuls et $f = c_1 \mathbb{1}_{A_1} + \ldots + c_N \mathbb{1}_{A_N}$. Montrer que f est intégrable si et seulement si tous les A_k sont de mesure finie, et qu'alors $\int f(x)\,dx = \sum_{k=1}^{N} c_k\,|A_k|$.

5.7. Soit $f : \mathbb{R}^n \to \mathbb{C}$ une fonction intégrable. Démontrer l'assertion suivante : $\forall \varepsilon > 0$, $\exists \delta > 0$: $\forall A \subset \mathbb{R}^n$, $|A| < \delta \Rightarrow \left|\int_A f(x)\,dx\right| < \varepsilon$ (on commencera par le cas où f est bornée, puis on se ramènera à ce cas en montrant que $\lim_{t\to\infty} \int_{A(t)} |f(x)|\,dx = 0$ lorsque l'on note $A(t) = \{\, x \in \mathbb{R}^n \,;\, |f(x)| > t \,\}$).

5.8. Soit f une fonction définie sur $\mathbb{R}^n$ à valeurs complexes. Le but des questions (a) et (b) ci-dessous est de montrer que f est une fonction mesurable si et seulement si $f^{-1}(X)$ est une partie mesurable de $\mathbb{R}^n$ pour tout ouvert X de $\mathbb{C}$. Les questions (c), (d) et (e) fournissent des applications de cette propriété.

(a) Soient X un ouvert de $\mathbb{C}$ et $d_X(c)$ la distance du point $c \in \mathbb{C}$ au complémentaire de X. Montrer que la formule $g_p(c) = \inf(p\,d_X(c), 1)$ définit une suite (g_p) de fonctions continues sur $\mathbb{C}$ qui converge partout vers $\mathbb{1}_X$, puis en déduire que si la fonction f est mesurable, alors $f^{-1}(X)$ est une partie mesurable de $\mathbb{R}^n$.

(b) On suppose maintenant que l'image réciproque par f de tout ouvert de $\mathbb{C}$ est une partie mesurable de $\mathbb{R}^n$. Montrer que l'image réciproque de tout fermé et de toute intersection finie d'ouverts et de fermés est aussi mesurable, puis en déduire que f est une fonction mesurable (on montrera que f est limite d'une suite de fonctions de la forme $\sum_{k=1}^{N} c_k \mathbb{1}_{A_k}$ où les A_k sont des parties mesurables de $\mathbb{R}^n$).

(c) Montrer que si f est à valeurs réelles, alors f est une fonction mesurable si et seulement si $f^{-1}([c, \infty[)$ est une partie mesurable de $\mathbb{R}^n$ pour tout $c \in \mathbb{R}$.

(d) Montrer qu'une fonction $f : \mathbb{R}^n \to \mathbb{C}$ continue presque partout est mesurable (on pourra remarquer que pour une partie $A \subset \mathbb{R}^n$ dont le complémentaire est négligeable, la fonction $f_{|A} : A \to \mathbb{C}$ est continue).

(e) Une fonction $f : \mathbb{R}^n \to \mathbb{R}$ est dite *semi-continue inférieurement* (resp. *supérieurement*) si pour tout $x_0 \in \mathbb{R}^n$ et tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que : $|x - x_0| \leq \delta \Rightarrow f(x) \geq f(x_0) - \varepsilon$ (resp. $f(x) \leq f(x_0) + \varepsilon$). Montrer que toute fonction semi-continue (inférieurement ou supérieurement) est mesurable.

5.9. Montrer que si $A \subset \mathbb{R}^n$ et $B \subset \mathbb{R}^m$ sont mesurables, alors $A \times B \subset \mathbb{R}^n \times \mathbb{R}^m$ est mesurable et $|A \times B| = |A|\,|B|$.

5.10. Soit $(A_p)_{p\in\mathbb{N}}$ une suite de parties mesurables de $\mathbb{R}^n$. Cette suite sera dite *croissante* si $A_p \subset A_{p+1}$ pour tout p, *décroissante* si $A_p \supset A_{p+1}$ pour tout p, et *monotone* si elle est croissante ou décroissante.

(a) On pose $\overline{\lim}\, A_p = \cap_{p\in\mathbb{N}} \cup_{q>p} A_q$ et $\underline{\lim}\, A_p = \cup_{p\in\mathbb{N}} \cap_{q>p} A_q$. Montrer que $\overline{\lim}\, A_p$ et $\underline{\lim}\, A_p$ sont mesurables, et que $\underline{\lim}\, A_p \subset \overline{\lim}\, A_p$.

(b) Montrer que ces deux limites sont égales lorsque la suite (A_p) est monotone. On notera alors $\lim A_p$ cette limite commune.

(c) Montrer que $|\lim A_p| = \lim |A_p|$ lorsque la suite (A_p) est croissante, mais que ce n'est plus vrai lorsque la suite (A_p) est décroissante (trouver un contre-exemple à l'aide de l'exercice **5.5** de la page 35).

5.11. *Approximation des parties mesurables.* Soit $A \subset \mathbb{R}^n$ une partie mesurable.

(a) Montrer qu'il existe une suite (D_p) de domaines pavables (voir définition 1.7) telle que la suite $(\mathbb{1}_{D_p})$ converge presque partout vers $\mathbb{1}_A$, puis que la formule $F_p = \cap_{q>p} \overline{D_q}$ définit une suite croissante (F_p) de compacts telle que la suite $(\mathbb{1}_{F_p})$ converge presque partout vers $\mathbb{1}_A$.

(b) Montrer que pour tout ensemble négligeable $B \subset \mathbb{R}^n$ et tout entier p, il existe un ouvert $X_p \supset B$ de mesure $|X_p| \leq 2^{-p}$ (on pourra prendre une réunion d'intérieurs de pavés).

(c) Montrer qu'il existe une suite croissante (K_p) de compacts contenus dans A telle que la suite $(\mathbb{1}_{K_p})$ converge presque partout vers $\mathbb{1}_A$, et en déduire que $|A| = \sup\{\, |K| \,;\, K \text{ compact} \subset A \,\}$.

(d) Montrer que l'on a aussi $|A| = \inf\{\, |X| \,;\, X \text{ ouvert} \supset A \,\}$ (on pourra commencer par traiter le cas où A est bornée).

6. Calcul des intégrales simples

Tout ce paragraphe est consacré au cas particulier de la dimension $n = 1$. Dans ce cas, on utilise souvent la notation $\int_a^b f(x)\,dx$ pour désigner $\int_{[a,b]} f(x)\,dx$ si $a \leq b$ et $-\int_{[b,a]} f(x)\,dx$ si $a \geq b$. Le principal avantage de cette convention réside dans la *relation de Chasles*

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$

valable pour tous réels a, b et c, ce qui résulte de l'identité $\mathbb{1}_{a,b} = \mathbb{1}_{a,c} + \mathbb{1}_{c,b}$ presque partout en notant $\mathbb{1}_{a,b} = \mathbb{1}_{[a,b]}$ ou $-\mathbb{1}_{[b,a]}$ suivant que $a \leq b$ ou $a > b$, laquelle se démontre en distinguant les différentes possibilités de positions relatives des points a, b et c. En revanche, il conviendra de se méfier de cette notation lorsque l'on voudra utiliser les propriétés de continuité, de positivité ou de croissance de l'intégrale.

Définition 6.1. *Soient $I \subset \mathbb{R}$ un intervalle ouvert, et $f : I \to \mathbb{C}$ et $F : I \to \mathbb{C}$ deux fonctions.*
(a) *On dit que f est* localement intégrable *dans I si f est intégrable sur tout intervalle compact $[a, b] \subset I$.*
(b) *Si f est localement intégrable, on dit que F est une* intégrale indéfinie *de f s'il existe $a \in I$ et $c \in \mathbb{C}$ tels que $F(x) = c + \int_a^x f(t)\,dt$.*
(c) *On dit que F est* lipschitzienne *sur I s'il existe une constante C, appelée* constante de Lipschitz *de F, telle que $|F(x) - F(y)| \leq C\,|x - y|$ pour tous x et $y \in I$.*

On a déjà observé dans les paragraphes précédents que toute fonction mesurable bornée est localement intégrable (critère du théorème 5.2) et que toute intégrale indéfinie est uniformément continue sur I (par convergence dominée).

Le théorème fondamental de l'analyse que nous montrerons plus loin repose sur le résultat difficile suivant.

Proposition 6.2. *Soient I un intervalle ouvert de $\mathbb{R}$ et $F : I \to \mathbb{C}$ une fonction lipschitzienne dans I. Alors :*
(a) *F est dérivable presque partout dans I.*
(b) *Si F est à valeurs réelles, si $A \subset I$ et si $F(A)$ est négligeable, alors $F'(t) = 0$ presque partout sur A.*

Pour démontrer cette proposition, nous établissons trois lemmes.

Lemme 6.3. *Soit $F :]a,b[\to \mathbb{R}$ une fonction continue. Alors l'ensemble*

$$X_F = \{ x \in]a,b[\,;\, \exists y > x \text{ tel que } F(y) > F(x) \}$$

est la réunion d'une suite d'intervalles disjoints $]a_j, b_j[$ vérifiant $F(a_j) \leq F(b_j)$ pour tout $j \in \mathbb{N}$.

Démonstration. Par continuité de F, l'ensemble X_F est clairement ouvert et il est donc la réunion d'une suite d'intervalles disjoints $]a_j, b_j[$. Il ne reste plus qu'à prouver que $F(a_j) \leq F(b_j)$ pour tout $j \in \mathbb{N}$, et pour cela, il sera utile de remarquer dès maintenant qu'aucun des b_j n'appartient à X_F.

Puisque F est continue, il suffit de prouver que pour tout $x \in]a_j, b_j[$, on a $F(x) \leq F(b_j)$ (on fait ensuite tendre x vers a_j). Supposons donc qu'il existe un $x \in]a_j, b_j[$ tel que $F(x) > F(b_j)$, puis définissons $y = \sup\{ z \in [x, b_j[\,;\, F(z) \geq F(x) \}$. Alors $F(y) \geq F(x)$ et $\forall z \in]y, b_j[$, $F(z) < F(y)$. Comme $y \in]a_j, b_j[\subset X_F$, il existe un $z > y$ tel que $F(z) > F(y)$, et ce point z vérifie donc $z > b_j$ et $F(z) > F(b_j)$. Il en résulte que $b_j \in X_F$, ce qui contredit la remarque faite plus haut. On peut donc conclure que $F(a_j) \leq F(b_j)$ pour tout $j \in \mathbb{N}$. ∎

Pour toute fonction continue $F :]a,b[\to \mathbb{R}$ et tout rationnel q, introduisons l'ensemble $A_d^+(q)$ (respectivement $A_d^-(q)$) formé des points $x \in]a,b[$ tels qu'il existe une suite (h_p) de réels strictement positifs tendant vers 0 vérifiant

$$\frac{F(x+h_p) - F(x)}{h_p} > q \quad \left(\text{respectivement } \frac{F(x+h_p) - F(x)}{h_p} < q \right)$$

pour tout $p \in \mathbb{N}$. De façon symétrique, on introduit les ensembles $A_g^+(q)$ et $A_g^-(q)$ définis de la même manière, mais avec des suites (h_p) de réels strictement négatifs tendant vers 0. Dans le lemme suivant, $A_*^+(q)$ désigne indifféremment $A_d^+(q)$ ou $A_g^+(q)$, et $A_*^-(q)$ indifféremment $A_d^-(q)$ ou $A_g^-(q)$.

Lemme 6.4. *Soit $F :]a,b[\to \mathbb{R}$ une fonction continue. Alors pour tout rationnel q, l'ensemble $A_*^+(q)$ (respectivement $A_*^-(q)$) peut être recouvert par une suite d'intervalles disjoints $]a_j, b_j[\subset]a,b[$ tels que $F(b_j) - F(a_j) \geq q\,(b_j - a_j)$ (respectivement $F(b_j) - F(a_j) \leq q\,(b_j - a_j)$).*

Démonstration. Introduisons la fonction continue $G(x) = F(x) - q\,x$. Alors l'ensemble $A_d^+(q)$ est contenu dans l'ensemble X_G défini au lemme 6.3 puisque les points $x + h_p$ vérifient $x + h_p > x$ et

$$G(x+h_p) - G(x) = h_p \,\frac{G(x+h_p) - G(x)}{h_p} = h_p \left(\frac{F(x+h_p) - F(x)}{h_p} - q \right) > 0\,.$$

L'ensemble $A_d^+(q)$ peut donc être recouvert par une suite d'intervalles disjoints $]a_j, b_j[$ vérifiant $G(a_j) \leq G(b_j)$ pour tout $j \in \mathbb{N}$, c'est-à-dire $F(a_j) - q\,a_j \leq F(b_j) - q\,b_j$, ce qui est la condition annoncée.

Pour traiter le cas de $A_g^+(q)$, on remarque que $-A_g^+(q) = \{ x \in]-b,-a[\,;\, -x \in A_g^+(q) \}$ est contenu dans l'ensemble X_G du lemme 6.3 pour la fonction continue $G(x) = -F(-x) - q\,x$ définie sur l'intervalle $]-b,-a[$, et on procède alors de la même manière que précédemment.

Enfin, les ensembles $A_d^-(q)$ et $A_g^-(q)$ se traitent aussi de la même manière en considérant les fonctions $G(x) = -F(x) + q\,x$ et $G(x) = F(-x) + q\,x$. ∎

Lemme 6.5. *Soient $I \subset \mathbb{R}$ un intervalle ouvert et $F : I \to \mathbb{R}$ une fonction lipschitzienne croissante. Alors F est dérivable presque partout dans I.*

Démonstration. Soient $q > r > 0$ deux rationnels, et $]a,b[\subset I$ un intervalle ouvert borné. Alors l'ensemble $]a,b[\cap A_*^-(r)$ (avec les mêmes notations qu'au lemme 6.4) peut être recouvert par une suite d'intervalles disjoints $]a_j, b_j[\subset]a,b[$ tels que

$$\sum_j \Big(F(b_j) - F(a_j) \Big) \leq r \sum_j (b_j - a_j) \leq r\,(b-a)\,.$$

Pour chacun de ces intervalles $]a_j, b_j[$, l'ensemble $]a_j, b_j[\cap A_*^+(q)$ peut être recouvert par une suite d'intervalles disjoints $]a_{jk}, b_{jk}[\subset]a_j, b_j[$ tels que

$$q \sum_k (b_{jk} - a_{jk}) \leq \sum_k \Big(F(b_{jk}) - F(a_{jk}) \Big) \leq F(b_j) - F(a_j)$$

où la dernière inégalité résulte de la croissance de F, les intervalles étant disjoints. En considérant la famille de tous les intervalles $]a_{jk}, b_{jk}[$ pour $(j, k) \in \mathbb{N}^2$, on voit que l'on a recouvert $]a, b[\cap A_*^-(r) \cap A_*^+(q)$ par une suite d'intervalles $]a_\ell, b_\ell[$ vérifiant $\sum_\ell (b_\ell - a_\ell) \leq \frac{r}{q}(b-a)$. Comme $\frac{r}{q} < 1$, il suffit de répéter cette construction sur chaque intervalle $]a_\ell, b_\ell[$ pour voir que l'ensemble $]a, b[\cap A_*^-(r) \cap A_*^+(q)$ peut être recouvert par une suite d'intervalles dont la somme des longueurs est arbitrairement petite, et comme l'intervalle $]a, b[$ est quelconque, il en résulte que $A_*^-(r) \cap A_*^+(q)$ est négligeable.

Soit maintenant $x \in I$ un point fixé. S'il existe une suite (h_p) de réels tendant vers 0 telle que la suite des quotients (positifs par croissance de F)

$$\frac{F(x+h_p) - F(x)}{h_p}$$

admette au moins deux valeurs d'adhérence distinctes, on peut trouver deux rationnels $q > r > 0$ situés strictement entre ces deux valeurs d'adhérence, et quitte à extraire des sous-suites pour lesquelles h_p reste de signe constant, on voit que $x \in A_*^-(r) \cap A_*^+(q)$.

Par conséquent, si x n'appartient pas à l'ensemble négligeable

$$A = \bigcup_{q>r>0} \Big((A_d^-(r) \cup A_g^-(r)) \cap (A_d^+(q) \cup A_g^+(q)) \Big),$$

la suite des quotients déjà cités admet au plus une valeur d'adhérence pour tout suite (h_p) de réels tendant vers 0. Comme ces suites de quotients sont des suites bornées par la constante de Lipschitz de F, elles sont convergentes, et cela prouve que F est dérivable en x.

On a ainsi démontré que F est dérivable presque partout. ∎

Démonstration de la proposition 6.2. Quitte à traiter séparément les parties réelle et imaginaire de F, on peut supposer que F est à valeurs réelles. Si C_F désigne une constante de Lipschitz de F, la fonction $G(x) = F(x) + C_F\,x$ est lipschitzienne et croissante, donc dérivable presque partout d'après le lemme 6.5, et il en est donc de même de la fonction $F(x) = G(x) - C_F\,x$ ce qui établit la propriété (a).

Si maintenant $F(A)$ est négligeable, considérons l'ensemble B formé des points $x \in A$ où F est dérivable et de dérivée $f(x) \neq 0$. Il suffit de démontrer que B est négligeable.

Si $x \in B$, il existe deux entiers positifs p et q tels que $|f(x)| \geq 2^{-p}$ et

$$(*) \qquad \forall y \in I, \qquad |y-x| \leq 2^{-q} \qquad \Rightarrow \qquad |F(y) - F(x) - f(x)(y-x)| \leq 2^{-p-1}|y-x|,$$

et il existe aussi un $r \in \mathbb{Z}$ tel que $x \in [2^{-q}r, 2^{-q}(r+1)]$. On en déduit que B est réunion dénombrable des ensembles B_{pqr} formés des points $x \in B \cap [2^{-q}r, 2^{-q}(r+1)]$ vérifiant $|f(x)| \geq 2^{-p}$ et $(*)$. On remarque alors que pour x et $y \in B_{pqr}$, on a $|y-x| \leq 2^{-q}$ et donc

$$|F(y) - F(x)| \geq |f(x)(y-x)| - |F(y) - F(x) - f(x)(y-x)| \geq 2^{-p-1}|y-x|.$$

Il en résulte que si $F(B_{pqr})$ est recouvert par une suite d'intervalles de longueurs ℓ_j, on peut recouvrir B_{pqr} par une suite d'intervalles de longueurs $2^{p+2}\ell_j$. Comme $F(B_{pqr}) \subset F(A)$ est négligeable, on en déduit que chaque B_{pqr} est négligeable, et donc que B est négligeable. ∎

Théorème 6.6 (fondamental de l'analyse). *Soit $I \subset \mathbb{R}$ un intervalle ouvert.*
(a) *Toute fonction lipschitzienne $F : I \to \mathbb{C}$ est dérivable presque partout, sa dérivée f est une fonction mesurable bornée dans I, et on a*

$$\int_a^b f(t)\,dt = F(b) - F(a)$$

pour tout intervalle compact $[a, b] \subset I$.

(b) *Réciproquement, si $f : I \to \mathbb{C}$ est une fonction mesurable bornée, toute intégrale indéfinie de f*

$$F(x) = c + \int_a^x f(t)\,dt$$

est lipschitzienne et de dérivée égale presque partout à f.

Démonstration. La fonction lipschitzienne F est dérivable presque partout d'après la proposition 6.2 (a). Sa dérivée est la limite presque partout de la suite de fonctions continues

$$f_p(t) = \frac{F(t+\frac{1}{p}) - F(t)}{\frac{1}{p}},$$

et puisque ces fonctions f_p sont uniformément bornées par la constante de Lipschitz C_F de F, leur limite f est une fonction mesurable bornée. On a $|\mathbb{1}_{[a,b]} f_p| \le C_F\,\mathbb{1}_{[a,b]}$, ce qui assure une majoration par une fonction intégrable indépendante de p, et par convergence dominée, on obtient $\int_{[a,b]} f(t)\,dt = \lim_{p\to\infty} \int_{[a,b]} f_p(t)\,dt$. En utilisant l'invariance par translation et la relation de Chasles, on a

$$\int_{[a,b]} f_p(t)\,dt = p\int_b^{b+\frac{1}{p}} F(t)\,dt - p\int_a^{a+\frac{1}{p}} F(t)\,dt\,,$$

et il ne reste plus qu'à montrer que le membre de droite tend vers $F(b) - F(a)$ quand p tend vers l'infini. Or l'identité

$$\left(p\textstyle\int_b^{b+\frac{1}{p}} F(t)\,dt\right) - F(b) = p\displaystyle\int_b^{b+\frac{1}{p}} \Big(F(t) - F(b)\Big)\,dt$$

permet d'obtenir la majoration

$$\left|\left(p\textstyle\int_b^{b+\frac{1}{p}} F(t)\,dt\right) - F(b)\right| \le p\displaystyle\int_b^{b+\frac{1}{p}} |F(t) - F(b)|\,dt \le \sup_{t\in[b,b+(1/p)]} |F(t) - F(b)|$$

qui permet de conclure que $\lim_{p\to\infty} p\int_b^{b+(1/p)} F(t)\,dt = F(b)$ grâce à la continuité de F en b. L'autre terme se traite de façon analogue.

Réciproquement, si f est mesurable et bornée par C, toute fonction de la forme $F(x) = c + \int_a^x f(t)\,dt$ vérifie

$$|F(x) - F(y)| = \left|\int_y^x f(t)\,dt\right| \le \int_{[y,x]} |f(t)|\,dt \le C\,(x-y)$$

pour tous $x > y$, ce qui prouve que F est lipschitzienne. D'après la première partie du théorème, elle est donc dérivable presque partout de dérivée mesurable bornée g, et on a

$$\int_{[a,b]} g(t)\,dt = F(b) - F(a) = \int_{[a,b]} f(t)\,dt$$

pour tout intervalle compact $[a,b]$. On conclut alors que $g - f = 0$ presque partout grâce au résultat suivant. ∎

Lemme 6.7. *Soit $f : I \to \mathbb{C}$ une fonction mesurable bornée vérifiant $\int_a^b f(t)\,dt = 0$ pour tout intervalle compact $[a,b] \subset I$. Alors $f = 0$ presque partout dans I.*

Démonstration. On a $\int \mathbb{1}_{[a,b]}(t)\, f(t)\, dt = 0$ pour tout intervalle compact $[a,b] \subset I$, et donc par combinaison linéaire, $\int f(t)\, g(t)\, dt = 0$ pour toute fonction g en escalier dans I. Pour tout intervalle compact $[a,b] \subset I$, la fonction intégrable $\mathbb{1}_{[a,b]}\, \overline{f}$ est limite en moyenne d'une suite (g_p) de fonctions en escalier, et comme f est bornée, le produit $\mathbb{1}_{[a,b]}\, |f|^2$ est donc limite en moyenne de la suite des produits $(f g_p)$, d'où $\int_{[a,b]} |f(t)|^2\, dt = 0$ par passage à la limite. On en déduit que la fonction positive $|f|^2$ s'annule presque partout sur $[a,b]$, puis que $f = 0$ presque partout dans I puisque l'intervalle $[a,b] \subset I$ est quelconque. ∎

Remarque. Le théorème 6.6 peut s'étendre de différentes façons. On peut par exemple remplacer les fonctions lipschitziennes par des fonctions localement lipschitziennes, c'est-à-dire lipschitziennes sur tout intervalle compact $[a,b] \subset I$ (avec une constante de Lipschitz dépendant de $[a,b]$) à condition de remplacer pareillement les fonctions bornées par des fonctions localement bornées. H. Lebesgue a démontré un résultat plus difficile où les fonctions bornées sont remplacées par les fonctions intégrables, et les fonctions lipschitziennes par des fonctions *absolument continues*.

Cela dit, ce théorème est utilisé le plus souvent pour des fonctions f continues et des fonctions F de classe C^1, c'est-à-dire que le calcul de l'intégrale d'une fonction continue f se réduit à la recherche d'une primitive F de f. En effet, si f est continue et si F en est une intégrale indéfinie, alors $F(x+h) - F(x) = \int_x^{x+h} f(t)\, dt$, ce qui permet d'écrire

$$\left| \frac{F(x+h) - F(x)}{h} - f(x) \right| = \left| \frac{1}{h} \int_x^{x+h} \big(f(t) - f(x)\big)\, dt \right| \leq \sup_{|t-x| \leq |h|} |f(t) - f(x)|\,,$$

puis de conclure que F est dérivable en tout point $x \in I$ et de dérivée égale à f.

Le théorème 6.6 nous permet de démontrer les deux résultats classiques suivants.

Corollaire 6.8 (Intégration par parties). *Soient $I \subset \mathbb{R}$ un intervalle ouvert, F et $G : I \to \mathbb{C}$ deux fonctions lipschitziennes de dérivées presque partout égales à f et g respectivement. Sur tout intervalle compact $[a,b] \subset I$, les fonctions $f\,G$ et $F\,g$ sont alors mesurables et bornées, et on a*

$$\int_a^b f(x)\, G(x)\, dx = F(b)\, G(b) - F(a)\, G(a) - \int_a^b F(x)\, g(x)\, dx\,.$$

Démonstration. Les fonctions F et G étant continues, elles sont mesurables et bornées sur tout intervalle compact, et donc les produits $f\,G$ et $F\,g$ aussi. De plus, la fonction $F\,G$ est lipschitzienne sur l'intervalle $[a,b]$ comme produit de deux fonctions lipschitziennes bornées. La formule d'intégration par parties s'obtient alors en appliquant le résultat du théorème 6.6 à cette fonction lipschitzienne $F\,G$ dont la dérivée presque partout vaut $f\,G + F\,g$. ∎

Corollaire 6.9 (Changement de variable). *Soient $I \subset \mathbb{R}$ un intervalle ouvert, $G : I \to \mathbb{R}$ une fonction lipschitzienne et g sa dérivée presque partout. Alors pour toute fonction f mesurable bornée sur l'intervalle $G(I)$, la fonction $(f \circ G)\, g$ est une fonction mesurable bornée sur I, et pour tout intervalle compact $[a, b] \subset I$ on a*

$$\int_{G(a)}^{G(b)} f(x)\, dx = \int_a^b \big(f \circ G(t)\big)\, g(t)\, dt\,.$$

Démonstration. Soit $F(y) = \int_c^y f(x)\, dx$ une intégrale indéfinie de f. Cette fonction est lipschitzienne d'après le théorème 6.6, la fonction $F \circ G$ est lipschitzienne comme composée de deux fonctions lipschitziennes, et la seule difficulté consiste à démontrer que cette fonction admet la fonction $(f \circ G)\, g$ pour dérivée presque partout.

Soient $A \subset G(I)$ et $B \subset I$ deux parties négligeables telles que $F' = f$ en dehors de A et $G' = g$ en dehors de B. Pour $t \notin B$ et $G(t) \notin A$, F est dérivable en $G(t)$ et G dérivable en t si bien que $F \circ G$ est dérivable en t, de dérivée $(F \circ G)' = (F' \circ G)\, G' = (f \circ G)\, g$. Sur $G^{-1}(A)$, on a $g = 0$ presque partout d'après la proposition 6.2 (b), c'est-à-dire qu'il existe une partie négligeable $C \subset I$ telle que G est de dérivée nulle sur $G^{-1}(A) \setminus C$. Alors l'inégalité

$$\left|\frac{F \circ G(t+h) - F \circ G(t)}{h}\right| \leq C_F \left|\frac{G(t+h) - G(t)}{h}\right|$$

montre, en prenant la limite pour $h \to 0$, que $F \circ G$ est dérivable en tout point $t \in G^{-1}(A) \setminus C$, de dérivée égale à $0 = \big(f \circ G(t)\big)\, g(t)$. On a donc montré que $(F \circ G)' = (f \circ G)\, g$ sur $I \setminus (B \cup C)$, c'est-à-dire presque partout dans I.

On peut alors écrire

$$\int_{G(a)}^{G(b)} f(x)\, dx = F \circ G(b) - F \circ G(a) = \int_a^b \big(f \circ G(t)\big)\, g(t)\, dt$$

par construction de F et d'après le théorème 6.6. ∎

Les résultats précédents fournissent des moyens de calcul des intégrales sur des intervalles compacts. Pour compléter ce paragraphe, il nous reste à discuter l'intégrabilité et le calcul des intégrales de fonctions localement intégrables (comme notamment les fonctions continues) sur des intervalles non compacts, ce qui nous amène à la notion d'intégrale généralisée, aussi appelée intégrale impropre ou intégrale de Cauchy.

Pour simplifier la discussion, nous nous plaçons sur l'intervalle $[a, b[$ (pour un $a \in \mathbb{R}$ et un $b \in \mathbb{R} \cup \{+\infty\}$) sur lequel nous nous donnons une fonction f localement intégrable, c'est-à-dire intégrable sur tout intervalle compact $[a, x] \subset [a, b[$ (on pourrait traiter de même le cas d'un intervalle $]b, a]$). Nous rappelons que l'intégrale de Cauchy d'une telle fonction $f : [a, b[\to \mathbb{C}$ est dite *convergente en* b si l'intégrale indéfinie $F(x) = \int_a^x f(t)\,dt$ possède une limite pour $x \to b$, et qu'elle est dite *absolument convergente en* b si l'intégrale indéfinie $G(x) = \int_a^x |f(t)|\,dt$ possède une limite pour $x \to b$.

Le lien entre l'intégrabilité de f sur $[a, b[$ et la convergence de l'intégrale de Cauchy de f en b est assuré par le résultat suivant.

Théorème 6.10. *Soit $f : [a, b[\to \mathbb{C}$ une fonction localement intégrable sur l'intervalle $[a, b[\subset \mathbb{R}$, où $b \in \mathbb{R} \cup \{+\infty\}$. Alors f est intégrable sur $[a, b[$ si et seulement si l'intégrale de Cauchy de f est absolument convergente en b. De plus, l'intégrale de Cauchy de f est alors convergente, et $\int_a^b f(t)\,dt = \lim_{x\to b} \int_a^x f(t)\,dt$.*

Démonstration. En effet, si f est intégrable sur $[a, b[$, alors $|f|$ aussi, et l'intégrale indéfinie $G(x) = \int_a^x |f(t)|\,dt$ est croissante et majorée par $\int_a^b |f(t)|\,dt$. Elle possède donc une limite (à gauche) en b.

Réciproquement, soit (b_p) une suite strictement croissante de réels tendant vers b. Si on suppose que la suite $\int_a^{b_p} |f(t)|\,dt$ est convergente, on déduit du théorème de convergence monotone (appliqué à la suite $g_p = \mathbb{1}_{[a,b_p]} |f|$) que $|f|$ est intégrable sur $[a, b[$. Alors pour toute suite (b_p) tendant vers b, la suite (f_p) définie par $f_p = \mathbb{1}_{[a,b_p]} f$ converge (partout) vers $\mathbb{1}_{[a,b[} f$, et vérifie la majoration $|f_p| \le \mathbb{1}_{[a,b[} |f|$. Par le théorème de convergence dominée, on en déduit donc que f est intégrable sur $[a, b[$, et que $\int_a^b f(t)\,dt = \lim_{p\to\infty} \int_a^{b_p} f(t)\,dt$. ∎

Remarque importante. Il résulte de ce théorème qu'il faudra toujours distinguer soigneusement les intégrales de Cauchy convergentes mais pas absolument convergentes, communément appelées *semi-convergentes*, des intégrales de Cauchy absolument convergentes qui seules sont de vraies intégrales (de Lebesgue). C'est pourquoi on évitera d'affirmer qu'une fonction f est intégrable sur $[a, b[$ si son intégrale de Cauchy n'est que semi-convergente, et on réservera la notation $\int_a^b f(t)\,dt$ aux intégrales des fonctions intégrables (intégrales de Cauchy absolument convergentes), les intégrales de Cauchy semi-convergentes restant notées $\lim_{x\to b} \int_a^x f(t)\,dt$.

Une autre conséquence du théorème, c'est que la convergence absolue de l'intégrale de Cauchy est un critère d'intégrabilité. Par exemple, si f est continue sur un intervalle I, on voit qu'elle y est intégrable si et seulement si $|f|$ possède une primitive bornée dans I puisque les primitives de fonctions positives sont croissantes. Si on ne connaît pas de primitive de $|f|$, on peut utiliser les critères de comparaison déjà étudiés en premier cycle, et qui découlent directement du théorème 5.2 : si $|f|$ est majoré par une fonction intégrable (une fonction continue possédant une primitive bornée par exemple), alors f est intégrable. La conséquence suivante du critère de comparaison est traditionnelle : lorsque l'on considère deux fonctions positives f et $g : [a, b[\to \mathbb{C}$ qui sont équivalentes au voisinage de b, il existe un voisinage de b dans lequel $f \le 2\,g$ et $g \le 2\,f$, ce qui permet de conclure que f est intégrable sur $[a, b[$ si et seulement si g est intégrable sur $[a, b[$.

Terminons en rappelant que ces critères de comparaison sont à réserver au cas des fonctions positives. On ne dispose pas de critère de ce type pour la convergence des intégrales de Cauchy semi-convergentes.

Exemple. Montrons que la fonction $g : x \mapsto e^{-x^2/2}$ est intégrable sur $\mathbb{R}$. Comme cette fonction n'a pas de primitive connue, on ne peut pas démontrer directement la convergence absolue de l'intégrale de Cauchy, mais il suffit de majorer g par une fonction intégrable. De l'inégalité $(|x| - 1)^2 \geq 0$ on tire les majorations $-x^2 \leq -2|x| + 1$ puis $|g(x)| = g(x) \leq e^{-|x|+(1/2)} = f(x)$. Sur $\mathbb{R}_+$, la fonction $|f| = f$ admet pour primitive la fonction bornée $F(x) = -e^{-x+(1/2)}$, tandis que sur $\mathbb{R}_-$, elle admet pour primitive la fonction bornée $F(x) = e^{x+(1/2)}$. L'intégrale de Cauchy de f est donc (absolument) convergente en $+\infty$ et en $-\infty$, ce qui permet d'affirmer que la fonction f est intégrable sur $\mathbb{R}$ tout entier, et qu'il en est de même de la fonction g.

Exercices

6.1. Montrer que la fonction $f(x) = \dfrac{\ln x}{1+x^2}$ est intégrable sur $]0, \infty[$, et calculer $\int_0^\infty f(x)\,dx$ en écrivant $\int_0^\infty = \int_0^1 + \int_1^\infty$.

6.2. On considère les fonctions $f(x) = \ln(|\sin x|)$, $g(x) = \ln(|\cos x|)$ et $h(x) = \ln(|\sin 2x|)$.

(a) Montrer que f et g sont périodiques de période π, et que h est périodique de période $\pi/2$.

(b) Montrer que f, g et h sont intégrables sur $[0, \pi]$ et qu'elles ont même intégrale, puis calculer la valeur de cette intégrale.

6.3. Soient $f :]0, \infty[\to \mathbb{R}$ une fonction intégrable et $a > 0$ un paramètre. Montrer que les fonctions $y \mapsto \big(1 + (a^2/y^2)\big)\, f\big(y - (a^2/y)\big)$ et $x \mapsto f(|x - (a^2/x)|)$ sont intégrables respectivement sur $]a, \infty[$ et sur $]0, \infty[$, et de mêmes intégrales que f.

6.4. Dans tout l'exercice, on considère des fonctions définies sur $]0, 1]$, mais qui se prolongent par continuité à l'intervalle $[0, 1]$.

(a) Pour tous entiers $p \geq 1$ et $q \geq 0$, montrer que les fonctions $x \mapsto x^p(\ln x)^q$ sont intégrables sur $[0, 1]$ et calculer les nombres $a_{p,q} = \int_{[0,1]} x^p(\ln x)^q\,dx$ (pour le calcul des intégrales, obtenir une relation de récurrence à l'aide d'une intégration par parties que l'on justifiera soigneusement).

(b) Montrer que la fonction $x \mapsto f(x) = x^x$ est intégrable sur $[0, 1]$.

(c) En écrivant la fonction f comme une série de fonctions du type des $x^p(\ln x)^q$, exprimer le nombre $\int_{[0,1]} f(x)\,dx$ comme la somme d'une série numérique, et en donner une valeur approchée à 10^{-5} près.

6.5. Montrer que pour toute fonction f lipschitzienne sur $[0, \pi]$, $\lim_{p\to\infty} \int_0^\pi f(x)\,e^{ipx}\,dx = 0$. Déterminer ensuite deux réels a et b tels que $\int_0^\pi (a\,x + b\,x^2)\cos(p\,x)\,dx = (1/p^2)$ pour tout $p \neq 0$, et en déduire la valeur de $\sum_{p=1}^\infty (1/p^2)$.

6.6. Soit $f_p(x) = \left(1 - \frac{x}{p}\right)^p \mathbb{1}_{[0,p]}(x)$.

(a) Pour $x > 0$ fixé, montrer que la fonction $\varphi(y) = y \ln\left(1 - \frac{x}{y}\right)$ est croissante sur $]x, +\infty[$, et en déduire que la suite $(f_p)_{p\in\mathbb{N}}$ est une suite croissante de fonctions intégrables.

(b) Déterminer $f(x) = \lim_{p\to\infty} f_p(x)$ et montrer que f est intégrable sur $[0, \infty[$.

(c) Montrer la formule $\int_{[0,\infty[} e^{-x} \sin x\,dx = \lim_{p\to\infty} \int_{[0,p]} \left(1 - \frac{x}{p}\right)^p \sin x\,dx$.

6.7. On note ici $|x| = \sqrt{x_1^2 + \ldots + x_n^2}$ la norme euclidienne d'un vecteur $x \in \mathbb{R}^n$. Soient $a > 0$, et f_a et g_a les fonctions définies sur $\mathbb{R}^n$ par $f_a(x) = (1 + |x|)^{-a}$, et $g_a(x) = |x|^{-a}\mathbb{1}_B(x)$ où $B = \{\, x \in \mathbb{R}^n \,;\, 0 < |x| \leq 1 \,\}$.

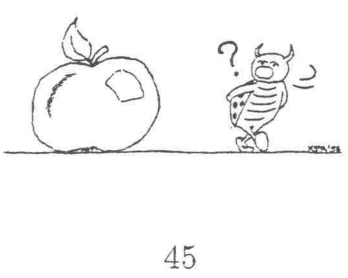
(a) Montrer que si $a > n$, alors f_a est intégrable sur $\mathbb{R}^n$ (on commencera par montrer que la fonction $t \mapsto (1 + t)^{-a}$ est intégrable sur $\mathbb{R}_+$ lorsque $a > 1$, puis on utilisera le résultat de l'exercice **3.8** de la page 26).

(b) Montrer que si $a < n$, alors g_a est intégrable sur $\mathbb{R}^n$ (on commencera par montrer que la fonction $t \mapsto t^{-a}$ est intégrable sur $]0, 1]$ lorsque $a < 1$).

Isaac Newton
(1642 - 1727)

Gottfried Leibniz
(1646 - 1716)

Les principaux créateurs du "calcul infinitésimal" qui fit apparaître l'intégration et la dérivation comme deux opérations inverses l'une de l'autre.



Augustin-Louis Cauchy
(1789 - 1857)

Bernhard Riemann
(1826 - 1866)

Les principaux créateurs des premières théories rigoureuses de l'intégration fondées sur des procédés de calcul : primitives et "sommes de Riemann".

7. Intégrales dépendant de paramètres

Soit $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{C}$. Le but de ce paragraphe est d'étudier les propriétés de continuité, de dérivabilité et d'intégrabilité de la fonction $g(y) = \int f(x,y)\,dx$ définie en tout point $y \in \mathbb{R}^m$ tel que $x \mapsto f(x,y)$ soit intégrable.

Théorème 7.1 (de continuité). *Soient Y un ouvert de $\mathbb{R}^m$, $y_0 \in Y$ et $f : \mathbb{R}^n \times Y \to \mathbb{C}$ une fonction vérifiant :*
(i) *Pour tout $y \in Y$, la fonction $x \mapsto f(x,y)$ est intégrable.*
(ii) *Pour tout $x \in \mathbb{R}^n$, la fonction $y \mapsto f(x,y)$ est continue en y_0.*
(iii) *Il existe une fonction $F : \mathbb{R}^n \to \mathbb{R}_+$ intégrable telle que $|f(x,y)| \leq F(x)$ pour tout $(x,y) \in \mathbb{R}^n \times Y$.*
Alors la fonction $g : Y \to \mathbb{C}$ définie par $g(y) = \int f(x,y)\,dx$ est continue en y_0.

Théorème 7.2 (de dérivation sous l'intégrale). *Soient Y un ouvert de $\mathbb{R}^m$, $j \leq m$ et $f : \mathbb{R}^n \times Y \to \mathbb{C}$ une fonction vérifiant :*
(i) *Pour tout $y \in Y$, la fonction $x \mapsto f(x,y)$ est intégrable.*
(ii) *Pour tout $x \in \mathbb{R}^n$, la fonction $y \mapsto f(x,y)$ est dérivable par rapport à y_j dans Y.*
(iii) *Il existe une fonction $F : \mathbb{R}^n \to \mathbb{R}_+$ intégrable telle que $|(\partial f/\partial y_j)(x,y)| \leq F(x)$ pour tout $(x,y) \in \mathbb{R}^n \times Y$.*
Alors la fonction $g : Y \to \mathbb{C}$ définie par $g(y) = \int f(x,y)\,dx$ est dérivable par rapport à y_j dans Y, et on a

$$\frac{\partial g}{\partial y_j}(y) = \int \frac{\partial f}{\partial y_j}(x,y)\,dx\,.$$

Démonstration. Pour démontrer le théorème 7.1, on utilisera qu'une fonction $g : \mathbb{R}^m \to \mathbb{C}$ est continue en y_0 si et seulement si $g(y_p)$ tend vers $g(y_0)$ pour toute suite (y_p) de points de Y tendant vers y_0. Si donc (y_p) est une telle suite, la suite de fonctions (f_p) définie par $f_p(x) = f(x,y_p)$ converge partout vers $f_0 : x \mapsto f(x,y_0)$ d'après l'hypothèse (ii), et vérifie la condition de domination $|f_p| \leq F$ d'après l'hypothèse (iii). On déduit alors du théorème de convergence dominée que $g(y_p)$ tend vers $g(y_0)$, ce qui prouve le théorème.

Pour démontrer le théorème 7.2, il suffit de le faire dans le cas où $m = 1$, les autres variables y n'étant que des paramètres ne jouant aucun rôle. La formule à démontrer s'écrit alors en y_0

$$\lim_{y \to y_0} \frac{g(y) - g(y_0)}{y - y_0} = \int \frac{\partial f}{\partial y}(x,y_0)\,dx\,,$$

et comme dans le théorème précédent, il suffit de remplacer la limite pour $y \to y_0$ par la limite quand $p \to \infty$ pour toute suite (y_p) tendant vers y_0. Or, si (y_p) est une telle suite, on a

$$\frac{g(y_p) - g(y_0)}{y_p - y_0} = \int h_p(x)\, dx \qquad \text{où} \quad h_p(x) = \frac{f(x, y_p) - f(x, y_0)}{y_p - y_0}\,.$$

Pour tout x fixé, la suite numérique $(h_p(x))$ converge vers $h(x) = (\partial f/\partial y)(x, y_0)$ avec la majoration

$$|h_p(x)| \leq \sup_{t \in [0,1]} \left| \frac{\partial f}{\partial y}\Big(x\,,\ t\, y_0 + (1-t)\, y_p\Big) \right| \leq F(x)$$

d'après les hypothèses (ii) et (iii), et l'inégalité des accroissements finis. On a donc le résultat grâce encore au théorème de convergence dominée. ∎

Remarque. Dans le théorème de convergence dominée, on peut se contenter de ne vérifier les hypothèses que presque partout. On pourrait donc affaiblir ici les hypothèses des théorèmes 7.1 et 7.2. Ainsi, l'étudiant pourra démontrer à titre d'exercice que les hypothèses (ii) et (iii) du théorème 7.1 (resp. du théorème 7.2) peuvent n'être vérifiées que pour $x \notin A(y)$ où $A(y) \subset \mathbb{R}^n$ est un ensemble négligeable pouvant dépendre de y (resp. pour $x \notin A$ où $A \subset \mathbb{R}^n$ est un ensemble négligeable indépendant de y). On notera aussi que la continuité n'est requise qu'en un point, tandis que la dérivabilité est requise dans tout un ouvert (pour pouvoir bénéficier des accroissements finis).

Exemple : la transformée de Fourier d'une fonction intégrable. Pour toute fonction intégrable f et tout $\xi \in \mathbb{R}^n$, on pose

$$\widehat{f}(\xi) = \int e^{-ix\cdot\xi} f(x)\, dx$$

en notant $x \cdot \xi = x_1\, \xi_1 + \ldots + x_n\, \xi_n$ le produit scalaire euclidien des vecteurs x et ξ. Cette formule définit une fonction $\widehat{f}$ sur $\mathbb{R}^n$ appelée *transformée de Fourier de* f, et nous allons montrer à l'aide des théorèmes 7.1 et 7.2 que c'est une fonction continue bornée, et qu'elle est de classe C^∞ si la fonction intégrable f s'annule en dehors d'un pavé.

Pour tout $\xi \in \mathbb{R}^n$ fixé, la fonction $f_\xi : x \mapsto e^{-ix\cdot\xi} f(x)$ est intégrable puisque c'est le produit d'une fonction intégrable par une fonction mesurable car continue, et qu'elle est majorée en module par la fonction intégrable $|f|$ (critère du théorème 5.2). Par continuité de l'intégrale, on obtient immédiatement la majoration $|\widehat{f}(\xi)| \leq \int |f(x)|\, dx$ qui prouve que $\widehat{f}$ est bornée (on peut aussi démontrer que la fonction $\widehat{f}$ tend vers 0 à l'infini, mais c'est un peu plus délicat).

Par ailleurs, à x fixé, la fonction $\xi \mapsto e^{-ix\cdot\xi} f(x)$ est de classe C^∞, et comme $|f_\xi|$ est majorée par la fonction intégrable $|f|$ indépendante de ξ, il résulte du théorème 7.1 que la fonction $\widehat{f}$ est continue. Lorsque la fonction f s'annule en dehors d'un pavé P, la fonction $x \mapsto x_j\, f(x)$ est aussi intégrable car $x \mapsto x_j$ est mesurable, et bornée sur le pavé P (critère du théorème 5.2). Alors l'identité

$$\left|\frac{\partial}{\partial \xi_j}\Big(e^{-ix\cdot\xi}\, f(x)\Big)\right| = |-i\, x_j\, e^{-ix\cdot\xi}\, f(x)| = |x_j\, f(x)|$$

permet de conclure, grâce au théorème 7.2, que $\widehat{f}$ est dérivable par rapport à chaque variable ξ_j. On obtient aussi dans ce cas que $(\partial\widehat{f}/\partial\xi_j)$ est la transformée de Fourier de la fonction intégrable $x \mapsto -i\, x_j\, f(x)$ qui s'annule en dehors de P, ce qui montre que $\widehat{f}$ est de classe C^1, puis par récurrence sur k, que la transformée de Fourier de toute fonction intégrable s'annulant hors de P est de classe C^k (pour tout k).

Théorème 7.3 (de Fubini, ou de l'intégrale itérée). *Soit f une fonction définie sur $\mathbb{R}^n \times \mathbb{R}^m$. Alors on a la formule*(1)

$$\int \Big(\int f(x,y)\, dx\Big)\, dy = \int f(z)\, dz = \int \Big(\int f(x,y)\, dy\Big)\, dx$$

dans les deux cas suivants :
(a) *Lorsque f est intégrable sur $\mathbb{R}^n \times \mathbb{R}^m$. Dans ce cas, pour presque tout $y \in \mathbb{R}^m$ (resp. pour presque tout $x \in \mathbb{R}^n$) la fonction $x \mapsto f(x,y)$ est intégrable sur $\mathbb{R}^n$ (resp. la fonction $y \mapsto f(x,y)$ est intégrable sur $\mathbb{R}^m$), et les fonctions (définies presque partout)*

$$g(y) = \int f(x,y)\, dx \qquad \text{et} \qquad h(x) = \int f(x,y)\, dy$$

sont intégrables respectivement sur $\mathbb{R}^m$ et sur $\mathbb{R}^n$, ce qui donne un sens à la formule annoncée.
(b) *Lorsque f est mesurable positive sur $\mathbb{R}^n \times \mathbb{R}^m$, auquel cas les fonctions g et h définies ci-dessus sont mesurables positives respectivement sur $\mathbb{R}^m$ et sur $\mathbb{R}^n$, ce qui donne encore un sens à la formule annoncée.*

Pour démontrer ce théorème, on commence par analyser finement la structure d'un ensemble négligeable dans $\mathbb{R}^n \times \mathbb{R}^m$ à l'aide du théorème de convergence monotone.

Lemme 7.4. *Soit $A \subset \mathbb{R}^n \times \mathbb{R}^m$. Pour tout $y \in \mathbb{R}^m$ on note $A(y)$ la section $\{x \in \mathbb{R}^n\,;\ (x,y) \in A\}$. Si A est négligeable dans $\mathbb{R}^n \times \mathbb{R}^m$, alors il existe un ensemble négligeable B dans $\mathbb{R}^m$ tel que pour tout $y \notin B$, $A(y)$ est négligeable dans $\mathbb{R}^n$.*

Démonstration. Pour tout entier $p \geq 1$, il existe une suite $(P_{p,q})_{q\in\mathbb{N}}$ de pavés de $\mathbb{R}^n \times \mathbb{R}^m$ recouvrant A et vérifiant $\sum_{q\in\mathbb{N}} |P_{p,q}| \leq 2^{-p}$. Comme l'ensemble des $(p,q) \in \mathbb{N}^2$ est dénombrable,

(1) On utilise aussi la notation $\int f(x,y)\, dx\, dy$ pour désigner l'intégrale $\int f(z)\, dz$ de la fonction f.

on peut renuméroter tous ces pavés en une unique suite $(P_q)_{q\in\mathbb{N}}$ telle que $\sum_{q\in\mathbb{N}} |P_q| \leq 1$ et que $A \subset \cup_{q\geq p} P_q$ pour tout entier p. Si on pose $P_q(y) = \{x \in \mathbb{R}^n ; (x,y) \in P_q\}$, on a donc aussi $A(y) \subset \cup_{q\geq p} P_q(y)$ pour tout entier p.

Pour tout $y \in \mathbb{R}^m$, l'ensemble $P_q(y)$ est un pavé de $\mathbb{R}^n$, sa mesure $|P_q(y)|$ est la fonction indicatrice d'un pavé de $\mathbb{R}^m$, et on a $\int |P_q(y)|\,dy = |P_q|$. Par conséquent, la suite (g_p) de fonctions en escalier dans $\mathbb{R}^m$ définie par $g_p(y) = \sum_{q\leq p} |P_q(y)|$ est croissante, et on a $\int g_p(y)\,dy = \sum_{q\leq p} |P_q| \leq 1$. Il résulte donc du théorème de convergence monotone que cette suite converge presque partout vers une fonction intégrable g, ce qui implique que pour un ensemble négligeable $B \subset \mathbb{R}^m$ et pour tout $y \notin B$, la suite $g_p(y) = \sum_{q\leq p} |P_q(y)|$ est convergente. Ainsi, pour tout $y \notin B$, la série $\sum_{q=0}^{\infty} |P_q(y)|$ est convergente, et il en résulte que pour tout $\varepsilon > 0$, on peut trouver un entier p (dépendant de y) tel que $\sum_{q\geq p} |P_q(y)| \leq \varepsilon$. Cela achève la démonstration puisque $A(y) \subset \cup_{q\geq p} P_q(y)$ pour tout entier p. ∎

Démonstration du théorème 7.3. Commençons par remarquer que le théorème est vrai pour les fonctions en escalier dans $\mathbb{R}^n \times \mathbb{R}^m$. En effet, si $f = \mathbb{1}_P$ est la fonction indicatrice d'un pavé P de $\mathbb{R}^n \times \mathbb{R}^m$, c'est en réalité un produit tensoriel $f(x,y) = \mathbb{1}_Q(x)\,\mathbb{1}_R(y)$, ce qui permet d'écrire

$$g(y) = \int f(x,y)\,dx = |Q|\,\mathbb{1}_R(y) \qquad \Rightarrow \qquad \int g(y)\,dy = |Q|\,|R| = |P|\,.$$

Par combinaisons linéaires, on en déduit que si f est une fonction en escalier dans $\mathbb{R}^n \times \mathbb{R}^m$, la fonction $g : y \mapsto \int f(x,y)\,dx$ est en escalier dans $\mathbb{R}^m$ avec $\int g(y)\,dy = \int f(z)\,dz$.

Soit maintenant une fonction f intégrable sur $\mathbb{R}^n \times \mathbb{R}^m$ (cas (a)). Comme les variables x et y jouent des rôles symétriques, il nous suffira de démontrer que la fonction $x \mapsto f(x,y)$ est intégrable pour presque tout $y \in \mathbb{R}^m$ et que la fonction $g : y \mapsto \int f(x,y)\,dx$ est intégrable et d'intégrale $\int g(y)\,dy = \int f(z)\,dz$.

Comme f est intégrable, elle est limite presque partout d'une suite de Cauchy (f_p) de fonctions en escalier, et il résulte de la remarque du début que chaque terme f_p de cette suite vérifie les conclusions du théorème de Fubini, c'est-à-dire que

$$\int g_p(y)\,dy = \int f_p(z)\,dz \qquad \text{où} \qquad g_p(y) = \int f_p(x,y)\,dx\,.$$

Il ne reste donc plus qu'à passer à la limite dans ces expressions, en justifiant en même temps l'intégrabilité des fonctions considérées.

Pour faciliter le raisonnement, nous remplaçons la suite (f_p) par une sous-suite vérifiant

$$\int |f_{p+1}(z) - f_p(z)|\,dz \leq 2^{-2p}$$

comme dans la démonstration du lemme d'Égorov. Si nous notons A_p l'ensemble des $y \in \mathbb{R}^m$ vérifiant $\int |f_{p+1}(x,y) - f_p(x,y)|\,dx \geq 2^{-p}$, alors $\mathbb{1}_{A_p}$ est une fonction en escalier dans $\mathbb{R}^m$ et on a

$$2^{-p}\,\mathbb{1}_{A_p}(y) \leq \int |f_{p+1}(x,y) - f_p(x,y)|\,dx \qquad \Rightarrow \qquad \int \mathbb{1}_{A_p}(y)\,dy \leq 2^p \int |f_{p+1}(z) - f_p(z)|\,dz \leq 2^{-p}$$

d'après la remarque du début puisque ce sont des fonctions en escalier. Si on note $B_q = \cup_{p>q} A_p$ puis $B = \cap_{q\in\mathbb{N}} B_q$, on voit par convergence monotone que $\mathbb{1}_{B_q}$ et $\mathbb{1}_B$ sont intégrables et que ces fonctions vérifient $\int \mathbb{1}_{B_q}(y)\,dy \leq 2^{-q}$ puis $\int \mathbb{1}_B(y)\,dy = 0$. On a donc $\mathbb{1}_B = 0$ presque partout d'après la proposition 3.5, c'est-à-dire que B est négligeable. Enfin, pour $y \notin B$, il existe un entier q tel que $y \notin B_q$, ce qui montre que la suite $(x \mapsto f_p(x,y))$ est une suite de Cauchy de fonctions en escalier puisque pour $r \geq p > q$, $\int |f_p(x,y) - f_r(x,y)|\,dx \leq 2^{-p}$.

Comme par ailleurs il résulte du lemme 7.4 qu'il existe une partie $B' \subset \mathbb{R}^m$ négligeable telle que pour tout $y \notin B'$ fixé, la suite de fonctions en escalier $x \mapsto f_p(x,y)$ converge presque partout vers la fonction $x \mapsto f(x,y)$, on peut conclure que pour $y \notin B \cup B'$, cette dernière fonction est intégrable et que son intégrale est la limite de la suite $\int f_p(x,y)\,dx$, c'est-à-dire que la suite de fonctions en escalier (g_p) converge presque partout vers la fonction $g : y \mapsto \int f(x,y)\,dx$. Enfin cette suite (g_p) est aussi de

Cauchy puisque, ces fonctions étant en escalier, $\int |g_p(y) - g_q(y)|\, dy \leq \int |f_p(z) - f_q(z)|\, dz$, et on en déduit que g est intégrable et d'intégrale égale à

$$\int g(y)\, dy = \lim_{p\to\infty} \int g_p(y)\, dy = \lim_{p\to\infty} \int f_p(z)\, dz = \int f(z)\, dz\ ,$$

ce qui achève la démonstration du théorème dans le cas (a).

Dans le cas d'une fonction mesurable positive f, considérons les suites (h_p) et (f_p) de fonctions intégrables définies par $h_p = p\, \mathbb{1}_{[-p,p]^{n+m}}$ et $f_p = \inf(f, h_p)$. Comme les fonctions f_p sont intégrables (par le critère du théorème 5.2), elles vérifient les conclusions du théorème, cas (a). Comme la suite (f_p) est croissante et qu'elle converge partout vers f, il résulte du théorème 5.4 (c) que les fonctions $g_p(y) = \int f_p(x,y)\, dx$ forment une suite croissante de fonctions intégrables positives qui converge partout vers $g(y) = \int f(x,y)\, dx$, ce qui permet de conclure que g est mesurable positive et que

$$\int g(y)\, dy = \lim_{p\to\infty} \int g_p(y)\, dy = \lim_{p\to\infty} \int f_p(z)\, dz = \int f(z)\, dz$$

grâce encore au théorème 5.4 (c). ∎

Remarque. On peut trouver des fonctions $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{C}$ ni intégrables ni mesurables positives telles que les intégrales itérées $\int \big(\int f(x,y)\, dx\big)\, dy$ et $\int \big(\int f(x,y)\, dy\big)\, dx$ soient toutes deux bien définies mais que leurs valeurs soient différentes. C'est pourquoi on prendra bien garde de vérifier l'intégrabilité de la fonction f avant d'intervertir l'ordre des intégrations, et pour établir cette intégrabilité, le plus commode consiste à montrer que $\int |f(z)|\, dz < \infty$ (critère du théorème 5.2) en utilisant la formule du théorème 7.3, cas (b).

Exemple : le produit de convolution de deux fonctions intégrables. Soient f et g deux fonctions intégrables sur $\mathbb{R}^n$. Nous allons montrer ici que la fonction $y \mapsto f(x-y)\, g(y)$ est intégrable pour presque tout x, et que la formule

$$f * g(x) = \int f(x-y)\, g(y)\, dy$$

définit (presque partout) une fonction $f * g$ intégrable appelée *produit de convolution de f et g* et dont la transformée de Fourier vaut $\widehat{f * g}(\xi) = \widehat{f}(\xi)\, \widehat{g}(\xi)$.

Pour établir cela, nous commençons par établir que si les fonctions f et g sont intégrables sur $\mathbb{R}^n$, alors la fonction $h : (x,y) \mapsto f(x-y)\, g(y)$ est intégrable sur $\mathbb{R}^n \times \mathbb{R}^n$. En effet, le théorème 7.3 appliqué à la fonction mesurable positive $|h|$ permet d'écrire avec $z = (x,y)$

$$\int |h(z)|\, dz = \int \Big(\int |f(x-y)\, g(y)|\, dx\Big)\, dy = \Big(\int |f(x)|\, dx\Big) \Big(\int |g(y)|\, dy\Big) < \infty\ ,$$

d'où l'intégrabilité de h.

Le théorème 7.3, cas (a), nous dit alors que pour presque tout x, la fonction $y \mapsto h(x,y)$ est intégrable, et que son intégrale $f * g(x) = \int f(x-y)\, g(y)\, dy$ est une fonction intégrable de x. De plus, la fonction $(x,y) \mapsto e^{-ix\cdot\xi} f(x-y)\, g(y)$ est aussi intégrable comme produit de la fonction intégrable h par une fonction mesurable bornée, et donc, à nouveau par le théorème de Fubini, nous pouvons écrire

$$\widehat{f*g}(\xi) = \int e^{-ix\cdot\xi}\,(f*g(x))\,dx = \int \Big(\int e^{-ix\cdot\xi} f(x-y)\,g(y)\,dy\Big)\,dx =$$
$$= \int \Big(\int e^{-ix\cdot\xi} f(x-y)\,g(y)\,dx\Big)\,dy = \int e^{-iy\cdot\xi}\,\widehat{f}(\xi)\,g(y)\,dy = \widehat{f}(\xi)\,\widehat{g}(\xi)\,.$$

Nous donnons ci-dessous une autre application du théorème de Fubini qui nous servira ultérieurement à démontrer le théorème de changement de variables.

On rappelle que l'on appelle *transformation affine* de $\mathbb{R}^n$ toute application $\phi : \mathbb{R}^n \to \mathbb{R}^n$ de la forme $\phi(y) = J_\phi y + x_0$ où J_ϕ est une matrice réelle de taille $n\times n$, et $x_0 \in \mathbb{R}^n$. On appelle ensuite *parallélépipède* toute image du cube unité $I^n = [0,1]^n$ par une transformation affine. Comme toute transformation affine est continue, tout parallélépipède P est compact, et sa mesure est bien définie puisque $\mathbb{1}_P$ est intégrable. On remarquera aussi que l'image d'un parallélépipède par une transformation affine est encore un parallélépipède et que tout pavé est un parallélépipède. Enfin, une application immédiate du théorème de Fubini montre que tout parallélépipède contenu dans un hyperplan est de mesure nulle.

Dans l'énoncé suivant, on pose $[\![J]\!] = |\text{dét}\, J|$ pour toute matrice réelle de taille $n\times n$.

Proposition 7.5. *Pour toute transformation affine ϕ et tout parallélépipède P, on a la relation $|\phi(P)| = [\![J_\phi]\!]\,|P|$.*

Démonstration. Il suffit de montrer que l'on a $|\phi(I^n)| = [\![J_\phi]\!]$ pour toute transformation affine ϕ. En effet, si $P = \psi(I^n)$, la relation précédente permet alors d'écrire

$$|\phi(P)| = |\phi\circ\psi(I^n)| = [\![J_{\phi\circ\psi}]\!] = [\![J_\phi]\!]\,[\![J_\psi]\!] = [\![J_\phi]\!]\,|\psi(I^n)| = [\![J_\phi]\!]\,|P|\,.$$

De plus, on peut supposer que $[\![J_\phi]\!] \neq 0$ puisque autrement $\phi(I^n)$ est contenu dans un hyperplan, d'où $|\phi(I^n)| = 0$.

Si la transformation ϕ est donnée par les formules $\phi_j(y) = b_j + \sum_{k=1}^n a_{jk}\,y_k$, on peut supposer que $a_{11}\neq 0$ quitte à changer l'ordre des y_k, ce qui revient à permuter deux coordonnées dans I^n, et on a alors $[\![J_\phi]\!] = [\![J_\psi]\!]$ pour la transformation ψ définie par les formules $\psi_1(y) = \phi_1(y)$, et pour $j>1$, $\psi_j(y) = \phi_j(y) - (a_{j1}/a_{11})\,\phi_1(y)$ par propriété des déterminants. Le parallélépipède $Q = \psi(I^n)$ est relié au parallélépipède P par l'identité $P_*(x_1) = Q_*(x_1) + \tau\,x_1$ pour tout $x_1\in\mathbb{R}$, en notant

$$P_*(x_1) = \{\,x_*\in\mathbb{R}^{n-1}\,;\,(x_1,x_*)\in P\,\}\,,\qquad Q_*(x_1) = \{\,x_*\in\mathbb{R}^{n-1}\,;\,(x_1,x_*)\in Q\,\}$$

et $\tau = -(a_{21},\dots,a_{n1})/a_{11}\in\mathbb{R}^{n-1}$. Comme l'intégrale est invariante par translation, on a par Fubini

$$|P| = \int \mathbb{1}_P(x)\,dx = \int |P_*(x_1)|\,dx_1 = \int |Q_*(x_1)|\,dx_1 = \int \mathbb{1}_Q(x)\,dx = |Q|\,,$$

et on est ainsi ramené à étudier des transformations affines du type de ψ, c'est-à-dire telles que $\psi_* = (\psi_2,\dots,\psi_n)$ ne dépende que des variables $y_* = (y_2,\dots,y_n)$.

Pour une telle transformation, le déterminant de J_ψ s'écrit aussi $[\![J_\psi]\!] = |a_{11}|\,[\![J_{\psi_*}]\!]$, et d'autre part, pour tout $x_*\in Q_* = \psi_*(I^{n-1})$, il existe un unique $y_*\in I^{n-1}$ tel que $x_* = \psi_*(y_*)$, ce qui permet de conclure que la section $Q_1(x_*) = \{\,x_1\in\mathbb{R}\,;\,(x_1,x_*)\in Q\,\}$ est un intervalle de mesure (c'est-à-dire de longueur) $|a_{11}|$. On peut alors utiliser à nouveau Fubini pour écrire

$$|Q| = \int \mathbb{1}_Q(x)\,dx = \int |Q_1(x_*)|\,dx_* = |a_{11}|\,|Q_*|\,,$$

et la proposition en découle par récurrence sur la dimension n. ∎

Exercices

7.1. Pour $y > 0$, on pose $g(y) = \int_{[0,\infty[} \frac{\sin x}{x} e^{-yx}\, dx$.

(a) Montrer que la fonction g est bien définie sur $]0, \infty[$.

(b) Montrer que la fonction g est de classe C^1 et calculer sa dérivée.

(c) Calculer $\lim_{y\to\infty} g(y)$, et en déduire une autre expression de g.

(d) Montrer que la fonction $\sin x/x$ est intégrable sur tout intervalle de la forme $[0, A]$. Est-elle intégrable sur $[0, \infty[$?

(e) À l'aide d'une intégration par parties, montrer que pour tout $y \geq 0$ et tout $A \geq \pi$

$$\int_{[0,A]} \frac{\sin x}{x} e^{-yx}\, dx = \int_{[0,\pi]} \frac{\sin x}{x} e^{-yx}\, dx + \frac{F(A,y)}{A} - \frac{F(\pi,y)}{\pi} + \int_{[\pi,A]} \frac{F(x,y)}{x^2}\, dx$$

pour une fonction F continue bornée sur $\mathbb{R}_+ \times \mathbb{R}_+$, puis en déduire que la fonction g se prolonge par continuité en 0 en posant $g(0) = \lim_{A\to\infty} \int_{[0,A]} \frac{\sin x}{x}\, dx$, et enfin établir que cette limite vaut $\pi/2$.

7.2. On note I l'ensemble des y pour lesquels la fonction $x \mapsto e^{-x} x^{y-1}$ est intégrable sur $]0, \infty[$, puis pour tout $y \in I$, on pose $\Gamma(y) = \int_{]0,\infty[} e^{-x} x^{y-1}\, dx$.

(a) Déterminer le domaine de définition I de la fonction Γ et montrer que Γ est de classe C^∞ sur I.

(b) Par une intégration par parties, montrer que pour tout $y \in I$, $\Gamma(y+1) = y\,\Gamma(y)$.

(c) Calculer $\Gamma(1)$, et en déduire $\Gamma(p)$ pour tout $p \in \mathbb{N}$.

7.3. Dans cet exercice, on note $D(z_0, R) = \{ z \in \mathbb{C} ; |z - z_0| < R \}$ le disque de centre z_0 et de rayon R dans $\mathbb{C}$.

(a) Soit $f(z) = \sum_{k=0}^{\infty} a_k z^k$ une série entière qui converge dans le disque $D(0, R)$. Montrer que pour tout $r < R$ et tout $k \in \mathbb{N}$, on a

$$a_k r^k = \frac{1}{2\pi} \int_{[0,2\pi]} f(r\, e^{i\theta})\, e^{-ik\theta}\, d\theta\,.$$

puis en déduire que pour tout $N \in \mathbb{N}$,

$$\sup_{|z|\leq r} \left| \sum_{k=0}^{N} a_k z^k \right| \leq \left(\frac{R}{R-r} \right) \sup_{|z|<R} |f(z)|\,.$$

(on commencera par le démontrer en remplaçant R par $R' \in]r, R[$).

(b) Soient I un intervalle de $\mathbb{R}$ et $f : I \times D(z_0, R) \to \mathbb{C}$ une fonction continue telle que pour chaque $x \in I$ fixé, la fonction $z \mapsto f(x, z)$ soit développable en série entière dans le disque $D(z_0, R)$. On suppose de plus qu'il existe une fonction h intégrable sur I telle que $|f(x, z)| \leq h(x)$ pour tous $x \in I$ et $z \in D(z_0, R)$. Montrer que la fonction $g(z) = \int_I f(x, z)\, dx$ est aussi développable en série entière dans le disque $D(z_0, R)$.

(c) Montrer que la fonction $\Gamma(z) = \int_{]0,\infty[} e^{-x} x^{z-1}\, dx$ (où $x^{z-1} = e^{(z-1)\ln x}$) est développable en série entière autour de tout point z_0 tel que $\mathcal{R}e\, z_0 > 0$.

7.4. Soit $h : [0, 1] \to \mathbb{R}$ une fonction continue. Pour $y \in [0, 1]$, on pose $f(x, y) = h(x)\, \mathbb{1}_{[0,y]}(x)$ et $g(y) = \int f(x, y)\, dx$. Montrer que pour tout $y \in [0, 1]$ fixé, $(\partial f/\partial y)(x, y)$ est définie pour presque tout x et vérifie $|(\partial f/\partial y)(x, y)| \leq F(x)$ presque partout pour une fonction F intégrable et indépendante de y. Par ailleurs, déterminer la dérivée $g'(y)$ de la fonction g. Qu'observe-t-on ?

7.5. On considère ici des fonctions intégrables définies sur $\mathbb{R}^n$ et à valeurs complexes.

(a) Calculer explicitement la transformée de Fourier de la fonction indicatrice du pavé $P = [a_1, b_1] \times \ldots \times [a_n, b_n]$. En déduire que la transformée de Fourier de toute fonction en escalier est continue et tend vers 0 à l'infini.

(b) Montrer que l'on a $\widehat{f}(\xi) = \frac{1}{2}\int \big(f(x) - f(x - (\pi\xi/|\xi|^2))\big)\, e^{-ix\cdot\xi}\, dx$ pour toute fonction f intégrable et tout $\xi \neq 0$, et en déduire que la transformée de Fourier de toute fonction continue à support compact est continue et tend vers 0 à l'infini.

(c) Plus généralement, montrer que la transformée de Fourier de toute fonction intégrable est continue et tend vers 0 à l'infini. Pour cela, on commencera par montrer que si (f_p) est une suite de fonctions intégrables qui converge en moyenne vers f, alors la suite $(\widehat{f_p})$ converge uniformément vers $\widehat{f}$, puis qu'une limite uniforme de fonctions continues tendant vers 0 à l'infini est une fonction continue tendant vers 0 à l'infini.

7.6. Soient $f(x, y) = x\, e^{-xy}$ et $g(x, y) = \sin x\, e^{-xy}$.

(a) Montrer que pour tout $A > 0$, les fonctions f puis g sont intégrables sur $[0, A] \times [0, \infty[$.

(b) En utilisant le théorème de Fubini pour la fonction g, retrouver la formule de l'exercice **7.1** de la page 53 : $\lim_{A\to\infty} \int_{[0,A]} \dfrac{\sin x}{x}\, dx = \pi/2$.

7.7. Soit $f : \mathbb{R}^n \to \mathbb{C}$ une fonction de classe C^1 à support compact. Montrer que f ainsi que ses dérivées partielles $f_j = (\partial f/\partial x_j)$ sont intégrables, puis en utilisant une intégration par parties, montrer que $\widehat{f_j}(\xi) = i\,\xi_j\,\widehat{f}(\xi)$. En déduire qu'il existe une constante C telle que $|\widehat{f}(\xi)| \leq C/(1 + |\xi|)$.

7.8. Soient I un intervalle ouvert de $\mathbb{R}$, f et g deux fonctions localement intégrables dans I, et F et G des intégrales indéfinies respectivement de f et g. Montrer que pour tous réels a et $b \in I$, on a

$$\int_a^b f(x)\, G(x)\, dx = F(b)\, G(b) - F(a)\, G(a) - \int_a^b F(x)\, g(x)\, dx$$

(ce résultat étend celui du corollaire 6.8 au cas où f et g ne sont pas nécessairement localement bornées).

7.9. Soit $f : \mathbb{R}^{n-1} \to \mathbb{R}_+$ une fonction mesurable positive, et F et G les parties de $\mathbb{R}^n$ définies par

$$F = \{\, x \in \mathbb{R}^n \,;\, 0 \leq x_n \leq f(x_1, \ldots, x_{n-1}) \,\} \qquad \text{et} \qquad G = \{\, x \in \mathbb{R}^n \,;\, x_n = f(x_1, \ldots, x_{n-1}) \,\}\,.$$

(a) Montrer que les fonctions $h(x) = x_n$ et $g(x) = f(x_1, \ldots, x_{n-1}) - x_n$ sont mesurables, et en déduire que F et G sont mesurables (on pourra utiliser les résultats des exercices **5.1** et **5.8** des pages 35 et 36).

(b) Montrer que $|F| = \int f(x_1, \ldots, x_{n-1})\, dx_1 \ldots dx_{n-1}$ et que G est négligeable.

7.10. Démontrer la réciproque suivante du lemme 7.4 : si $A \subset \mathbb{R}^n \times \mathbb{R}^m$ est mesurable, si $B \subset \mathbb{R}^m$ est négligeable et si $\forall y \notin B$, l'ensemble $A(y) = \{\, x \in \mathbb{R}^n \,;\, (x, y) \in A \,\}$ est négligeable, alors A est négligeable.

7.11. Soient $I =]0, 1]$ et F et $f : I \times I \to \mathbb{R}$ les fonctions définies par $F(x, y) = \dfrac{x}{x^2 + y^2}$ et $f(x, y) = \dfrac{x^2 - y^2}{(x^2 + y^2)^2}$. Calculer $\partial F/\partial x$, $\int_I \big(\int_I f(x, y)\, dy\big)\, dx$ et $\int_I \big(\int_I f(x, y)\, dx\big)\, dy$. La fonction f est-elle intégrable sur $I \times I$?

8. Changement de variables

Soient X et Y deux ouverts de $\mathbb{R}^n$, et ϕ une bijection de Y sur X. Soit enfin f une fonction intégrable sur X. Le but de ce paragraphe est d'établir une formule permettant de calculer l'intégrale $\int_X f(x)\,dx$ "en posant $x = \phi(y)$", c'est-à-dire à l'aide d'une intégrale sur Y faisant intervenir la fonction $f \circ \phi$. On commence par rappeler la définition suivante.

Définition 8.1. *Étant donnés deux ouverts X et Y de $\mathbb{R}^n$, une bijection ϕ de Y sur X est un* difféomorphisme *de Y sur X si c'est une application de classe C^1 et si sa réciproque est aussi de classe C^1. Pour un difféomorphisme de Y sur X, on notera $J_\phi(y) = \big((\partial\phi_j/\partial y_k)(y)\big)_{1\leq j,k\leq n}$ et $[\![J_\phi(y)]\!]$ respectivement la matrice de ses dérivées premières, appelée* matrice jacobienne *de ϕ, et la valeur absolue du déterminant de $J_\phi(y)$, appelée* jacobien *de ϕ.*

Pour vérifier qu'une bijection $\phi : Y \to X$ est un difféomorphisme, on utilise en pratique le critère fourni par le *théorème d'inversion locale* : la bijection $\phi : Y \to X$ de classe C^1 est un difféomorphisme si et seulement si son jacobien ne s'annule en aucun point. Le principal résultat de ce paragraphe s'énonce comme suit.

Théorème 8.2 (du changement de variables). *Soient X et Y deux ouverts de $\mathbb{R}^n$, ϕ un difféomorphisme de Y sur X et $f : X \to \mathbb{C}$. Alors on a la formule*

$$\int_X f(x)\,dx = \int_Y \Big(f \circ \phi(y)\Big)\,[\![J_\phi(y)]\!]\,dy$$

dans les deux cas suivants :
(a) *Lorsque f est intégrable sur X, ce qui se produit si et seulement si la fonction $y \mapsto \big(f \circ \phi(y)\big)\,[\![J_\phi(y)]\!]$ est intégrable sur Y.*
(b) *Lorsque f est mesurable positive sur X, ce qui se produit si et seulement si la fonction $y \mapsto \big(f \circ \phi(y)\big)\,[\![J_\phi(y)]\!]$ est mesurable positive sur Y.*

La démonstration de ce théorème repose sur les propriétés énoncées dans le lemme suivant.

Lemme 8.3. *Soient X et Y deux ouverts de $\mathbb{R}^n$, et ϕ un difféomorphisme de Y sur X. Alors :*
(a) *Pour toute suite (P_q) de pavés de X contenant un même point $\phi(y_0)$, homothétiques à un même pavé P_0 de mesure > 0 et de diamètres tendant vers 0, on a*

$$[\![J_\phi(y_0)]\!] = \lim_{q\to\infty} \frac{|P_q|}{|\phi^{-1}(P_q)|}\,.$$

(b) *Pour tous pavés $Q \subset P \subset X$, on a $|Q| \leq \big(\sup_{y\in\phi^{-1}(P)} [\![J_\phi(y)]\!]\big)\,|\phi^{-1}(Q)|$.*
(c) *Une partie $A \subset Y$ est négligeable si et seulement si $\phi(A)$ est négligeable.*

Démonstration. Pour démontrer la propriété (a), fixons quelques notations. Pour tout parallélépipède R et tout réel $a > 0$, notons $a\,R$ le parallélépipède de même centre et homothétique à R de rapport a. Le point $y_0 \in Y$ étant fixé, notons $x_0 = \phi(y_0)$, $\psi = \phi^{-1}$, puis $\psi_0(x) = y_0 + J_\psi(x_0)(x - x_0)$ la transformation affine tangente à ψ en x_0, et $\phi_0 = \psi_0^{-1}$ la transformation affine tangente à ϕ en y_0. Par différentiabilité de ψ en x_0 et de ϕ en y_0, il existe pour tout $\varepsilon > 0$, un voisinage U_ε de x_0 et un voisinage V_ε de y_0 tels que $|\psi(x) - \psi_0(x)| \leq \varepsilon\,|x - x_0|$ si $x \in U_\varepsilon$ et que $|\phi(y) - \phi_0(y)| \leq \varepsilon\,|y - y_0|$ si $y \in V_\varepsilon$. Soient alors $P \subset U_\varepsilon$ un pavé contenant x_0 tel que $Q = \psi_0(P) \subset V_\varepsilon$, et $a \in]0,1[$ un réel qui sera choisi ultérieurement. On a $\phi_0(a\,Q) = a\,P$. En notant $\delta(P)$ le diamètre de P et $A + x = \{\, z + x\,;\ z \in A\,\}$, on voit donc que

$$\phi(a\,Q) \subset \bigcup_{|x| \leq \varepsilon\,\delta(Q)} \Big(a\,P + x\Big) \quad \text{et} \quad \psi(P) \subset \bigcup_{|y| \leq \varepsilon\,\delta(P)} \Big(Q + y\Big)\,.$$

On a $\delta(Q) \leq C_0\,\delta(P)$ et $\delta(P) \leq C_0\,\delta(Q)$ pour une constante C_0 ne dépendant que de ψ_0. Comme $|P_0| > 0$, alors pour tout pavé P homothétique à P_0 on a $|P| > 0$, et aussi $|Q| = [\![J_\psi(x_0)]\!]\,|P| > 0$ d'après la proposition 7.5, ce qui permet de voir qu'en choisissant $a = (1 + C\,\varepsilon)^{-1}$ avec $C > 0$ assez grande, on a $\cup_{|x| \leq \varepsilon\,\delta(Q)}(a\,P + x) \subset P$ et $\cup_{|y| \leq \varepsilon\,\delta(P)}(Q + y) \subset a^{-1}Q$, d'où les inclusions $a\,Q = \psi(\phi(a\,Q)) \subset \psi(P) \subset a^{-1}Q$. En transformant ces inclusions en inégalités entre les mesures des ensembles indiqués, en divisant par $|P|$ et en utilisant l'homogénéité de la mesure, on obtient donc

$$a^n\,[\![J_\psi(x_0)]\!] \leq \frac{|\psi(P)|}{|P|} \leq a^{-n}\,[\![J_\psi(x_0)]\!]\,.$$

On voit ainsi que le quotient $|\psi(P)|/|P|$ peut être rendu arbitrairement proche de $[\![J_\psi(x_0)]\!]$ pourvu que a soit assez proche de 1 c'est-à-dire ε assez proche de 0, c'est-à-dire pourvu que $P \subset U_\varepsilon$ et $Q \subset V_\varepsilon$ pour ε assez petit. C'est bien la propriété (a) puisque $[\![J_\psi(x_0)]\!] = [\![J_\phi(y_0)]\!]^{-1}$.

La propriété (b) se démontre par l'absurde, et quitte à remplacer le pavé P par le pavé Q, on peut donc supposer qu'il existe un pavé $P_0 = P \subset X$ tel que $|P| > C\,|\phi^{-1}(P)|$ pour une constante $C > \sup_{y \in \phi^{-1}(P)} [\![J_\phi(y)]\!]$. Alors il existe aussi une constante $c < 1/2$ telle que $(2\,c)^n\,|P| > C\,|\phi^{-1}(P)|$, et on considère, pour chaque sommet s_k du pavé P, le pavé Q_k image de P par l'homothétie de centre s_k et de rapport c. Les pavés Q_k sont contenus dans P, et comme $c < 1/2$, ils sont deux à deux disjoints. Alors les ensembles $\phi^{-1}(Q_k)$ sont contenus dans $\phi^{-1}(P)$ et sont deux à deux disjoints, ce qui permet d'écrire

$$C \sum_{k=1}^{2^n} |\phi^{-1}(Q_k)| \leq C\,|\phi^{-1}(P)| < 2^n\,(c^n|P|)\,,$$

d'où l'on déduit que l'un au moins des pavés Q_k vérifie $C\,|\phi^{-1}(Q_k)| < c^n|P| = |Q_k|$. Ce pavé Q_k étant noté P_1, on recommence à partir de P_1 la construction effectuée à partir de $P_0 = P$. On obtient de la sorte une suite (P_q) de pavés homothétiques à P_0, dont les diamètres tendent vers 0, et vérifiant $|P_q| > C\,|\phi^{-1}(P_q)|$ pour tout q. Comme c'est une suite décroissante de compacts non vides contenus dans P, leur intersection consiste en un point $\phi(y_0)$ pour un $y_0 \in P$. La construction de cette suite conduit donc à une contradiction puisque d'après la propriété (a) démontrée ci-dessus,

$$[\![J_\phi(y_0)]\!] \leq \sup_{\phi^{-1}(P)} [\![J_\phi]\!] < C \leq \lim_{q\to\infty} \frac{|P_q|}{|\phi^{-1}(P_q)|} = [\![J_\phi(y_0)]\!]\,.$$

On démontrerait de la même manière (avec $c = 1/2$, les pavés Q_k devant alors recouvrir P au lieu d'être deux à deux disjoints) l'inégalité $|Q| \geq \big(\inf_{\phi^{-1}(P)} [\![J_\phi]\!]\big)\,|\phi^{-1}(Q)|$ pour tous pavés $Q \subset P \subset X$.

Enfin, supposons que $A \subset Y$ possède une image $\phi(A)$ négligeable. Comme X est réunion de la famille dénombrable des pavés $P_q \subset X$ dont les sommets appartiennent à $\mathbb{Q}^n$, l'ensemble A est réunion de la famille dénombrable $(A_q)_{q\in\mathbb{N}}$ définie par $A_q = A \cap \phi^{-1}(P_q)$. Pour démontrer que A est négligeable, il suffit donc de montrer que chaque A_q est négligeable, ce qui revient à supposer que $\phi(A)$ est contenu dans un pavé $P \subset X$. Comme $\phi(A)$ est négligeable, il existe pour tout $\varepsilon > 0$ une suite (Q_p) de pavés (que l'on peut supposer contenus dans P) recouvrant $\phi(A)$ et vérifiant $\sum_{p\in\mathbb{N}} |Q_p| \leq \varepsilon$. Alors $A \subset \cup_{p\in\mathbb{N}}\,\phi^{-1}(Q_p)$, ce qui permet d'écrire

$$\Big(\inf_{\phi^{-1}(P)} [\![J_\phi]\!]\Big)\,|A| \leq \Big(\inf_{\phi^{-1}(P)} [\![J_\phi]\!]\Big) \sum_{p\in\mathbb{N}} |\phi^{-1}(Q_p)| \leq \sum_{p\in\mathbb{N}} |Q_p| \leq \varepsilon\,,$$

puis de conclure que $|A| = 0$ ($\Leftrightarrow$ A est négligeable) puisque cette inégalité est vraie pour tout $\varepsilon > 0$ et que $\inf_{\phi^{-1}(P)} [\![J_\phi]\!] > 0$ grâce à la compacité de P. On a donc prouvé que $\phi(A)$ négligeable implique A négligeable, et pour obtenir la réciproque, il suffit de remplacer le difféomorphisme ϕ par le difféomorphisme ϕ^{-1}. ∎

Démonstration du théorème 8.2. Il suffit de démontrer le théorème dans un sens, puisque l'on obtient l'autre sens en réappliquant le théorème avec ϕ^{-1} à la place de ϕ.

On commence par démontrer la formule dans le cas de la fonction indicatrice d'un pavé $P \subset X$, ce qui résulte immédiatement du lemme 8.3 (c) lorsque le pavé P est de mesure nulle. Dans le cas d'un pavé P de mesure non nulle, on procède alors à la construction familière suivante : on pose $\mathscr{P}_0 = \{P\}$, puis on note $\mathscr{P}_1$ la famille de pavés homothétiques à P obtenue en divisant chaque côté de P en deux, $\mathscr{P}_2$ la famille obtenue en répétant cette opération sur chaque élément de $\mathscr{P}_1$, et ainsi de suite. Pour tout $q \in \mathbb{N}$, on peut écrire

$$\int_X \mathbb{1}_P(x)\,dx = |P| = \sum_{Q\in\mathscr{P}_q} \frac{|Q|}{|\phi^{-1}(Q)|}\,|\phi^{-1}(Q)| = \sum_{Q\in\mathscr{P}_q} \frac{|Q|}{|\phi^{-1}(Q)|} \int_Y \left(\mathbb{1}_Q \circ \phi(y)\right) dy\,,$$

soit $\int_X \mathbb{1}_P(x)\,dx = \int_Y g_q(y)\,dy$ en notant g_q la fonction

$$g_q(y) = \sum_{Q\in\mathscr{P}_q} \frac{|Q|}{|\phi^{-1}(Q)|} \left(\mathbb{1}_Q \circ \phi(y)\right).$$

La suite $(g_q)_{q\in\mathbb{N}}$ est une suite de fonctions intégrables sur le compact $K = \phi^{-1}(P)$. Comme $\mathbb{1}_Q \circ \phi \leq 1$, il résulte du lemme 8.3 (b) que ces fonctions g_q sont uniformément bornées (presque partout) par la constante $\sup_{\phi^{-1}(P)} [\![J_\phi]\!]$, et d'après le lemme 8.3 (a), on voit que la suite $(g_q)_{q\in\mathbb{N}}$ tend presque partout vers la fonction $(\mathbb{1}_P \circ \phi)\,[\![J_\phi]\!]$. On obtient donc le résultat par convergence uniformément bornée sur un compact (convergence dominée).

La formule de changement de variables étant acquise pour les fonctions indicatrices de pavés, elle l'est aussi pour les fonctions f en escalier par combinaisons linéaires. Dans le cas d'une fonction f intégrable, celle-ci est limite en moyenne, et aussi presque partout d'après le théorème d'Égorov, d'une suite $(f_p)_{p\in\mathbb{N}}$ de fonctions en escalier dans X. Alors la suite $g_p(y) = (f_p \circ \phi(y))\,[\![J_\phi(y)]\!]$ est une suite de fonctions intégrables sur Y, qui est de Cauchy pour la convergence en moyenne puisque la formule est vraie pour les fonctions en escalier $|f_p - f_q|$, et qui converge vers $(f \circ \phi)\,[\![J_\phi]\!]$ presque partout puisque les ensembles négligeables de X et de Y se correspondent d'après le lemme 8.3 (c). On en déduit donc l'intégrabilité de $(f \circ \phi)\,[\![J_\phi]\!]$, ainsi que la formule de changement de variables, par passage à la limite dans les intégrales.

Enfin, dans le cas d'une fonction f mesurable positive sur X, on écrit X comme la réunion d'une suite croissante (K_p) de compacts, puis on considère la suite (f_p) de fonctions intégrables définie par $f_p = \inf(f, p\,\mathbb{1}_{K_p})$. Comme les fonctions f_p sont intégrables (par le critère du théorème 5.2), elles vérifient les conclusions du théorème 8.2, cas (a), et donc $g_p = (f_p \circ \phi)\,[\![J_\phi]\!]$ est aussi intégrable sur Y, de même intégrale que f_p. Comme par ailleurs la suite (f_p) est croissante et qu'elle converge partout vers f, la suite (g_p) est aussi croissante et converge partout vers $g = (f\circ\phi)[\![J_\phi]\!]$ (qui est donc mesurable positive), et on obtient alors le résultat en passant à la limite grâce au théorème 5.4 (c). ∎

Exemple. Le paramètre $x \in \mathbb{R}^n$ étant fixé, considérons le changement de variables $\phi : \mathbb{R}^n \to \mathbb{R}^n$ défini par $y = \phi(z) = x - z$. Sa matrice jacobienne vaut $J_\phi(z) = -I_n$, d'où $[\![J_\phi(z)]\!] = |(-1)^n| = 1$. Alors si f et g sont deux fonctions intégrables sur $\mathbb{R}^n$, leur produit de convolution $f * g(x) = \int f(x-y)\,g(y)\,dy$ défini au paragraphe précédent s'écrit aussi $f * g(x) = \int f(z)\,g(x-z)\,dz$, c'est-à-dire que le produit de convolution est commutatif.

Corollaire 8.4 (Coordonnées polaires dans le plan). *Une fonction f est intégrable sur $\mathbb{R}^2$ si et seulement si la fonction $(r,\theta) \mapsto r\, f(r\cos\theta, r\sin\theta)$ est intégrable sur $Y =]0,\infty[\, \times\,]-\pi,\pi[$, et alors on a*

$$\int f(x_1,x_2)\, dx_1\, dx_2 = \int_Y f(r\cos\theta, r\sin\theta)\, r\, dr\, d\theta\,.$$

Démonstration. Notons X le complémentaire dans $\mathbb{R}^2$ de l'ensemble négligeable $\{\,(x_1,0)\,;\, x_1 \leq 0\,\}$ (demi-axe des x_1 négatifs). La transformation $\phi : Y \to X$ définie par $\phi(r,\theta) = (r\cos\theta, r\sin\theta)$ est clairement de classe C^1, et son jacobien, qui vaut

$$[\![J_\phi(r,\theta)]\!] = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r\,,$$

ne s'annule en aucun point de Y. De plus, cette transformation est bijective : sa réciproque est $\phi^{-1}(x_1,x_2) = \big(\sqrt{x_1^2+x_2^2},\, 2\arctan(x_2/(x_1+\sqrt{x_1^2+x_2^2}))\big)$, et c'est donc un difféomorphisme. Il résulte alors du théorème 8.2 que f est intégrable sur X si et seulement si $(f\circ\phi)\,[\![J_\phi]\!]$ est intégrable sur Y, auquel cas

$$\int_X f(x_1,x_2)\, dx_1\, dx_2 = \int_Y f(r\cos\theta, r\sin\theta)\, r\, dr\, d\theta\,.$$

Enfin, comme $f = \mathbb{1}_X f$ presque partout puisque le complémentaire de X est négligeable, on peut remplacer X par $\mathbb{R}^2$ dans ces conclusions. ∎

Application : les fonctions gaussiennes. On appelle (*fonction*) *gaussienne* l'exponentielle d'une forme quadratique définie négative. Nous en étudions ci-dessous quelques propriétés, en notant $|x|^2 = x_1^2 + \ldots + x_n^2$ le carré de la norme euclidienne de $x \in \mathbb{R}^n$.

(a) *Intégrale d'une gaussienne :* $\int e^{-|x|^2/2}\, dx = (2\pi)^{n/2}$. La fonction $g(x) = e^{-|x|^2/2}$ est intégrable sur $\mathbb{R}^n$ puisque c'est le produit (tensoriel) des fonctions $x_j \mapsto e^{-x_j^2/2}$ d'une seule variable, et que ces fonctions sont intégrables sur $\mathbb{R}$ comme on l'a vu au paragraphe 6. De plus, par tensorialité, son intégrale vaut

$$\int_{\mathbb{R}^n} e^{-|x|^2/2}\, dx = \left(\int_{\mathbb{R}} e^{-t^2/2}\, dt\right)^n,$$

et il suffit, pour démontrer la formule annoncée, de l'établir pour une seule valeur de n. Pour $n = 2$, la formule du changement de variables en coordonnées polaires du corollaire 8.4 donne

$$\left(\int_{\mathbb{R}} e^{-t^2/2}\, dt\right)^2 = \int_{\mathbb{R}^2} e^{-|x|^2/2}\, dx = \int_Y e^{-r^2/2}\, r\, dr\, d\theta = 2\pi \int_{]0,\infty[} r\, e^{-r^2/2}\, dr\,.$$

Or la fonction continue positive $r \mapsto r\, e^{-r^2/2}$ admet pour primitive la fonction $F(r) = -e^{-r^2/2}$, ce qui permet d'écrire $\int_{]0,\infty[} r\, e^{-r^2/2}\, dr = \lim_{p\to\infty}(F(p) - F(0))$ $= 1$. La formule annoncée est ainsi démontrée pour $n = 2$, donc pour tout n.

(b) *Transformée de Fourier d'une gaussienne : la fonction* $g^{\varepsilon}(x) = e^{-\varepsilon^2|x|^2/2}$ *admet pour transformée de Fourier la fonction* $\widehat{g^{\varepsilon}}(\xi) = (2\pi)^{n/2}\,\varepsilon^{-n}\,e^{-|\xi|^2/2\varepsilon^2}$. Comme dans le calcul précédent, la fonction g^{ε} étant un produit de fonctions d'une seule variable, il suffit par tensorialité d'établir le résultat pour une valeur de n, et nous choisirons ici $n=1$. D'après le théorème 7.2, la fonction $f : \xi \mapsto e^{\xi^2/2\varepsilon^2}\,\widehat{g^{\varepsilon}}(\xi)$, qui s'écrit aussi $f(\xi) = \int e^{-(\varepsilon x + (i\xi/\varepsilon))^2/2}\,dx$, est de classe C^1 et de dérivée

$$\begin{aligned} f'(\xi) &= -\frac{i}{\varepsilon}\int (\varepsilon x + (i\xi/\varepsilon))\,e^{-(\varepsilon x+(i\xi/\varepsilon))^2/2}\,dx = \\ &= \frac{i}{\varepsilon^2}\lim_{q\to\infty}\int_{[-q,q]}\frac{\partial}{\partial x}\Big(e^{-(\varepsilon x+(i\xi/\varepsilon))^2/2}\Big)\,dx = \\ &= \frac{i}{\varepsilon^2}\lim_{q\to\infty}\Big(e^{-(\varepsilon q+(i\xi/\varepsilon))^2/2} - e^{-(-\varepsilon q+(i\xi/\varepsilon))^2/2}\Big) = 0\,. \end{aligned}$$

La fonction f est donc constante, soit $\widehat{g^{\varepsilon}}(\xi) = f(0)\,e^{-\xi^2/2\varepsilon^2}$, et on trouve la constante $f(0) = \sqrt{2\pi}/\varepsilon$ en se ramenant au calcul du (a) par le changement de variable $x = \phi(y) = y/\varepsilon$ pour lequel on a $J_{\phi}(y) = \phi'(y) = 1/\varepsilon = [\![J_{\phi}(y)]\!]$.

(c) *Convolution par une gaussienne : le produit de convolution* $f * g_{\varepsilon}$ *d'une fonction intégrable* f *par la gaussienne* $g_{\varepsilon}(x) = (2\pi)^{-n/2}\,\varepsilon^{-n}\,e^{-|x|^2/2\varepsilon^2}$ *converge en moyenne vers* f *lorsque* ε *tend vers* 0. Il suffit de démontrer ce résultat lorsque f est continue à support compact. En effet, si f est intégrable et $\delta > 0$, il existe une fonction continue à support compact h telle que $\int |h(x) - f(x)|\,dx \le \delta/3$. Comme l'intégrale de g_{ε} vaut 1, on a

$$\int |f*g_{\varepsilon}(x) - h*g_{\varepsilon}(x)|\,dx \le \int |f(x-y) - h(x-y)|\,g_{\varepsilon}(y)\,dx\,dy = \int |f(x) - h(x)|\,dx\,,$$

d'où, par inégalité triangulaire,

$$\begin{aligned} &\int |f*g_{\varepsilon}(x) - f(x)|\,dx \le \\ &\quad\le \int |f*g_{\varepsilon}(x) - h*g_{\varepsilon}(x)|\,dx + \int |h*g_{\varepsilon}(x) - h(x)|\,dx + \int |h(x) - f(x)|\,dx \le \\ &\quad\le \frac{2\delta}{3} + \int |h*g_{\varepsilon}(x) - h(x)|\,dx\,, \end{aligned}$$

ce qui peut être rendu inférieur à δ pour ε assez petit si le résultat est établi pour h. La fonction f étant donc supposée continue à support compact, on utilise le changement de variables $y = \phi(z) = \varepsilon z$, pour lequel $J_{\phi}(z) = \varepsilon\,I_n \Rightarrow [\![J_{\phi}(z)]\!] = \varepsilon^n$, pour écrire $f*g_{\varepsilon}(x) = \int f(x-y)\,g_{\varepsilon}(y)\,dy = \int f(x-\varepsilon z)\,g_1(z)\,dz$ puis comme plus haut

$$\begin{aligned} &\int |f*g_{\varepsilon}(x) - f(x)|\,dx \le \int |f(x-\varepsilon z) - f(x)|\,g_1(z)\,dx\,dz \le \\ &\qquad\le \int_{|z|\le q} |f(x-\varepsilon z) - f(x)|\,g_1(z)\,dx\,dz + 2\int |f(x)|\,dx \int_{|z|\ge q} g_1(z)\,dz\,. \end{aligned}$$

Par convergence dominée, le dernier terme tend vers 0 pour $q \to \infty$ tandis que le terme précédent, à q fixé, tend vers 0 pour $\varepsilon \to 0$.

Utilisation : la formule d'inversion de Fourier. Si f est une fonction intégrable telle que $\widehat{f}$ soit aussi intégrable, alors on a $f(x) = (2\pi)^{-n} \int e^{ix\cdot\xi}\, \widehat{f}(\xi)\, d\xi$ pour presque tout x.

Pour le démontrer, commençons par remplacer $\widehat{f}$ par sa définition dans cette formule : on trouve l'intégrale $(2\pi)^{-n} \int e^{i(x-y)\cdot\xi}\, f(y)\, dy\, d\xi$ qui n'a aucun sens puisque, d'après le théorème de Fubini pour les fonctions mesurables positives (théorème 7.3 (b)), la fonction $(y,\xi) \mapsto e^{i(x-y)\cdot\xi}\, f(y)$ n'est pas intégrable. En revanche, ce même théorème montre que pour tout $\varepsilon > 0$ la fonction $h_\varepsilon : (y,\xi) \mapsto e^{i(x-y)\cdot\xi}\, f(y)\, e^{-\varepsilon^2|\xi|^2/2}$ est intégrable, et en l'intégrant d'abord en y on trouve

$$(2\pi)^{-n} \int h_\varepsilon(y,\xi)\, dy\, d\xi = (2\pi)^{-n} \int e^{ix\cdot\xi}\, \widehat{f}(\xi)\, e^{-\varepsilon^2|\xi|^2/2}\, d\xi\,,$$

ce qui tend vers $(2\pi)^{-n} \int e^{ix\cdot\xi}\, \widehat{f}(\xi)\, d\xi$ par convergence dominée lorsque $\varepsilon \to 0$. En intégrant h_ε d'abord en ξ, on trouve

$$(2\pi)^{-n} \int h_\varepsilon(y,\xi)\, dy\, d\xi = (2\pi)^{-n} \int f(y)\, \widehat{g^\varepsilon}(y-x)\, dy = f * g_\varepsilon(x)$$

en notant $g^\varepsilon(\xi) = e^{-\varepsilon^2|\xi|^2/2}$ et $g_\varepsilon(x) = (2\pi)^{-n/2}\, \varepsilon^{-n}\, e^{-|x|^2/2\varepsilon^2}$ comme précédemment, et cette fonction de x converge en moyenne vers la fonction f lorsque $\varepsilon \to 0$ d'après la propriété (c) des gaussiennes. Il ne reste plus qu'à utiliser le théorème d'Égorov (théorème 3.8) pour conclure par extraction d'une sous-suite qui converge presque partout.

Exercices

8.1. La fonction Γ ayant été définie à l'exercice **7.2** de la page 53, calculer $\Gamma(1/2)$ (utiliser le changement de variable $x = y^2/2$), puis en déduire $\Gamma(p + (1/2))$ pour tout entier $p \in \mathbb{N}$.

8.2. Pour tout entier $n \geq 1$ et tout réel $r > 0$ on note $B_n(r) = \{\, x \in \mathbb{R}^n \,;\, |x| \leq r \,\}$ la boule euclidienne de rayon r dans $\mathbb{R}^n$, puis $|B_n(r)| = \int \mathbb{1}_{B_n(r)}(x)\, dx$ son volume.

(a) Montrer que l'on a $|B_n(r)| = r^n\, |B_n(1)|$.

(b) Calculer $|B_1(1)|$ et $|B_2(1)|$, puis établir la relation de récurrence $|B_n(1)| = \frac{2\pi}{n}\, |B_{n-2}(1)|$ pour tout $n > 2$. En déduire la valeur de $|B_n(r)|$ pour tout entier $n \geq 1$, et montrer que l'on a $|B_n(r)| = r^n \pi^{n/2} / \Gamma(1 + (n/2))$ (pour la définition de la fonction Γ, voir les exercices **7.2** et **8.1** des pages 53 et 60).

(c) Soit $a = (a_1, \ldots, a_n)$ un n-uplet de réels strictement positifs. Calculer le volume $|E(a)| = \int \mathbb{1}_{E(a)}(x)\, dx$ de l'ellipsoïde défini par $E(a) = \{\, x \in \mathbb{R}^n \,;\, \sum_{j=1}^n (x_j/a_j)^2 \leq 1 \,\}$.

8.3. Dans tout cet exercice, A désigne une matrice réelle de taille $n \times n$ et ${}^t\!A$ désigne sa transposée.

(a) Montrer que l'on a $(Ax) \cdot \xi = x \cdot ({}^t\!A\xi)$ pour tous vecteurs x et $\xi \in \mathbb{R}^n$.

(b) Soient $f : \mathbb{R}^n \longrightarrow \mathbb{C}$ une fonction intégrable et $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ une transformation affine inversible. Calculer la transformée de Fourier de $f \circ \phi$ en fonction de celle de f (on écrira $\phi(y) = x_0 + Ay$ avec $x_0 = \phi(0)$ et $A = J_\phi$).

(c) On suppose que la matrice A est symétrique définie positive, et on note $q : \mathbb{R}^n \longrightarrow \mathbb{R}_+$ la forme quadratique de matrice A. Montrer qu'il existe une matrice inversible B (reliée à la matrice A par une formule que l'on donnera) telle que $q(x) = |Bx|^2$, et calculer la transformée de Fourier de la gaussienne $g_A(x) = e^{-q(x)/2}$ en fonction de la matrice A.

8.4. Soient f et $g : \mathbb{R}^n \longrightarrow \mathbb{C}$ deux fonctions intégrables, et $f * g(x) = \int f(x-y)\, g(y)\, dy = \int f(z)\, g(x-z)\, dz$ leur produit de convolution.

(a) Montrer que si f est continue à support compact, alors $f * g$ est continue, et que si f est de classe C^1 à support compact, alors $f * g$ est de classe C^1 et vérifie $\partial(f * g)/\partial x_j = (\partial f/\partial x_j) * g$ pour tout $j \leq n$. Plus généralement, montrer que si f est de classe C^k à support compact et g de classe C^ℓ à support compact, alors $f * g$ est de classe $\mathrm{C}^{k+\ell}$ à support compact.

(b) Soient φ et $\psi : \mathbb{R}^n \longrightarrow \mathbb{R}_+$ deux fonctions de classe C^1 à support compact telles que $\int \varphi(x)\, dx = 1$ et $\psi(x) = 1$ pour $|x| \leq 1$. On pose ensuite $\varphi_\varepsilon(x) = \varepsilon^{-n} \varphi(x/\varepsilon)$, $\psi_\varepsilon(x) = \psi(\varepsilon x)$ et $f_\varepsilon = \varphi_\varepsilon * (f \psi_\varepsilon)$. Montrer que si f est différentiable avec $f_j = (\partial f/\partial x_j)$ intégrable pour tout $j \leq n$, alors les fonctions f_ε et $(\partial f_\varepsilon/\partial x_j)$ convergent en moyenne respectivement vers f et f_j lorsque ε tend vers 0, puis utiliser l'exercice **7.7** de la page 54 pour montrer que $\widehat{f_j}(\xi) = i\,\xi_j\, \widehat{f}(\xi)$.

8.5. Pour tout $a > 0$, on considère la gaussienne $g_a : \mathbb{R} \longrightarrow \mathbb{R}_+$ définie par $g_a(x) = (a\sqrt{2\pi})^{-1}\, e^{-x^2/(2a^2)}$. Calculer le produit de convolution $g_a * g_b(x)$ (on pourra utiliser le changement de variable $z \mapsto y = \alpha z + \beta x$ pour des coefficients α et β bien choisis, puis on posera $c = \sqrt{a^2 + b^2}$ pour simplifier les calculs). Plus généralement dans $\mathbb{R}^n$, que vaudra le produit de convolution de deux gaussiennes ?

8.6. On se propose de calculer les intégrales $I(n,k) = \int_{\mathbb{R}^n} \dfrac{dx}{(1+|x|^2)^{k/2}}$ pour tout entier $k > n$.

(a) Calculer $I(1,2)$ et $I(1,3)$ (on utilisera le changement de variable $x = \sinh t$ pour calculer $I(1,3)$).

(b) En partant de l'expression $I(1,k) = \int_{\mathbb{R}} \dfrac{dx}{(1+x^2)^{(k+2)/2}} + \int_{\mathbb{R}} \dfrac{x^2\, dx}{(1+x^2)^{(k+2)/2}}$, démontrer une formule de récurrence pour $I(1,k)$, et en déduire la valeur de $I(1,k)$ pour tout k.

(c) Pour $n > 1$ et $x \in \mathbb{R}^n$, on pose $x = (x_*, x_n)$ avec $x_* = (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}$. En effectuant le changement de variables $x_n = t\,(1 + |x_*|^2)^{1/2}$ puis en intégrant d'abord en t par Fubini, montrer que $I(n,k) = I(1,k)\, I(n-1, k-1)$, puis en déduire la valeur de $I(n,k)$ pour tout $k > n$.

8.7. Redémontrer le résultat de l'exercice **8.5** ci-dessus en utilisant la formule d'inversion de Fourier.

Émile Borel
(1871 - 1956)

Henri Lebesgue
(1875 - 1941)

Les principaux créateurs de la théorie moderne de l'intégration fondée sur la notion de mesure.

9. Espaces $\mathcal{L}^p$

Dans tout ce paragraphe, X désigne une partie mesurable de $\mathbb{R}^n$ (généralement un ouvert) avec éventuellement $X = \mathbb{R}^n$, et on étudiera des fonctions f mesurables sur X, c'est-à-dire telles que $\mathbb{1}_X f$ soit mesurable.

Définition 9.1. *Soit X une partie mesurable de $\mathbb{R}^n$.*
(a) *Les nombres réels $p > 1$ et $q > 1$ sont dits (exposants)* conjugués *si $p+q = pq$, c'est-à-dire si $(1/p) + (1/q) = 1$ ou encore si $q = p/(p-1)$. On dira aussi que les exposants 1 et ∞ sont conjugués.*
(b) *Pour toute fonction $f : X \to \mathbb{C}$ mesurable sur X et tout réel $p \geq 1$, on pose $\|f\|_p = \left(\int_X |f(x)|^p \, dx \right)^{1/p}$. On pose aussi $\|f\|_\infty = \inf\{\, c \in \mathbb{R}_+ \,;\, |f(x)| \leq c \ \text{ pp} \,\}$ (pour $p = \infty$).*
(c) *Pour $1 \leq p \leq \infty$, on note $\mathcal{L}^p(X)$ l'ensemble de toutes les fonctions $f : X \to \mathbb{C}$ mesurables sur X telles que $\|f\|_p < \infty$.*

D'après le théorème 5.2, l'ensemble $\mathcal{L}^1(X)$ est l'ensemble de toutes les fonctions intégrables sur X, et plus généralement pour $p < \infty$, l'ensemble $\mathcal{L}^p(X)$ est l'ensemble de toutes les fonctions de puissance $p^{\text{ième}}$ intégrable sur X.

Pour $p = \infty$, on observe que l'ensemble des points x où $|f(x)| > \|f\|_\infty$ est réunion pour $r \in \mathbb{N}$ des ensembles $A_r = \{\, x \,;\, |f(x)| \geq \|f\|_\infty + 2^{-r} \,\}$ qui sont tous de mesure nulle, et on voit donc que l'on a toujours $|f(x)| \leq \|f\|_\infty$ presque partout. Ainsi, l'ensemble $\mathcal{L}^\infty(X)$ est l'ensemble des fonctions qui coïncident presque partout avec une fonction bornée ; ses éléments sont parfois appelés fonctions *essentiellement bornées.*

Théorème 9.2. *Soient p et q deux exposants conjugués, f et g deux fonctions mesurables sur X.*
(a) *Si $p, q < \infty$ et $a, b \geq 0$, alors $a\,b \leq (a^p/p) + (b^q/q)$ et $(a+b)^p \leq 2^{p-1}(a^p + b^p)$.*
(b) *Inégalité de* Hölder *: $\int_X |f(x)\,g(x)|\,dx \leq \|f\|_p \, \|g\|_q$. En particulier, si $f \in \mathcal{L}^p(X)$ et $g \in \mathcal{L}^q(X)$, alors le produit fg est intégrable sur X et on a $\left| \int_X f(x)\,g(x)\,dx \right| \leq \|f\|_p \, \|g\|_q$.*
(c) *Inégalité de* Minkowski *: $\|f + g\|_p \leq \|f\|_p + \|g\|_p$. En particulier, si f et $g \in \mathcal{L}^p(X)$, alors $f + g \in \mathcal{L}^p(X)$.*
(d) *L'ensemble $\mathcal{L}^p(X)$ est un espace vectoriel, et l'application $f \mapsto \|f\|_p$ est une semi-norme sur cet espace.*

Démonstration. Pour démontrer la première inégalité (a), on utilise la propriété suivante des fonctions convexes : si $f : [0,1] \to \mathbb{R}$ est une fonction de classe C^2, vérifiant $f(0) = f(1) = 0$ et $f'' \geq 0$, alors f est négative sur $[0,1]$. En effet, par le théorème de Rolle il existe un $t_0 \in]0,1[$ tel que $f'(t_0) = 0$, ce qui implique, puisque $f'' \geq 0$, que f' est négative et f décroissante sur $[0, t_0]$ et f' positive et f croissante sur $[t_0, 1]$. Cette propriété s'applique à la fonction $f(t) = a^{pt}\, b^{q(1-t)} - t\, a^p - (1-t)\, b^q$ puisque $f''(t) = a^{pt}\, b^{q(1-t)} \left(\ln(a^p/b^q)\right)^2 \geq 0$, et on peut donc conclure que $f(1/p) \leq 0$, ce qui est l'inégalité annoncée (le cas $a\, b = 0$ se traite directement). De la même façon, on obtient la deuxième inégalité en prenant $t = a/(a+b)$ dans l'inégalité : $\forall t \in [0,1]$, $g(t) = t^p + (1-t)^p \geq 2^{1-p}$, laquelle résulte de ce que le minimum de g sur $[0,1]$ est atteint en $t = 1/2$ (ici aussi, le cas $a + b = 0$ se traite directement).

Démontrons ensuite l'inégalité de Hölder. Lorsque $p = \infty$ et $q = 1$, on a $|f(x)\, g(x)| \leq \|f\|_\infty\, |g(x)|$ presque partout, ce qui fournit l'inégalité voulue après intégration sur X. On peut donc supposer que $p < \infty$ et $q < \infty$. Si $\|f\|_p\, \|g\|_q = 0$, c'est que l'un des deux facteurs est nul, soit $f = 0$ ou $g = 0$ pp, auquel cas le membre de gauche vaut 0 et l'inégalité est bien vérifiée, et bien évidemment elle l'est aussi si $\|f\|_p\, \|g\|_q = \infty$. On peut donc supposer que $\|f\|_p$ et $\|g\|_q$ appartiennent à $]0, \infty[$. Posons alors $a(x) = |f(x)|/\|f\|_p$ et $b(x) = |g(x)|/\|g\|_q$. D'après la propriété (a),

$$a(x)\, b(x) \leq \frac{|f(x)|^p}{p\, \|f\|_p^p} + \frac{|g(x)|^q}{q\, \|g\|_q^q}\,, \qquad \text{d'où} \qquad \int_X a(x)\, b(x)\, dx \leq \frac{1}{p} + \frac{1}{q} = 1\,,$$

ce qui prouve l'inégalité de Hölder puisque

$$\int_X |f(x)\, g(x)|\, dx = \|f\|_p\, \|g\|_q \int_X \frac{|f(x)|}{\|f\|_p}\, \frac{|g(x)|}{\|g\|_q}\, dx = \|f\|_p\, \|g\|_q \int_X a(x)\, b(x)\, dx\,.$$

Si donc $f \in \mathcal{L}^p(X)$ et $g \in \mathcal{L}^q(X)$, on a $\int_X |f(x)\, g(x)|\, dx < \infty$, ce qui prouve l'intégrabilité du produit fg, et la majoration de l'intégrale provient alors de la continuité de l'intégrale.

Pour l'inégalité de Minkowski, on traite encore le cas $p = \infty$ à part : puisque $|f| \leq \|f\|_\infty$ en dehors d'un ensemble négligeable A et que $|g| \leq \|g\|_\infty$ en dehors d'un ensemble négligeable B, on a $|f + g| \leq |f| + |g| \leq \|f\|_\infty + \|g\|_\infty$ en dehors de $A \cup B$, d'où $\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$. Si $p < \infty$, on écrit

$$\int_X |f(x) + g(x)|^p\, dx \leq \int_X |f(x)|\, |f(x) + g(x)|^{p-1}\, dx + \int_X |g(x)|\, |f(x) + g(x)|^{p-1}\, dx\,,$$

d'où en utilisant le conjugué q de p et l'inégalité de Hölder

$$\|f + g\|_p^p = \int_X |f(x) + g(x)|^p\, dx \leq \|f\|_p\, \||f + g|^{p-1}\|_q + \|g\|_p\, \||f + g|^{p-1}\|_q =$$
$$\leq \left(\|f\|_p + \|g\|_p\right) \|f + g\|_p^{p-1}\,.$$

Si $\|f+g\|_p = 0$ ou ∞, il n'y a rien à démontrer car dans ce deuxième cas, il résulte de la propriété (a) que $\int_X |f(x)|^p\,dx + \int_X |g(x)|^p\,dx \geq 2^{1-p} \int_X |(f+g)(x)|^p\,dx = \infty$, d'où $\|f\|_p + \|g\|_p = \infty$ aussi. Si donc $\|f+g\|_p \in]0,\infty[$, il suffit de diviser l'inégalité écrite ci-dessus par ce terme élevé à la puissance $p-1$.

Enfin, si f et $g \in \mathcal{L}^p(X)$ et $c \in \mathbb{C}$, alors les fonctions $f+g$ et cf appartiennent à $\mathcal{L}^p(X)$ puisque $\|f+g\|_p \leq \|f\|_p + \|g\|_p$ comme on vient de le voir, et que $\|cf\|_p = |c|\,\|f\|_p$. Il en résulte que $\mathcal{L}^p(X)$ est un espace vectoriel et que $f \mapsto \|f\|_p$ est une semi-norme sur cet espace. ∎

Remarque. La semi-norme $\|.\|_p$ n'est pas une norme, mais on peut caractériser les fonctions f vérifiant $\|f\|_p = 0$: d'après la proposition 3.5, ce sont les fonctions f telles que $|f|^p = 0$ presque partout, c'est-à-dire les fonctions nulles presque partout (ce qui est vrai aussi dans le cas $p = \infty$). Mais cela n'empêche pas de définir la convergence en moyenne d'ordre p, c'est-à-dire la convergence au sens de la semi-norme d'indice p, qui généralise la convergence en moyenne (correspondant au cas $p = 1$). Lorsque $p = \infty$, c'est la convergence uniforme sauf sur un ensemble négligeable. Comme dans le cas de la convergence en moyenne, il n'y a pas unicité de la limite, mais presque : si f est une limite en moyenne d'ordre p d'une suite (f_q), alors g en est aussi une limite si et seulement si $f = g$ presque partout. Pour avoir une vraie norme, et ainsi unicité des limites, on préfère souvent considérer les espaces vectoriels $\mathrm{L}^p(X) = \mathcal{L}^p(X)/\mathcal{N}(X)$ quotients de $\mathcal{L}^p(X)$ par le sous-espace $\mathcal{N}(X)$ des fonctions nulles presque partout dans X.

Le cas $p = 2$ est un peu à part car l'exposant 2 est son propre conjugué. Les éléments de $\mathcal{L}^2(X)$ s'appellent les fonctions de carré intégrable dans X, et la convergence pour la semi-norme d'indice 2 s'appelle convergence en moyenne quadratique. L'inégalité de Hölder s'appelle dans ce cas inégalité *de Cauchy-Schwarz*, et on l'écrit plus volontiers sous la forme

$$\left| \int_X f(x)\,\overline{g(x)}\,dx \right| \leq \|f\|_2\,\|g\|_2\,,$$

l'application $(f,g) \mapsto \int_X f(x)\,\overline{g(x)}\,dx$ étant une forme sesquilinéaire hermitienne sur $\mathcal{L}^2(X)$.

L'inégalité de Hölder permet de démontrer les résultats suivants.

Proposition 9.3. *Soient $1 \leq p < q \leq \infty$ deux exposants, et X une partie mesurable de $\mathbb{R}^n$.*
(a) *Pour toute fonction $f \in \mathcal{L}^p(X) \cap \mathcal{L}^q(X)$ et tout exposant $r \in]p,q[$, on a $f \in \mathcal{L}^r(X)$ avec $\|f\|_r \leq \|f\|_p^\alpha\,\|f\|_q^\beta$ en posant $\alpha = \frac{pq-pr}{qr-pr}$ et $\beta = \frac{qr-pq}{qr-pr}$ si $q < \infty$, et $\alpha = p/r$ et $\beta = 1-(p/r)$ si $q = \infty$.*
(b) *Si $|X| < \infty$, alors $\mathcal{L}^q(X) \subset \mathcal{L}^p(X)$.*

Démonstration. Si $q < \infty$, les exposants P et Q définis par

$$\frac{1}{P} = \frac{q-r}{q-p} \qquad \text{et} \qquad \frac{1}{Q} = \frac{r-p}{q-p}$$

sont solution du système $(1/P)+(1/Q)=1$ et $(p/P)+(q/Q)=r$. L'inégalité de Hölder permet donc d'écrire pour toute f mesurable sur X

$$\|f\|_r^r = \int_X |f(x)|^{p/P}\,|f(x)|^{q/Q}\,dx \leq \|(|f|^{p/P})\|_P\,\|(|f|^{q/Q})\|_Q = \|f\|_p^{r\alpha}\,\|f\|_q^{r\beta}$$

(et de même si $q=\infty$, $\int_X |f(x)|^r\,dx \leq \|f\|_\infty^{r-p}\int |f(x)|^p\,dx$). La propriété (a) découle de cette inégalité élevée à la puissance $1/r$.

Par ailleurs, comme $q/p>1$, l'inégalité de Hölder permet d'écrire pour toute f mesurable sur X

$$\int_X |f(x)|^p\,dx = \int_X |f(x)|^p\,\mathbb{1}_X(x)\,dx \leq \|(|f|^p)\|_{q/p}\,\|\mathbb{1}_X\|_r = \|f\|_q^p\,|X|^{1/r},$$

en notant r l'exposant conjugué de q/p. Si donc $f\in\mathcal{L}^q(X)$ et $|X|<\infty$, on a $f\in\mathcal{L}^p(X)$, ce qui prouve la propriété (b). ∎

Beaucoup de théorèmes établis dans le cas des fonctions intégrables peuvent s'étendre au cas des fonctions de puissance $p^{\text{ième}}$ intégrable. C'est notamment le cas des énoncés suivants.

Théorème 9.4 (de complétude, d'Égorov et de densité). *Pour tout exposant $1\leq p\leq\infty$, l'espace $\mathcal{L}^p(X)$ muni de la semi-norme $f\mapsto\|f\|_p$ est complet. Plus précisément, si $(f_q)_{q\in\mathbb{N}}$ est une suite de Cauchy pour cette semi-norme, il existe $f\in\mathcal{L}^p(X)$ telle que $\lim_{q\to\infty}\|f_q-f\|_p=0$, et il existe une sous-suite qui converge presque partout vers f. Enfin, pour $p<\infty$ et $f\in\mathcal{L}^p(\mathbb{R}^n)$, il existe une suite (f_q) de fonctions continues à support compact qui converge vers f en moyenne d'ordre p.*

Théorème 9.5. *Soient $1\leq p\leq\infty$ un exposant, $f\in\mathcal{L}^p(\mathbb{R}^n)$ et $g\in\mathcal{L}^1(\mathbb{R}^n)$. Alors le produit de convolution*

$$f*g(x)=\int f(x-y)\,g(y)\,dy=\int f(z)\,g(x-z)\,dz$$

est bien défini pour presque tout $x\in\mathbb{R}^n$. En outre, la fonction $f*g$ appartient à $\mathcal{L}^p(\mathbb{R}^n)$ et on a $\|f*g\|_p\leq\|f\|_p\|g\|_1$. De plus, si $p<\infty$ et si $g_\varepsilon(x)=(2\pi)^{-n/2}\,\varepsilon^{-n}\,e^{-|x|^2/2\varepsilon^2}$ est la gaussienne introduite à la fin du paragraphe 8, alors pour toute $f\in\mathcal{L}^p(\mathbb{R}^n)$, le produit de convolution $f*g_\varepsilon$ tend vers f en moyenne d'ordre p quand ε tend vers 0.

Démonstration du théorème 9.4. Dans le cas $p=\infty$, la complétude provient de ce que toute suite de Cauchy (f_q) converge uniformément en dehors de la réunion (dénombrable) des $A_{q,r}=\{x\in X\,;\,|f_q(x)-f_r(x)|>\|f_q-f_r\|_\infty\}$ qui sont tous de mesure nulle, et que la limite f ainsi obtenue est clairement bornée en dehors des $A_{q,r}$. On achève la démonstration en remarquant que $\|f_q-f\|_\infty$ tend vers 0 quand q tend vers l'infini puisque $\sup_{x\notin\cup A_{q,r}}|f_q(x)-f(x)|$ tend vers 0.

Pour $p < \infty$, soit (f_q) une suite de Cauchy de $\mathcal{L}^p(X)$. On peut en extraire une sous-suite (g_q) telle que $\|g_{q+1} - g_q\|_p \leq 2^{-2q}$, et comme (f_q) est de Cauchy dans $\mathcal{L}^p(X)$, il suffit de montrer que (g_q) est convergente dans $\mathcal{L}^p(X)$. Si on pose $A_q = \{ x \in X : |g_{q+1}(x) - g_q(x)| \geq 2^{-q} \}$, on a

$$|A_q| \leq 2^{pq} \|g_{q+1} - g_q\|_p^p \leq 2^{-pq}.$$

L'ensemble des points x où la suite $g_q(x)$ ne converge pas est contenu dans $\cup_{q \geq N} A_q$ quel que soit $N \in \mathbb{N}$, et cette réunion peut être rendue de mesure aussi petite que l'on veut en prenant N assez grand. La suite (g_q) converge donc presque partout vers une fonction f et nous allons montrer que $f \in \mathcal{L}^p(X)$ et que (g_q) converge vers f dans $\mathcal{L}^p(X)$.

La fonction mesurable positive $h_q(x) = \sum_{r \geq q} |g_{r+1}(x) - g_r(x)|$ majore les fonctions $|g_s - g_q|$ pour tout $s \geq q$, et par la propriété (c) de la proposition 5.4 (convergence croissante),

$$\begin{aligned} \|h_q\|_p^p = \lim_{s \to \infty} \int_X \Big(\sum_{r=q}^{s} |g_{r+1}(x) - g_r(x)| \Big)^p dx = \lim_{s \to \infty} \Big\| \sum_{r=q}^{s} |g_{r+1} - g_r| \Big\|_p^p \leq \\ \leq \lim_{s \to \infty} \Big(\sum_{r=q}^{s} \|g_{r+1} - g_r\|_p \Big)^p \leq \Big(\frac{4}{3} 2^{-2q} \Big)^p \end{aligned}$$

grâce à l'inégalité de Minkowski, si bien que ${h_q}^p$ est intégrable sur X et que par convergence dominée on a

$$\int_X |f(x) - g_q(x)|^p \, dx = \lim_{s \to \infty} \int_X |g_s(x) - g_q(x)|^p \, dx \leq \|h_q\|_p^p \leq \Big(\frac{4}{3} 2^{-2q} \Big)^p$$

d'où $\|f - g_q\|_p \leq (4/3)\, 2^{-2q}$, ce qui prouve à la fois que $f \in \mathcal{L}^p(X)$ et que (g_q) converge vers f en moyenne d'ordre p.

Pour montrer que toute fonction $f \in \mathcal{L}^p(\mathbb{R}^n)$ peut être approchée par une fonction continue à support compact, on montre d'abord qu'elle peut être approchée par une fonction bornée s'annulant en dehors d'un pavé. Il suffit en effet pour cela de poser $f_q = \mathbb{1}_{A_q} f$ en notant A_q l'ensemble mesurable $\{ x \in [-q, q]^n ; |f(x)| \leq q \}$. Alors les fonctions f_q sont bornées (par q) et nulles en dehors d'un pavé, et les fonctions $|f_q - f|^p$ sont majorées par la fonction intégrable $|f|^p$ et tendent vers 0 presque partout, si bien que par convergence dominée, il existe pour q assez grand des fonctions f_q arbitrairement proches de f en moyenne d'ordre p.

On montre ensuite que toute fonction f bornée et s'annulant en dehors d'un pavé P peut être approchée par une fonction continue à support compact. Pour cela, on commence par considérer la fonction $h : \mathbb{R} \to \mathbb{R}$ nulle en dehors de $[-1, 1]$, valant $h(t) = 1 + t$ sur $[-1, 0]$ et $h(t) = 1 - t$ sur $[0, 1]$, ce qui permet de construire $h_0(x) = \prod_{j=1}^{n} h(x_j)$ puis $h_q(x) = 2^{nq} h_0(2^q x)$ qui sont des fonctions continues à support compact d'intégrales égales à 1. Alors la fonction $f * h_q$ est continue à support compact puisqu'elle peut s'écrire $f * h_q(x) = \int f(z)\, h_q(x - z)\, dz$, ce qui définit une fonction continue d'après le théorème 7.1, et que cette fonction s'annule en dehors du pavé $P + [-1, 1]^n$. De plus, $f * h_q$ vérifie

$$\|f * h_q\|_\infty \leq \sup_{x \in \mathbb{R}^n} |f * h_q(x)| \leq \|f\|_\infty \int h_q(y)\, dy = \|f\|_\infty ,$$

et grâce à l'inégalité de la proposition 9.3 (a), on a donc

$$\|f - (f * h_q)\|_p \leq \|f - (f * h_q)\|_1^{1/p} \, \|f - (f * h_q)\|_\infty^{1-(1/p)} \leq (2\, \|f\|_\infty)^{1-(1/p)} \, \|f - (f * h_q)\|_1^{1/p} .$$

Pour conclure, il ne reste plus qu'à observer que $f * h_q$ converge vers f en moyenne, ce qui se démontre comme à la page 59 pour la convolution par les gaussiennes g_ε en approchant la fonction intégrable f (pour la convergence en moyenne) par une fonction continue à support compact. ∎

Démonstration du théorème 9.5. Le passage de l'une à l'autre formule se fait par le changement de variables $y = \phi(z) = x - z$ dont le jacobien vaut $[\![J_\phi(z)]\!] = |(-1)^n|$. Si $p = \infty$, les affirmations sont immédiates puisque $|f(x - y)| \leq \|f\|_\infty$ pour presque tout $y \in \mathbb{R}^n$, et si $p = 1$, le résultat a été établi comme application du théorème de Fubini.

Enfin, si $p \in]1, \infty[$, on utilise le réel q conjugué de p ainsi que l'inégalité de Hölder pour écrire

$$\int |f(x-y)|\,|g(y)|\,dy = \int |f(x-y)|\,|g(y)|^{1/p}\,|g(y)|^{1/q}\,dy \le \Big(\int |f(x-y)|^p\,|g(y)|\,dy\Big)^{1/p}\Big(\int |g(y)|\,dy\Big)^{1/q},$$

ce qui implique que

$$\Big(\int |f(x-y)|\,|g(y)|\,dy\Big)^p \le \|g\|_1^{p/q}\int |f(x-y)|^p\,|g(y)|\,dy\,.$$

On en déduit comme dans le cas $p = 1$ par Fubini que $\int |f(x-y)|\,|g(y)|\,dy < \infty$ pour presque tout $x \in \mathbb{R}^n$, puis que

$$\|f*g\|_p^p = \int \Big|\int f(x-y)\,g(y)\,dy\Big|^p dx \le \int \Big(\int |f(x-y)|\,|g(y)|\,dy\Big)^p dx \le \|g\|_1^{(p/q)+1}\,\|f\|_p^p = (\|f\|_p\,\|g\|_1)^p,$$

ce qui établit le résultat principal.

Pour étudier la convolution avec la gaussienne g_ε, soit $\delta > 0$. D'après le théorème 9.4, il existe une fonction h continue à support compact telle que $\|h - f\|_p \le \delta/3$, et on raisonne alors comme dans le cas $p = 1$ en utilisant l'inégalité de Hölder et la majoration de $\|f * g\|_p$ démontrée ci-dessus. En effet, puisque $\|g_\varepsilon\|_1 = 1$, on a

$$\begin{aligned}\|(h*g_\varepsilon) - h\|_p^p &\le \|(h*g_\varepsilon) - h\|_1\,\|(h*g_\varepsilon) - h\|_\infty^{p-1} \le \\ &\le \|(h*g_\varepsilon) - h\|_1\Big(\|h*g_\varepsilon\|_\infty + \|h\|_\infty\Big)^{p-1} \le \Big(2\,\|h\|_\infty\Big)^{p-1}\,\|(h*g_\varepsilon) - h\|_1\,,\end{aligned}$$

ce qui peut être rendu inférieur à $(\delta/3)^p$ pour $\varepsilon > 0$ assez petit puisque $\|(h*g_\varepsilon) - h\|_1$ tend vers 0 avec ε. Comme $\|(f*g_\varepsilon) - (h*g_\varepsilon)\|_p = \|(f-h)*g_\varepsilon\|_p \le \|h - f\|_p$, on en déduit la majoration

$$\|(f*g_\varepsilon) - f\|_p \le \|(f*g_\varepsilon) - (h*g_\varepsilon)\|_p + \|(h*g_\varepsilon) - h\|_p + \|h - f\|_p \le 2\,\|h - f\|_p + (\delta/3) \le \delta$$

qui prouve le théorème. ∎

Exemple : la transformée de Fourier d'une fonction de carré intégrable. Il s'agit ici de définir une application $f \mapsto \widehat{f}$ sur $\mathcal{L}^2(\mathbb{R}^n)$ prolongeant la transformation de Fourier des fonctions intégrables, et d'en étudier quelques propriétés.

(a) *La formule de Plancherel pour les fonctions intégrables : pour toute fonction $f \in \mathcal{L}^1(\mathbb{R}^n) \cap \mathcal{L}^2(\mathbb{R}^n)$, on a $\widehat{f} \in \mathcal{L}^2(\mathbb{R}^n)$ avec $\|\widehat{f}\|_2 = (2\pi)^{n/2}\,\|f\|_2$.* En effet, puisque f est intégrable et $\widehat{f}$ bornée, on peut calculer l'intégrale de $|\widehat{f}|^2$ par le théorème de Fubini à condition de multiplier par la gaussienne g^ε pour assurer l'intégrabilité. Nous obtenons ainsi

$$\begin{aligned}\int |\widehat{f}(\xi)|^2\,e^{-\varepsilon^2|\xi|^2/2}\,d\xi &= \int e^{-ix\cdot\xi}\,f(x)\,e^{iy\cdot\xi}\,\overline{f(y)}\,e^{-\varepsilon^2|\xi|^2/2}\,dx\,dy\,d\xi = \\ &= (2\pi)^n\int f(x)\,\overline{f(y)}\,g_\varepsilon(x-y)\,dx\,dy = (2\pi)^n\int f(x)\,\overline{f*g_\varepsilon(x)}\,dx\,,\end{aligned}$$

en notant $(2\pi)^n\,g_\varepsilon$ la transformée de Fourier de la gaussienne g^ε comme au paragraphe précédent. Il résulte alors de l'inégalité de Cauchy-Schwarz que

$$\left| \int |\widehat{f}(\xi)|^2 \, e^{-\varepsilon^2|\xi|^2/2} \, d\xi - (2\pi)^n \int |f(x)|^2 \, dx \right| =$$

$$= (2\pi)^n \left| \int f(x) \left(\overline{f*g_\varepsilon(x)} - \overline{f(x)} \right) dx \right| \leq (2\pi)^n \, \|f\|_2 \, \|f*g_\varepsilon - f\|_2 \, ,$$

et puisque $f*g_\varepsilon$ converge en moyenne quadratique vers f d'après le théorème 9.5, on peut en conclure en passant à la limite pour $\varepsilon \to 0$ que $\|\widehat{f}\|_2 = (2\pi)^{n/2} \, \|f\|_2$.

(b) *La transformée de Fourier d'une fonction de carré intégrable.* Soient $f \in \mathcal{L}^2(\mathbb{R}^n)$, et (f_p) une suite d'éléments de $\mathcal{L}^1(\mathbb{R}^n) \cap \mathcal{L}^2(\mathbb{R}^n)$ qui converge en moyenne quadratique vers f (il existe toujours de telles suites : par exemple la suite de fonctions continues à support compact fournie par le théorème 9.4, ou de façon plus constructive, la suite définie par $f_p = \mathbb{1}_{[-p,p]^n} f$). La suite (f_p) est de Cauchy en moyenne quadratique car convergente, donc la suite $(\widehat{f_p})$ aussi puisque $\|\widehat{f_p} - \widehat{f_q}\|_2 = (2\pi)^{n/2} \, \|f_p - f_q\|_2$ d'après la formule de Plancherel du (a), et l'existence d'une limite $\widehat{f} \in \mathcal{L}^2(\mathbb{R}^n)$ de la suite $(\widehat{f_p})$ résulte donc du théorème de complétude. Pour que cette définition soit cohérente, il faut encore vérifier que la limite $\widehat{f}$ obtenue ne dépend que de f, mais pas de la suite (f_p) utilisée. Or, si (g_p) est une autre suite d'éléments de $\mathcal{L}^1(\mathbb{R}^n) \cap \mathcal{L}^2(\mathbb{R}^n)$ tendant vers f en moyenne quadratique, la suite (h_p) définie par $h_{2p} = f_p$, $h_{2p+1} = g_p$ converge encore en moyenne quadratique vers f, ce qui permet d'affirmer que la suite $(\widehat{h_p})$ converge en moyenne quadratique, c'est-à-dire que les suites $(\widehat{f_p})$ et $(\widehat{g_p})$ ont même limite (à l'égalité presque partout près).

Ce mode de construction permet en outre de démontrer les propriétés de la transformation de Fourier par passage à la limite. C'est ce que nous proposons dans un troisième point.

(c) *Pour toutes fonctions f et $g \in \mathcal{L}^2(\mathbb{R}^n)$ et $h \in \mathcal{L}^1(\mathbb{R}^n)$, on a*

(Formule de Plancherel) $$\int \widehat{f}(\xi) \, \overline{\widehat{g}(\xi)} \, d\xi = (2\pi)^n \int f(x) \, \overline{g(x)} \, dx \, ,$$

(Inversion de Fourier) $$\widehat{\widehat{f}}(x) = (2\pi)^n \, f(-x) \qquad \text{pour presque tout } x \in \mathbb{R}^n \, ,$$

(Convolution et produit) $$\widehat{f*h} = \widehat{f}\,\widehat{h} \qquad \text{et} \qquad \widehat{fg} = (2\pi)^{-n} \, (\widehat{f} * \widehat{g}) \qquad \text{presque partout.}$$

En effet, si $f \in \mathcal{L}^2(\mathbb{R}^n)$ et si (f_p) est une suite d'éléments de $\mathcal{L}^1(\mathbb{R}^n) \cap \mathcal{L}^2(\mathbb{R}^n)$ qui converge vers f en moyenne quadratique, on peut écrire

$$\|\widehat{f}\|_2 = \lim_{p \to \infty} \|\widehat{f_p}\|_2 = \lim_{p \to \infty} (2\pi)^n \, \|f_p\|_2 = (2\pi)^n \, \|f\|_2 \, ,$$

ce qui étend la formule de Plancherel du (a) aux fonctions de carré intégrable. Pour f et $g \in \mathcal{L}^2(\mathbb{R}^n)$, les produits $\widehat{f}\,\overline{\widehat{g}}$ et $f\,\overline{g}$ sont intégrables d'après le théorème 9.2 (b). Par ailleurs, pour tous nombres complexes z et w, on a

$$4 \, z \, \overline{w} = |z + w|^2 - |z - w|^2 + i \, |z + i \, w|^2 - i \, |z - i \, w|^2 \, ,$$

et en appliquant cette identité aux nombres $z = \widehat{f}(\xi)$ et $w = \widehat{g}(\xi)$ puis en intégrant en ξ, on obtient

$$\begin{aligned}4\int \widehat{f}(\xi)\,\overline{\widehat{g}(\xi)}\,d\xi &= \|\widehat{f}+\widehat{g}\|_2^2 - \|\widehat{f}-\widehat{g}\|_2^2 + i\,\|\widehat{f}+i\,\widehat{g}\|_2^2 - i\,\|\widehat{f}-i\,\widehat{g}\|_2^2 = \\ &= (2\pi)^n\left(\|f+g\|_2^2 - \|f-g\|_2^2 + i\,\|f+i\,g\|_2^2 - i\,\|f-i\,g\|_2^2\right) = \\ &= 4\,(2\pi)^n \int f(x)\,\overline{g(x)}\,dx\,,\end{aligned}$$

ce qui prouve la formule de Plancherel annoncée.

Prouvons ensuite la formule d'inversion de Fourier pour une fonction f de carré intégrable. Pour tout $\delta > 0$, il existe une fonction $h \in \mathcal{L}^1(\mathbb{R}^n) \cap \mathcal{L}^2(\mathbb{R}^n)$ telle que $\|f-h\|_2 \le \delta/2$, puis pour $\varepsilon > 0$ assez petit, il résulte du théorème 9.5 que $\|h-(h*g_\varepsilon)\|_2 \le \delta/2$, ce qui montre que f est limite en moyenne quadratique d'une suite (f_p) de fonctions de la forme $f_p = h_p * g_{\varepsilon_p}$ avec $h_p \in \mathcal{L}^1(\mathbb{R}^n)\cap\mathcal{L}^2(\mathbb{R}^n)$. Comme h_p et g_{ε_p} sont intégrables, on peut écrire

$$\widehat{f_p}(\xi) = \widehat{h_p}(\xi)\,\widehat{g_{\varepsilon_p}}(\xi) = \widehat{h_p}(\xi)\,e^{-\varepsilon_p^2|\xi|^2/2}\,,$$

et comme $\widehat{h_p}$ est bornée comme transformée de Fourier d'une fonction intégrable, il en résulte que f_p et $\widehat{f_p}$ sont toutes deux intégrables. On a démontré la formule d'inversion dans ce cas à la fin du paragraphe 8, si bien que l'on a $\widehat{\widehat{f_p}}(x) = (2\pi)^n f_p(-x)$ pour presque tout $x \in \mathbb{R}^n$. Comme la suite (f_p) converge en moyenne quadratique vers f, la suite $(\widehat{\widehat{f_p}})$ converge en moyenne quadratique vers $\widehat{\widehat{f}}$, et en extrayant successivement deux sous-suites, ces suites convergent aussi presque partout. La formule d'inversion pour f en résulte.

Enfin pour établir les dernières formules, on commence par remarquer que si $f,\ g \in \mathcal{L}^2(\mathbb{R}^n)$ et $h \in \mathcal{L}^1(\mathbb{R}^n)$, alors $f*h$ et $\widehat{f}\,\widehat{h} \in \mathcal{L}^2(\mathbb{R}^n)$, et $fg \in \mathcal{L}^1(\mathbb{R}^n)$ et $\widehat{f}*\widehat{g} \in \mathcal{L}^\infty(\mathbb{R}^n)$ d'après le théorème 9.2 (b). Si (f_p) est une suite d'éléments de $\mathcal{L}^1(\mathbb{R}^n)\cap\mathcal{L}^2(\mathbb{R}^n)$ convergeant en moyenne quadratique vers f, on obtient la formule $\widehat{f*h} = \widehat{f}\,\widehat{h}$ par passage à la limite en moyenne quadratique dans l'identité $\widehat{f_p*h} = \widehat{f_p}\,\widehat{h}$. Et pour la dernière formule, on commence par choisir deux suites $(\widehat{f_p})$ et $(\widehat{g_p})$ d'éléments de $\mathcal{L}^1(\mathbb{R}^n)\cap\mathcal{L}^2(\mathbb{R}^n)$ qui convergent en moyenne quadratique respectivement vers $\widehat{f}$ et $\widehat{g}$. Alors

$$(2\pi)^n\,\widehat{\widehat{f_p}*\widehat{g_p}}(x) = (2\pi)^n\,\widehat{\widehat{f_p}}(x)\widehat{\widehat{g_p}}(x) = (2\pi)^n\,(f_p\,g_p)(-x) = \widehat{\widehat{f_p\,g_p}}(x)\,,$$

d'où $\widehat{f_p\,g_p} = (2\pi)^n\,(\widehat{f_p}*\widehat{g_p})$ presque partout. Pour obtenir la formule $\widehat{fg} = (2\pi)^n\,(\widehat{f}*\widehat{g})$, il ne reste plus alors qu'à remarquer que

$$\begin{aligned}&\|\widehat{f_p\,g_p} - \widehat{fg}\|_\infty \le \|f_p\,g_p - fg\|_1 \le \|f_p - f\|_2\,\|g_p\|_2 + \|f\|_2\,\|g_p - g\|_2 \\ \text{et}\quad &\|\widehat{f_p}*\widehat{g_p} - \widehat{f}*\widehat{g}\|_\infty \le \|\widehat{f_p} - \widehat{f}\|_2\,\|\widehat{g_p}\|_2 + \|\widehat{f}\|_2\,\|\widehat{g_p} - \widehat{g}\|_2\,,\end{aligned}$$

ce qui prouve que les suites $(\widehat{f_p\,g_p})$ et $(\widehat{f_p}*\widehat{g_p})$ convergent (essentiellement) uniformément vers $\widehat{fg}$ et $\widehat{f}*\widehat{g}$.

Exercices

9.1. Soient $f : \mathbb{R}^n \to \mathbb{C}$ une fonction mesurable, pour tout $t \ge 0$, $m(t) = |A(t)|$ la mesure de l'ensemble $A(t) = \{\,x \in \mathbb{R}^n\,;\ |f(x)| > t\,\}$, et p un réel ≥ 1.

(a) Montrer l'*inégalité de Chebyshev* : $\forall t > 0$, $m(t) \le \left(\|f\|_p/t\right)^p$.

(b) Montrer l'égalité $\|f\|_p^p = p\int_{[0,\infty[} t^{p-1}\,m(t)\,dt$.

9.2. Donner un exemple de fonction $f \in \mathcal{L}^2(\mathbb{R})$ qui n'est pas intégrable.

9.3. Le but de cet exercice est de proposer une autre démonstration de l'inégalité de Hölder dans le cas où $p = 2$. Soient f et $g : \mathbb{R}^n \to \mathbb{R}$ deux fonctions de carré intégrable, c'est-à-dire f et $g \in \mathcal{L}^2(\mathbb{R}^n)$.

(a) Justifier de façon élémentaire l'inégalité $|f(x)\, g(x)| \leq (f(x)^2 + g(x)^2)/2$, et en déduire que le produit fg est intégrable.

(b) En recherchant le minimum de la fonction $h(t) = \int (f(x) + t\, g(x))^2\, dx$, montrer que l'on a l'*inégalité de Cauchy-Schwarz* : $\left| \int f(x)\, g(x)\, dx \right| \leq \|f\|_2\, \|g\|_2$.

9.4. Soient p, q et r trois exposants vérifiant $1 \leq p < q < r \leq \infty$. Montrer que l'on a $\mathcal{L}^q(\mathbb{R}^n) \subset \mathcal{L}^p(\mathbb{R}^n) + \mathcal{L}^r(\mathbb{R}^n) = \{ f = g + h \,;\, g \in \mathcal{L}^p(\mathbb{R}^n) \text{ et } h \in \mathcal{L}^r(\mathbb{R}^n) \}$ (on pourra distinguer les points $x \in \mathbb{R}^n$ pour lesquels $|f(x)| \leq 1$ de ceux pour lesquels $|f(x)| > 1$).

9.5. Soit $f : \mathbb{R}^n \to \mathbb{C}$ une fonction intégrable bornée. Montrer que l'on a $\|f\|_\infty = \lim_{p\to\infty} \|f\|_p$ (on pourra considérer les normes d'indice p de la fonction $g_t(x) = f(x)/t$ avec $t > \|f\|_\infty$ ou $t < \|f\|_\infty$). Vérifier cette propriété sur un exemple en calculant les normes d'indice p de la gaussienne $g(x) = e^{-|x|^2/2}$.

9.6. Dans cet exercice, toutes les fonctions sont supposées mesurables, ou même mesurables positives si l'on préfère.

(a) Établir que si $(1/p) + (1/q) = (1/r)$, $f \in \mathcal{L}^p(\mathbb{R}^n)$ et $g \in \mathcal{L}^q(\mathbb{R}^n)$, alors $fg \in \mathcal{L}^r(\mathbb{R}^n)$ avec $\|fg\|_r \leq \|f\|_p\, \|g\|_q$.

(b) En déduire que si $\sum_{j=1}^N (1/p_j) = 1$ et $f_j \in \mathcal{L}^{p_j}(\mathbb{R}^n)$, alors $\prod_{j=1}^N f_j$ est intégrable avec $\|\prod_{j=1}^N f_j\|_1 \leq \prod_{j=1}^N \|f_j\|_{p_j}$.

(c) Montrer que si $(1/p) + (1/q) = 1 + (1/r)$, $f \in \mathcal{L}^p(\mathbb{R}^n)$ et $g \in \mathcal{L}^q(\mathbb{R}^n)$, alors le produit de convolution $f * g$ est bien défini et appartient à $\mathcal{L}^r(\mathbb{R}^n)$. Indication : en notant p' et q' les exposants conjugués de p et q, prouver que l'on a $(1/p') + (1/q') + (1/r) = 1$ et

$$|f(x-y)\, g(y)| = |f(x-y)|^{p/q'} \Big(|f(x-y)|^p\, |g(y)|^q \Big)^{1/r} |g(y)|^{q/p'} .$$

9.7. Soient $p \in [1, \infty[$ un exposant, $f \in \mathcal{L}^p(\mathbb{R}^n)$ une fonction à valeurs complexes, $x_0 \in \mathbb{R}^n$ et $f_\varepsilon(x) = f(x - \varepsilon\, x_0)$. Montrer que la fonction f_ε converge en moyenne d'ordre p vers f lorsque ε tend vers 0 (on commencera par le faire dans le cas d'une fonction f continue à support compact).

9.8. Soient $1 \leq p \leq q < \infty$ deux exposants, et $f : \mathbb{R}^{n+m} \to \mathbb{R}_+$ une fonction mesurable positive. On notera x les variables dans $\mathbb{R}^n$ et y les variables dans $\mathbb{R}^m$, et on cherche à démontrer l'inégalité

$$\left(\int \Big(\int f(x,y)^p\, dx \Big)^{q/p} dy \right)^{1/q} \leq \left(\int \Big(\int f(x,y)^q\, dy \Big)^{p/q} dx \right)^{1/p} .$$

On posera $r = q/p$.

(a) Montrer l'identité

$$\int \Big(\int f(x,y)^p\, dx \Big)^r dy = \int \left(\int f(z,y)^p \Big(\int f(x,y)^p\, dx \Big)^{r-1} dy \right) dz .$$

(b) En appliquant l'inégalité de Hölder à l'intégrale en y du membre de droite ci-dessus, montrer l'inégalité annoncée dans le cas où f est bornée et s'annule en dehors d'un pavé P.

(c) En déduire le résultat pour toute f mesurable positive.

9.9. Soit (f_p) une suite de fonctions continues à support compact dans $\mathbb{R}^n$ qui converge au sens de la semi-norme $\|\,.\,\|_\infty$ vers une fonction $f : \mathbb{R}^n \to \mathbb{C}$. Montrer que f est (égale presque partout à une fonction) continue et tend vers 0 à l'infini. Réciproquement, montrer que toute fonction $f : \mathbb{R}^n \to \mathbb{C}$ continue tendant vers 0 à l'infini est limite au sens de la semi-norme $\|\,.\,\|_\infty$ d'une suite de fonctions continues à support compact.

9.10. Soient $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction mesurable et $f : \mathbb{R} \to \mathbb{C}$ une fonction intégrable.

(a) Montrer que la formule $g(y) = \int e^{i\varphi(x)y} f(x)\,dx$ définit une fonction continue bornée sur $\mathbb{R}$.

(b) Soient (f_p) une suite de fonctions intégrables qui converge en moyenne vers f, et $g_p(y) = \int e^{i\varphi(x)y} f_p(x)\,dx$. Montrer que la suite (g_p) converge uniformément vers g.

(c) Montrer que si p et q sont des exposants conjugués, $\varphi \in \mathcal{L}^p(\mathbb{R})$ et $f \in \mathcal{L}^q(\mathbb{R})$, alors la fonction g est dérivable sur $\mathbb{R}$.

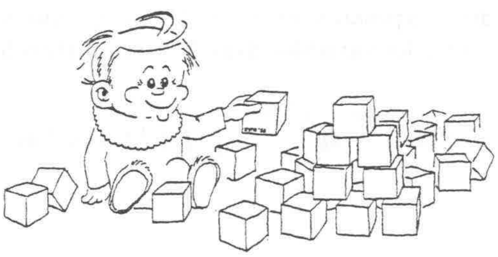
9.11. Soit I un intervalle ouvert de $\mathbb{R}$.

(a) Montrer que toute fonction $f \in \mathcal{L}^2(I)$ est intégrable sur tout intervalle compact $[a, b] \subset I$.

*Dans la suite de l'exercice, on dira que la fonction $F : I \to \mathbb{C}$ appartient à l'*espace de Sobolev $\mathrm{H}^1(I)$ *s'il existe $f \in \mathcal{L}^2(I)$ et $x_0 \in I$ tels que $F \in \mathcal{L}^2(I)$ et $F(x) = F(x_0) + \int_{x_0}^{x} f(t)\,dt$. On pourra utiliser les résultats de l'exercice* **7.8** *de la page 54.*

(b) Montrer que si $F \in \mathrm{H}^1(I)$, alors $|F(x)|$ tend vers des limites lorsque x tend vers les bornes de l'intervalle I.

(c) Montrer que le produit de deux éléments de $\mathrm{H}^1(I)$ est encore dans $\mathrm{H}^1(I)$.



10. Séries de Fourier

Dans les paragraphes précédents, nous avons étudié la transformation de Fourier qui permet de représenter une fonction f comme la superposition d'exponentielles $x \mapsto e^{ix\cdot\xi}$, la moyenne en ξ étant calculée avec un poids $(2\pi)^{-n}\widehat{f}(\xi)$ que l'on sait déterminer à partir de la fonction f. Dans ce paragraphe, nous étudierons un analogue discret de ce type de représentation, et pour ne pas trop compliquer les choses, nous nous limiterons à la dimension $n = 1$.

Le nombre réel $\omega > 0$ étant fixé, on considère les puissances entières $e^{ip\omega t}$, pour $p \in \mathbb{Z}$, de la fonction oscillante $t \mapsto e^{i\omega t}$. Puis on appelle *polynôme trigonométrique* (de degré $\leq q$) toute combinaison linéaire $\sum_{|p|\leq q} c_p\, e^{ip\omega t}$ à coefficients complexes de ces puissances, et on appelle *série trigonométrique* toute somme infinie $\sum_{p\in\mathbb{Z}} c_p\, e^{ip\omega t}$, dont on définira la convergence comme celle (en un sens à préciser) de la suite de fonctions (S_q) formée des sommes partielles symétriques $S_q(t) = \sum_{|p|\leq q} c_p\, e^{ip\omega t}$. On rappelle que l'on appelle fonction *périodique* de période T, ou fonction *T-périodique*, toute fonction f vérifiant $f(t+T) = f(t)$ pour tout $t \in \mathbb{R}$. La fonction $t \mapsto e^{i\omega t}$ est périodique de période $T = 2\pi/\omega$, ce qui implique la même propriété pour les polynômes et séries trigonométriques, et on a le résultat élémentaire suivant.

Proposition 10.1. *Soit $(c_p)_{p\in\mathbb{Z}}$ une suite de nombres complexes telle que $\sum_{p\in\mathbb{Z}} |c_p| < \infty$. Alors la suite (S_q) définie par $S_q(t) = \sum_{|p|\leq q} c_p\, e^{ip\omega t}$ converge uniformément vers une fonction continue T-périodique f, et les coefficients c_p sont reliés à cette fonction par la formule $c_p = \frac{1}{T}\int_0^T e^{-ip\omega t} f(t)\, dt$.*

Démonstration. Posons $u_0 = c_0$ et pour $p > 0$, $u_p(t) = c_p\, e^{ip\omega t} + c_{-p}\, e^{-ip\omega t}$. La convergence de la suite (S_q) équivaut alors à celle de la série de fonctions de terme général u_p, et comme $|u_p(t)| \leq |c_p| + |c_{-p}|$, on voit que cette série converge normalement, donc uniformément. Les u_p étant des fonctions continues T-périodiques, il en est de même, par convergence uniforme, de la somme f de la série, et on a

$$\frac{1}{T}\int_0^T e^{-ip\omega t} f(t)\, dt = \lim_{q\to\infty} \frac{1}{T}\int_0^T e^{-ip\omega t} S_q(t)\, dt$$

avec $$\frac{1}{T}\int_0^T e^{-ip\omega t} S_q(t)\, dt = \sum_{|r|\leq q} c_r\, \frac{1}{T}\int_0^T e^{-ip\omega t}\, e^{ir\omega t}\, dt\,.$$

La dernière intégrale écrite vaut T si $r = p$, et vaut 0 si $r \neq p$ car dans ce dernier cas la fonction $t \mapsto e^{i(r-p)\omega t}$ possède une primitive T-périodique. On a donc $\frac{1}{T}\int_0^T e^{-ip\omega t} S_q(t)\, dt = c_p$ pour tout $q \geq |p|$, d'où le résultat annoncé. ∎

Remarque. Par périodicité et invariance par translation, l'intégrale sur $[0, T]$ utilisée ici pourrait être remplacée par l'intégrale sur n'importe quel intervalle $[a, a+T]$ de longueur T. Cette remarque est valable chaque fois que l'on intègre une fonction T-périodique, et nous l'utiliserons plus loin.

Dans le reste du paragraphe, nous allons examiner réciproquement à quelles conditions une fonction $f : \mathbb{R} \to \mathbb{C}$ peut être représentée par une série trigonométrique. Il est bien clair que l'on ne pourra représenter ainsi que des fonctions T-périodiques, et c'est pourquoi la fonction f sera désormais supposée T-périodique. On écrira $f \in \mathcal{L}^1[0, T]$ si la fonction f est intégrable sur $[0, T]$, et on introduit les notions suivantes.

Définition 10.2. *Soit $f \in \mathcal{L}^1[0, T]$ une fonction T-périodique. Alors on appelle :*
(a) Coefficients de Fourier *de f les nombres complexes $c_p(f) = \frac{1}{T}\int_0^T e^{-ip\omega t} f(t)\,dt$.*
(b) Série de Fourier *de f la série trigonométrique $\sum_{p\in\mathbb{Z}} c_p(f)\, e^{ip\omega t}$, et* sommes partielles *de f les sommes partielles symétriques $S_q f(t) = \sum_{|p|\leq q} c_p(f)\, e^{ip\omega t}$.*
(c) Sommes de Cesàro *de f les moyennes $\mathfrak{S}_q f(t) = \frac{1}{q}\sum_{p=0}^{q-1} S_p f(t)$ des q premières sommes partielles de f.*

Les principales propriétés des coefficients de Fourier sont décrites dans l'énoncé suivant.

Proposition 10.3. *Soit $f \in \mathcal{L}^1[0, T]$ une fonction T-périodique. Alors :*
(a) *$\lim_{p\to\pm\infty} c_p(f) = 0$ (lemme de* Riemann-Lebesgue*).*
(b) *Si $f \in \mathcal{L}^2[0, T]$, alors $\sum_{p\in\mathbb{Z}} |c_p(f)|^2 \leq \frac{1}{T}\int_0^T |f(t)|^2\,dt$ (inégalité de* Bessel*).*
(c) *Si f est lipschitzienne et si f' est sa dérivée presque partout (qui est mesurable bornée), alors $c_p(f') = i\,p\,\omega\, c_p(f)$ pour tout $p \in \mathbb{Z}$.*
(d) *Les sommes partielles et les sommes de Cesàro de f s'expriment par les formules*

$$S_q f(t) = \frac{1}{T}\int_{-T/2}^{T/2} D_q(s)\, f(t-s)\,ds \qquad \text{et} \qquad \mathfrak{S}_q f(t) = \frac{1}{T}\int_{-T/2}^{T/2} \mathfrak{F}_q(s)\, f(t-s)\,ds$$

où les noyaux de Dirichlet *D_q et de* Fejér *$\mathfrak{F}_q$ valent*

$$D_q(s) = \frac{\sin\left((q+\frac{1}{2})\,\omega\, s\right)}{\sin\left(\frac{1}{2}\,\omega\, s\right)} \qquad \text{et} \qquad \mathfrak{F}_q(s) = \frac{\sin^2\left(\frac{1}{2}\,q\,\omega\, s\right)}{q\sin^2\left(\frac{1}{2}\,\omega\, s\right)}$$

et vérifient $\frac{1}{T}\int_{-T/2}^{T/2} D_q(s)\,ds = \frac{1}{T}\int_{-T/2}^{T/2} \mathfrak{F}_q(s)\,ds = 1$ pour tout $q \geq 1$.
(e) *Pour tout $t \in \mathbb{R}$, la convergence de la suite $(S_q(t))$ entraîne celle de $(\mathfrak{S}_q(t))$.*

Démonstration. Comme la fonction $g = \mathbb{1}_{[0,T]} f$ est intégrable par hypothèse, il suffit pour établir le lemme de Riemann-Lebesgue de montrer que l'intégrale $\int e^{-ip\omega t} g(t)\,dt$ tend vers 0 lorsque p tend vers l'infini pour toute fonction intégrable g. Commençons par le montrer pour $g = \mathbb{1}_{[a,b]}$: on a dans ce cas

$$\left|\int e^{-ip\omega t} g(t)\,dt\right| = \left|\frac{e^{-ip\omega b} - e^{-ip\omega a}}{-ip\omega}\right| \leq \frac{2}{|p|\,\omega}$$

si $p \neq 0$, d'où le résultat. Par combinaisons linéaires, cela reste vrai pour toute fonction g en escalier, et si g est limite en moyenne d'une suite (g_q) de fonctions en escalier, on peut trouver pour tout $\varepsilon > 0$ un indice q tel que $\|g - g_q\|_1 \leq \varepsilon/2$ (ce qui implique que $\left| \int e^{-ip\omega t} \big(g(t) - g_q(t)\big)\, dt \right| \leq \varepsilon/2$ pour tout p), puis un indice N tel que $\left| \int e^{-ip\omega t} g_q(t)\, dt \right| \leq \varepsilon/2$ pour tout $p \geq N$. On a donc montré que si $|p| \geq N$, alors $\left| \int e^{-ip\omega t} g(t)\, dt \right| \leq \varepsilon$, et c'est ce qu'il fallait démontrer.

Pour démontrer l'inégalité de Bessel, il est commode de noter $(f|g) = \frac{1}{T}\int_0^T f(t)\,\overline{g(t)}\, dt$, et $e_p(t) = e^{ip\omega t}$. En effet, on voit alors que l'on peut écrire $c_p(f) = (f|e_p)$, $S_q f = \sum_{|p|\leq q} c_p(f)\, e_p$, et

$$\frac{1}{T}\int_0^T |f(t)|^2\, dt = (f|f) = (f - S_q f | f - S_q f) + 2\mathscr{Re}\,(f - S_q f | S_q f) + (S_q f | S_q f)\,.$$

En utilisant que $(e_p|e_r) = \frac{1}{T}\int_0^T e^{i(p-r)\omega t}\, dt = 1$ ou 0 suivant que $r = p$ ou que $r \neq p$, on voit que le dernier terme s'écrit $\sum_{|p|\leq q,\, |r|\leq q} c_p(f)\,\overline{c_r(f)}\,(e_p|e_r) = \sum_{|p|\leq q} |c_p(f)|^2$. Le terme précédent est nul car $(f - S_q f | S_q f) = \sum_{|p|\leq q} \overline{c_p(f)}\,(f - S_q f | e_p)$ et que $(f - S_q f|e_p) = c_p(f) - \sum_{|r|\leq q} c_r(f)\,(e_r|e_p) = 0$ si $|p| \leq q$. Enfin, le premier terme étant positif, on a donc prouvé que $\sum_{|p|\leq q} |c_p(f)|^2 \leq \frac{1}{T}\int_0^T |f(t)|^2\, dt$ pour tout $q \in \mathbb{N}$, d'où l'inégalité de Bessel.

La propriété (c) résulte d'une simple intégration par parties puisque

$$c_p(f') = \frac{1}{T}\int_0^T e^{-ip\omega t} f'(t)\, dt = \frac{f(T) - f(0)}{T} - \frac{1}{T}\int_0^T \left(e^{-ip\omega t}\right)' f(t)\, dt =$$
$$= \frac{ip\omega}{T}\int_0^T e^{-ip\omega t} f(t)\, dt = ip\omega\, c_p(f)\,,$$

le terme intégré $f(T) - f(0)$ étant nul par périodicité.

Pour les propriétés du noyau de Dirichlet, on écrit $c_p(f)\, e^{ip\omega t} = \frac{1}{T}\int_0^T e^{ip\omega(t-s)} f(s)\, ds$, d'où

$$S_q f(t) = \frac{1}{T}\int_0^T \Big(\sum_{|p|\leq q} e^{ip\omega(t-s)}\Big) f(s)\, ds = \frac{1}{T}\int_0^T D_q(t-s)\, f(s)\, ds$$

en posant $D_q(s) = \sum_{|p|\leq q} e^{ip\omega s}$. Par changement de variable $s = \phi(u) = t - u$ puis invariance par translation puisque les fonctions considérées sont T-périodiques, on obtient

$$S_q f(t) = \frac{1}{T}\int_{t-T}^t D_q(u)\, f(t-u)\, du = \frac{1}{T}\int_{-T/2}^{T/2} D_q(s)\, f(t-s)\, ds$$

comme annoncé. Il ne reste plus qu'à expliciter D_q qui vérifie clairement $\frac{1}{T}\int_0^T D_q(s)\, ds = 1$. Or on a

$$D_q(s) = \sum_{|p|\leq q} e^{ip\omega s} = e^{-iq\omega s}\,\frac{1 - e^{i(2q+1)\omega s}}{1 - e^{i\omega s}} = \frac{e^{-i(q+\frac{1}{2})\omega s} - e^{i(q+\frac{1}{2})\omega s}}{e^{-i\frac{1}{2}\omega s} - e^{i\frac{1}{2}\omega s}} = \frac{\sin\left(\left(q + \frac{1}{2}\right)\omega s\right)}{\sin\left(\frac{1}{2}\omega s\right)}\,,$$

ce qui complète les résultats sur le noyau de Dirichlet. Les propriétés du noyau de Fejér s'en déduisent puisque $\mathfrak{F}_q(s) = \frac{1}{q}\sum_{p=0}^{q-1} D_p(s)$, ce qui permet d'écrire

$$\mathfrak{F}_q(s) = \frac{1}{q\,\sin\left(\frac{1}{2}\omega s\right)}\sum_{p=0}^{q-1} \sin\left(\left(q + \frac{1}{2}\right)\omega s\right) = \frac{1}{q\,\sin\left(\frac{1}{2}\omega s\right)}\,\mathscr{Im}\left(e^{i\frac{1}{2}\omega s}\,\frac{1 - e^{iq\omega s}}{1 - e^{i\omega s}}\right) = \frac{\sin^2\left(\frac{1}{2}q\,\omega s\right)}{q\,\sin^2\left(\frac{1}{2}\omega s\right)}\,.$$

Enfin la propriété (e) est classique. Supposons que la suite $(S_p f(t))$ converge vers une limite $\ell \in \mathbb{C}$ en un point $t \in \mathbb{R}$. Alors la suite $(S_p f(t) - \ell)$ qui tend vers 0 est bornée par un nombre M, et pour tout $\varepsilon > 0$, on peut trouver un rang N tel que $|S_p f(t) - \ell| \leq \varepsilon/2$ pour $p \geq N$. On écrit alors pour $q > N$

$$|\mathfrak{S}_q f(t) - \ell| \leq \frac{1}{q}\sum_{p=0}^{q-1} |S_p f(t) - \ell| = \frac{1}{q}\sum_{p=0}^{N-1} |S_p f(t) - \ell| + \frac{1}{q}\sum_{p=N}^{q-1} |S_p f(t) - \ell| \leq$$
$$\leq \frac{NM}{q} + \frac{q-N}{q}\,\frac{\varepsilon}{2} \leq \frac{NM}{q} + \frac{\varepsilon}{2}\,,$$

ce qui peut être rendu inférieur à ε pour q assez grand. ∎

Le théorème de Dirichlet ci-dessous établit la convergence de la série de Fourier d'une fonction f en un point t lorsque f garde un comportement "raisonnable" au voisinage de t, par exemple lorsque f vérifie la propriété suivante : on dit qu'une fonction T-périodique f est lipschitzienne *par morceaux* s'il existe un nombre fini de points $0 = T_0 < T_1 < \ldots < T_{N-1} < T_N = T$ tels que f soit lipschitzienne sur chaque intervalle $]T_{k-1}, T_k[$. Une telle fonction lipschitzienne par morceaux possède des limites à droite et à gauche en tout point, et nous noterons $\widetilde{f}(t)$ la demi-somme des limites à droite et à gauche de f en t.

Théorème 10.4 (de Dirichlet). *Soit $f \in \mathcal{L}^1[0, T]$ une fonction T-périodique.*
(a) *On suppose que pour un $t \in \mathbb{R}$, il existe $\ell \in \mathbb{C}$ et $\varepsilon > 0$ tels que*

$$\lim_{s \to 0^+} \frac{f(t-s) + f(t+s) - 2\,\ell}{s^\varepsilon} = 0\,.$$

Alors $\lim_{q \to \infty} S_q f(t) = \ell$.
(b) *Si f est lipschitzienne par morceaux, alors (S_q) converge en tout point vers $\widetilde{f}$.*
(c) *Si f est lipschitzienne, alors la suite (S_q) converge uniformément vers f.*

Démonstration. Commençons par la propriété (a). Il résulte de la proposition 10.3 (d) que l'on peut écrire

$$S_q f(t) - \ell = \frac{1}{T} \int_{-T/2}^{T/2} D_q(s) \Big(f(t-s) - \ell \Big)\, ds =$$

$$= \frac{1}{2T} \int_{-T/2}^{T/2} D_q(s) \Big(f(t-s) + f(t+s) - 2\,\ell \Big)\, ds$$

puisque le noyau de Dirichlet est pair. En développant $\sin\big((q + \frac{1}{2})\,\omega\, s\big)$ par les formules trigonométriques, cela s'écrit aussi

$$S_q f(t) - \ell = \frac{1}{2T} \int_{-T/2}^{T/2} \Big(\sin(q\,\omega\, s)\, g(s) + \cos(q\,\omega\, s)\, h(s) \Big)\, ds$$

en notant g et h les fonctions T-périodiques définies sur $[-T/2, T/2]$ par

$$g(s) = \frac{f(t-s) + f(t+s) - 2\,\ell}{\tan\big(\frac{1}{2}\,\omega\, s\big)} \qquad \text{et} \qquad h(s) = f(t-s) + f(t+s) - 2\,\ell\,.$$

La fonction h est clairement intégrable sur $[-T/2, T/2]$, et la fonction g aussi grâce aux arguments suivants : comme sur $[-T/2, T/2]$, $\big|\tan\big(\frac{1}{2}\,\omega\, s\big)\big| \geq \frac{1}{2}\,\omega\,|s|$, on peut majorer par hypothèse la fonction $|g|$ par la fonction intégrable $s \mapsto |s|^{\varepsilon - 1}$ sur un voisinage $]-\delta, \delta[$ de 0, tandis que pour $|s| \in [\delta, T/2]$ elle se majore par la fonction intégrable $|h|/\tan\big(\frac{1}{2}\omega\,\delta\big)$. On peut donc écrire

$$S_q f(t) - \ell = \frac{c_q(g) - c_{-q}(g)}{4\,i} + \frac{c_q(h) + c_{-q}(h)}{4}\,,$$

ce qui permet de conclure que $\lim_{q\to\infty} S_q f(t) = \ell$ d'après le lemme de Riemann-Lebesgue.

Pour obtenir la propriété (b), il suffit d'observer que les hypothèses du (a) sont vérifiées en tout point $t \in \mathbb{R}$ avec $\ell = \widetilde{f}(t)$. Notons $]T_{k-1}, T_k[$ les intervalles où f est lipschitzienne, et C_k la constante de Lipschitz de f sur cet intervalle. Si $t \in]T_{k-1}, T_k[$, on a $|f(t-s)+f(t+s)-2\,f(t)| \le 2\,C_k\,|s|$, ce qui prouve la propriété voulue avec $\varepsilon = 1/2$ par exemple. En un point de discontinuité $t = T_k$, la fonction f possède des limites à droite et à gauche en T_k puisqu'elle est uniformément continue dans chaque intervalle où elle est lipschitzienne, et on peut aussi écrire $|f(t-s)+f(t+s)-2\,\widetilde{f}(t)| \le (C_k + C_{k+1})\,|s|$, ce qui permet de conclure de la même façon.

Enfin, dans le cas d'une fonction f lipschitzienne, on a d'après la proposition 10.3 (c) la majoration

$$|c_p(f)| = \frac{1}{|p|\,\omega}\,|c_p(f')| \le \frac{1}{2\,\omega}\left(\frac{1}{p^2} + |c_p(f')|^2\right)$$

pour $p \neq 0$ puisque la différence entre le terme de droite et le terme central est un carré, et comme $f' \in \mathcal{L}^2[0,T]$ puisqu'elle est mesurable et bornée, on peut déduire de la proposition 10.3 (b) que $\sum_{p\in\mathbb{Z}} |c_p(f)| < \infty$. Il résulte alors de la proposition 10.1 que la série de Fourier de f converge uniformément, et d'après le cas (b) démontré ci-dessus, elle converge vers $\widetilde{f} = f$. ∎

Exemple. Soit f la fonction 2π-périodique qui vaut $f(t) = \pi\,|t|$ sur $[-\pi,\pi]$. Ses coefficients de Fourier, pour $p \neq 0$, valent

$$c_p(f) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ipt}\,\pi\,|t|\,dt = \int_0^{\pi} \frac{e^{ipt}+e^{-ipt}}{2}\,t\,dt = \int_0^{\pi} t\,\cos(pt)\,dt = \frac{(-1)^p - 1}{p^2}$$

puisque la fonction $t \mapsto (t\sin(pt)/p) + (\cos(pt)/p^2)$ est une primitive de la fonction $t \mapsto t\cos(pt)$, et on a $c_0(f) = \pi^2/2$. Comme f est lipschitzienne, le théorème de Dirichlet permet donc d'écrire

$$0 = f(0) = \frac{\pi^2}{2} + \sum_{p\in\mathbb{Z}^*} \frac{(-1)^p - 1}{p^2} = \frac{\pi^2}{2} - 4\sum_{p\in\mathbb{N}} \frac{1}{(2p+1)^2},$$

d'où l'on peut déduire que la somme des inverses des carrés impairs vaut $\pi^2/8$. En observant que la somme des inverses des carrés pairs vaut $\sum_{p>0} 1/(2p)^2 = (1/4)\sum_{p>0} 1/p^2$, on en déduit les identités suivantes

$$\sum_{p>0,\,p\ \text{impair}} \frac{1}{p^2} = \frac{\pi^2}{8}, \qquad \sum_{p>0} \frac{1}{p^2} = \frac{\pi^2}{6}, \qquad \sum_{p>0,\,p\ \text{pair}} \frac{1}{p^2} = \frac{\pi^2}{24}.$$

On se gardera de penser que la série de Fourier d'une fonction continue converge nécessairement : on sait même construire des fonctions continues dont la série de Fourier diverge en tout point d'une partie dense de $\mathbb{R}$. On dispose cependant du résultat (plus faible) suivant.

Théorème 10.5 (de Fejér). *Pour toute fonction T-périodique $f \in \mathcal{L}^1[0,T]$ et tout point $t \in \mathbb{R}$ tel que $\ell = \lim_{s\to 0^+} \frac{1}{2}\big(f(t-s)+f(t+s)\big)$ existe, les sommes de Cesàro $(\mathfrak{S}_q f(t))$ convergent vers ℓ. En outre, la suite $(\mathfrak{S}_q f)$ converge uniformément vers f lorsque f est continue.*

Démonstration. On procède essentiellement comme pour le théorème de Dirichlet. En effet, d'après la proposition 10.3 (d), on peut écrire

$$\begin{aligned}\mathfrak{S}_q f(t) - \ell &= \frac{1}{T}\int_{-T/2}^{T/2} \mathfrak{F}_q(s)\left(f(t-s)-\ell\right) ds = \\ &= \frac{1}{T}\int_0^{T/2} \mathfrak{F}_q(s)\Big(f(t-s)+f(t+s)-2\,\ell\Big)\,ds\end{aligned}$$

puisque le noyau de Fejér $\mathfrak{F}_q$ est pair. On utilise maintenant la positivité de ce noyau $\mathfrak{F}_q$ pour écrire pour tout $\delta > 0$,

$$\begin{aligned}|\mathfrak{S}_q f(t) - \ell| \le \frac{1}{T}\int_0^{\delta} \mathfrak{F}_q(s)\,&|f(t-s)+f(t+s)-2\,\ell|\,ds + \\ &+ \frac{1}{T}\int_{\delta}^{T/2} \mathfrak{F}_q(s)\,|f(t-s)+f(t+s)-2\,\ell|\,ds \le \\ \le \Big(\sup_{0<s<\delta}\Big|\frac{f(t-s)+f(t+s)}{2} - \ell\Big|\Big)&\,\frac{2}{T}\int_0^{T/2}\mathfrak{F}_q(s)\,ds + \\ &+ \Big(\sup_{\delta<s<T/2}\mathfrak{F}_q(s)\Big)\left(|\ell| + \frac{1}{T}\int_0^T |f(s)|\,ds\right).\end{aligned}$$

Soit alors $\varepsilon > 0$. Si $\frac{1}{2}\left(f(t-s)+f(t+s)\right)$ tend vers ℓ lorsque s tend vers 0, il existe un nombre $\delta > 0$ tel que le premier terme soit inférieur à $\varepsilon/2$, puis il existe un rang q à partir duquel le dernier terme est aussi inférieur à $\varepsilon/2$ puisque $\sup_{[\delta,T/2]}\mathfrak{F}_q \le 1/\big(q\sin^2\left(\frac{1}{2}\omega\,\delta\right)\big)$. Dans le cas d'une fonction f continue, f est uniformément continue et bornée sur l'intervalle compact $[-T,T]$, et donc aussi sur $\mathbb{R}$ par périodicité. On reprend alors le même raisonnement avec $\ell = f(t)$ en remarquant que le nombre $\delta > 0$ puis le rang q peuvent être choisis indépendamment de t. ∎

Le théorème de Fejér a de nombreuses conséquences. Par exemple, il montre qu'une fonction continue T-périodique est entièrement caractérisée par ses coefficients de Fourier puisque deux fonctions continues possédant les mêmes coefficients de Fourier possèdent aussi les mêmes sommes de Cesàro. On en déduit aussi les résultats d'approximation suivants.

Corollaire 10.6. *Toute fonction continue T-périodique est limite uniforme d'une suite de polynômes trigonométriques, et toute fonction continue sur un intervalle compact $[a,b]$ y est limite uniforme d'une suite de polynômes algébriques (théorèmes de* Weierstraß*). En outre si $1 \le p < \infty$, toute fonction $f \in \mathcal{L}^p[0,T]$ est limite en moyenne d'ordre p sur $[0,T]$ d'une suite de polynômes trigonométriques.*

Démonstration. Une fonction continue T-périodique est limite uniforme de polynômes trigonométriques d'après le théorème de Fejér puisque les sommes de Cesàro $\mathfrak{S}_q f$ sont des polynômes trigonométriques.

Si f est une fonction continue sur l'intervalle compact $[a,b]$, elle s'étend en une fonction continue sur $[a,b+1]$ vérifiant $f(b+1) = f(a)$ en choisissant $f(t) = f(b) + \big(f(a) - f(b)\big)\,(t-b)$ pour $t \in [b, b+1]$, puis en une fonction continue T-périodique avec $T = b-a+1$, et elle est donc limite uniforme de polynômes trigonométriques. Un polynôme trigonométrique étant somme d'une série entière de rayon de convergence infini comme combinaison linéaire d'exponentielles complexes, il est à son tour limite uniforme sur tout compact, en particulier sur $[a,b]$, d'une suite de polynômes algébriques, ce qui prouve le deuxième théorème de Weierstraß.

Enfin, si $f \in \mathcal{L}^p[0,T]$ pour un $p < \infty$, la fonction $\mathbb{1}_{[0,T]} f$ est limite en moyenne d'ordre p d'une suite (f_q) de fonctions continues à support compact dans $\mathbb{R}$ (théorème 9.4). Soit alors h_r la fonction définie par

$$h_r(t) = \begin{cases} 2^r\, t & \text{pour } t \in [0, 2^{-r}]\,, \\ 1 & \text{pour } t \in [2^{-r}, T - 2^{-r}]\,, \\ 2^r\,(T-t) & \text{pour } t \in [T - 2^{-r}, T]\,, \\ 0 & \text{pour } t \notin [0,T]\,. \end{cases}$$

Pour tous q et $r \in \mathbb{N}$, la fonction $f_{q,r} = h_r\, f_q$ est continue et vérifie $f_{q,r}(0) = f_{q,r}(T) = 0$. De plus, comme $f_{q,r} = \mathbb{1}_{[0,T]}\, f_{q,r}$, on a

$$\begin{aligned} \|f_{q,r} - \mathbb{1}_{[0,T]} f\|_p &\le \|\mathbb{1}_{[0,T]}\,(f_{q,r} - f_q)\|_p + \|\mathbb{1}_{[0,T]}\,(f_q - f)\|_p \le \\ &\le \Big(\sup_{[0,T]} |f_q|\Big)\, \|\mathbb{1}_{[0,T]}\,(h_r - 1)\|_p + \|f_q - f\|_p\,. \end{aligned}$$

Pour tout $\varepsilon > 0$, il existe un entier q tel que $\|f_q - f\|_p \le \varepsilon/2$, puis il existe un entier r tel que $\|\mathbb{1}_{[0,T]}\,(h_r - 1)\|_p \le \varepsilon/(2\sup_{[0,T]} |f_q|)$ puisque $\|\mathbb{1}_{[0,T]}\,(h_r - 1)\|_p \le 2^{-r/p}$. On a ainsi montré que la fonction f est limite en moyenne d'ordre p sur $[0,T]$ d'une suite de fonctions continues $(f_{q,r})$ vérifiant $f_{q,r}(0) = f_{q,r}(T) = 0$.

Une telle fonction continue se prolonge en une fonction continue T-périodique, et elle est donc à son tour limite uniforme d'une suite de polynômes trigonométriques, ce qui achève la démonstration du corollaire puisque la majoration

$$\|f - g\|_p = \Big(\int_0^T |f(t) - g(t)|^p\, dt\Big)^{1/p} \le \Big(T \sup_{[0,T]} |f - g|^p\Big)^{1/p} = T^{1/p} \sup_{[0,T]} |f - g|$$

montre que la convergence uniforme entraîne la convergence en moyenne d'ordre p sur $[0,T]$. ∎

Jusqu'à présent, nous n'avons étudié que la convergence simple ou uniforme des séries de Fourier. Pour la convergence en moyenne d'ordre p, le résultat le plus plaisant est le cas $p = 2$, c'est-à-dire le cas des fonctions de carré intégrable.

Théorème 10.7 (de Riesz-Fischer). *Toute fonction T-périodique $f \in \mathcal{L}^2[0,T]$ est somme en moyenne quadratique sur $[0,T]$ de sa série de Fourier, et on a la* formule de Parseval

$$\sum_{p\in\mathbb{Z}} |c_p(f)|^2 = \frac{1}{T}\int_0^T |f(t)|^2\, dt\,.$$

Réciproquement, pour toute suite $(c_p)_{p\in\mathbb{Z}}$ de nombres complexes vérifiant $\sum_{p\in\mathbb{Z}} |c_p|^2 < \infty$, la série trigonométrique $\sum_{p\in\mathbb{Z}} c_p\, e^{ip\omega t}$ converge en moyenne quadratique sur $[0,T]$ vers une fonction $f \in \mathcal{L}^2[0,T]$ dont les coefficients de Fourier sont les nombres c_p donnés.

Démonstration. Comme dans la démonstration de la proposition 10.3 (b), nous utiliserons les notations suivantes : pour f et $g \in \mathcal{L}^2[0,T]$, on pose $(f|g) = \frac{1}{T}\int_0^T f(t)\,\overline{g(t)}\, dt$, $\|f\|^2 = (f|f)$ et $e_p(t) = e^{ip\omega t}$, ce qui permet d'écrire $c_p(f) = (f|e_p)$ et $S_q f = \sum_{|p|\le q} c_p(f)\, e_p$, et nous rappelons l'identité

$$\|f\|^2 = \|f - S_q f\|^2 + \|S_q f\|^2 = \|f - S_q f\|^2 + \sum_{|p|\le q} |c_p(f)|^2$$

qui avait été établie dans cette démonstration. Il en résulte que la suite numérique $(\|f - S_q f\|)_{q\in\mathbb{N}}$ est décroissante, et que l'on aura prouvé la première partie du théorème si l'on montre qu'elle possède une sous-suite tendant vers 0.

Par ailleurs, si $f \in \mathcal{L}^2[0,T]$ et si g est un polynôme trigonométrique de degré inférieur ou égal à q, on a

$$\|f - g\|^2 = \|f - S_q f\|^2 + 2\,\mathscr{Re}\,(f - S_q f|S_q f - g) + \|S_q f - g\|^2\,,$$

et le terme central étant une combinaison linéaire, pour $|p| \le q$, de termes $(f - S_q f|e_p) = (f|e_p) - \sum_{|r|\le q} c_r(f)\,(e_r|e_p) = 0$, on en déduit que $\|f - g\| \ge \|f - S_q f\|$. Or d'après le corollaire 10.6, f est limite en moyenne quadratique sur $[0,T]$ d'une suite (g_r) de polynômes trigonométriques : cela fournit la sous-suite $(\|f - S_{q_r} f\|)_{r\in\mathbb{N}}$ cherchée en notant q_r le degré de g_r.

Pour établir la réciproque, notons $f_q = \sum_{|p|\le q} c_p\, e_p$. La suite (f_q) vérifie

$$\|f_q - f_r\|^2 = \Big\|\sum_{r<|p|\le q} c_p\, e_p\Big\|^2 = \sum_{r<|p|\le q} |c_p|^2$$

pour $q \ge r$, ce qui permet de conclure qu'elle est de Cauchy pour la convergence en moyenne quadratique sur $[0,T]$. Elle converge donc (toujours en moyenne quadratique) vers une fonction $f \in \mathcal{L}^2[0,T]$ d'après le théorème 9.4, et l'inégalité de Cauchy-Schwarz permet d'écrire

$$|c_p(f) - c_p| = \Big|(f - f_q, e_p) + (f_q, e_p) - c_p\Big| \le \|f - f_q\| + |(f_q, e_p) - c_p|\,.$$

On en déduit que $c_p(f) = c_p$ en faisant tendre q vers l'infini, puisque $\|f - f_q\|$ tend alors vers 0 et que $(f_q|e_p) = c_p$ pour tout $q \ge |p|$. ∎

Exemple. Reprenons l'exemple proposé après le théorème 10.4, c'est-à-dire l'exemple de la fonction 2π-périodique f valant $f(t) = \pi\,|t|$ sur $[-\pi, \pi]$. On calcule que $\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)|^2\,dt = \pi^4/3$, et comme ses coefficients de Fourier ont été calculés plus haut, la formule de Parseval s'écrit ici

$$\frac{\pi^4}{3} = \frac{\pi^4}{4} + \sum_{p\in\mathbb{Z}^*} \frac{\left((-1)^p - 1\right)^2}{p^4} = \frac{\pi^4}{4} + 8\sum_{p\in\mathbb{N}} \frac{1}{(2p+1)^4}\,.$$

On en déduit que la somme des inverses des puissances quatrièmes des entiers impairs vaut $\pi^4/96$, et comme on a aussi $\sum_{p>0} 1/(2p)^4 = (1/16)\sum_{p>0} 1/p^4$, on en déduit les identités suivantes

$$\sum_{p>0,\,p\ \text{impair}} \frac{1}{p^4} = \frac{\pi^4}{96}\,, \qquad \sum_{p>0} \frac{1}{p^4} = \frac{\pi^4}{90}\,, \qquad \sum_{p>0,\,p\ \text{pair}} \frac{1}{p^4} = \frac{\pi^4}{1440}\,.$$

Terminons en évoquant un aspect des séries de Fourier que nous n'avons pas étudié dans ce paragraphe : c'est le lien entre la régularité de la fonction f et la vitesse de décroissance de ses coefficients de Fourier. En effet, il résulte de la proposition 10.3 (a) et (c) que les coefficients de Fourier d'une fonction de classe C^k décroissent plus vite que $|p|^{-k}$. En particulier, si f est de classe C^∞, ses coefficients de Fourier sont *à décroissance rapide*, c'est-à-dire qu'ils vérifient $\lim_{p\to\pm\infty} p^k\, c_p(f) = 0$ pour tout entier k. Réciproquement, il résulte des théorèmes usuels sur la dérivation terme à terme des séries de fonctions que si les coefficients de Fourier vérifient $\sum_{p\in\mathbb{Z}} |p^k\, c_p(f)| < \infty$, alors f est de classe C^k.

Exercices

10.1. Soient $a \in\,]0,1[$ et $\theta \in\,]0,\pi[$ deux paramètres réels.

(a) Calculer les coefficients de Fourier des fonctions 2π-périodiques définies sur l'intervalle $]-\pi,\pi[$ par $f(t) = e^{iat}$ et $g(t) = \mathbb{1}_{[-\theta,\theta]}(t)$.

(b) À l'aide du théorème de Dirichlet et de la formule de Parseval, en déduire les identités suivantes

$$\frac{\pi}{\sin(a\pi)} = \frac{1}{a} + 2a\sum_{p=1}^{\infty} \frac{(-1)^p}{a^2 - p^2}\,, \qquad \pi\,\mathrm{cotan}\,(a\pi) = \frac{1}{a} + 2a\sum_{p=1}^{\infty} \frac{1}{a^2 - p^2}\,,$$

$$\frac{\pi^2}{\sin^2(a\pi)} = \frac{1}{a^2} + 2\sum_{p=1}^{\infty} \frac{a^2 + p^2}{(a^2 - p^2)^2}\,, \qquad \theta\,(\pi - \theta) = 2\sum_{p=1}^{\infty} \frac{\sin^2(p\theta)}{p^2}$$

$$\pi - \theta = 2\sum_{p=1}^{\infty} \frac{\sin(p\theta)}{p}\,, \qquad \theta = 2\sum_{p=1}^{\infty} (-1)^{p+1}\frac{\sin(p\theta)}{p}\,, \qquad \frac{\pi}{2} - \theta = \sum_{p=1}^{\infty} \frac{\sin(2p\theta)}{p}\,.$$

10.2. Soient f et $g \in \mathcal{L}^1[0,T]$ deux fonctions T-périodiques à valeurs complexes.

(a) Montrer que pour tout $s \in \mathbb{R}$ on a $\int_0^T f(t)\,dt = \int_0^T f(s+t)\,dt = \int_0^T f(s-t)\,dt$.

(b) Montrer que la fonction $f * g : s \mapsto \frac{1}{T}\int_0^T f(s-t)\,g(t)\,dt$ est une fonction T-périodique vérifiant $f*g(s) = \frac{1}{T}\int_0^T f(t)\,g(s-t)\,dt$ et $f*g \in \mathcal{L}^1[0,T]$.

(c) Montrer que les coefficients de Fourier des fonctions f, g et $f*g$ vérifient $c_p(f*g) = c_p(f)\,c_p(g)$ pour tout $p \in \mathbb{Z}$.

10.3. Soit $f : \mathbb{R} \to \mathbb{C}$ une fonction de classe C^1 vérifiant $|f(t)| + |f'(t)| \leq C\,(1+|t|)^{-a}$ pour des constantes $C > 0$ et $a > 1$ indépendantes de $t \in \mathbb{R}$. Soient aussi $\omega > 0$ un nombre réel et $T = 2\pi/\omega$.

(a) Montrer que f est intégrable. On notera $\widehat{f}$ sa transformée de Fourier.

(b) Montrer que la formule $g(t) = \sum_{q\in\mathbb{Z}} f(t+qT)$ définit une fonction T-périodique g de classe C^1.

(c) Calculer les coefficients de Fourier de g en fonction de $\widehat{f}$, et en déduire la *formule sommatoire de Poisson* : $\sum_{q\in\mathbb{Z}} f(qT) = \frac{1}{T}\sum_{p\in\mathbb{Z}} \widehat{f}(p\,\omega)$.

10.4. Soit $f : [0,\ell] \to \mathbb{R}$ une fonction lipschitzienne vérifiant $f(0) = f(\ell) = 0$. On veut déterminer la température en tout point $x \in [0,\ell]$ d'une barre de longueur ℓ et en tout instant $t \geq 0$ lorsque la répartition des températures est donnée par f à l'instant $t = 0$ et que les extrémités de la barre sont maintenues à température constante $= 0$. Ce problème consiste à chercher une fonction continue $u : \overline{D} = [0,\infty[\, \times [0,\ell] \to \mathbb{R}$ de classe C^2 dans $D =]0,\infty[\, \times\,]0,\ell[$ et à dérivées premières bornées dans tout domaine $[t,T] \times]0,\ell[$ avec $T > t > 0$, et solution du problème

$$\frac{\partial u}{\partial t}(t,x) = \frac{\partial^2 u}{\partial x^2}(t,x) \text{ dans } D \qquad (\textit{équation de la chaleur}),$$
$$u(t,0) = u(t,\ell) = 0 \text{ pour } t \geq 0, \qquad u(0,x) = f(x) \text{ pour } x \in [0,\ell].$$

(a) Soit u une solution du problème. Montrer que pour tout $t \geq 0$ fixé, la fonction $x \mapsto u(t,x)$ prolongée à $\mathbb{R}$ par imparité et 2ℓ-périodicité (on notera alors $\omega = 2\pi/2\ell$) est continue, puis montrer que ses coefficients de Fourier $c_p(t)$ sont continus sur $[0,\infty[$ et de classe C^1 sur $]0,\infty[$, et qu'ils vérifient l'équation différentielle $c_p'(t) + p^2\omega^2 c_p(t) = 0$.

(b) Résoudre l'équation différentielle trouvée en (a) avec une donnée convenable en $t = 0$, montrer que la formule $u(t,x) = \sum_{p\in\mathbb{Z}} c_p(t)\,e^{ip\omega x}$ définit alors une fonction u qui est l'unique solution du problème posé, et que cette solution est de classe C^∞ dans D.

(c) Montrer qu'en général, le même problème ne peut pas être résolu du côté $t \leq 0$.

10.5. Pour tout entier $q \geq 1$, soit $G_q(t) = \sum_{p=1}^{q} p^{-1}\sin(pt)$ le *noyau de Gibbs*.

(a) En notant D_q le noyau de Dirichlet, montrer que $G_q(t) = \int_0^t D_q(s)\,ds - (t/2)$, puis montrer l'inégalité $|G_q(t)| \leq 1 + (\pi/2)$ pour tout $q \geq 1$ et tout $t \in \mathbb{R}$.

(b) Pour tout entier $q \geq 1$, on pose $q' = 2^{q^3}$ et $f_q(t) = e^{2iq't}\,G_{q'}(t)$. Montrer que f_q est continue 2π-périodique, et calculer ses coefficients de Fourier $c_p(f_q)$.

(c) Montrer que la série de fonctions $f = \sum_{q=1}^{\infty} q^{-2} f_q$ converge uniformément, que sa somme f est continue 2π-périodique, et que $c_p(f) = \sum_{q=1}^{\infty} q^{-2}\,c_p(f_q)$ pour tout $p \in \mathbb{Z}$.

(d) Montrer que $|S_{2q'}f(0)| \geq (q\ln 2)/3$ pour tout q assez grand (on utilisera l'équivalent classique $\sum_{p=1}^{q'} p^{-1} \sim \ln q'$), et en déduire que la série de Fourier de la fonction continue f ne converge pas en 0.

10.6. Soient $\omega > 0$, $T = 2\pi/\omega$ et $f : [0,T] \to \mathbb{R}$ une fonction croissante. On appelle *variation totale* de f le réel positif $V(f) = f(T) - f(0)$, et on notera $c_p(f) = \frac{1}{T}\int_0^T e^{-ip\omega t} f(t)\,dt$ les coefficients de Fourier de la fonction $f_{|[0,T[}$ prolongée à $\mathbb{R}$ par T-périodicité.

(a) Soit $\varepsilon > 0$. Montrer qu'il existe une subdivision $0 = a_1 < a_2 < \ldots < a_{2N+1} = T$ de l'intervalle $[0,T]$ telle que pour tout $k = 1, \ldots, N$ et tout $t \in]a_{2k-1}, a_{2k+1}[$, on ait $|f(t) - f(a_{2k})| \leq \varepsilon$.

(b) On prend $p \neq 0$. En écrivant, avec les notations de la question (a), $f(t) = f(a_{2k}) + f(t) - f(a_{2k})$ sur l'intervalle $]a_{2k-1}, a_{2k+1}[$, démontrer que $|c_p(f)| \leq \varepsilon + \frac{1}{\pi\,|p|}\left(f(a_{2N}) - f(a_2)\right)$, et en déduire que $|c_p(f)| \leq V(f)/(\pi\,|p|)$.

10.7. Soit (u_p) une suite de nombres complexes vérifiant, pour tout $p > 0$, une majoration $|u_p| \leq C/p$ avec une constante C indépendante de p. On pose $s_q = \sum_{p=0}^{q} u_p$ et $\sigma_q = q^{-1} \sum_{p=0}^{q-1} s_p$.

(a) Pour tous entiers $p > q > r$, justifier l'identité

$$(p-q)\, s_p = \sum_{k=q}^{p-1} s_k + \sum_{k=q+1}^{p} (k-q)\, u_k \, .$$

et en déduire pour tout $\ell \in \mathbb{C}$ la majoration

$$|s_p - \ell| \leq \frac{p+q}{p-q} \sup_{k \geq r} |\sigma_k - \ell| + \frac{C}{2} \frac{p-q}{q} \, .$$

(b) Montrer que pour tout $\varepsilon \in]0, \frac{1}{2}[$ et tout $p > 2\,(r + \varepsilon^{-1})$, l'intervalle $[p\,(1-\varepsilon), p\,(1-\varepsilon^2)]$ contient au moins un entier $q > r$, et en déduire que

$$|s_p - \ell| \leq 2\,\varepsilon^{-2} \sup_{k \geq r} |\sigma_k - \ell| + C\,\varepsilon \, .$$

(c) Démontrer le théorème *de Hardy-Landau* : si (σ_p) converge vers un nombre $\ell \in \mathbb{C}$, alors (s_p) converge aussi vers ℓ.

10.8. Soient $\omega > 0$ et $T = 2\pi/\omega$. Une fonction $f : [0, T] \to \mathbb{R}$ est dite *à variation bornée* s'il existe deux fonctions croissantes g et $h : [0, T] \to \mathbb{R}$ telles que $f = g - h$.

(a) Montrer qu'une fonction lipschitzienne par morceaux est à variation bornée, et qu'une fonction monotone par morceaux est à variation bornée.

(b) Montrer qu'une fonction à variation bornée admet des limites à droite et à gauche en tout point. On notera $\widetilde{f}(t)$ la demi-somme des limites à droite et à gauche en t de la fonction f.

(c) En utilisant l'exercice **10.6** de la page 82, montrer que les coefficients de Fourier d'une fonction T-périodique à variation bornée sur $[0, T]$ vérifient $|c_p(f)| \leq V(f)/(\pi\,|p|)$ en notant $V(f) = V(g) + V(h)$ pour une décomposition de $f = g - h$ en différence de deux fonctions g et h croissantes sur $[0, T]$.

(d) En utilisant l'exercice **10.7** ci-dessus, démontrer le théorème *de Jordan* : si f est une fonction T-périodique à variation bornée dans $[0, T]$, alors sa série de Fourier converge en tout point vers $\widetilde{f}$, et la convergence est uniforme lorsque f est de plus continue.

10.9. Dans tout cet exercice, on identifie le plan affine euclidien $\mathbb{R}^2$ à $\mathbb{C}$. Ainsi, l'image d'une application $g : [a, b] \to \mathbb{C}$ continue, injective sur $[a, b[$ et vérifiant $g(a) = g(b)$ sera appelée *courbe de Jordan.* Dans tout l'exercice $C = g([a, b])$ désigne une courbe de Jordan admettant un paramétrage g lipschitzien.

(a) Soient $c \in \mathbb{C}$, $r > 0$ et $f : [0, 2\pi] \to \mathbb{C}$ la fonction définie par $f(t) = c + r\, e^{it}$. Montrer que f est un paramétrage lipschitzien d'une courbe de Jordan dont on donnera le nom, et calculer les nombres

$$\ell = \int_0^{2\pi} |f'(t)|\, dt \qquad \text{et} \qquad a = \left| \int_a^b \tfrac{1}{2}\, \mathscr{I}m\, (f'(t)\, \overline{f(t)})\, dt \right| .$$

(b) On suppose que $g = f \circ h$ pour deux fonctions lipschitziennes $f : [0, T] \to \mathbb{C}$ et $h : [a, b] \to [0, T]$ avec h monotone surjective. Montrer que

$$\int_a^b |g'(s)|\, ds = \int_0^T |f'(t)|\, dt \qquad \text{et} \qquad \left| \int_a^b \tfrac{1}{2}\, \mathscr{I}m\, (g'(s)\, \overline{g(s)})\, ds \right| = \left| \int_0^T \tfrac{1}{2}\, \mathscr{I}m\, (f'(t)\, \overline{f(t)})\, dt \right| .$$

Ces deux nombres seront notés respectivement $\ell(C)$ et $a(C)$, et appelés *longueur* de C et *aire du domaine* délimité par C.

(c) Montrer que la fonction $h : s \mapsto h(s) = \int_a^s |g'(r)|\, dr$ est lipschitzienne et strictement monotone surjective de $[a, b]$ sur $[0, T]$ avec $T = \ell(C)$. On pose ensuite $\widetilde{g'}(s) = g'(s)/|g'(s)|$ si $g'(s) \neq 0$, $\widetilde{g'}(s) = 1$ si $g'(s) = 0$ ou si g n'est pas dérivable en s, et $f(t) = g(a) + \int_0^t \widetilde{g'} \circ h^{-1}(r)\, dr$ pour $t \in [0, T]$. Montrer que f est lipschitzienne sur $[0, T]$ avec $|f'| = 1$ presque partout, et que $g = f \circ h$. Dans la suite, on utilisera le paramétrage f de la courbe C.

(d) En remarquant que $|f'| = |f'|^2$ presque partout et en utilisant la formule de Parseval, montrer que $\ell(C)^2 = 4\pi^2 \sum_{p\in\mathbb{Z}} p^2\, |c_p(f)|^2$ et que $a(C) = \pi \sum_{p\in\mathbb{Z}} p\, |c_p(f)|^2$ en notant $c_p(f)$ les coefficients de Fourier de la fonction T-périodique égale à f sur $[0, T[$.

(e) En déduire l'*inégalité isopérimétrique* : $a(C) \leq \ell(C)^2/4\pi$, l'égalité étant vérifiée si et seulement si la courbe C est un cercle.

10.10. Dans cet exercice, on se donne un *polynôme trigonométrique réel f de degré d*, c'est-à-dire une fonction à valeurs réelles de la forme $f(t) = \sum_{|p|\leq d} c_p\, e^{ipt}$ (on a donc $c_{-p} = \overline{c_p}$ pour tout p).

(a) Soient $c \in \mathbb{R}$ vérifiant $|c| > d \sup_{\mathbb{R}} |f|$, $t_0 \in \mathbb{R}$ et $g(t) = d\, f(t) - c \sin\big(d(t - t_0)\big)$. Montrer que la fonction g est 2π-périodique et que g' s'annule au moins $2d$ fois par période.

(b) Soit $t_0 \in \mathbb{R}$ tel que $|f'(t_0)| = \sup_{\mathbb{R}} |f'|$. Montrer que si $|f'(t_0)| > d \sup_{\mathbb{R}} |f|$, alors on peut trouver un polynôme trigonométrique g de degré d tel que g'' s'annule au moins $2d + 1$ fois par période, puis montrer que cela conduit à une contradiction.

(c) En déduire l'*inégalité de Bernstein* : $\sup_{\mathbb{R}} |f'| \leq d \sup_{\mathbb{R}} |f|$.

10.11. En notant $\mathfrak{F}_q$ le noyau de Fejér, on définit le *noyau de Jackson* par la formule $\mathscr{J}_q(t) = 2\pi\Big(\int_{-\pi}^{\pi} \mathfrak{F}_q(s)^2\, ds\Big)^{-1} \mathfrak{F}_q(t)^2$.

(a) Montrer que $\mathscr{J}_q$ est un polynôme trigonométrique de degré $2q$ (voir l'exercice **10.10** ci-dessus) vérifiant $\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathscr{J}_q(t)\, dt = 1$ et $\frac{1}{2\pi}\int_{-\pi}^{\pi} |t|\, \mathscr{J}_q(t)\, dt \leq \pi^5/6q$.

(b) Pour $k \in \mathbb{N}$ et $r \in]0, 1[$, on dit que la fonction $f : \mathbb{R} \to \mathbb{R}$ est *höldérienne d'ordre $k + r$* si f est de classe C^k et vérifie $|f^{(k)}(s) - f^{(k)}(t)| \leq C\, |s - t|^r$ pour une constante C indépendante de s et t. Montrer que si f est une fonction 2π-périodique höldérienne d'ordre $k + r$, il existe une suite (f_q) de polynômes trigonométriques dont les degrés d_q tendent vers l'infini telle que $\sup_{\mathbb{R}} |f_q - f| \leq C/d_q^{k+r}$ (on démontrera par récurrence sur k qu'il existe une telle suite vérifiant de plus $\frac{1}{2\pi}\int_{-\pi}^{\pi} f_q(t)\, dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\, dt$, et pour $k = 0$ on choisira le produit de convolution $f_q = f * \mathscr{J}_q$ défini à l'exercice **10.2** de la page 82).

(c) Réciproquement supposons qu'il existe un réel $\ell > k \in \mathbb{N}$ et une suite (f_q) de polynômes trigonométriques dont les degrés d_q tendent vers l'infini telle que $\sup_{\mathbb{R}} |f_q - f| \leq C/d_q^{\ell}$. Montrer que f est de classe C^k (on pourra utiliser l'inégalité de Bernstein de l'exercice **10.10** ci-dessus).

10.12. On rappelle que l'on appelle *fonction entière* la somme d'une série entière de rayon de convergence infini.

(a) Soit $f(z) = \sum_{p=0}^{\infty} a_p\, z^p$ une série entière de rayon de convergence $R > r \geq 0$. Montrer que la fonction d'une variable réelle $t \mapsto f(r\, e^{it})$ est continue et 2π-périodique, calculer ses coefficients de Fourier, et en déduire que $\sum_{p=0}^{\infty} |a_p|^2\, r^{2p} = \frac{1}{2\pi}\int_0^{2\pi} |f(r\, e^{it})|^2\, dt$.

(b) Démontrer le *théorème de Liouville* : toute fonction entière bornée est constante.

(c) En admettant que si f est une fonction entière qui ne s'annule pas sur $\mathbb{C}$, alors $1/f$ est aussi une fonction entière, en déduire le *théorème de d'Alembert-Gauß* : tout polynôme non constant à coefficients complexes possède au moins une racine complexe.

Appendice – Préliminaires indispensables

Les mathématiques de licence se situent à un niveau nettement plus abstrait que celles du premier cycle. On y fait usage, dès les premiers cours, de théorie des ensembles et de raisonnements logiques qui déconcertent souvent les étudiants. C'est pourquoi il nous a semblé indispensable de présenter ici de façon concise les connaissances exigées des étudiants dans ces domaines. Ces connaissances relèvent le plus souvent du simple bon sens, mais l'étudiant est vivement encouragé à les étudier méticuleusement afin d'acquérir l'aisance parfaite qui le rendra à même de profiter pleinement des enseignements de licence.

De plus, il va sans dire que les mathématiques du premier cycle seront supposées connues. En particulier, pour les besoins de l'intégration, il sera essentiel de bien maîtriser les quelques propriétés des fonctions de plusieurs variables réelles rappelées ci-dessous. Là encore, l'étudiant est chaleureusement invité à consacrer du temps à ces rappels afin de ne pas risquer d'être perdu dès les premiers cours.

Théorie des ensembles et logique mathématique

1. Théorie naïve[1] des ensembles et logique formelle

Ensembles et parties d'un ensemble. Au niveau le plus fondamental, le discours mathématique porte sur des objets regroupés au sein d'*ensembles*. Un ensemble est entièrement caractérisé par la liste de ses *éléments*, ce qui signifie que deux ensembles sont égaux si et seulement s'ils ont les mêmes éléments. L'ensemble qui ne possède aucun élément est appelé *l'ensemble vide* et on le note $\emptyset$.

Si E est un ensemble, on considérera des propriétés $p(x)$ qui dépendent d'un paramètre $x \in E$, propriétés qui sont donc vraies pour certains éléments x de E et fausses pour les autres. Les éléments pour lesquels la propriété $p(x)$ est vraie forment un nouvel ensemble P appelé *partie de* E et noté $P = \{\, x \in E\,;\, p(x) \,\}$. Si deux propriétés $p(x)$ et $q(x)$ définissent la même partie de E, on dit alors qu'elles sont *équivalentes*, et on peut remplacer partout l'une par l'autre.

Logique formelle. Étant données deux propriétés $p(x)$ et $q(x)$ dépendant d'un paramètre $x \in E$, on peut considérer de nouvelles propriétés $(\text{non}\, p)(x)$, $(p \,\text{et}\, q)(x)$ et $(p \,\text{ou}\, q)(x)$ définies comme suit : les éléments $x \in E$ pour lesquels la propriété $(\text{non}\, p)(x)$ est vraie sont ceux pour lesquels la propriété $p(x)$ est fausse ; de même, les éléments $x \in E$ pour lesquels la propriété $(p \,\text{et}\, q)(x)$ est vraie sont ceux pour lesquels les propriétés $p(x)$ et $q(x)$ sont toutes deux vraies ; enfin, les éléments $x \in E$ pour lesquels la propriété $(p \,\text{ou}\, q)(x)$ est vraie sont ceux pour lesquels l'une au moins des deux propriétés $p(x)$ et $q(x)$ est vraie. Si $P = \{\, x \in E\,;\, p(x) \,\}$ et $Q = \{\, x \in E\,;\, q(x) \,\}$, les ensembles $\{\, x \in E\,;\, (\text{non}\, p)(x) \,\}$, $\{\, x \in E\,;\, (p \,\text{et}\, q)(x) \,\}$ et $\{\, x \in E\,;\, (p \,\text{ou}\, q)(x) \,\}$ s'appellent respectivement le *complémentaire* de P (noté $E \setminus P$), l'*intersection* de P et Q (notée $P \cap Q$) et la *réunion* de P et Q (notée $P \cup Q$).

Il est facile de voir que la propriété $(\text{non}\,(\text{non}\, p))(x)$ est équivalente à la propriété $p(x)$. Lorsque plusieurs propriétés interviennent, il est commode de dresser un tableau récapitulant toutes les

[1] Cette expression s'oppose à celle de théorie *axiomatique* des ensembles. Dans la théorie axiomatique, on bâtit une théorie mathématique à partir d'axiomes énonçant les règles selon lesquelles certaines collections d'objets mathématiques peuvent être considérées comme ensembles (il est classique par exemple de démontrer dans cette théorie que la collection de tous les ensembles ne forme pas un ensemble). Ici, il s'agit simplement de passer en revue quelques-unes de ces règles en s'appuyant sur la notion intuitive d'ensemble.

situations possibles. Dans le cas de trois propriétés par exemple, ce tableau, appelé *table de vérité*, se présente sous la forme suivante :

p	q	r	p et q	p et r	q et r	q ou r	p et $(q$ et $r)$	$(p$ et $q)$ et $(p$ et $r)$	p et $(q$ ou $r)$	$(p$ et $q)$ ou $(p$ et $r)$
V	V	V	V	V	V	V	V	V	V	V
V	V	F	V	F	F	V	F	F	V	V
V	F	V	F	V	F	V	F	F	V	V
V	F	F	F	F	F	F	F	F	F	F
F	V	V	F	F	V	V	F	F	F	F
F	V	F	F	F	F	V	F	F	F	F
F	F	V	F	F	F	V	F	F	F	F
F	F	F	F	F	F	F	F	F	F	F

p	q	r	p ou q	p ou r	q et r	q ou r	p ou $(q$ et $r)$	$(p$ ou $q)$ et $(p$ ou $r)$	p ou $(q$ ou $r)$	$(p$ ou $q)$ ou $(p$ ou $r)$
V	V	V	V	V	V	V	V	V	V	V
V	V	F	V	V	F	V	V	V	V	V
V	F	V	V	V	F	V	V	V	V	V
V	F	F	V	V	F	F	V	V	V	V
F	V	V	V	V	V	V	V	V	V	V
F	V	F	V	F	F	V	F	F	V	V
F	F	V	F	V	F	V	F	F	V	V
F	F	F	F	F	F	F	F	F	F	F

Dans ces tableaux, la troisième ligne par exemple représente le cas des éléments $x \in E$ pour lesquels les propriétés $p(x)$ et $r(x)$ sont vraies tandis que $q(x)$ est fausse. Les groupes de colonnes à droite des tableaux établissent l'équivalence de certaines propriétés composées. Ainsi on constate que :

$(p$ et $(q$ et $r))(x)$ est équivalente à $((p$ et $q)$ et $(p$ et $r))(x)$

$(p$ ou $(q$ et $r))(x)$ est équivalente à $((p$ ou $q)$ et $(p$ ou $r))(x)$

$(p$ et $(q$ ou $r))(x)$ est équivalente à $((p$ et $q)$ ou $(p$ et $r))(x)$

$(p$ ou $(q$ ou $r))(x)$ est équivalente à $((p$ ou $q)$ ou $(p$ ou $r))(x)$.

De la même façon, on voit que (non $(p$ et $q))(x)$ est équivalente à $(($non $p)$ ou (non $q))(x)$, et que (non $(p$ ou $q))(x)$ est équivalente à $(($non $p)$ et (non $q))(x)$ en dressant le tableau suivant :

p	q	non p	non q	p et q	p ou q	non $(p$ et $q)$	(non $p)$ ou (non $q)$	non $(p$ ou $q)$	(non $p)$ et (non $q)$
V	V	F	F	V	V	F	F	F	F
V	F	F	V	F	V	V	V	F	F
F	V	V	F	F	V	V	V	F	F
F	F	V	V	F	F	V	V	V	V

Comme deux propriétés équivalentes définissent la même partie de E, toutes les équivalences signalées ci-dessus se traduisent par des identités correspondantes pour les complémentaires, intersections et réunions (voir notamment l'exercice **0.2** de la page 89).

2. Quantificateurs et démonstrations mathématiques

Quantificateurs. Soit $p(x)$ une propriété dépendant d'un paramètre $x \in E$. Lorsque la propriété $p(x)$ est vraie pour tous les éléments $x \in E$ sans exception, on a donc $\{\, x \in E\,;\, p(x)\,\} = E$, et cette propriété se note $\forall x \in E,\ p(x)$, ce qui se lit : "quel que soit l'élément x de E, $p(x)$ est vraie". De même, lorsque la propriété $p(x)$ est vraie pour au moins un élément $x \in E$, on a donc $\{\, x \in E\,;\, p(x)\,\} \neq \emptyset$, et cette propriété se note $\exists x \in E : p(x)$, ce qui se lit : "il existe un élément x de E tel que $p(x)$ est vraie". Le symbole $\forall$ est le *quantificateur universel*, et le symbole $\exists$ est le *quantificateur existentiel*. Dans ces expressions, la lettre x n'a de sens qu'à l'intérieur de cet énoncé, et pourrait être remplacée par n'importe quelle autre lettre à condition d'effectuer ce remplacement partout : on dit que c'est une *variable muette*.

Dire que l'affirmation $\forall x \in E, p(x)$ est fausse, c'est dire que l'affirmation $\exists x \in E : (\text{non}\, p)(x)$ est vraie. En effet, la première signifie que $\{ x \in E ; p(x) \} = E$ tandis que la seconde s'écrit $\{ x \in E ; (\text{non}\, p)(x) \} \neq \emptyset$, et elles affirment le contraire l'une de l'autre puisque $\emptyset$ est le complémentaire de E. On démontre de la même façon que dire que $\exists x \in E : p(x)$ est fausse, c'est dire que $\forall x \in E, (\text{non}\, p)(x)$ est vraie. L'étudiant trouvera dans l'exercice **0.1** de la page 89 de quoi s'exercer à énoncer le contraire d'affirmations comportant un ou plusieurs quantificateurs.

Soient maintenant $p(x)$ et $q(x)$ deux propriétés dépendant d'un paramètre $x \in E$, et notons $P = \{ x \in E ; p(x) \}$ et $Q = \{ x \in E ; q(x) \}$ les parties correspondantes. Si la partie Q contient au moins tous les éléments de P (et éventuellement d'autres), on dit que P est *inclus dans* Q, ce qui se note $P \subset Q$. On vérifie facilement que l'affirmation $P \subset Q$ équivaut à chacune des affirmations suivantes : $(E \setminus Q) \subset (E \setminus P)$, $P \cap Q = P$, $P \cup Q = Q$, $Q \cup (E \setminus P) = E$, $P \cap (E \setminus Q) = \emptyset$. Si on note $(p \Rightarrow q)(x)$ la propriété $(q\, \text{ou}\, (\text{non}\, p))(x)$, on voit également que cela équivaut à affirmer que $\forall x \in E, (p \Rightarrow q)(x)$ est vraie, ou que $\forall x \in E, ((\text{non}\, q) \Rightarrow (\text{non}\, p))(x)$ est vraie, ou encore que $\exists x \in E : (p\, \text{et}\, (\text{non}\, q))(x)$ est fausse, et on dit alors que la propriété $p(x)$ *implique* la propriété $q(x)$. Si on note $(p \Leftrightarrow q)(x)$ la propriété $((p \Rightarrow q)\, \text{et}\, (q \Rightarrow p))(x)$, on voit aussi que l'affirmation $\forall x \in E, (p \Leftrightarrow q)(x)$ signifie que l'on a à la fois $P \subset Q$ et $Q \subset P$, ce qui revient à dire que $P = Q$, ou encore que les propriétés $p(x)$ et $q(x)$ sont équivalentes.

Différents types de démonstration. Une démonstration en mathématiques consiste le plus souvent à établir une affirmation du type $\forall x \in E, (h \Rightarrow c)(x)$, où la propriété $h(x)$ s'appelle l'*hypothèse* et la propriété $c(x)$ s'appelle la *conclusion*. Classiquement, il existe principalement trois façons de procéder :

(a) *La démonstration directe.* Dans ce type de démonstration, on commence par fixer un élément $x \in E$ pour lequel $h(x)$ est vraie. L'élément $x \in E$ est fixé, mais on s'interdit d'utiliser d'autres informations que la propriété $h(x)$, en sorte que cette démonstration soit valable pour tous les $x \in E$ pour lesquels $h(x)$ est vraie. On procède alors par déductions successives, en distinguant plusieurs cas si c'est utile, jusqu'à ce que l'on obtienne que la propriété $c(x)$ est vraie, auquel cas la démonstration est terminée.

(b) *La démonstration par la contraposée.* La propriété $((\text{non}\, c) \Rightarrow (\text{non}\, h))(x)$ est appelée *contraposée* de la propriété $(h \Rightarrow c)(x)$. La démonstration par la contraposée consiste à établir une démonstration directe de l'affirmation $\forall x \in E, ((\text{non}\, c) \Rightarrow (\text{non}\, h))(x)$, ce qui, comme on l'a indiqué plus haut, revient à établir l'affirmation habituelle $\forall x \in E, (h \Rightarrow c)(x)$. Ainsi, on commence par fixer un élément $x \in E$ pour lequel $c(x)$ est fausse, puis on procède par déductions successives jusqu'à obtenir que $h(x)$ est fausse.

(c) *La démonstration par l'absurde.* Dans ce type de démonstration, on cherche plutôt à établir que l'affirmation $\exists x \in E : (h\, \text{et}\, (\text{non}\, c))(x)$ est fausse, ce qui, comme on l'a là encore indiqué plus haut, revient à établir que l'affirmation $\forall x \in E, (h \Rightarrow c)(x)$ est vraie. En pratique, on commence donc par fixer un élément $x \in E$ pour lequel $h(x)$ est vraie et $c(x)$ est fausse, puis on procède par déductions successives jusqu'à obtenir une contradiction qui montre qu'un tel élément $x \in E$ ne peut pas exister.

À ces trois types de démonstration, il convient d'en ajouter un quatrième d'un genre assez différent :

(d) *La démonstration par récurrence.* Pour établir une affirmation de la forme $\forall n \in \mathbb{N}, p(n)$ où $\mathbb{N}$ désigne l'ensemble des entiers naturels, on dispose d'un procédé spécifique appelé *récurrence*. Une démonstration par récurrence comprend deux parties : une *initialisation*, et une preuve de la propriété de *transmission*. L'initialisation consiste à montrer que $p(0)$ est vraie (ou parfois $p(n_0)$ pour un $n_0 > 0$), tandis que prouver la propriété de transmission, c'est établir qu'à chaque fois que la propriété est vraie à un certain rang n, alors elle est aussi vraie au rang suivant $n + 1$; de façon plus formalisée, c'est démontrer l'affirmation $\forall n \in \mathbb{N}, (p \Rightarrow p_1)(n)$ en notant $p_1(n)$ la propriété vraie par définition si et seulement si $p(n + 1)$ est vraie. Il est souvent plus parlant de noter une propriété $(p \Rightarrow q)(x)$ sous la forme $(p(x) \Rightarrow q(x))$; avec cette convention, la propriété de transmission s'écrirait aussi $\forall n \in \mathbb{N}, (p(n) \Rightarrow p(n + 1))$.

3. Propriétés dépendant de plusieurs paramètres

Produits cartésiens d'ensembles, relations d'équivalence et relations d'ordre. Si E et F

sont deux ensembles, les couples (x, y) formés d'un élément $x \in E$ et d'un élément $y \in F$ constituent un ensemble noté $E \times F$ et appelé *produit cartésien* des ensembles E et F. Si une propriété $p(x, y)$ dépend de deux paramètres $x \in E$ et $y \in F$, on peut aussi bien la considérer comme dépendant d'un seul paramètre $(x, y) \in E \times F$, et la partie de $E \times F$ correspondante s'appelle aussi le *graphe* de p. Si $p(x, y)$ est une telle propriété, on note $p^{-1}(y, x)$ la propriété dépendant du paramètre $(y, x) \in F \times E$ qui est vraie lorsque $p(x, y)$ est vraie, et si $q(y, z)$ est une propriété dépendant du paramètre $(y, z) \in F \times G$, on note $q \circ p(x, z)$ la propriété dépendant du paramètre $(x, z) \in E \times G$ qui est vraie lorsqu'il existe un $y \in F$ tel que $p(x, y)$ et $q(y, z)$ soient toutes deux vraies. Ces propriétés s'appellent propriété *réciproque* et propriété *composée*. Enfin, si E est un ensemble, nous noterons ici $i(x, y)$ la propriété d'égalité sur E, qui est la propriété dépendant du paramètre $(x, y) \in E \times E$ vraie si et seulement si $x = y$.

Si E est un ensemble, une propriété $p(x, y)$ dépendant du paramètre $(x, y) \in E \times E$ est une *relation d'équivalence sur E* si les trois affirmations suivantes sont vraies : $\forall x \in E, p(x, x)$ (réflexivité), $\forall (x, y) \in E \times E, (p \Rightarrow p^{-1})(x, y)$ (symétrie), $\forall (x, y) \in E \times E, ((p \circ p) \Rightarrow p)(x, y)$ (transitivité). Avec la convention plus parlante introduite plus haut, la propriété de symétrie s'énonce aussi $\forall (x, y) \in E \times E, (p(x, y) \Rightarrow p(y, x))$ tandis que la propriété de transitivité peut aussi s'énoncer $\forall (x, y, z) \in E \times E \times E, (((p(x, z) \text{ et } p(z, y)) \Rightarrow p(x, y))$. Le prototype de la relation d'équivalence, c'est l'égalité, et on s'y ramène en considérant les *classes d'équivalence* pour la relation d'équivalence $p(x, y)$, qui sont les parties de E définies par la formule $\dot{x} = \{ y \in E ; p(x, y) \}$. Si $p(x, y)$ est vraie, alors $\dot{x} = \dot{y}$, tandis que si $p(x, y)$ est fausse, alors $\dot{x} \cap \dot{y} = \emptyset$. Les classes d'équivalence forment alors un nouvel ensemble appelé *ensemble quotient* de E par p, et noté E/p.

Si E est un ensemble, une propriété $p(x, y)$ dépendant du paramètre $(x, y) \in E \times E$ est une *relation d'ordre sur E* si les trois affirmations suivantes sont vraies : $\forall x \in E, p(x, x)$ (réflexivité), $\forall (x, y) \in E \times E, ((p \text{ et } p^{-1}) \Rightarrow i)(x, y)$ (antisymétrie), $\forall (x, y) \in E \times E, ((p \circ p) \Rightarrow p)(x, y)$ (transitivité). Avec la convention plus parlante déjà introduite plus haut, la propriété d'antisymétrie peut aussi s'énoncer $\forall (x, y) \in E \times E, ((p(x, y) \text{ et } p(y, x)) \Rightarrow x = y)$. De plus, on dit que l'ordre est *total* si l'affirmation $\forall (x, y) \in E \times E, (p \text{ ou } p^{-1})(x, y)$ est vraie, ou si l'on préfère, si $\forall (x, y) \in E \times E, (p(x, y) \text{ ou } p(y, x))$ est vraie. Si $p(x, y)$ est une relation d'ordre sur E et si A est une partie de E, on appelle *majorant* de A tout élément $a \in E$ pour lequel la propriété $\forall x \in A, p(x, a)$ est vraie. Si la partie A contient l'un de ses majorants, ce dernier est appelé *le plus grand élément de A*, et A n'en contient pas d'autre. On appelle de même *minorant* de A tout élément $a \in E$ pour lequel la propriété $\forall x \in A, p(a, x)$ est vraie, et le minorant de A qui appartient à A, s'il existe, est appelé *le plus petit élément de A*. Enfin, lorsqu'ils existent, on appelle *borne supérieure de A* le plus petit élément de l'ensemble des majorants de A, et *borne inférieure de A* le plus grand élément de l'ensemble des minorants de A.

Fonctions. On dit que la propriété $p(x, y)$ dépendant du paramètre $(x, y) \in E \times F$ définit une *fonction* $f : E \to F$ si pour chaque élément $x \in E$ fixé, l'ensemble $\{ y \in F ; p(x, y) \}$ possède exactement un élément. Cet élément $y \in F$ déterminé par x est alors noté $f(x)$ et appelé *image de x par f*, et si A est une partie de E et B une partie de F, on appellera plus généralement *image par f de A* la partie de F définie par $f(A) = \{ f(x) \in F ; x \in A \}$ et *image réciproque par f de B* la partie de E définie par $f^{-1}(B) = \{ x \in E ; f(x) \in B \}$. L'ensemble E s'appelle alors l'ensemble de *départ* de f, et l'ensemble F s'appelle l'ensemble d'*arrivée*. On dit aussi que f est *définie sur E* et *à valeurs dans F*. Si les propriétés $p(x, y)$ et $q(y, z)$ dépendant des paramètres $(x, y) \in E \times F$ et $(y, z) \in F \times G$ définissent des fonctions $f : E \to F$ et $g : F \to G$, alors la propriété $q \circ p(x, z)$ définit sur E une fonction à valeurs dans G qui est notée $g \circ f$ et qui vérifie $g \circ f(x) = g\big(f(x)\big)$.

Soit $f : E \to F$ une fonction définie par une propriété $p(x, y)$. On dit que f est *injective* si l'équation $y = f(x)$ ne possède jamais plus d'une solution $x \in E$ pour $y \in F$ fixé, et qu'elle est *surjective* si cette équation possède toujours au moins une solution $x \in E$ pour tout $y \in F$ fixé. La fonction f est donc à la fois injective et surjective si et seulement si l'équation $y = f(x)$ possède exactement une solution $x \in E$ pour tout $y \in F$, c'est-à-dire si la propriété $p^{-1}(y, x)$ définit elle-même une fonction. Dans ce cas, la fonction f est dite *bijective*, et la fonction définie par la propriété $p^{-1}(y, x)$ est alors notée $f^{-1} : F \to E$ et appelée *réciproque* de f.

Étant données une fonction $f : E \to F$ définie par une propriété $p(x, y)$, une partie A de E et une partie B de F, on peut considérer la propriété $p_{|A \times B}(x, y)$ dépendant du paramètre $(x, y) \in A \times B$ vraie par définition si $p(x, y)$ est vraie. Cette propriété définira sur A une fonction à valeurs dans B pourvu que $f(A) \subset B$. Dans ce cas, la fonction définie par $p_{|A\times B}(x,y)$ est notée $f_{|A}$ et appelée *restriction* de f à A. On prendra bien garde que les propriétés d'injectivité et de surjectivité peuvent être modifiées lorsque l'on passe d'une fonction f à une restriction $f_{|A}$.

Familles de parties d'un ensemble. La collection de toutes les parties d'un ensemble E forme un nouvel ensemble noté $\mathscr{P}(E)$. On appelle *famille de parties de E indicée par l'ensemble I* la donnée d'une fonction définie sur I et à valeurs dans $\mathscr{P}(E)$, qui donc à chaque élément $i \in I$ associe une partie de E que nous noterons A_i, la famille étant alors notée $(A_i)_{i\in I}$. Étant donnée une telle famille $(A_i)_{i\in I}$ de parties de E, on appelle *intersection* de la famille la partie de E définie par $\bigcap_{i\in I} A_i = \{x \in E; \forall i \in I,\ x \in A_i\}$, et de même on définit la *réunion* de la famille par $\bigcup_{i\in I} A_i = \{x \in E; \exists i \in I:\ x \in A_i\}$. On appellera *intersection finie* de parties de E l'intersection d'une famille $(A_i)_{i\in I}$ dont l'ensemble I des indices n'a qu'un nombre fini d'éléments, et de même on utilisera l'expression *réunion finie* pour désigner la réunion d'une telle famille. Diverses propriétés des familles de parties sont considérées dans l'exercice **0.3** ci-dessous.

4. Exercices

0.1. Énoncer le contraire des affirmations suivantes.

(a) Dans tous les amphis de licence, il y a un étudiant gentil mais pas fort en maths.
Solution : il y a un amphi de licence où tous les étudiants sont méchants ou forts en maths.

(b) Dans tous les gouvernements, il y a un ministre dont une responsabilité antérieure l'a mêlé à une affaire pas claire ou illégale.

(c) Il y a des millionnaires dont tous les héritiers sont soit impatients et intéressés, soit ignorants de leur fortune et floués par les autres héritiers.

(d) Tous ceux qui diront la messe sur cet autel le feront gratis et pour l'amour de Dieu et sans aucun espoir de récompense temporelle, sans quoi ils ne pourront bénéficier du privilège de retirer une âme du Purgatoire de la même façon que s'ils la disaient sur les autels privilégiés de Rome (d'après St François-Xavier, *lettre 16*).

(e) Ne vous imaginez pas être différente de ce qu'il eût pu sembler à autrui que vous fussiez ou eussiez pu être en restant identique à ce que vous fûtes sans jamais paraître autre que vous n'étiez avant d'être devenue ce que vous êtes (d'après L. Carroll, *Alice's adventures in Wonderland*).

(f) Jamais une femme n'a été capable de faire progresser les sciences mathématiques ni les sciences expérimentales (d'après J. Verne, *Sans dessus dessous*).

0.2. Soient E et F deux ensembles et $f : E \to F$ une fonction. Soient A_1 et A_2 deux parties de E, B_1 et B_2 deux parties de F. On rappelle les définitions $f(A) = \{y \in F;\ y = f(x) \text{ pour au moins un } x \in A\}$ et $f^{-1}(B) = \{x \in E;\ f(x) \in B\}$. Montrer les relations suivantes.

(a) $E \setminus (A_1 \cup A_2) = (E \setminus A_1) \cap (E \setminus A_2)$ et $E \setminus (A_1 \cap A_2) = (E \setminus A_1) \cup (E \setminus A_2)$.

(b) $f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$ et $f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2)$ sans qu'il y ait toujours égalité.
Existe-t-il une relation entre $f(E \setminus A_1)$ et $F \setminus f(A_1)$? Et si on suppose f injective ou surjective ?

(c) $f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$, $f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$,
et $f^{-1}(F \setminus B_1) = E \setminus f^{-1}(B_1)$.

(d) $f\big(f^{-1}(B_1)\big) \subset B_1$ avec égalité si f est surjective, et $f^{-1}\big(f(A_1)\big) \supset A_1$ avec égalité si f est injective.

(e) $(A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \cap A_2) \times (B_1 \cap B_2)$,
$(A_1 \times B_1) \cup (A_2 \times B_2) \subset (A_1 \cup A_2) \times (B_1 \cup B_2)$ sans qu'il y ait toujours égalité,
et $(E \times F) \setminus (A_1 \times B_1) = \big((E \setminus A_1) \times F\big) \cup \big(E \times (F \setminus B_1)\big)$.

0.3. Soient E et F deux ensembles, A une partie de E et $f : E \to F$ une fonction. Soient encore $(A_i)_{i\in I}$ une famille de parties de E et $(B_i)_{i\in I}$ une famille de parties de F. Montrer les relations suivantes.

(a) $A \cap (\bigcup_{i\in I} A_i) = \bigcup_{i\in I} (A \cap A_i)$, $A \cup (\bigcap_{i\in I} A_i) = \bigcap_{i\in I} (A \cup A_i)$,
$E \setminus (\bigcup_{i\in I} A_i) = \bigcap_{i\in I} (E \setminus A_i)$ et $E \setminus (\bigcap_{i\in I} A_i) = \bigcup_{i\in I} (E \setminus A_i)$.

(b) $f^{-1}(\bigcup_{i\in I} B_i) = \bigcup_{i\in I} f^{-1}(B_i)$ et $f^{-1}(\bigcap_{i\in I} B_i) = \bigcap_{i\in I} f^{-1}(B_i)$.

(c) $\bigcap_{i\in I} (A_i \times B_i) = (\bigcap_{i\in I} A_i) \times (\bigcap_{i\in I} B_i)$ et $(\bigcup_{i\in I} A_i) \times (\bigcup_{i\in I} B_i) = \bigcup_{i\in I, j\in I} (A_i \times B_j)$.

Rappels de premier cycle : fonctions de plusieurs variables réelles

Voici, avec peut-être quelques compléments, les résultats de premier cycle qu'il sera bon de conserver en mémoire pour aborder convenablement le cours d'intégration.

5. Voisinages, ouverts, fermés, compacts et pavés

Normes et boules. On note ici $\mathbb{R}^n$ l'espace vectoriel des n-uplets $x = (x_1, \ldots, x_n)$ de nombres réels. On appelle *norme sur* $\mathbb{R}^n$ toute application $N : \mathbb{R}^n \to \mathbb{R}_+$ vérifiant les trois propriétés suivantes : $\forall (c, x) \in \mathbb{R} \times \mathbb{R}^n$, $N(cx) = |c| N(x)$ (homogénéité), $\forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, $N(x + y) \leq N(x) + N(y)$ (inégalité de convexité), et $N(x) = 0 \;\Rightarrow\; x = 0$ (séparation). L'exemple le plus classique est celui de la norme *euclidienne* $x \mapsto |x| = \sqrt{x_1^2 + \ldots + x_n^2}$, mais les applications $x \mapsto |x|_1 = \sum_{j=1}^n |x_j|$ et $x \mapsto |x|_\infty = \max_{1 \leq j \leq n} |x_j|$ (par exemple) en sont deux autres d'un usage très courant. Étant donnée une norme N sur $\mathbb{R}^n$, on appelle *boule* (ouverte) de centre $x_0 \in \mathbb{R}^n$ et de rayon $r > 0$ la partie de $\mathbb{R}^n$ définie par $B_N(x_0, r) = \{\, x \in \mathbb{R}^n \,;\, N(x - x_0) < r \,\}$ (nous ne considérerons ici que des boules ouvertes, et de rayon > 0). On prendra garde que si les boules euclidiennes ont bien la rondeur suggérée par leur nom, les boules pour la norme ∞ par exemple sont des cubes !

Voisinages, ouverts et fermés. On appelle *voisinage* d'un point $x \in \mathbb{R}^n$ toute partie V de $\mathbb{R}^n$ contenant une boule de centre x. On appelle ensuite *ouvert de* $\mathbb{R}^n$ toute partie de $\mathbb{R}^n$ qui est voisinage de chacun de ses points, et on appelle *fermé de* $\mathbb{R}^n$ toute partie de $\mathbb{R}^n$ dont le complémentaire est un ouvert de $\mathbb{R}^n$. Intuitivement, ces définitions signifient qu'un voisinage de x est une partie de $\mathbb{R}^n$ contenant au moins tous les points "très proches" de x, qu'aucun point d'un ouvert n'est situé "au bord" de cet ouvert, ou si l'on préfère, que tous les points d'un ouvert sont situés "à l'intérieur" de cet ouvert, et symétriquement que tous les points situés "au bord" d'un fermé appartiennent automatiquement à ce fermé. On peut montrer que la collection des ouverts de $\mathbb{R}^n$ définit une *topologie*, c'est-à-dire qu'elle vérifie les trois propriétés suivantes : $\emptyset$ et $\mathbb{R}^n$ sont des ouverts, toute intersection finie d'ouverts est un ouvert, toute réunion (quelconque) d'ouverts est un ouvert (c'est l'exercice **0.4** de la page 93). Les normes explicitées plus haut vérifient les inégalités $|x|_\infty \leq |x| \leq |x|_1 \leq n\,|x|_\infty$, ce qui a pour conséquence que toute boule pour l'une de ces normes contient des boules de même centre pour les deux autres, et cette propriété permet d'établir qu'on obtient les mêmes voisinages et les mêmes collections d'ouverts et de fermés lorsque l'on remplace dans la définition donnée plus haut les boules euclidiennes par des boules pour l'une ou l'autre de ces normes (on exprime cette propriété en disant que ces trois normes sont *équivalentes*).

Étant donnée une partie A de $\mathbb{R}^n$, on appelle *intérieur de* A l'ensemble $\overset{\circ}{A}$ obtenu par réunion de tous les ouverts contenus dans A, et on appelle *adhérence de* A l'ensemble $\overline{A}$ obtenu par intersection de tous les fermés contenant A. On peut montrer que $\overset{\circ}{A}$ est le plus grand ouvert contenu dans A et que $\overline{A}$ est le plus petit fermé contenant A, et on appelle *frontière* ou *bord* de A le fermé $\overline{A} \setminus \overset{\circ}{A}$ (voir l'exercice **0.5** de la page 93).

Compacts et pavés de $\mathbb{R}^n$. Il y a lieu de distinguer une catégorie de fermés qui jouissent de propriétés particulières : ce sont les fermés *bornés*, c'est-à-dire les fermés de $\mathbb{R}^n$ contenus dans une grande boule. En effet, ce sont les parties *compactes* de $\mathbb{R}^n$, c'est-à-dire qu'ils vérifient la propriété *de Borel-Lebesgue* qui définit les compacts : si un compact est contenu dans la réunion d'une famille d'ouverts, alors il est aussi contenu dans la réunion d'une sous-famille *finie* de ces ouverts. Autrement dit, si K est une partie fermée et bornée de $\mathbb{R}^n$ vérifiant $K \subset \cup_{i \in I} U_i$ pour une famille $(U_i)_{i \in I}$ d'ouverts de $\mathbb{R}^n$, alors il existe une partie finie J de I telle que $K \subset \cup_{i \in J} U_i$. On vérifie facilement que toute réunion finie de compacts est encore un compact.

Nous utiliserons les notions ci-dessus principalement dans le cas des pavés de $\mathbb{R}^n$. On appelle *pavé (compact)* de $\mathbb{R}^n$ tout produit cartésien $P = [a_1, b_1] \times \ldots \times [a_n, b_n]$ d'intervalles fermés bornés de $\mathbb{R}$, ce qui s'écrit aussi $P = \{ (x_1, \ldots, x_n) \in \mathbb{R}^n ; a_j \leq x_j \leq b_j \text{ pour } j \leq n \}$. Nous admettrons qu'un tel pavé est compact (montrer à titre d'exercice qu'il est fermé et borné), et que son intérieur vaut

$$\overset{\circ}{P} = \begin{cases} \emptyset & \text{si } a_j = b_j \text{ pour au moins un indice } j, \\ ]a_1, b_1[\times \ldots \times]a_n, b_n[& \text{si } a_j < b_j \text{ pour tous les indices } j. \end{cases}$$

On en déduit que le bord de ce pavé est la réunion de la famille finie de pavés d'intérieurs vides $\prod_{j<k}[a_j, b_j] \times \{a_k\} \times \prod_{j>k}[a_j, b_j]$ et $\prod_{j<k}[a_j, b_j] \times \{b_k\} \times \prod_{j>k}[a_j, b_j]$ pour $1 \leq k \leq n$.

6. Fonctions continues

Continuité. Soient $f : \mathbb{R}^n \to \mathbb{R}^m$ une fonction et $x_0 \in \mathbb{R}^n$. On dit que f est *continue au point* x_0 si toute boule[2] de centre $f(x_0)$ contient l'image d'une boule de centre x_0. Si dans cette définition on note ε le rayon de la boule de centre $f(x_0)$ et δ le rayon de la boule de centre x_0, cela revient à affirmer que $\forall \varepsilon > 0, \exists \delta > 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon$. On dit ensuite que f est *continue* (tout court) si elle est continue en tout point $x \in \mathbb{R}^n$. On peut démontrer (le faire en exercice) qu'une fonction f est continue si et seulement si l'image réciproque par f de tout ouvert de $\mathbb{R}^m$ est un ouvert de $\mathbb{R}^n$, et de même si et seulement si l'image réciproque par f de tout fermé de $\mathbb{R}^m$ est un fermé de $\mathbb{R}^n$. Si f n'est définie que sur une partie A de $\mathbb{R}^n$, on dira que f est *continue sur* A si l'image réciproque par f de tout ouvert de $\mathbb{R}^m$ est égale à l'intersection avec A d'un ouvert de $\mathbb{R}^n$, ce qui peut aussi s'exprimer à l'aide des fermés ou des boules comme pourra le montrer l'étudiant.

Si $f : \mathbb{R}^n \to \mathbb{R}^m$ est une fonction continue, alors pour tout $x \in \mathbb{R}^n$ et tout $\varepsilon > 0$, la boule $B(f(x), \varepsilon)$ contient l'image par f d'une boule $B(x, \delta)$. Bien entendu, le rayon $\delta > 0$ dépend du rayon $\varepsilon > 0$ que l'on s'est donné, mais il peut dépendre aussi du point $x \in \mathbb{R}^n$ considéré. S'il n'en dépend pas, on dira que la fonction f est *uniformément continue*. Cette propriété peut s'écrire $\forall \varepsilon > 0, \exists \delta > 0 : \forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n, |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$, ou encore $\lim_{\delta \to 0} \sup_{|x-y|<\delta} |f(x) - f(y)| = 0$. On dira plus généralement qu'une fonction f est *uniformément continue sur une partie* A de $\mathbb{R}^n$ si $\forall \varepsilon > 0$, $\exists \delta > 0 : \forall (x, y) \in A \times A, |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$.

Fonctions continues sur un compact. Lorsqu'une fonction $f : K \to \mathbb{R}^m$ est continue sur une partie compacte K de $\mathbb{R}^n$, alors d'une part son image $f(K)$ est un compact de $\mathbb{R}^m$, et d'autre part f est uniformément continue sur K. Ces deux propriétés se démontrent à l'aide de la propriété de Borel-Lebesgue. En particulier lorsque $m = 1$, la compacité de $f(K)$ entraîne que cette partie de $\mathbb{R}$ possède un plus grand élément et un plus petit élément, c'est-à-dire qu'il existe $x_1 \in K$ et $x_2 \in K$ tels que $f(x_1) = \sup_{x \in K} f(x)$ et $f(x_2) = \inf_{x \in K} f(x)$. On exprime ces propriétés en disant que *f est bornée et atteint ses bornes sur K*.

7. Suites

Suites d'éléments d'un ensemble. On appelle *suite d'éléments de E* la donnée d'une fonction $f : \mathbb{N} \to E$ où $\mathbb{N}$ désigne l'ensemble des entiers naturels. On appelle *terme de rang p* de la suite l'image x_p de l'entier p par f, et la suite est alors désignée par $(x_p)_{p \in \mathbb{N}}$.

Suites de points de $\mathbb{R}^n$. Soit $(x_p)_{p \in \mathbb{N}}$ une suite de points de $\mathbb{R}^n$. On dit que cette suite *converge vers le point* $\ell \in \mathbb{R}^n$, et on écrit $\lim_{p \to \infty} x_p = \ell$, si toute boule de centre ℓ contient tous les termes de la suite à partir d'un certain rang. En notant ε le rayon d'une boule de centre ℓ, cette propriété s'énonce aussi $\forall \varepsilon > 0, \exists N \in \mathbb{N} : \forall p > N, |x_p - \ell| < \varepsilon$. On appelle *suite convergente* toute suite qui converge vers une limite, et on rappelle le *critère de Cauchy* : une suite $(x_p)_{p \in \mathbb{N}}$ de points de $\mathbb{R}^n$ est convergente si et seulement si $\forall \varepsilon > 0, \exists N \in \mathbb{N} : \forall p, q > N, |x_p - x_q| < \varepsilon$. Les suites servent notamment à caractériser la continuité d'une fonction ou les bornes d'une partie de $\mathbb{R}$.

[2] On prend ici des boules euclidiennes, mais comme précédemment, l'utilisation d'une autre norme ne changerait pas la notion définie.

En effet, on peut démontrer qu'une fonction $f : \mathbb{R}^n \to \mathbb{R}^m$ est continue en un point $x_0 \in \mathbb{R}^n$ si et seulement si on a $\lim_{p\to\infty} f(x_p) = f(x_0)$ pour toute suite $(x_p)_{p\in\mathbb{N}}$ de points de $\mathbb{R}^n$ qui converge vers x_0.

On sait par ailleurs que toute partie de $\mathbb{R}$ non vide et majorée possède une borne supérieure, et que toute partie de $\mathbb{R}$ non vide et minorée possède une borne inférieure. Pour éviter les exceptions, on convient aussi que l'ensemble vide a pour borne inférieure $+\infty$ et pour borne supérieure $-\infty$, et qu'une partie non majorée a pour borne supérieure $+\infty$ tandis qu'une partie non minorée a pour borne inférieure $-\infty$. Avec ces conventions, toute partie de $\mathbb{R}$ possède une borne inférieure et une borne supérieure dans $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$. On démontre alors que si A est une partie non vide de $\mathbb{R}$, il existe une suite d'éléments de A qui converge vers sa borne supérieure, et une suite d'éléments de A qui converge vers sa borne inférieure.

Suites de fonctions à valeurs dans $\mathbb{R}^m$. On considère maintenant une suite $(f_p)_{p\in\mathbb{N}}$ de fonctions définies sur $\mathbb{R}^n$ et à valeurs dans $\mathbb{R}^m$. On dit que cette suite *converge simplement* vers la fonction $f : \mathbb{R}^n \to \mathbb{R}^m$ si pour tout $x \in \mathbb{R}^n$ fixé, la suite $(f_p(x))_{p\in\mathbb{N}}$ de points de $\mathbb{R}^m$ converge vers $f(x)$.

La convergence simple d'une suite de fonctions n'implique pas que cette convergence a lieu partout "à la même vitesse". Autrement dit, pour un $\varepsilon > 0$ fixé, le rang N à partir duquel les termes de la suite $(f_p(x))_{p\in\mathbb{N}}$ appartiennent à la boule $B(f(x), \varepsilon)$ peut dépendre non seulement de ε, mais aussi du point x. Lorsque ce rang ne dépend que de ε mais pas du point $x \in \mathbb{R}^n$, ou si l'on préfère, lorsque l'on peut trouver une vitesse de convergence minimum commune à tous les points $x \in \mathbb{R}^n$, on dit que la convergence de la suite $(f_p)_{p\in\mathbb{N}}$ est *uniforme.* Cette propriété peut s'écrire $\forall\varepsilon > 0, \exists N \in \mathbb{N} : \forall x \in \mathbb{R}^n$, $\forall p > N, |f_p(x) - f(x)| < \varepsilon$. On parle aussi de convergence *uniforme sur une partie A de $\mathbb{R}^n$* si on a $\forall\varepsilon > 0, \exists N \in \mathbb{N} : \forall x \in A, \forall p > N, |f_p(x) - f(x)| < \varepsilon$.

Comme pour les suites de points, la convergence des suites de fonctions peut être caractérisée par des critères de Cauchy qui ne font pas intervenir explicitement la fonction limite. Ainsi la suite $(f_p)_{p\in\mathbb{N}}$ converge simplement sur la partie A de $\mathbb{R}^n$ (vers une fonction f non précisée) si et seulement si on peut écrire $\forall x \in A, \forall\varepsilon > 0, \exists N \in \mathbb{N} : \forall p, q > N, |f_p(x) - f_q(x)| < \varepsilon$. Et de même, cette suite converge uniformément sur A si et seulement si $\forall\varepsilon > 0, \exists N \in \mathbb{N} : \forall p, q > N, \sup_{x\in A} |f_p(x) - f_q(x)| < \varepsilon$.

Citons encore le résultat suivant : si $(f_p)_{p\in\mathbb{N}}$ est une suite de fonctions continues sur une partie A de $\mathbb{R}^n$ et à valeurs dans $\mathbb{R}^m$, et si cette suite converge vers f uniformément sur A, alors la fonction limite f est continue sur A.

8. Calcul différentiel

Dérivée et accroissements finis. Une fonction $f : I \to \mathbb{R}$ définie sur un intervalle I de $\mathbb{R}$ est dite *dérivable en $a \in \mathring{I}$* si le taux d'accroissement $(f(a+h) - f(a))/h$ possède une limite $f'(a)$ lorsque h tend vers 0, et elle est dite dérivable *dans* I si elle est dérivable en tout point de I. La fonction $x \mapsto f'(x)$ est alors appelée *dérivée* de la fonction f. Pour toute fonction $f : [x, y] \to \mathbb{R}$ continue sur $[x, y]$, dérivable dans $]x, y[$, il existe (au moins) un point $z \in]x, y[$ tel que $f(y) - f(x) = f'(z)\,(y - x)$: c'est le théorème des *accroissements finis.* Lorsque f est à valeurs complexes ou à valeurs dans $\mathbb{R}^m$, le théorème subsiste, mais seulement sous forme d'une inégalité : sous les mêmes hypothèses, on obtient alors $|f(y) - f(x)| \leq \sup_{z\in]x,y[} |f'(z)|\,(y - x)$ où les valeurs absolues désignent le module ou la norme euclidienne selon le cas.

Fonctions de classe C^1. Soient U un ouvert de $\mathbb{R}^n$, $f : U \to \mathbb{C}$ une fonction à valeurs complexes et $a = (a_1, \ldots, a_n) \in U$. On dit que la fonction f est *dérivable par rapport à x_j en a* si la fonction d'une seule variable réelle $x_j \mapsto f(a_1, \ldots, a_{j-1}, x_j, a_{j+1}, \ldots, a_n)$ est dérivable en a_j, et on note alors $(\partial f/\partial x_j)(a)$ la dérivée de cette fonction, appelée *dérivée partielle* de f par rapport à x_j en a. Lorsque f est dérivable par rapport à chacune des variables x_j et que les fonctions $x \mapsto (\partial f/\partial x_j)(x)$ sont continues dans U, on dit que f est *de classe* C^1 *dans* U. Intuitivement, cela signifie que près de chaque point $a \in U$, f est bien approchée par une fonction affine.

Si maintenant on considère une fonction $f : U \to \mathbb{R}^m$, on dira que cette fonction est de classe C^1 dans U si chacune de ses composantes $f_k : U \to \mathbb{R}$ l'est. Ses dérivées partielles peuvent alors être disposées en une matrice $J_f(a) = \big((\partial f_k/\partial x_j)(a)\big)_{1\leq j\leq n,\, 1\leq k\leq m}$ appelée *matrice jacobienne de f en a*, qui est, dans les bases canoniques, la matrice d'une application linéaire $f'(a) : \mathbb{R}^n \to \mathbb{R}^m$ appelée *différentielle de f en a*. On montre qu'avec ces notations, la fonction f ne diffère de la fonction affine $x \mapsto f(a) + f'(a)(x-a)$ que d'un $o(|x-a|)$. Le théorème des accroissements finis s'étend aux fonctions $f : U \to \mathbb{R}^m$ de classe C^1 sous une forme appropriée que nous n'explicitons pas.

Si $f : U \to \mathbb{R}^m$ et $g : V \to \mathbb{R}^\ell$ sont de classe C^1 avec $f(U) \subset V$, alors la fonction $g \circ f : U \to \mathbb{R}^\ell$ est de classe C^1, et si $b = f(a)$, on a $(g \circ f)'(a) = g'(b) \circ f'(a)$, ce qui s'écrit aussi $J_{g \circ f}(a) = J_g(b)\, J_f(a)$ avec les matrices jacobiennes. La matrice jacobienne de la fonction $x \mapsto x$, étant la matrice identité I_n, on voit que si la fonction f admet une réciproque f^{-1} de classe C^1, on a $I_n = J_{f^{-1}}(b)\, J_f(a) = J_f(a)\, J_{f^{-1}}(b)$, c'est-à-dire que $J_{f^{-1}}(b)$ doit être la matrice inverse de $J_f(a)$. Réciproquement, le théorème *d'inversion locale* affirme que si $f : U \to \mathbb{R}^n$ est une fonction de classe C^1 dont la matrice jacobienne en a est inversible, alors il existe un voisinage V de a et un voisinage W de $b = f(a)$ tels que $f_{|V}$ soit une bijection de V sur W de classe C^1 et dont la réciproque est aussi de classe C^1. Une telle application f s'appelle un *difféomorphisme de V sur W*.

Régularité d'ordres supérieurs. Lorsque la fonction $f : U \to \mathbb{R}^m$ est de classe C^1 et que les fonctions $x \mapsto (\partial f_k / \partial x_j)(x)$ sont elles-mêmes de classe C^1 (resp. de classe C^ℓ pour un entier $\ell \geq 1$), on dit que f est de classe C^2 (resp. de classe $\mathrm{C}^{\ell+1}$). On dit en outre que f est de classe C^∞ si elle est de classe C^ℓ pour tout $\ell \in \mathbb{N}$. Intuitivement, une fonction de classe C^ℓ est une fonction qui peut être bien approchée près de tout point $a \in U$ par des polynômes (de plusieurs variables) de degré ℓ modulo une erreur en $o(|x-a|^\ell)$: c'est la formule de *Taylor-Young*.

Pour vérifier qu'une fonction f est de classe C^ℓ, il faut donc établir la continuité de toutes ses dérivées partielles par rapport aux différentes variables jusqu'à l'ordre ℓ. Par exemple pour montrer que $f : U \to \mathbb{R}$ est de classe C^2, il faut regarder les n^2 dérivées partielles $(\partial^2 f / \partial x_j\, \partial x_k)$ pour $1 \leq j, k \leq n$. Mais une fois que l'on sait que f est de classe C^2, le théorème *de Schwarz* affirme que l'on a les identités $(\partial^2 f / \partial x_j\, \partial x_k) = (\partial^2 f / \partial x_k\, \partial x_j)$, ce qui réduit le nombre des dérivées partielles du deuxième ordre distinctes à $n(n+1)/2$.

9. Exercices

0.4. Les ouverts de $\mathbb{R}^n$ étant définis comme ci-dessus au paragraphe **5** de la page 90, montrer que la famille des ouverts de $\mathbb{R}^n$ définit une topologie, c'est-à-dire que $\mathbb{R}^n$ et $\emptyset$ sont des ouverts, qu'une intersection finie d'ouverts est un ouvert, et qu'une réunion quelconque d'ouverts est un ouvert.

0.5. Pour toute partie $A \subset \mathbb{R}^n$, on note $\mathring{A}$ la réunion de tous les ouverts contenus dans A, $\overline{A}$ l'intersection de tous les fermés contenant A, et enfin $\partial A = \overline{A} \setminus \mathring{A}$. L'ensemble $\mathring{A}$ s'appelle l'*intérieur* de A, $\overline{A}$ l'*adhérence* de A, et ∂A la *frontière* ou le *bord* de A.

(a) Montrer que $\mathring{A}$ est un ouvert contenu dans A, et que si U est un ouvert contenu dans A, alors $U \subset \mathring{A}$ (on dit alors que $\mathring{A}$ est *le plus grand* ouvert contenu dans A). De même, montrer que $\overline{A}$ est un fermé contenant A, et que si F est un fermé contenant A, alors $F \supset \overline{A}$ (on dit alors que $\overline{A}$ est *le plus petit* fermé contenant A). Montrer aussi que ∂A est un fermé.

(b) Montrer que si B est le complémentaire de A, alors $\mathring{B}$ est le complémentaire de $\overline{A}$ et $\overline{B}$ est le complémentaire de $\mathring{A}$.

Montrer que si $A \subset B$, alors $\mathring{A} \subset \mathring{B}$ et $\overline{A} \subset \overline{B}$.

Montrer que $\overline{A \cup B} = \overline{A} \cup \overline{B}$, mais que $\overline{A \cap B} \subset \overline{A} \cap \overline{B}$ sans qu'il y ait toujours égalité.

Quelles sont les formules correspondantes pour les intérieurs ?

0.6. Soit $f : I \to \mathbb{R}$ une fonction définie sur un intervalle ouvert I de $\mathbb{R}$. On suppose que f est dérivable sur I et que sa dérivée est bornée. Montrer que f est uniformément continue sur I.

0.7. Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction continue *périodique de période T*, c'est-à-dire vérifiant $f(x+T) = f(x)$ pour tout $x \in \mathbb{R}$.

(a) Montrer que $f(\mathbb{R})$ est compact.

(b) Montrer que $f_{|[-T,T]}$ est uniformément continue, puis en déduire que f est uniformément continue sur $\mathbb{R}$.

0.8. On appelle *courbe de Jordan* l'image de toute application continue $g : [a, b] \longrightarrow \mathbb{R}^2$ injective sur $[a, b[$ et vérifiant $g(a) = g(b)$.

(a) Montrer qu'une courbe de Jordan est un compact de $\mathbb{R}^2$.

(b) Soit $f : C \longrightarrow \mathbb{R}$ une fonction continue définie sur une courbe de Jordan C. Montrer que f atteint sa borne supérieure M et sa borne inférieure m, puis en déduire que si $m < M$, alors f prend au moins deux fois chaque valeur $y \in]m, M[$.

Quelques conseils de rédaction

Pour l'étudiant en mathématiques, la rédaction d'un devoir représente souvent une grande difficulté. Ce paragraphe a pour but d'énoncer quelques règles qui vous serviront à améliorer la qualité de vos rédactions.

Tout d'abord, il faut se convaincre de l'utilité d'apprendre à rédiger des mathématiques. Rédiger est un travail indispensable lorsque l'on veut communiquer quelque chose, et c'est vrai aussi en mathématiques à quelque niveau que ce soit. Ensuite, il faut bien comprendre que plus le niveau est élevé, moins il faut écrire de détails de peur de noyer l'essentiel : le but est avant tout de convaincre la personne qui vous lira que vous avez repéré les difficultés et que vous savez comment les traiter. C'est pourquoi le soin dans la présentation contribue de façon sensible à la qualité de la rédaction même s'il n'en constitue pas l'ingrédient principal.

Voici donc quelques conseils de rédaction groupés par thèmes.

10. Comment aborder un exercice ?

(a). Pour chercher un exercice, il faut commencer par **regarder la question posée**, c'est-à-dire **la conclusion**, et non pas raisonner au hasard à partir de l'hypothèse. Ainsi dans l'exercice **0.6** de la page 93, la question posée, c'est de montrer que la fonction f est uniformément continue, c'est-à-dire que $\forall \varepsilon > 0, \exists \delta > 0 : |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$. À nouveau, pour démontrer cette affirmation, on commencera par regarder la conclusion, c'est-à-dire par chercher à majorer $|f(x) - f(y)|$ en utilisant ce qui est connu : $|x - y| < \delta$ (pour un δ qui reste à construire en fonction de ε et f), et l'hypothèse de l'exercice (f dérivable et à dérivée bornée).

(b). Pour démontrer (à nouveau comme dans l'exercice **0.6** de la page 93) une affirmation du type $\forall \varepsilon > 0, \exists \delta > 0$: etc., il ne saurait être question de traîner le $\forall \varepsilon > 0$ tout au long du raisonnement, car le nombre δ va certainement dépendre du nombre ε et celui qui conviendra pour un ε ne conviendra pas pour un autre. Il faut donc **fixer le** $\varepsilon > 0$ en début de démonstration, construire le δ à partir de ce ε fixé, et, à la fin seulement, observer que la construction est valable **pour tout** $\varepsilon > 0$, ce qui prouve que l'on a bien démontré ce qu'il fallait.

(c). Il faut préférer le raisonnement **par implications** plutôt que par équivalences. En effet, rares sont les situations où l'on peut passer à une condition vraiment équivalente, et de toutes façons, il est beaucoup plus naturel d'utiliser les implications, c'est-à-dire de progresser des causes vers les conséquences. Si l'on doit montrer que deux propriétés $p(x)$ et $q(x)$ sont équivalentes, il sera préférable de découper le raisonnement en deux parties, l'une pour prouver que $p(x) \Rightarrow q(x)$ et l'autre pour prouver que $q(x) \Rightarrow p(x)$. Ainsi, dans l'exercice **2.1** de la page 17, on commencera par observer que s'il existe une bijection de $\mathbb{N}$ sur E, alors E est infini dénombrable puisqu'une bijection est surjective, et on attaquera ensuite en deuxième partie la démonstration de la réciproque : si E est dénombrable, alors il existe une bijection de $\mathbb{N}$ sur E.

11. Quels détails doit-on choisir d'inclure dans la rédaction ?

(a). **Chaque question posée mérite une réponse**. Si la réponse vous paraît évidente, justifiez-la quand même par une courte phrase plutôt que par une expression méprisante du style *Il est trivial que...* Ainsi, n'écrivez jamais *La fonction f est clairement intégrable* (l'intégrabilité d'une fonction n'est d'ailleurs jamais évidente), mais plutôt *La fonction f est intégrable car elle est continue à support compact* (par exemple). De même, n'écrivez pas *Le nombre a est trivialement positif* mais plutôt *Le nombre a est positif car il est égal à une somme de carrés.*

(b). Repérez ce qui, dans votre réponse, vous paraît **le plus important**, et insistez là-dessus. Un argument précis vaut mieux que trois arguments vagues. Ainsi, n'écrivez pas *La fonction f est intégrable car c'est une fonction paire uniformément continue et égale presque partout à un produit de deux fonctions intégrables*, car ces arguments, même ajoutés les uns aux autres, n'impliquent pas l'intégrabilité de f ; écrivez plutôt *La fonction f est intégrable car elle est continue et majorée en module par la fonction intégrable* $g(x) = e^{-|x|^2/2}$, ce qui constitue un argument convaincant et sans rien en trop.

C'est à vous de faire le choix de ce qui est important, mais il y a cependant des affirmations qu'il est obligatoire de justifier : par exemple, il n'est jamais évident qu'une fonction est intégrable ; par ailleurs, vous devez toujours signaler, conformément à la règle suivante, les théorèmes importants du cours que vous voulez utiliser.

(c). Lorsque vous utilisez **un théorème important du cours**, respectez toujours les trois règles suivantes : (a) Ne citez pas l'énoncé du théorème (tout le monde est censé le connaître) ; (b) Établissez soigneusement que les hypothèses en sont bien vérifiées dans le contexte où vous voulez l'utiliser (c'est ainsi que le correcteur verra que vous connaissez l'énoncé du théorème) ; (c) Concluez de préférence en citant le nom du théorème. Ainsi, pour calculer la limite de la suite numérique $(\int f_p(x)\,dx)$, on pourra calculer la limite simple (ou presque partout) de la suite de fonctions (f_p), puis construire un majorant F des $|f_p|$ et enfin montrer que F est une fonction *intégrable* et *indépendante de p*. Ces vérifications faites, on pourra alors conclure en nommant le théorème de convergence dominée.

12. Comment présenter la rédaction ?

(a). Essayez de **résumer en une phrase** qui explique tout ce que vous seriez tentés d'établir à l'aide de longs calculs fastidieux. Ainsi, si vous devez démontrer que l'application $L(f) = \int f(x)\,g(x)\,dx$ est linéaire (la fonction g étant donnée et fixée), écrivez *L'application L est linéaire par linéarité de l'intégrale* plutôt que de vous lancer dans le calcul de $L(c_1 f_1 + c_2 f_2)$ qui occupera une page, ou au moins une demi-page de votre copie.

(b). **Structurez vos démonstrations** en en distinguant les différentes étapes, en leur donnant des titres, en annonçant ce que vous allez démontrer. Par exemple dans une démonstration par récurrence, on doit toujours distinguer nettement les deux parties (initialisation, et propriété de transmission).

(c). **Renoncez à toute abréviation, et soignez l'orthographe et la grammaire**. S'il est vrai que dans un cours on utilise quelques abréviations pour gagner du temps, vous devez y renoncer dans vos copies pour assurer la clarté de votre discours ; en effet on lit parfois *Il suffit d'utiliser une ipp et CS pour mq f est int ssi...* là où on aimerait plutôt lire *Il suffit d'utiliser une intégration par parties et l'inégalité de Cauchy-Schwarz pour montrer que f est intégrable si et seulement si...*

Quant à l'orthographe et la grammaire, il suffit, pour se convaincre de leur importance, de comparer les deux versions suivantes de la même phrase : *On va montrer que...* et *On -a montré que...* (dans cette deuxième version, le *v* est illisible et il y a une faute d'orthographe sur le verbe montrer).

Bibliographie

Nous donnons ci-dessous les titres de quelques ouvrages pour compléter ce cours d'intégration. Certains d'entre eux traitent le sujet d'une façon proche de la nôtre, mais les autres adoptent un point de vue très différent, et complémentaire.

Deux ouvrages classiques mais dont la lecture difficile ne pourra profiter qu'aux très bons étudiants :

[1] Lebesgue H., *Leçons sur l'intégration et la recherche des fonctions primitives*, Gauthier-Villars, Paris 1928.

[2] Riesz F. & Nagy B.-Sz., *Leçons d'analyse fonctionnelle*, Akadémiai Kiado, Budapest 1952.

Trois ouvrages plus accessibles, le Dixmier étant très proche du cours présenté ici, le Rudin (chapitre 1) et le Gapaillard en étant plus éloignés (construction plus classique de l'intégrale directement dans le cadre de la théorie de la mesure) :

[3] Dixmier J., *L'intégrale de Lebesgue*, CDU – Les cours de Sorbonne, Paris 1966.

[4] Rudin W., *Real and complex analysis*, McGraw Hill, New-York 1966 (paru en édition française, *Analyse réelle et complexe*, chez Masson, Paris 1975).

[5] Gapaillard J., *Intégration pour la licence*, Masson, Paris 1997.

Un ouvrage pour ceux qui voudraient en savoir plus sur l'intégration dans $\mathbb{R}^n$, mais dont la lecture exige d'être très motivé :

[6] Evans L. & Gariepy R., *Measure theory and fine properties of functions*, CRC Press, Boca Raton (Florida) 1992.

Trois ouvrages d'exercices d'intégration qui ont le défaut de proposer des solutions (rappelons qu'il est absolument nécessaire de chercher un exercice avant d'en regarder la solution) :

[7] Verley J.-L., *Théorie élémentaire de l'intégration avec 60 exercices*, CDU – Les cours de Sorbonne, Paris 1972.

[8] Heurteaux Y., Hulin D., Picard F. & Queffélec H., *Exercices et problèmes corrigés de calcul intégral*, Public. Univ. Scientif., Orsay 1991.

[9] Benoist J. & Salinier A., *Exercices de Calcul intégral avec rappels de cours*, Masson, Paris 1997.

Index

Table des Matières

Ce polycopié peut être commandé à

l'IREM des Pays de Loire
Université de Nantes — Faculté des Sciences
2 rue de la Houssinière
BP 92208
44322 Nantes cedex 3

TITRE : Intégrales simples et multiples.

I.R.E.M. : des Pays de Loire (Nantes).

AUTEUR : Xavier Saint Raymond, professeur à l'Université de Nantes.

DATE : novembre 1998.

NIVEAU : de la licence à l'agrégation de mathématiques.

PUBLIC CONCERNÉ : Étudiants et enseignants
de deuxième cycle universitaire.

MOTS-CLÉS : Convolution,
Fourier (analyse de),
Intégrale,
Intégration.

RÉSUMÉ : Ce cours d'intégration propose une construction géométrique directe des intégrales multiples sur des domaines de $\mathbb{R}^n$ sans commencer par la construction de la mesure de Lebesgue.
On s'y efforce d'établir les principaux résultats de la théorie de Lebesgue dans ce cadre, et de faire le lien avec les méthodes usuelles de calcul des intégrales (intégration par parties, changement de variables).
On y présente en outre un certain nombre d'applications dans le domaine de l'analyse de Fourier (transformation de Fourier, convolution, séries de Fourier).

FORMAT	NOMBRE DE PAGES	PRIX	3^e TIRAGE
$21 \times 29{,}7$	104	5,50 €	100 ex.

N° I.S.B.N. : 2 – 86300 – 026 – 8