

INTRODUCTION

Ce cours est principalement consacré à deux grands outils de l'Analyse, l'analyse de Fourier et la théorie des distributions, ainsi qu'à diverses applications à des équations de la physique mathématique. Il s'agit de domaines ayant des racines très anciennes, mais qui font toujours l'objet de recherches actives, et qui ont connu dans une période récente des développements très importants dont nous essaierons de donner une idée.

L'analyse de Fourier. — Dite encore analyse harmonique, cette branche est dominée par les concepts de série et surtout de transformation de Fourier, ainsi que par l'opération de convolution qui leur est étroitement reliée.

On peut, sous des hypothèses très générales, représenter une fonction f définie dans \mathbb{R}^n sous la forme suivante

$$f(x) = (2\pi)^{-n} \int_{\mathbb{R}^n} e^{ix \cdot \xi} \widehat{f}(\xi) d\xi,$$

où la fonction \widehat{f} définie dans \mathbb{R}^n et qui se déduit de f par une formule analogue s'appelle la transformée de Fourier de f .

Il faut voir cette formule de la manière suivante. Elle permet d'écrire une fonction quelconque f comme une "superposition" de fonctions oscillantes simples : les applications $x \mapsto e^{ix \cdot \xi}$, chacune d'elles ayant une amplitude $|\widehat{f}(\xi)|$, et un déphasage $\arg \widehat{f}(\xi)$.

Lorsqu'on veut analyser qualitativement ou quantitativement une fonction, l'idée la plus simple est l'analyse "en amplitude" fondée sur les valeurs ponctuelles : en quels points x la fonction est-elle nulle, "petite", "grande"... ? L'analyse de Fourier permet d'y superposer une analyse "en fréquence" : quelles sont les fréquences ξ qui contribuent à l'écriture de f ci-dessus, y contribuent-elles peu ou beaucoup... ? Ce point de vue revient à faire l'analyse "en amplitude" de \widehat{f} .

Peut-être d'apparence moins naturelle, l'analyse en fréquence s'est révélée d'une importance aussi grande que l'analyse en amplitude (leurs rôles sont symétriques en mécanique quantique par exemple). Elle prend même un caractère prédominant dans l'étude d'un certain nombre de questions, notamment :

— Les phénomènes oscillants, cela va (presque) de soi.

— Les phénomènes régis par des équations aux dérivées partielles linéaires à coefficients constants. Pour une équation d'évolution par exemple, il existe des solutions particulièrement simples, de la forme $e^{ix \cdot \xi} e^{\alpha t}$, où le nombre complexe α se détermine à partir de ξ par un simple calcul algébrique. L'analyse de Fourier permet d'écrire toute évolution, aussi compliquée soit-elle, comme une superposition de ces évolutions simples.

— Les problèmes de régularité. Comme nous le verrons, un “principe” correspondant à de nombreux théorèmes assure qu'une fonction est d'autant plus régulière que sa transformée de Fourier est petite à l'infini. Cela permet d'utiliser des méthodes éprouvées (majorations, développements asymptotiques...) sur \hat{f} pour étudier la régularité de la fonction f .

La transformation de Fourier a néanmoins un inconvénient, son caractère global. Pour le traitement du signal, ou pour l'étude des équations aux dérivées partielles à coefficients variables, l'analyse en fréquence reste indispensable, mais il faut pouvoir la mener localement en les variables de temps ou d'espace. C'est l'objet de l'analyse microlocale, branche née vers 1970 dont nous ne donnerons qu'un très bref aperçu, dans la section 9.7.

La théorie des distributions. — C'est une extension de la notion de fonction, qui a joué un rôle très important dans le développement de l'Analyse. Bien que son introduction par L. Schwartz soit encore relativement récente, elle a permis de tels progrès en théorie des équations aux dérivées partielles et en analyse harmonique que l'on ne saurait plus parler de ces deux branches sans y avoir recours.

L'ensemble du cours, et notamment l'introduction du chapitre 4, montrera l'intérêt de cette généralisation de la notion de fonction. Cela dit, l'aspect déductif de l'exposition risque de donner une idée fautive du développement historique des mathématiques, qui est tout sauf déductif. L'introduction des distributions est aussi l'aboutissement d'un processus s'étalant sur plus d'un demi-siècle, en mathématiques et en physique⁽¹⁾.

⁽¹⁾Le lecteur intéressé trouvera au chapitre 6 de l'ouvrage de Laurent Schwartz “Un Mathématicien aux prises avec le siècle” (Odile Jacob, 1997) l'histoire de l'invention des distributions assortie d'une analyse du contexte scientifique.

Le calcul symbolique de Heaviside (1893), et surtout le formalisme introduit par P. Dirac (1926) pour les besoins de la mécanique quantique, la célèbre “fonction δ ” notamment, posaient un problème intéressant. La définition de ces concepts comme des fonctions au sens mathématique du terme était parfaitement contradictoire. Néanmoins, utilisés par Dirac lui-même ou d’autres bons physiciens, ils se révélaient efficaces et féconds. Des situations de ce type ne sont pas rares, il en existe actuellement, et elles signifient en général que des progrès mathématiques sont à l’ordre du jour.

En mathématiques, une multitude de concepts et de résultats, parfaitement rigoureux mais un peu épars, développés pendant la première partie de ce siècle n’ont trouvé leur unification que dans le cadre des distributions.

Ainsi, la notion de dérivée au sens des distributions permet de définir des solutions d’équations aux dérivées partielles qui ne sont pas suffisamment dérivables pour être des solutions au sens usuel du terme. En fait, pour presque chaque type d’équation aux dérivées partielles, on avait été amené à définir une ou des notions de solutions généralisées (ou solutions “faibles”). Ces définitions rentrent maintenant dans un cadre commun et beaucoup plus général

De même, comme nous le verrons, beaucoup d’opérateurs permettant de résoudre ces équations s’exprimeront pour nous en termes de convolution par une distribution, alors que l’introduction de ces opérateurs est souvent bien antérieure. Pour l’équation des ondes par exemple, Hadamard les avait exprimés par une formule ressemblant à celle d’une convolution, mais où des intégrales divergentes devaient être remplacées par leur “partie finie”.

Enfin, un pas très important avait été franchi par Sobolev en introduisant les espaces qui portent son nom, nous aurons l’occasion d’en voir l’intérêt, et en contribuant à clarifier le concept de solution faible.

La théorie des équations aux dérivées partielles. — C’est un thème qui ne sera qu’effleuré dans ce cours. Aucun développement systématique ne lui sera consacré, il est clair qu’un sujet d’une telle ampleur nécessite un enseignement spécifique.

Par contre, nous nous sommes efforcés de montrer l’efficacité et la puissance des outils théoriques introduits en les appliquant systématiquement à des équations de la physique mathématique, principalement aux équations de Laplace, de Schrödinger, et aux équations de la propagation des ondes et de la chaleur.

Les trois premiers chapitres doivent être considérés comme des préliminaires, présentant des outils d’utilisation constante dans la suite. Le premier est consacré à *l’intégrale de Lebesgue*. Un exposé complet de la théorie, avec

toutes les démonstrations, aurait été trop long pour le cadre horaire limité de ce cours. Par contre, nous avons pu en présenter les résultats les plus utiles, qui sont relativement peu nombreux, faciles à mémoriser, et considérablement plus efficaces que leurs homologues en théorie de l'intégrale de Riemann.

L'*analyse fonctionnelle* aura une place relativement réduite dans ce cours. Le chapitre 2 est destiné à consolider, et à développer sur quelques points, les connaissances du lecteur sur les espaces métriques, de Banach et de Hilbert. Cela nous suffira pour la suite, où nous n'introduirons pas, sauf pour les espaces normés, de topologie sur les espaces fonctionnels que nous aurons à utiliser.

En ce qui concerne l'espace des distributions, il s'agit d'un choix délibéré : la notion de suite convergente est tout à fait suffisante dans la pratique. En ce qui concerne les espaces de fonctions dérivables, il aurait par contre été raisonnable d'introduire leur structure d'espace métrisable. Si nous avons préféré, sans même parler de semi-norme, écrire explicitement chaque condition de continuité, c'est uniquement pour réaliser, au prix d'un petit nombre de répétitions, une économie de temps et de pensée.

Cela dit, pour ceux de nos lecteurs qui seraient suffisamment pourvus de ces deux denrées, nous avons consacré l'appendice C à l'étude des espaces de Fréchet. Nous y avons aussi donné les démonstrations d'un certain nombre de résultats admis dans le cours. Ils reposent sur le théorème de Banach-Steinhaus, qui repose lui-même sur la théorie des espaces de Baire développée dans l'appendice B.

Enfin nous aurons besoin de quelques compléments de *calcul différentiel*, principalement à plusieurs variables. Il s'agit d'un outil indispensable, les équations de la physique mathématique étant posées en dimension 3 ou 4. Le chapitre 3 rappelle les propriétés fondamentales des fonctions de classe C^k dans un ouvert de \mathbb{R}^n , cadre qui sera suffisant pour la suite, et y ajoute des résultats importants sur l'approximation par des fonctions différentiables.

L'appendice A développe un point de vue plus intrinsèque, et donne quelques compléments géométriques ainsi qu'une présentation complète de l'intégrale de surface. Il n'aborde toutefois que quelques aspects de la géométrie différentielle, et ne saurait se substituer à un enseignement consacré à ce sujet.

Le cours proprement dit se compose des chapitres 1 à 10, à l'exclusion des appendices et des textes en petits caractères. Ceux-ci sont destinés à apporter des compléments et selon les cas à satisfaire, ou à piquer, la curiosité du lecteur.

Les chapitres sont divisés en sections au sein desquelles tous les énoncés, des théorèmes aux exercices (qui ne sont pas moins importants) sont numérotés linéairement. Les formules sont numérotées au sein de chaque chapitre.

On trouvera enfin une petite bibliographie à la page 249.

*
* *

Cet ouvrage reprend, presque sans changement, le texte d'un cours trimestriel enseigné à l'École Polytechnique de 1986 à 1996. Il doit beaucoup aux cours d'Analyse de mes prédécesseurs : Laurent Schwartz, le regretté Charles Goulaouic et Yves Meyer. Le cadre trimestriel, certainement trop restreint, a eu sans doute un avantage : nous contraindre à aller droit à l'essentiel. Néanmoins, le sujet mériterait d'être traité à un rythme moins soutenu et d'être précédé d'un enseignement de calcul différentiel et intégral.

CHAPITRE 1

L'INTÉGRALE DE LEBESGUE

La théorie de l'intégrale de Riemann est à bien des égards insuffisante pour les besoins de l'Analyse, et notamment pour étudier les notions (analyse de Fourier, théorie des distributions) qui joueront un rôle central dans ce cours. Ce n'est que dans le cadre de la théorie de l'intégration introduite en 1900 par H. Lebesgue que l'on dispose des "bons" théorèmes de passage à la limite dans les intégrales, et que les espaces de fonctions sommables jouissent de bonnes propriétés.

Pour des raisons de temps, il ne serait pas possible de donner dans le cadre de ce cours un exposé complet, avec toutes les démonstrations, de la théorie de l'intégration. La solution adoptée consiste à admettre d'emblée un petit nombre de résultats, dont la démonstration est longue mais dont l'énoncé est simple, et d'en déduire les théorèmes fondamentaux de la théorie. Nous admettrons notamment dès le début l'existence et les principales propriétés de l'intégrale des fonctions positives — ce qui dans un exposé déductif nécessiterait toute la construction de la mesure de Lebesgue — ainsi que les théorèmes 1.4.1 et 1.5.2.

Pour les mêmes raisons, nous ferons délibérément "comme si" toutes les fonctions étaient mesurables ; nous nous expliquons sur ce point dans la section 1.6.

1.1. Intégrale des fonctions positives

Nous considérerons en fait des fonctions à valeurs dans $\overline{\mathbb{R}_+} = [0, \infty]$. Dans cet ensemble, auquel on étend de façon évidente l'addition, la multiplication par les éléments de $]0, \infty[$ et la relation d'ordre, toute suite croissante est convergente.

Nous admettrons qu'il existe une application

$$f \mapsto \int_{\mathbb{R}^n} f(x) dx$$

(que l'on notera parfois $\int f(x) dx$ voire $\int f$) qui, à toute fonction f , définie sur \mathbb{R}^n et à valeurs dans $\overline{\mathbb{R}_+}$ fait correspondre un élément de $\overline{\mathbb{R}_+}$, et qui vérifie les quatre propriétés suivantes.

(a) *Linéarité* : pour f et g comme ci-dessus, et pour $\lambda, \mu \in]0, \infty[$, on a $\int(\lambda f + \mu g) = \lambda \int f + \mu \int g$.

(b) *Croissance* : si $f(x) \leq g(x)$ pour tout x , on a $\int f \leq \int g$. Cette propriété est en fait conséquence de la précédente.

(c) *Normalisation* : Pour un pavé $P = [a_1, b_1] \times \dots \times [a_n, b_n]$, si on appelle $\mathbf{1}_P$ la fonction égale à 1 sur P et à 0 ailleurs, on a

$$\int \mathbf{1}_P(x) dx = \prod_{i=1}^n (b_i - a_i).$$

(d) *Le théorème fondamental suivant* :

Théorème 1.1.1 (de Beppo Levi ou de la convergence monotone)

Si f_j est une suite croissante de fonctions définies sur \mathbb{R}^n et à valeurs dans $\overline{\mathbb{R}_+}$, on a

$$\int \lim_{j \rightarrow \infty} f_j(x) dx = \lim_{j \rightarrow \infty} \int f_j(x) dx. \quad (1.1)$$

En particulier, pour des fonctions u_j positives, prenant éventuellement la valeur $+\infty$, on a toujours

$$\int \sum_{j=1}^{\infty} u_j(x) dx = \sum_{j=1}^{\infty} \int u_j(x) dx. \quad (1.2)$$

Le second point résulte facilement du premier en considérant la suite croissante constituée des sommes partielles $S_p(x) = \sum_{j=1}^p u_j(x)$.

1.1.2. Mesure des ensembles. — Si A est un sous ensemble de \mathbb{R}^n , on note $\mathbf{1}_A$ la fonction égale à 1 sur A et à 0 ailleurs, et on pose

$$\mu(A) = \int \mathbf{1}_A(x) dx.$$

L'application μ s'appelle la mesure de Lebesgue. C'est une application de l'ensemble des parties de \mathbb{R}^n dans $\overline{\mathbb{R}_+}$ qui vérifie les deux propriétés suivantes.

Croissance. — si $A \subset B$, on a $\mu(A) \leq \mu(B)$. Cela résulte de la croissance de l'intégrale (ou si l'on veut de $\mu(B) = \mu(A) + \mu(B \setminus A)$).

Additivité dénombrable. — si des sous-ensembles A_j , $j \in \mathbb{N}$ sont deux à deux disjoints, on a $\mu(\bigcup A_j) = \sum \mu(A_j)$. En effet, en posant $A = \bigcup_j A_j$, on a alors $\mathbf{1}_A = \sum_j \mathbf{1}_{A_j}$, et le résultat découle de (1.2).

Si on ne suppose pas les A_j disjoints deux à deux, on a seulement la sous-additivité : $\mu(\bigcup A_j) \leq \sum \mu(A_j)$. On a également la relation $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$.

Les ensembles de mesure nulle, on dit aussi *ensembles négligeables*, jouent un rôle important dans la théorie. D'après la sous-additivité, *la réunion d'une infinité dénombrable d'ensembles de mesure nulle est de mesure nulle*. Un point (et plus généralement un pavé plat) étant de mesure nulle, il en résulte qu'un ensemble dénombrable est de mesure nulle.

Exercice 1.1.3. — Démontrer que le graphe d'une application continue de \mathbb{R} dans \mathbb{R} est un sous-ensemble de mesure nulle dans \mathbb{R}^2 (quels que soient ε et N , on montrera que la portion du graphe comprise entre les abscisses $-N$ et N est contenue dans une réunion de rectangles dont la somme des mesures est $\leq \varepsilon$).

Exercice 1.1.4. — Soit A un ensemble de mesure nulle dans \mathbb{R}^n . Démontrer que son complémentaire est partout dense.

On dit qu'une propriété $P(x)$ dépendant d'un point x est vérifiée *presque partout* (en abrégé p.p.) si l'ensemble $\{x \mid \text{non } P(x)\}$ est de mesure nulle. Si chacune des propriétés $P_j(x)$, $j \in \mathbb{N}$, est vraie presque partout, il en est donc de même de $(\forall j, P_j(x))$.

Proposition 1.1.5. — Soit f définie dans \mathbb{R}^n et à valeurs dans $\overline{\mathbb{R}_+}$. On a

$$\int f(x) dx = 0 \iff f(x) = 0 \text{ p.p.}$$

Posons $A = \{x \mid f(x) \neq 0\}$. On a

$$f(x) \leq \lim_{j \rightarrow \infty} j \mathbf{1}_A(x).$$

Si $\mu(A) = 0$, on obtient donc en utilisant (1.1)

$$\int f(x) dx \leq \lim_{j \rightarrow \infty} j \int \mathbf{1}_A(x) dx = 0.$$

Réciproquement, si l'intégrale de f est nulle et en remarquant que $\mathbf{1}_A(x) \leq \lim_{j \rightarrow \infty} j f(x)$, on obtient

$$\mu(A) = \int \mathbf{1}_A(x) dx \leq \lim_{j \rightarrow \infty} j \int f(x) dx = 0.$$

Théorème 1.1.6 (Fatou). — Soit f_j une suite (non nécessairement monotone) de fonctions positives telle que pour chaque $x \in \mathbb{R}^n$, la suite $f_j(x)$ soit convergente, et que la suite $\int f_j(x) dx$ soit également convergente. On a

$$\int \lim_{j \rightarrow \infty} f_j(x) dx \leq \lim_{j \rightarrow \infty} \int f_j(x) dx.$$

Remarquons d'abord que, si une suite u_j d'éléments de $\overline{\mathbb{R}_+}$ converge vers $u \in \overline{\mathbb{R}_+}$, la suite $v_j = \inf_{k \geq j} u_k$ converge également vers u . En effet, dans le cas $u < \infty$, pour tout $\varepsilon > 0$ donné, tous les u_j appartiennent à $[u - \varepsilon, u + \varepsilon]$ à partir d'un certain rang, et tous les v_j appartiennent au même intervalle à partir du même rang. Le raisonnement est analogue si $u = +\infty$, avec des intervalles du type $[A, +\infty]$.

Soit g_j la fonction définie par $g_j(x) = \inf_{k \geq j} f_k(x)$. En posant $f(x) = \lim_{j \rightarrow \infty} f_j(x)$, on a d'après l'argument ci-dessus $f(x) = \lim_{j \rightarrow \infty} g_j(x)$ et de même $\lim_{j \rightarrow \infty} \int f_j(x) dx = \lim_{j \rightarrow \infty} (\inf_{k \geq j} \int f_k(x) dx)$.

D'autre part, l'intégrale de g_j est inférieure à l'intégrale de f_k pour chaque $k \geq j$, et on a donc $\int g_j(x) dx \leq \inf_{k \geq j} \int f_k(x) dx$. En appliquant le théorème de Beppo Levi à la suite croissante des g_j qui converge vers f , on obtient

$$\int f(x) dx = \lim_{j \rightarrow \infty} \int g_j(x) dx \leq \lim_{j \rightarrow \infty} \left(\inf_{k \geq j} \int f_k(x) dx \right) = \lim_{j \rightarrow \infty} \int f_j(x) dx,$$

ce qui est le résultat cherché.

Remarque 1.1.7. — Dans le cas particulier où la limite des intégrales est nulle, on peut en déduire que l'intégrale de la limite $\int \lim_{j \rightarrow \infty} f_j(x) dx$ est nulle.

En général, ce théorème ne fournit qu'une inégalité, et il est important de bien comprendre les deux raisons, illustrées par les deux exemples suivants, qui peuvent empêcher l'intégrale de passer à la limite. Pour en avoir une représentation plus imagée, on pourra penser que les f_j sont des densités de masses positives.

— La fonction f_j est égale à 1 sur l'intervalle $[j, j + 1]$ et à 0 dans le complémentaire. On a $f_j(x) \rightarrow 0$ en tout point x alors que $\lim \int f_j(x) dx = 1$. On voit que "la masse disparaît à l'infini".

— La fonction f_j est égale à j sur l'intervalle $]0, 1/j[$ et à 0 dans le complémentaire. On a encore $f_j(x) \rightarrow 0$ en tout point x et $\lim \int f_j(x) dx = 1$. On

peut dire que “la masse se concentre sur un ensemble de mesure nulle”, en l’occurrence l’ensemble réduit à l’origine.

Le lecteur pourra construire beaucoup d’exemples analogues. Nous verrons plus loin que le théorème de Lebesgue, dont l’hypothèse signifie grosso modo que l’on se prémunit contre ces deux types d’accident, permettra de passer à la limite.

Exercice 1.1.8. — Rappelons d’abord la définition de la *limite inférieure* d’une suite de nombres réels bornée inférieurement

$$\liminf_{j \rightarrow \infty} x_j = \lim_{j \rightarrow \infty} \left(\inf_{k \geq j} x_k \right) \in] - \infty, +\infty],$$

les bornes inférieures du membre de droite étant bien définies, et formant une suite qui croît avec j .

Démontrer la forme plus générale suivante du théorème de Fatou : si f_j est une suite quelconque de fonctions positives, on a

$$\int \liminf_{j \rightarrow \infty} f_j(x) dx \leq \liminf_{j \rightarrow \infty} \int f_j(x) dx.$$

1.2. Fonctions sommables

Définition 1.2.1. — Soit f une fonction définie dans \mathbb{R}^n à valeurs dans \mathbb{C} . On dit que f est sommable (ou intégrable au sens de Lebesgue) si on a

$$\int_{\mathbb{R}^n} |f(x)| dx < \infty.$$

L’espace des fonctions sommables est noté $\mathcal{L}^1(\mathbb{R}^n)$.

Dans la pratique, pour montrer qu’une fonction est sommable, il suffit de montrer qu’elle est majorée en module par une fonction positive d’intégrale finie.

Dans le cas où une fonction sommable f est à valeurs réelles, on peut l’écrire sous la forme $f = g - h$ où g et h sont positives et d’intégrales finies. On peut par exemple prendre $f = f_+ - f_-$, en posant $f_{\pm}(x) = \max\{\pm f(x), 0\}$, et les fonctions f_{\pm} sont majorées par la fonction $|f|$ qui est d’intégrale finie.

On définit alors l’intégrale de f par

$$\int_{\mathbb{R}^n} f(x) dx = \int_{\mathbb{R}^n} g(x) dx - \int_{\mathbb{R}^n} h(x) dx,$$

le résultat ne dépendant pas de la décomposition choisie. En posant en effet $r = g - f_+$, on a également $h - f_- = r$, la fonction r étant positive et d’intégrale

finie (elle est majorée par g). On a alors $\int g = \int f_+ + \int r$, $\int h = \int f_- + \int r$ et donc $\int g - \int h = \int f_+ - \int f_-$.

Si $f = f_1 + if_2$ est sommable et à valeurs complexes, les fonctions f_j sont alors sommables, et on pose $\int f = \int f_1 + i \int f_2$.

À partir des propriétés correspondantes pour les fonctions positives, on montre facilement les propriétés élémentaires suivantes.

Linéarité. — L'espace $\mathcal{L}^1(\mathbb{R}^n)$ est un espace vectoriel, et l'application $f \mapsto \int f$ est linéaire de $\mathcal{L}^1(\mathbb{R}^n)$ dans \mathbb{C} .

Croissance. — Pour f et g sommables à valeurs réelles vérifiant $f(x) \leq g(x)$ pour tout x , on a $\int f \leq \int g$.

Pour f et g sommables à valeurs complexes, on a

$$\left| \int f(x) dx \right| \leq \int |f(x)| dx, \quad (1.3)$$

et

$$\int |f(x) + g(x)| dx \leq \int |f(x)| dx + \int |g(x)| dx. \quad (1.4)$$

Démontrons par exemple l'additivité pour des fonctions réelles. Si $f = g - h$ et $f' = g' - h'$ sont des décompositions de f et f' en différences de fonctions positives d'intégrale finie, on a la décomposition $f + f' = (g + g') - (h + h')$ et donc $\int(f + f') = \int(g + g') - \int(h + h')$, ce qui entraîne le résultat voulu.

Montrons également (1.3), qui entraîne facilement (1.4). Soit θ un nombre complexe de module 1 tel que $\theta \int f$ soit réel positif. On a alors

$$\left| \int f(x) dx \right| = \int \operatorname{Re}(\theta f(x)) dx \leq \int (\operatorname{Re}(\theta f(x)))_+ dx \leq \int |f(x)| dx.$$

1.2.2. Intégration sur un sous-ensemble. — Considérons un sous-ensemble A de \mathbb{R}^n et f une fonction définie sur A . On notera f_A la fonction égale à $f(x)$ pour $x \in A$ et à 0 sinon.

On dit que la fonction f est sommable sur A si la fonction f_A est sommable dans \mathbb{R}^n . On note $\mathcal{L}^1(A)$ l'espace des fonctions sommables sur A , et on pose alors

$$\int_A f(x) dx = \int_{\mathbb{R}^n} f_A(x) dx.$$

Le passage de f à f_A permet d'associer à tout résultat sur l'espace $\mathcal{L}^1(\mathbb{R}^n)$ un résultat correspondant sur $\mathcal{L}^1(A)$, que nous laisserons au lecteur le soin d'énoncer.

Remarque 1.2.3. — Si deux fonctions f et g sont égales presque partout, et si l'une d'elles est sommable, l'autre l'est aussi, et elles ont la même intégrale. On a en effet $|\int f - \int g| \leq \int |f - g|$, quantité qui est nulle d'après la proposition 1.1.5. En particulier, l'intégrale d'une fonction sur un ensemble de mesure nulle est toujours nulle. De même, en dimension 1, il est équivalent d'intégrer une fonction f sur $[a, b]$ ou sur $]a, b[$, le résultat étant habituellement noté $\int_a^b f(x) dx$.

Intégrale d'une fonction définie presque partout. — Soit f une fonction définie seulement presque partout dans \mathbb{R}^n . Si l'un de ses prolongements \tilde{f} à l'espace entier est sommable, il résulte de la remarque précédente que tous les autres prolongements le sont aussi et qu'ils ont tous la même intégrale. Par abus de langage et de notation, on dit encore dans ce cas que la fonction f est sommable dans \mathbb{R}^n et on note $\int_{\mathbb{R}^n} f(x) dx$ l'intégrale de f .

Le théorème suivant est d'une importance capitale. La supériorité de l'intégrale de Lebesgue sur l'intégrale de Riemann est due en partie au fait que l'on peut intégrer plus de fonctions, mais elle est surtout due au fait que l'on dispose de théorèmes beaucoup plus efficaces. On comparera l'énoncé suivant, et le théorème de dérivation sous le signe somme qui en découle, aux résultats analogues fondés sur la convergence uniforme.

Théorème 1.2.4 (de Lebesgue ou de la convergence dominée)

Soit f_j une suite de fonctions qui converge presque partout vers une fonction f . On suppose qu'il existe une fonction positive sommable fixe h telle que l'on ait $|f_j(x)| \leq h(x)$ p.p. pour tout j . On a alors

$$\int |f(x) - f_j(x)| dx \rightarrow 0 \quad (1.5)$$

et

$$\int f_j(x) dx \rightarrow \int f(x) dx. \quad (1.6)$$

Considérons les ensembles $A_j = \{x \mid |f_j(x)| > h(x)\}$ et l'ensemble B constitué des points x où la suite $f_j(x)$ ne converge pas vers $f(x)$. Ces ensembles sont de mesure nulle par hypothèse, et il en est donc de même de $N = B \cup \left(\bigcup_{j=0}^{\infty} A_j\right)$.

En posant $\tilde{f}_j(x) = f_j(x)$ pour $x \notin N$ et $\tilde{f}_j(x) = 0$ pour $x \in N$, et en définissant de même \tilde{f} , il suffit de démontrer le théorème en remplaçant les fonctions f_j par les \tilde{f}_j (les intégrales sont inchangées), et nous nous sommes ainsi ramenés au cas où les majorations et la convergence ont lieu partout.

En posant $g_j(x) = \left| \tilde{f}_j(x) - \tilde{f}(x) \right|$, puis $h_j(x) = \sup_{k \geq j} g_k(x)$, on voit facilement que h_j est une suite décroissante de fonctions positives qui tend vers 0 en tout point, et que l'on a $h_j(x) \leq 2h(x)$ en tout point. Il nous reste à prouver que l'intégrale de h_j tend vers 0, ce qui entraîne immédiatement (1.5) et (1.6).

En appliquant le théorème de la convergence monotone à la suite croissante de fonctions positives $2h - h_j$, qui tend vers $2h$ en tout point, on obtient

$$\lim_{j \rightarrow \infty} \int (2h(x) - h_j(x)) dx = 2 \int h(x) dx.$$

Comme la fonction h est sommable, l'intégrale figurant au premier membre est la différence des intégrales de $2h$ et de h_j . Cela prouve que $\int h_j(x) dx \rightarrow 0$ et achève la démonstration.

Exercice 1.2.5. — Soit A_j une suite croissante de sous-ensembles de \mathbb{R}^n et posons $A = \bigcup_j A_j$. Démontrer que, pour toute fonction $f \in \mathcal{L}^1(A)$, on a $\int_A f(x) dx = \lim_{j \rightarrow \infty} \int_{A_j} f(x) dx$. Démontrer qu'une fonction f définie sur A appartient à $\mathcal{L}^1(A)$ si et seulement si $\sup_j \int_{A_j} |f(x)| dx < \infty$.

Théorème 1.2.6 (de dérivation sous le signe somme)

Soient I un intervalle de \mathbb{R} et A un sous-ensemble de \mathbb{R}^n . On se donne une fonction f définie sur $A \times I$ vérifiant les trois hypothèses suivantes.

- (a) Pour tout $\lambda \in I$, la fonction $x \mapsto f(x, \lambda)$ est sommable sur A
- (b) La dérivée partielle $\partial f / \partial \lambda(x, \lambda)$ existe en tout point de $A \times I$
- (c) Il existe une fonction h positive et sommable sur A telle que l'on ait $|\partial f / \partial \lambda(x, \lambda)| \leq h(x)$ quels que soient x et λ .

Alors la fonction F définie par

$$F(\lambda) = \int_A f(x, \lambda) dx \tag{1.7}$$

est dérivable dans I , et on a

$$F'(\lambda) = \int_A \frac{\partial f}{\partial \lambda}(x, \lambda) dx. \tag{1.8}$$

1.2.7. Mode d'emploi. — Ce résultat ne prend toute sa force que si on l'accompagne des deux remarques suivantes.

(a) — Avant d'appliquer le théorème, on peut retirer du domaine d'intégration un ensemble de mesure nulle, ce qui ne change pas les intégrales dans (1.7) et (1.8), et donc ne vérifier les hypothèses que dans l'ensemble A' ainsi obtenu.

Par contre, il ne suffirait pas que, pour chaque λ , les hypothèses soient satisfaites sauf sur un sous-ensemble de A , fût-il réduit à un point, qui dépend de λ (voir l'exercice 1.2.10).

(b) — La dérivabilité est une propriété locale. Pour prouver que F est dérivable dans I , il suffit de montrer que F est dérivable dans tout intervalle compact $[c, d] \subset I$. Il suffira donc de trouver des fonctions positives sommables h_{cd} qui majorent $\partial f / \partial \lambda$ en module lorsque λ parcourt $[c, d]$.

Remarque 1.2.8. — Il existe un théorème de *continuité sous le signe somme* pour $F(\lambda) = \int_A f(x, \lambda) dx$, lorsque λ parcourt un ouvert de \mathbb{R}^p ou même un espace métrique quelconque, et le lecteur pourra l'énoncer s'il le désire. Il faut prouver que pour toute suite λ_j tendant vers un point λ_0 , on a $F(\lambda_j) \rightarrow F(\lambda_0)$, et le théorème de Lebesgue fournit immédiatement les conditions voulues.

Lorsque λ parcourt un ouvert $\omega \subset \mathbb{R}^p$, pour prouver que F est de classe C^1 on procède généralement en deux temps. On prouve d'abord l'existence des $\partial F / \partial \lambda_i$ en gelant les autres variables λ_j et en appliquant le théorème 1.2.6 à la fonction de x et λ_i ainsi obtenue. Les $\partial F / \partial \lambda_i$ étant exprimées par une intégrale, on prouve ensuite leur continuité dans ω à l'aide de l'argument ci-dessus.

1.2.9. *Démonstration du théorème 1.2.6.* — On a, λ étant fixé,

$$\frac{1}{l} (F(\lambda + l) - F(\lambda)) = \int_A g_l(x) dx,$$

où on a posé

$$g_l(x) = \frac{1}{l} (f(x, \lambda + l) - f(x, \lambda)).$$

La fonction g_l converge en tout point vers $\partial f / \partial \lambda$. D'autre part, on a $|g_l(x)| \leq \sup_{0 \leq \theta \leq 1} |\partial f / \partial \lambda(x, \lambda + \theta l)| \leq h(x)$ d'après le théorème des accroissements finis, et le résultat est conséquence immédiate du théorème de Lebesgue.

Exercice 1.2.10. — Soit φ une fonction continue sur $[0, 1]$. On considère, dans $[0, 1] \times [0, 1]$ la fonction f définie par $f(x, \lambda) = \varphi(x)$ si $x \leq \lambda$ et $f(x, \lambda) = 0$ sinon. On pose $F(\lambda) = \int_0^1 f(x, \lambda) dx$. Pour chaque λ la dérivée partielle $\partial f / \partial \lambda$ existe sauf en un point, et elle est majorée par une fonction sommable fixe : la fonction 0. Déterminer la dérivée de F .

Exercice 1.2.11. — (a) Soit f une fonction sommable sur $]0, \infty[$. Pour $\lambda \in [0, \infty[$, on définit sa transformée de Laplace F par $F(\lambda) = \int_0^\infty e^{-\lambda t} f(t) dt$. Démontrer que F est continue sur $[0, \infty[$.

(b) On suppose maintenant que f , toujours définie sur $]0, \infty[$, est telle que $(1+t)^{-N} f(t)$ est sommable pour un N convenable. Démontrer que $F(\lambda)$ est bien définie par la formule ci-dessus pour $\lambda \in]0, \infty[$, et que la fonction F

est de classe C^∞ sur cet intervalle. Déterminer la transformée de Laplace de $t \mapsto t^k f(t)$.

(c) Sous les hypothèses de (b), pour $z = x + iy$ complexe vérifiant $x > 0$, on pose $F(z) = \int_0^\infty e^{-zt} f(t) dt$. Démontrer que F est une fonction de classe C^1 (et même C^∞) des variables x, y dans le demi-plan $x > 0$. Démontrer que l'on a $\partial F/\partial x + i\partial F/\partial y = 0$. En déduire que, pour tout z dans ce demi-plan, il existe $A_z \in \mathbb{C}$ tel que l'on ait, pour h complexe, $F(z+h) = F(z) + A_z h + hr(h)$ avec $\lim_{h \rightarrow 0} r(h) = 0$ (de telles fonctions sont dites holomorphes).

Exercice 1.2.12. — (a) Pour $x > 0$, on pose $\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt$ (fonction Γ d'Euler). Démontrer que la fonction Γ est de classe C^∞ sur $]0, \infty[$.

(b) Démontrer que $\Gamma(x) \rightarrow +\infty$ pour $x \rightarrow 0$. (Pourquoi utiliser des ε ? La convergence monotone sur $]0, 1]$ est tellement plus simple.)

(c) Pour $\operatorname{Re} z > 0$, on pose $\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt$. Démontrer que Γ est holomorphe (voir l'exercice précédent) dans le demi-plan $\operatorname{Re} z > 0$.

Exercice 1.2.13. — Calculer la dérivée à droite en 0 de la fonction

$$\lambda \mapsto \int_0^1 (g(x) + \lambda^2)^{1/2} dx$$

où g est une fonction bornée, positive ou nulle sur $]0, 1[$.

1.3. Cas de la dimension 1

Théorème 1.3.1. — Soit f une fonction continue sur un intervalle $[a, b]$ et soit F une primitive de f . On a alors

$$\int_{[a,b]} f(x) dx = F(b) - F(a).$$

Considérons la fonction $G(x) = \int_{[a,x]} f(t) dt$. On a, pour $l > 0$ (le raisonnement serait le même pour $l < 0$)

$$\begin{aligned} \frac{1}{l} (G(x+l) - G(x)) &= \frac{1}{l} \int_{[x,x+l]} f(t) dt \\ &= \frac{1}{l} \int_{[x,x+l]} f(x) dt + \frac{1}{l} \int_{[x,x+l]} (f(t) - f(x)) dt. \end{aligned} \quad (1.9)$$

D'après le propriété (c) de la section 1.1, l'intégrale d'une constante M sur un intervalle $[a, b]$ est égale à $M(b-a)$. Le membre de droite de (1.9) est la somme de $f(x)$ et d'une quantité majorée en module par $\max_{[x,x+l]} |f(t) - f(x)|$ qui tend vers 0 avec l . Il en résulte que G est dérivable, et que $G'(x) = f(x)$. La fonction $F - G$ est donc constante, et on a $G(b) - G(a) = F(b) - F(a)$ ce qui est le résultat voulu.

On voit donc que l'intégrale de Lebesgue est un prolongement de l'intégrale de Riemann des fonctions continues. Cela dit, l'étude préalable de l'intégrale de Riemann n'est nullement nécessaire. Pour le calcul des intégrales de Lebesgue, on peut à la fois utiliser le calcul des primitives dans les cas simples et disposer de théorèmes puissants sur les passages à la limite.

Exercice 1.3.2. — Soit f une fonction intégrable au sens de Riemann sur $[a, b]$. Démontrer que f est sommable et que les deux définitions de l'intégrale de f coïncident. On pourra, pour chaque entier m , poser $h = (b - a)/2^m$ et introduire les fonctions φ_m [resp. ψ_m] constantes sur chaque intervalle $[kh, (k+1)h[$ et y prenant comme valeur le sup [resp. inf] de f sur le même intervalle. On montrera ensuite que les suites (φ_m) et (ψ_m) sont monotones et, en utilisant la proposition 1.1.5, que l'on a $\inf \varphi_m = \sup \psi_m$ p.p.

Bien entendu, on ne se privera pas de la notation classique $\int_a^b f(x) dx$ pour désigner, selon les cas, l'intégrale sur $]a, b[$ ou l'opposé de l'intégrale sur $]b, a[$.

1.3.3. Qu'est-il advenu des intégrales absolument convergentes ... — À la différence de celle de Riemann, l'intégrale de Lebesgue est directement définie de manière globale sur \mathbb{R} . Cela rend beaucoup plus nette la distinction entre deux concepts de nature très différente mais qui sont souvent présentés ensemble (sous le nom plus que discutable d'“intégrales impropres”) en théorie de l'intégrale de Riemann.

Quoi qu'il en soit, il est bien sûr intéressant d'avoir des relations entre les intégrales sur \mathbb{R} entier, et celles sur les intervalles finis. Pour $a < b$, on notera f_{ab} la fonction égale à f pour $x \in [a, b]$ et à 0 ailleurs.

Si f est une fonction positive, et si a_n et b_n sont des suites respectivement décroissant vers $-\infty$ et croissant vers $+\infty$, la suite des fonctions $f_{a_n b_n}$ converge en croissant vers la fonction f , et on a convergence des intégrales d'après le théorème de Beppo Levi. On a donc

$$\forall f \geq 0, \quad \int_{\mathbb{R}} f(x) dx = \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow +\infty}} \int_a^b f(x) dx \leq +\infty$$

formule dans laquelle on peut remplacer \lim par \sup .

En particulier, en appliquant ce qui précède à la fonction $|f|$, on obtient le résultat suivant : *pour qu'une fonction f à valeurs complexes soit sommable, il faut et il suffit que $\sup \int_a^b |f(x)| dx < +\infty$.*

Si la fonction f est sommable, la suite des fonctions $f_{a_n b_n}$ définies comme ci-dessus converge vers f en chaque point, et est majorée en module par la

fonction sommable fixe $|f(\cdot)|$. On a donc, en vertu du théorème de Lebesgue,

$$\forall f \in \mathcal{L}^1(\mathbb{R}), \quad \int_{\mathbb{R}} f(x) dx = \lim_{\substack{a \rightarrow -\infty \\ b \rightarrow +\infty}} \int_a^b f(x) dx.$$

On voit que le concept d'intégrale d'une fonction sommable généralise celui d'"intégrale de Riemann absolument convergente".

Bien entendu ce qui précède s'applique également aux rapports entre $\int_a^b f(x) dx$ et les $\int_{a_n}^{b_n} f(x) dx$ lorsque a_n et b_n tendent respectivement vers a et b .

1.3.4. ... et des intégrales semi-convergentes? — Il arrive qu'une fonction f , sommable sur tout compact, ne soit pas sommable sur \mathbb{R} mais que les intégrales $\int_a^b f(x) dx$ tendent vers une limite pour $a \rightarrow -\infty$ et $b \rightarrow +\infty$. Un exemple classique est la fonction $\sin x/x$. Il est traditionnel de noter encore $\int_{-\infty}^{+\infty} f(x) dx$ la limite et de parler d'intégrale semi-convergente, mais *de telles expressions ne sont pas vraiment des intégrales — aucun théorème de la théorie de l'intégration ne s'applique à elles — ce sont des limites d'intégrales et il n'y a guère plus à en dire*. Pour les étudier, on applique la théorie de l'intégration aux vraies intégrales $\int_a^b f(x) dx$, et on passe à la limite avec les moyens du bord.

1.4. Intégrales multiples

Théorème 1.4.1 (Fubini). — Soit $f(x, y)$ une fonction définie dans le produit $\mathbb{R}^p \times \mathbb{R}^q$.

(a) Si f est à valeurs dans $\overline{\mathbb{R}_+}$, on a l'égalité suivante, où les trois membres définissent un élément de $\overline{\mathbb{R}_+}$

$$\iint_{\mathbb{R}^{p+q}} f(x, y) dx dy = \int_{\mathbb{R}^p} \left\{ \int_{\mathbb{R}^q} f(x, y) dy \right\} dx = \int_{\mathbb{R}^q} \left\{ \int_{\mathbb{R}^p} f(x, y) dx \right\} dy \quad (1.10)$$

(b) Si f est sommable dans \mathbb{R}^{p+q} , les trois membres de (1.10) ont un sens dans \mathbb{C} et sont égaux. Pour être précis, dire que le troisième membre a un sens signifie les deux choses suivantes :

- pour presque chaque y , la fonction $x \mapsto f(x, y)$ est sommable dans \mathbb{R}^p ,
- la fonction $\varphi(y) = \int f(x, y) dx$ qui est ainsi définie presque partout est sommable dans \mathbb{R}^q .

1.4.2. Mode d'emploi. — On désire souvent intervertir les signes d'intégration, et prouver l'égalité des second et troisième membres de (1.10) pour une fonction f à valeurs complexes. On procède en deux temps.

(a) — On prouve d'abord que f est sommable dans \mathbb{R}^{p+q} . Pour cela, on utilise la forme (a) du théorème de Fubini pour calculer ou majorer l'intégrale $\iint |f(x, y)| dx dy$ par intégration successive dans l'ordre que l'on désire.

(b) — Une fois établi le fait que f est sommable "du couple", on utilise à nouveau le théorème de Fubini sous la forme (b).

On prendra garde au fait suivant : il peut très bien arriver (pour une fonction non sommable dans \mathbb{R}^{p+q}) que les second et troisième membres de (1.10) aient un sens dans \mathbb{C} , et qu'ils soient différents.

Remarque 1.4.3. — Pour calculer l'intégrale double d'une fonction f sur un sous-ensemble A de \mathbb{R}^{p+q} , on se ramène au cas de l'intégration de la fonction f_A . Le troisième membre, par exemple, de (1.10) devient

$$\int_{\pi(A)} \left\{ \int_{A_y} f(x, y) dx \right\} dy$$

où A_y est la "tranche" $\{x \in \mathbb{R}^p \mid (x, y) \in A\}$, et où $\pi(A)$ est la projection de A sur \mathbb{R}^q .

Exercice 1.4.4. — Soit A le graphe d'une application de \mathbb{R}^{n-1} dans \mathbb{R} . Démontrer que A est de mesure nulle dans \mathbb{R}^n .

Théorème 1.4.5 (de changement de variable). — Soient Ω_1 et Ω_2 deux ouverts de \mathbb{R}^n , et Φ un difféomorphisme de Ω_1 sur Ω_2 . On notera $J_\Phi(x)$ le déterminant de la matrice jacobienne de Φ au point x

$$J_\Phi = \begin{vmatrix} \partial\Phi_1/\partial x_1 & \cdots & \partial\Phi_1/\partial x_n \\ \cdots & \cdots & \cdots \\ \partial\Phi_n/\partial x_1 & \cdots & \partial\Phi_n/\partial x_n \end{vmatrix}.$$

Soit f une fonction définie sur Ω_2 .

(a) Si la fonction f est à valeurs dans $\overline{\mathbb{R}_+}$, on a l'égalité suivante, où les deux membres ont un sens dans $\overline{\mathbb{R}_+}$

$$\int_{\Omega_2} f(y) dy = \int_{\Omega_1} f(\Phi(x)) |J_\Phi(x)| dx. \quad (1.11)$$

(b) Si f est à valeurs complexes, elle est sommable dans Ω_2 si et seulement si la fonction $f(\Phi(x)) |J_\Phi(x)|$ est sommable dans Ω_1 et les deux membres de (1.11) sont alors égaux.

Remarque 1.4.6. — Il est souvent utile, pour se trouver dans les conditions d'application du théorème, de retirer des domaines d'intégration des ensembles de mesure nulle, ce qui ne change aucune des intégrales. Par exemple, les coordonnées polaires ne fournissent pas de difféomorphisme du plan sur un ouvert. Par contre, si on retire du plan l'origine et le demi-axe des x négatif, on obtient un difféomorphisme de $]0, \infty[\times]-\pi, \pi[$ sur l'ouvert ainsi obtenu, et donc la formule classique (le jacobien étant égal à r)

$$\iint_{\mathbb{R}^2} f(x, y) dx dy = \iint_{]0, \infty[\times]-\pi, \pi[} f(r \cos \theta, r \sin \theta) r dr d\theta,$$

valable pour f positive ou sommable.

On utilisera parfois la formule analogue dans \mathbb{R}^n . Ce n'est pas exactement une application directe du théorème, mais on s'y ramène facilement en prenant des équations locales de portions de la sphère unité \mathbb{S}^{n-1} de \mathbb{R}^n . On notera θ le point courant de celle-ci, et $d\sigma_\theta$ la mesure de surface (voir la section A.3).

$$\int_{\mathbb{R}^n} f(x) dx = \int_0^\infty \int_{\mathbb{S}^{n-1}} f(r\theta) r^{n-1} dr d\sigma_\theta.$$

1.5. Espaces \mathcal{L}^1 , \mathcal{L}^2 , \mathcal{L}^∞

1.5.1. L'espace $\mathcal{L}^1(\mathbb{R}^n)$. — Nous l'avons défini comme l'espace des fonctions sommables. Si on pose $\|f\| = \int |f(x)| dx$, on voit facilement que $\|\lambda f\| = |\lambda| \|f\|$, et (1.4) exprime que $\|f + g\| \leq \|f\| + \|g\|$. Ce n'est néanmoins pas une norme sur $\mathcal{L}^1(\mathbb{R}^n)$, car la relation $\|f\| = 0$ équivaut d'après la proposition 1.1.5 à $f = 0$ p.p. et non pas à $f = 0$. Pour la même raison, l'application $(f, g) \mapsto \int |f(x) - g(x)| dx$ vérifie l'inégalité du triangle mais n'est pas une distance.

Cela n'empêche pas de définir la *convergence en moyenne*. On dit qu'une suite f_j d'éléments de $\mathcal{L}^1(\mathbb{R}^n)$ converge en moyenne vers f si $\int |f_j(x) - f(x)| dx$ tend vers 0 pour $j \rightarrow \infty$. On n'a pas unicité de la limite f , mais on a le résultat suivant : pour que f_j converge également en moyenne vers une fonction \tilde{f} , il faut et il suffit que $f = \tilde{f}$ p.p.. Il est clair que la condition est suffisante, les intégrales de $|f_j(x) - f(x)|$ et de $|f_j(x) - \tilde{f}(x)|$ étant les mêmes. Réciproquement, si f_j converge en moyenne vers f et \tilde{f} , on a $\int |f - \tilde{f}| \leq \int |f - f_j| + \int |\tilde{f} - f_j|$ pour tout j . Le premier membre est donc nul, ce qui entraîne $f = \tilde{f}$ p.p. d'après la proposition 1.1.5.

Tous ces résultats suggèrent de passer au quotient par la relation d'équivalence $f = g$ p.p., ce que nous ferons dans la section 2.6. Le résultat suivant, que nous admettrons, est important.

Théorème 1.5.2. — Soit f un élément de $\mathcal{L}^1(\mathbb{R}^n)$. Il existe une suite f_j de fonctions continues à support compact (c'est-à-dire nulles en dehors d'un ensemble borné) qui converge vers f en moyenne.

Exercice 1.5.3. — Soit $f \in \mathcal{L}^1(\mathbb{R}^n)$. On définit la transformée de Fourier \widehat{f} de f par $\widehat{f}(\xi) = \int e^{-ix \cdot \xi} f(x) dx$ pour tout $\xi \in \mathbb{R}^n$.

- (a) Montrer que la fonction \widehat{f} est continue et bornée et que l'on a $\sup_{\xi \in \mathbb{R}^n} |\widehat{f}(\xi)| \leq \int |f(x)| dx$.
- (b) Démontrer que, si φ est continue à support compact, la fonction $\widehat{\varphi}$ tend vers 0 à l'infini (on montrera, en posant $\theta = \pi/\xi$, que $\widehat{f}(\xi) = - \int e^{-ix \cdot \xi} f(x+\theta) dx$).
- (c) En déduire que, pour $f \in \mathcal{L}^1(\mathbb{R}^n)$, la fonction \widehat{f} tend vers 0 à l'infini.

1.5.4. L'espace $\mathcal{L}^2(\mathbb{R}^n)$. — C'est l'espace des fonctions de carré sommable, c'est-à-dire telles que $\int |f(x)|^2 dx < \infty$. À partir des relations $|f(x) + g(x)|^2 \leq 2(|f(x)|^2 + |g(x)|^2)$ et $|f(x)g(x)| \leq 1/2(|f(x)|^2 + |g(x)|^2)$, on montre facilement que $\mathcal{L}^2(\mathbb{R}^n)$ est un espace vectoriel et que, pour deux éléments de cet espace, la quantité

$$(f | g) = \int_{\mathbb{R}^n} f(x) \overline{g(x)} dx \quad (1.12)$$

définit un élément de \mathbb{C} , la fonction à intégrer étant sommable.

Il est facile de voir que cette expression possède toutes les propriétés d'un produit scalaire, à l'exception de la suivante : la nullité de $(f | f)$ équivaut à $f = 0$ p.p. et non à $f = 0$. Cela nous suggère encore de passer au quotient par la relation d'équivalence $f = g$ p.p. Nous le ferons dans la section 2.6 où l'on trouvera d'autres propriétés des fonctions de carré sommable.

La quantité $(f - g | f - g)^{1/2} = \left(\int |f(x) - g(x)|^2 dx \right)^{1/2}$ est appelée *écart quadratique moyen* de f et g et on dit que f_j converge vers f en moyenne quadratique si l'écart quadratique moyen de f et f_j tend vers 0.

1.5.5. L'espace $\mathcal{L}^\infty(\mathbb{R}^n)$. — Il s'agit de l'espace des fonctions qui sont presque partout égales à une fonction bornée.

Théorème et Définition 1.5.6. — Soit f une fonction à valeurs réelles définie sur \mathbb{R}^n . On dit que $M \in \mathbb{R}$ est un presque majorant de f si on a $f(x) \leq M$ p.p. Si f est presque majorée, l'ensemble de ses presque majorants possède un plus petit élément, que l'on appelle la borne supérieure essentielle de f et que l'on note $\sup \text{ess } f(x)$.

Soit μ la borne inférieure de l'ensemble des presque majorants de f . Il suffit de montrer que μ est encore un presque majorant de f . Pour chaque entier

j , le nombre $\mu + 1/j$ est un presque majorant et il existe donc un ensemble négligeable A_j en dehors duquel on a $f(x) \leq \mu + 1/j$. L'ensemble $A = \bigcup A_j$ est de mesure nulle, et on a $f(x) \leq \mu$ dans le complémentaire de A ce qui achève la démonstration.

On appelle $\mathcal{L}^\infty(\mathbb{R}^n)$ l'espace des fonctions essentiellement bornées, c'est-à-dire telles que la fonction $|f(x)|$ possède un presque majorant. On vérifie facilement que l'application $f \mapsto \text{supess}|f(x)|$ possède toutes les propriétés d'une norme, à l'exception du fait que la nullité de $\text{supess}|f(x)|$ implique seulement $f = 0$ p.p.

1.6. Sur la construction de l'intégrale

Le but de cette section est de donner un aperçu très superficiel de la construction de l'intégrale de Lebesgue et de dire quelques mots des questions de mesurabilité. Le lecteur intéressé par la théorie pourra se reporter par exemple à l'ouvrage de W. Rudin cité dans la bibliographie.

On définit d'abord la mesure des ensembles pavables, c'est à dire des ensembles qui sont réunion finie de pavés d'intérieur disjoint, comme étant égale à la somme des volumes de ces pavés — après avoir vérifié que le résultat ne dépend pas de la décomposition choisie.

Un premier passage à la limite définit la mesure d'un ouvert Ω de \mathbb{R}^n : c'est la borne supérieure des mesures des ensembles pavables contenus dans Ω . Un second passage à la limite définit la mesure d'un compact : c'est la borne inférieure des mesures des ouverts qui le contiennent. On dit alors qu'un ensemble A est *mesurable (pour la mesure de Lebesgue)* si, pour tout $\varepsilon > 0$, on peut trouver une suite de compacts K_j et une suite d'ouverts Ω_j tels que l'on ait $K_j \subset \Omega_j$,

$$\bigcup_{j=0}^{\infty} K_j \subset A \subset \bigcup_{j=0}^{\infty} \Omega_j \quad \text{et} \quad \sum_0^{\infty} \mu(\Omega_j \setminus K_j) \leq \varepsilon.$$

On définit alors la mesure de A par

$$\mu(A) = \inf_{\substack{\Omega \supset A \\ \Omega \text{ ouvert}}} \mu(\Omega) = \sup_{\substack{K \subset A \\ K \text{ compact}}} \mu(K) \leq +\infty.$$

Les résultats les plus importants que l'on obtient au terme de cette construction sont les suivants.

— Les ensembles mesurables forment une *tribu* c'est-à-dire une famille de parties de \mathbb{R}^n qui est stable par passage au complémentaire et par réunion

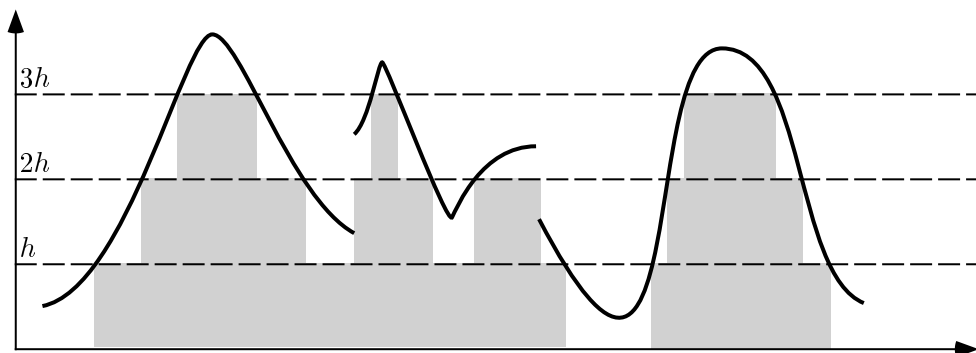
dénombrable (et donc par intersection dénombrable). Cette tribu contient notamment les ouverts et les fermés.

— La mesure μ est une application croissante et dénombrablement additive (voir le n° 1.1.2) de la tribu des ensembles mesurables dans $\overline{\mathbb{R}}_+$.

On dit ensuite qu'une application f de \mathbb{R}^n dans $\overline{\mathbb{R}}_+$ est mesurable si l'image réciproque de tout intervalle est un ensemble mesurable, et on définit enfin l'intégrale de f par la formule

$$\int f(x) dx = \sup_{h>0} \sum_{j=0}^{\infty} \mu(f^{-1}([jh, (j+1)h[)) jh \quad (1.13)$$

(sauf dans le cas où f prend la valeur $+\infty$ sur un ensemble de mesure > 0 , auquel cas on pose bien entendu $\int f = +\infty$). On peut d'ailleurs remplacer $\sup_{h>0}$ par $\lim_{h \rightarrow 0}$ dans la formule ci-dessus. On interprète souvent cette



La fonction f_h et son intégrale

formule en disant que l'intégrale de Lebesgue découpe f selon l'axe des y , alors que l'intégrale de Riemann la découpe selon l'axe des x . L'expression (1.13) exprime effectivement $\int f(x) dx$ comme limite d'intégrales $\int f_h(x) dx$, où la fonction f_h est égale à kh lorsque $kh \leq f(x) < (k+1)h$.

On voit les avantages de cette méthode : les fonctions f_h approchent f à h près en norme uniforme, même pour des fonctions f très irrégulières, alors que l'approximation uniforme par des fonctions constantes sur des intervalles ne fonctionne que pour des fonctions très particulières (continues ou plus généralement réglées). En outre, l'approximation est "presque" croissante : on voit facilement que l'on a $f_{h/2} \geq f_h$, et la suite des $f_{2^{-n}}$ par exemple est croissante. Cela rend plus naturel le théorème de Beppo Levi, dont la démonstration n'est effectivement plus très difficile à ce stade.

Le prix à payer est bien sûr le fait que, pour définir l'intégrale de f_h , il faut disposer de la mesure des ensembles $f^{-1}([kh, (k+1)h[)$, ensembles qui peuvent être très compliqués, ce qui nécessite toute la construction précédente. Au contraire, l'intégrale de Riemann n'utilisait que le concept élémentaire de mesure d'un intervalle.

1.6.1. Existe-t-il des ensembles et des fonctions non mesurables?

L'expérience suggère la réponse *non*. En effet, les ensembles mesurables forment une tribu contenant les ouverts et il en résulte que l'espace des fonctions mesurables contient les fonctions continues et est stable par toutes les opérations dénombrables usuelles : limite d'une suite (ou somme d'une série) de fonctions qui converge en chaque point, sup ou inf dénombrable, ... À titre d'exemple, le lecteur pourra voir dans l'exercice B.2.4 que la fonction égale à 1 en tout point rationnel et à 0 en tout point irrationnel — le type même de la fonction non intégrable au sens de Riemann, alors que c'est une excellente fonction sommable d'intégrale nulle — est limite d'une suite de fonctions dont chacune est limite d'une suite de fonctions continues. On peut bien sûr faire beaucoup plus compliqué, mais on n'arrive jamais à construire une fonction non mesurable sans faire appel à l'axiome du choix.

La véritable réponse à la question posée est : *cela dépend des axiomes mis à la base des mathématiques*. On a en effet les deux résultats suivants.

— Si on adjoint l'axiome du choix aux axiomes usuels de la théorie des ensembles, on peut prouver effectivement qu'il existe des ensembles non mesurables (voir l'exercice 1.6.2).

— Par contre, un résultat relativement récent de logique mathématique (Solovay, 1966) assure que l'on peut adjoindre à ces mêmes axiomes, sans introduire de contradiction, les formes dénombrables de l'axiome du choix et l'axiome "tout sous-ensemble de \mathbb{R}^n est mesurable".

Dans la pratique cela signifie que, à moins de le faire exprès à l'aide de l'axiome du choix, il est exclu que l'on ait à considérer des fonctions non mesurables. C'est pourquoi ce cours a été écrit comme si toutes les fonctions étaient mesurables. La véritable raison est bien sûr une question de temps, il y a mieux à faire que de démontrer, par des méthodes répétitives, des résultats dont on sait d'avance qu'ils sont toujours vrais. Le lecteur n'aura qu'à ajouter mentalement l'adjectif "mesurable" chaque fois qu'il rencontrera le mot "ensemble" ou "fonction".

Cela dit, le lecteur excessivement scrupuleux qui serait choqué par cette façon de faire pourra se placer dans le système d'axiomes autorisé par Solovay. C'est un cadre dans lequel on peut développer toute l'analyse classique, et où tous les énoncés de ce chapitre sont effectivement des théorèmes.

Ce qui précède s'applique à la mesure de Lebesgue, et il ne faudrait pas en conclure que toutes les questions de mesurabilité sont sans intérêt. En théorie des probabilités, on introduit fréquemment plusieurs tribus (dépendant par exemple du temps), la mesurabilité d'une variable aléatoire X par rapport à telle ou telle tribu ayant un contenu probabiliste précis. Dans un tel contexte, la démonstration de la mesurabilité d'une variable aléatoire peut être un résultat important, et éventuellement difficile.

Exercice 1.6.2. — Soit Q_1 l'ensemble des nombres rationnels contenus dans $[-1, 1]$, dont on numérotera les éléments : $Q_1 = \{r_1, r_2, \dots\}$. On considère, dans l'ensemble $[0, 1]$, la relation d'équivalence $(x - y) \in Q_1$. On forme un ensemble A en choisissant un point et un seul dans chaque classe d'équivalence (on notera que l'ensemble des classes d'équivalences n'est pas dénombrable). Démontrer que A n'est pas mesurable. (On introduira les ensembles translatés $A_n = A + r_n$, on montrera que ces ensembles sont disjoints et que l'on a $[0, 1] \subset \bigcup_n A_n \subset [-1, 2]$. On montrera ensuite, en supposant A mesurable, que les hypothèses $\mu(A) = 0$ et $\mu(A) > 0$ conduisent toutes deux à une contradiction.)

1.7. Les quatre opérations

On dit souvent que l'intégration est l'opération inverse de la dérivation et il semble, nous y reviendrons, qu'il y ait quelque apparence de vérité dans cette assertion en dimension 1. Qu'en est-il en dimension supérieure?

Dans \mathbb{R}^n , il faut remplacer la dérivation par la différentiation, opération qui à une fonction f (de classe C^1 pour fixer les idées) associe sa différentielle $df = \sum(\partial f / \partial x_i) dx_i$. On connaît bien l'opération inverse, qui est la résolution des équations dites "aux différentielles totales" : les fonctions α_i étant données (disons de classe C^1 dans un ouvert Ω), déterminer f telle que $df = \sum \alpha_i dx_i$. Le lecteur n'ignore pas que les conditions $\partial \alpha_i / \partial x_j = \partial \alpha_j / \partial x_i$ sont nécessaires pour avoir l'existence d'une solution, et qu'elles sont suffisantes dans un "bon" ouvert (convexe, étoilé ou plus généralement "simplement connexe" conviennent).

Quoi qu'il en soit, l'opération inverse de la différentiation n'a rien à voir avec l'intégration dans \mathbb{R}^n , et on peut se demander quelle est l'opération inverse de celle-ci.

Avant de poursuivre, nous allons examiner un type de raisonnement courant en Physique. Par exemple, dans un ouvrage d'électrostatique où on demande de calculer le potentiel U créé à l'origine par une répartition de charges de densité $\rho(\mathbf{r})$, ($\mathbf{r} = (x, y, z)$), on pourra trouver le raisonnement suivant.

Considérons autour du point \mathbf{r} un parallélépipède infinitésimal de côtés dx, dy, dz . Il porte une charge $\rho(\mathbf{r}) dx dy dz$ qui crée à l'origine un potentiel

$$dU = \frac{1}{4\pi\epsilon_0 r} \rho(\mathbf{r}) dx dy dz, \quad r = |\mathbf{r}|. \quad (1.14)$$

On a donc

$$U = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\mathbf{r})}{r} dx dy dz. \quad (1.15)$$

Ce texte pose quelques questions : d'où vient le *donc* ci-dessus ; s'il y a une différentielle dU existerait-il une fonction U , et si oui une fonction de quoi ?

Pour répondre sur le plan mathématique aux interrogations qui précèdent, il faut en fait introduire deux concepts et énoncer deux théorèmes.

On appelle *mesure* (nous continuons à faire comme si tous les ensembles étaient mesurables) une application $A \mapsto m(A)$ qui à tout sous-ensemble A de \mathbb{R}^n associe un élément de $\overline{\mathbb{R}}_+$, qui vérifie $m(\emptyset) = 0$ et la propriété d'additivité dénombrable : si des A_j , $j \in \mathbb{N}$, sont deux à deux disjoints, on a $m(\cup A_j) = \sum m(A_j)$.

On appelle, lorsqu'elle existe, *densité au point x de la mesure m par rapport à la mesure de Lebesgue μ* , la quantité $f(x)$ définie comme suit

$$f(x) = \frac{dm}{d\mu}(x) = \lim_{r \rightarrow 0} \frac{m(B(x, r))}{\mu(B(x, r))},$$

où $B(x, r)$ désigne la boule (que l'on pourrait remplacer par un cube) de centre x et de rayon r .

Cette quatrième opération, qui à une mesure associe sa densité, va être l'opération inverse de l'intégration. C'est ce que montrent les deux résultats suivants, que nous ne ferons qu'énoncer.

Théorème 1.7.1 (de dérivation de Lebesgue). — *Soit f une fonction positive et localement sommable dans \mathbb{R}^n . Alors la fonction d'ensemble m définie par $m(A) = \int_A f(x) dx$ est une mesure qui possède presque partout une densité et on a $\frac{dm}{d\mu}(x) = f(x)$ p.p.*

Théorème 1.7.2 (Radon-Nikodym). — *Soit m une mesure telle que l'on ait $m(K) < \infty$ pour tout compact K et $m(N) = 0$ pour tout ensemble N de mesure de Lebesgue nulle. Il existe alors une fonction f , sommable sur tout compact, telle que l'on ait $m(A) = \int_A f(x) dx$.*

Les mathématiques sous-jacentes au raisonnement électrostatique ci-dessus sont alors claires. On considère la fonction d'ensemble U qui à chaque $A \subset \mathbb{R}^n$ associe le potentiel $U(A)$ créé à l'origine par les charges contenues dans A . Il est raisonnable de penser que c'est une mesure, l'additivité (finie) étant explicitement énoncée dans les ouvrages d'électrostatique sous le nom de principe de superposition. Il reste à en calculer la densité $dU/d\mu$, ce qu'évoque assez bien l'argument infinitésimal et la formule (1.14), et à appliquer les deux théorèmes ci-dessus pour calculer $U = U(\mathbb{R}^3)$, ce qui est résumé dans le *donc*.

Ce type de raisonnement est très courant en physique et chaque fois que, pour une quantité Q , on écrit que dQ est proportionnel à l'élément de volume, on fait appel implicitement à la notion de mesure. Il s'agit d'un concept dont l'importance est comparable à celle du concept de fonction, et l'absence de référence explicite tient sans doute au fait qu'il n'est apparu historiquement qu'au début du XX^e siècle.

L'hypothèse du théorème de Radon-Nikodym ($\mu(N) = 0 \Rightarrow m(N) = 0$) est importante. Elle n'est pas respectée dans le cas de charges portées par des surfaces, des courbes ou des points, et ces cas sont étudiés séparément dans les ouvrages d'électrostatique (la formule (1.15) donnerait toujours $U = 0$).

Les quatre opérations existent en dimension 1, mais on ne les perçoit pas toujours comme différentes. Étant donné une mesure m (de masse finie pour simplifier) sur \mathbb{R} , on peut lui associer la fonction croissante F définie par $F(x) = m(]-\infty, x])$. Réciproquement, à une telle fonction F on associe facilement la fonction additive d'intervalles définie par $m(]a, b]) = F(b) - F(a)$ (on peut en fait la prolonger en une mesure) et il est clair que, en un

point où F est dérivable, la mesure m admet une densité égale à $F'(x)$. L'opération inverse de la dérivation est le calcul des primitives, tandis que le calcul des "intégrales définies" est plutôt l'inverse du calcul de la densité d'une mesure.