Chapitre 1

Rappels d'intégration

Le but de ces notes est de rediscuter rapidement le contenu du cours d'intégration de Lebesgue. Ce résumé rapide est bien évidemment très incomplet et ne saurait se substituer à un cours plus sérieux. Il s'agit plutôt d'un guide de relecture du cours de licence, dans le but de se rafraichir la mémoire avant d'aller plus loin.

Quelques exemples de bonnes lectures sur le sujet :

— Th. Gallouët, R. Herbin, *Mesure, intégration, probabilité.*

— M. Briane, G. Pagès, *Théorie de l'intégration, convolution et transformée de Fourier.*

— W. Rudin, *Analyse réelle et complexe.*

1.1 Intégrales d'une fonction continue par morceaux sur un segment

On commence par rappeler une définition et les principales propriétés de l'intégrale de Riemann pour une fonction d'une variable réelle. On se contentera d'énoncer les résultats pour des fonctions continues ou continues par morceaux. Considérer en toute généralité les fonctions intégrables au sens de Riemann est un raffinement intéressant mais pas vraiment indispensable, surtout quand on dispose par ailleurs de l'intégrale de Lebesgue.

Néanmoins, quitte à faire ces rappels, on en profite pour observer que tous les résultats (à l'exception de la proposition 1.7 à propos de la relation d'ordre) sont valables pour des fonctions de $\mathbb{R}$ dans un espace de Banach quelconque, ce qui ne sera pas le cas pour l'intégrale de Lebesgue telle qu'on la construit ici.

Dans tout ce paragraphe on considère donc un espace de Banach E sur $\mathbb{R}$ ou sur $\mathbb{C}$ (on notera $\mathbb{K}$ le corps en question). Pour conserver des notations proches de celles du cas usuel ($E = \mathbb{R}$), on notera simplement $|\cdot|$ la norme de E.

On commence par rappeler la définition et les propriétés de base des fonctions continues par morceaux.

Définition 1.1. Soient $[a, b]$ un segment de $\mathbb{R}$ et f une fonction de $[a, b]$ dans E. On dit que f est continue par morceaux s'il existe $n \in \mathbb{N}^*$ et $x_1, \dots, x_n \in [a, b]$ tels que

$$a = x_0 < x_1 < \cdots < x_n = b,$$

et pour tout $j \in [\![1, n]\!]$ la restriction de f à $]x_{j-1}, x_j[$ est continue et admet des limites finies en x_{j-1} et x_j.

L'ensemble des fonctions continues par morceaux est stable par somme, produit et multiplication par un scalaire. D'autre part, une fonction continue par morceaux sur un segment est bornée.

On rappelle par ailleurs qu'une fonction continue sur un segment est uniformément continue (théorème de Heine).

Quelle que soit la définition choisie, l'idée de l'intégrale de Riemann est d'approcher une fonction par des fonctions en escalier (constantes par morceaux) de plus en plus fines.

Définition 1.2. Soit $[a, b]$ un segment de $\mathbb{R}$. On appelle subdivision pointée σ de $[a, b]$ la donnée
— d'un entier $n \in \mathbb{N}^*$,
— de $x_0, \ldots, x_n \in [a, b]$ tels que $a = x_0 < x_1 < \cdots < x_n = b$,
— et de $\xi_1 \in [x_0, x_1]$, $\xi_2 \in [x_1, x_2]$, ..., $\xi_N \in [x_{n-1}, x_N]$.
En outre on dit que le pas de cette subdivision σ est inférieur ou égal à $\delta > 0$ si

$$\forall j \in [\![1, n]\!], \quad x_j - x_{j-1} \leqslant \delta.$$

Étant donnée une fonction f de $[a, b]$ dans E on note alors

$$S_\sigma(f) = \sum_{j=1}^{n} f(\xi_j)(x_j - x_{j-1}).$$

Cela revient à approcher f par la fonction qui vaut $f(\xi_j)$ sur $]x_{j-1}, x_j[$ pour tout $j \in [\![1, n]\!]$.

Proposition 1.3. *Soit f une fonction continue par morceaux de $[a, b]$ dans E. Alors il existe un unique élément $I(f)$ de E vérifiant la propriété suivante. Pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que pour toute subdivision pointée σ de $[a, b]$ de pas inférieur ou égal à δ on a*

$$|S_\sigma(f) - I(f)| \leqslant \varepsilon.$$

On note alors

$$I(f) = \int_a^b f(x)\, dx.$$

Typiquement, un choix naturel pour avoir une subdivision de plus en plus fine du segment $[a, b]$ est de choisir une subdivision uniforme. Étant donné $N \in \mathbb{N}^*$ on pose, pour tout $j \in [\![0, n]\!]$

$$x_j = a + \frac{j}{n}(b - a).$$

Et des choix naturels pour choisir ξ_j dans $[x_{j-1}, x_j]$ sont de prendre $\xi_j = x_{j-1}$ ou $\xi_j = x_j$. On obtient alors, pour tout fonction f continue par morceaux sur $[a, b]$:

$$\frac{b-a}{n} \sum_{j=0}^{n-1} f\left(a + \frac{j}{n}(b - a)\right) \xrightarrow[n \to +\infty]{} \int_a^b f(x)\, dx$$

ou

$$\frac{b-a}{n}\sum_{j=1}^{n} f\left(a+\frac{j}{n}(b-a)\right) \xrightarrow[n\to+\infty]{} \int_a^b f(x)\,dx.$$

Si $a, b \in \mathbb{R}$ sont tels que $a < b$ alors on pose

$$\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx.$$

Proposition 1.4 (Relation de Chasles)**.** *Soient $[a,b]$ un segment de $\mathbb{R}$ et f une fonction continue par morceaux de $[a,b]$ dans E. Soit $c \in [a,b]$. Alors on a*

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

Proposition 1.5 (Linéarité de l'intégrale)**.** *Soient $[a,b]$ un segment de $\mathbb{R}$, f et g deux fonctions continues par morceaux de $[a,b]$ dans E et $\lambda \in \mathbb{K}$. Alors on a*

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

et

$$\int_a^b (\lambda f)(x)\,dx = \lambda\int_a^b f(x)\,dx.$$

Proposition 1.6 (Inégalité triangulaire)**.** *Soient $[a,b]$ un segment de $\mathbb{R}$ et f une fonction continue par morceaux de $[a,b]$ dans E. Alors on a*

$$\left|\int_a^b f(x)\,dx\right| \leqslant \int_a^b |f(x)|\,dx.$$

Proposition 1.7 (Positivité de l'intégrale)**.** *Soient $[a,b]$ un segment de $\mathbb{R}$ et f et g deux fonctions continues par morceaux de $[a,b]$ dans $\mathbb{R}$. Si*

$$\forall x \in [a,b], \quad f(x) \leqslant g(x),$$

alors

$$\int_a^b f(x)\,dx \leqslant \int_a^b g(x)\,dx.$$

Proposition 1.8. *Soient I un intervalle de $\mathbb{R}$, $a \in I$ et f une fonction continue par morceaux de $[a,b]$ dans E. Pour $x \in I$ on pose*

$$F(x) = \int_a^x f(t)\,dt. \tag{1.1}$$

Alors F est continue sur I.

Théorème 1.9 (Primitive d'une fonction continue)**.** *Soient I un intervalle de $\mathbb{R}$, $a \in I$ et f une fonction continue de I dans E. Alors la fonction F définie par (1.1) est de classe C^1 sur I et $F' = f$.*

Théorème 1.10 (Théorème fondamental de l'analyse)**.** *Soient I un intervalle de $\mathbb{R}$, $a \in I$ et f une fonction continue de I dans E. Alors f admet une infinité de primitives sur I, et si F est l'une d'entre elles alors pour tous $a, b \in I$ on a*

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

Proposition 1.11 (Intégration par parties)**.** *Soient $[a, b]$ un segment de $\mathbb{R}$ et u et v deux fonctions de classe C^1 de I dans E. Alors on a*

$$\int_a^b u'(x)v(x)\,dx = u(b)v(b) - u(a)v(a) - \int_a^b u(x)v'(x)\,dx.$$

Théorème 1.12 (Changement de variables)**.** *Soient $[a, b]$ un segment de $\mathbb{R}$ et φ une fonction de classe C^1 de $[a, b]$ dans un intervalle I de $\mathbb{R}$. Soit f une fonction continue par morceaux de I dans $\mathbb{R}$. Alors on a*

$$\int_a^b f(\varphi(t))\ \varphi'(t)\,dt = \int_{\varphi(a)}^{\varphi(b)} f(x)\,dx.$$

1.2 Intégrales généralisées

On considère maintenant des fonctions définies sur des intérvalles qui ne sont pas nécessairement des segments. On rappelle qu'une fonction continue par morceaux sur un intervalle I de $\mathbb{R}$ est une fonction dont la restriction à n'importe quel segment de I est continue par morceaux.

Définition 1.13. — Soient $a \in \mathbb{R}$ et $b \in]a, +\infty[\cup \{+\infty\}$. Soit f une fonction continue par morceaux de $[a, b[$ dans E. On dit que l'intégrale généralisée $\int_a^b f(t)\,dt$ est convergente si la fonction

$$x \mapsto \int_a^x f(t)\,dt$$

(bien définie sur $[a, b[$) admet une limite quand x tend vers b. Dans ce cas on note

$$\int_a^b f(t)\,dt = \lim_{x \to b} \int_a^x f(t)\,dt.$$

— On a une définition analogue pour une fonction sur $]a, b]$ avec $b \in \mathbb{R}$ et $a \in]-\infty, b[\cup \{-\infty\}$.

— Soient $a \in \mathbb{R} \cup \{-\infty\}$ et $b \in]a, +\infty[\cup \{+\infty\}$. Soit $c \in]a, b[$ Soit f une fonction continue par morceaux de $]a, b[$ dans E. On dit que l'intégrale généralisée $\int_a^b f(t)\,dt$ est convergente si les intégrales $\int_a^c f(t)\,dt$ et $\int_c^b f(t)\,dt$ le sont. Dans ce cas on pose

$$\int_a^b f(t)\,dt = \int_a^c f(t)\,dt + \int_c^b f(t)\,dt$$

La convergence de l'intégrale et, le cas échéant, sa valeur, ne dépendent pas du choix de c.

Proposition 1.14 (Intégrales de Riemann). — *L'intégrale généralisée* $\int_1^{+\infty} \frac{1}{t^\alpha}\,dt$ *est convergente si et seulement si*

$$\alpha > 1,$$

et dans ce cas on a

$$\int_1^{+\infty} \frac{1}{t^\alpha}\,dt = \frac{1}{\alpha - 1}.$$

— *L'intégrale généralisée* $\int_0^1 \frac{1}{t^\alpha}\,dt$ *est convergente si et seulement si*

$$\alpha < 1,$$

et dans ce cas on a

$$\int_0^1 \frac{1}{t^\alpha}\,dt = \frac{1}{1 - \alpha}.$$

Proposition 1.15 (Comparaison d'intégrales généralisées). *Soient* $a \in \mathbb{R}$ *et* $b \in]a, +\infty[\cup \{+\infty\}$. *Soient* f *une fonction continue par morceaux de* $[a, b[$ *dans* E *et* g *une fonction continue par morceaux de* $[a, b[$ *dans* $\mathbb{R}_+$. *On suppose que pour tout* $x \in [a, b[$ *on a*

$$|f(x)| \leqslant g(x).$$

Alors si l'intégrale $\int_a^b g(t)\,dt$ *est convergente, l'intégrale* $\int_a^b f(t)\,dt$ *l'est également.*

On dit que l'intégrale $\int_a^b f(t)\,dt$ est absolument convergente si l'intégrale $\int_a^b |f(t)|\,dt$ est convergente. La proposition précédente implique en particulier qu'une intégrale absolument convergente est en particulier convergente. La réciproque est fausse, par exemple l'intégrale

$$\int_1^{+\infty} \frac{\cos(x)}{x}\,dx$$

est convergente sans être absolument convergente.

1.3 Mesure de Lebesgue

Le point de départ de la théorie de l'intégration est le calcul d'aires de parties du plan $\mathbb{R}^2$ ou de volumes de parties de l'espace $\mathbb{R}^3$. Cette notion d'aire est intuitive tant que l'on se contente d'ensembles simples. Le but de la mesure de Lebesgue est de définir ce qu'on appelle volume d'une partie de $\mathbb{R}^d$ (ou longueur en dimension 1 et aire en dimension 2) de façon générale. Ce sera utile en particulier pour définir l'intégrale de Lebesgue.

Le cahier des charges minimal pour une notion raisonnable de mesure est le suivant :

(i) La mesure d'une partie de $\mathbb{R}^d$ est un réel positif ou $+\infty$.

(ii) La mesure coïncide avec la définition intuitive sur les ensembles de base (on se basera sur les pavés de $\mathbb{R}^d$). En dimension 2, cela signifie que la mesure d'un rectangle de la forme $[a, b] \times [c, d]$ (avec $a < b$ et $c < d$) doit être $(b - a)(d - c)$.

(iii) La mesure est invariante par translation.

(iv) La mesure d'une union disjointe de parties de $\mathbb{R}^d$ doit être la somme des mesures des parties considérées.

Cette dernière condition cache une subtilité importante et doit être précisée. On ne peut pas demander que la mesure d'une union quelconque de parties de $\mathbb{R}^d$ soit la somme des mesures des parties en question. Déjà, il faudrait être clair sur ce qu'est une somme quelconque. Mais surtout, un singleton est une partie de mesure nulle et toute partie de $\mathbb{R}^d$ est union de singletons deux à deux disjoints, donc toute partie de $\mathbb{R}^d$ serait de mesure nulle.

On peut restreindre la condition aux unions finies. Cela fonctionne, mais cela ne donne pas une définition suffisament souple (en gros, cela définit la mesure de Jordan, qui correspond en fait à l'intégrale de Riemann).

Le bon compromis, qui typiquement donnera de bonnes propriétés vis-à-vis des passages à la limite, est d'imposer la troisième condition pour les unions *dénombrables* de parties de $\mathbb{R}^d$ deux à deux disjointes. La sommes des mesures de ces parties n'est alors rien d'autre qu'une série à termes positifs.

Malheureusement, il n'existe pas d'application sur l'ensemble des parties de $\mathbb{R}^d$ qui vérifie ces conditions. La seule concession que l'on peut encore faire est donc de renoncer à définir la mesure de toute partie de $\mathbb{R}^d$.

Une façon usuelle de construire la mesure de Lebesgue (disons sur $\mathbb{R}^2$ qui est le cas le plus simple à visualiser) est d'approcher un ensemble par une l'union d'une suite de petits rectangles et d'approcher la mesure de l'ensemble par la somme des aires des rectangles. Plus précisément on définit

$$\lambda^*(A) = \inf \sum_{n\in\mathbb{N}} \mathsf{Aire}(R_n),$$

où $(R_n)_{n\in\mathbb{N}}$ est une suite de rectangles recouvrant A, et l'infimum est pris sur l'ensemble de ces suites.

Malheureusement l'application λ^* ainsi définie ne convient pas, on peut avoir $A\cap B = \emptyset$ et $\lambda^*(A) + \lambda^*(B) > \lambda^*(A \sqcup B)$. On restreint alors l'application λ^* à certaines parties de $\mathbb{R}^d$ seulement, qui constituent ce qu'on appelle la tribu de Lebesgue.

On retient donc qu'on n'est pas capable d'étendre de façon raisonnable la notion d'aire ou de volume à toutes les parties de $\mathbb{R}^d$. Mais l'ensemble des parties mesurables est en fait tellement large qu'en pratique on ne rencontre jamais de partie de $\mathbb{R}^d$ qui ne l'est pas.

1.4 Mesures

La surface ou le volume d'un ensemble n'est pas forcément une donnée pertinante pour mesurer sa « taille ». Par exemple, si on veut comparer la taille des pays du monde, on peut s'intéresser à leur surface (on n'est plus sur un plan, mais la notion de surface s'étend aux parties de surfaces plus générales) mais aussi au nombre d'habitants, au PIB, ou encore au nombre de victoires en coupe du monde. Certaines sont plus pertinantes que d'autres, cela dépend du contexte. Cependant, beaucoup de propriétés sont communes à toutes les façons de mesurer, ce qui explique l'introduction d'un cadre abstrait.

Définition 1.16. Soit X un ensemble. On appelle tribu sur X une famille $\mathcal{M}$ de parties de X vérifiant les propriétés suivantes.

(i) $X \in \mathcal{M}$.

(ii) Si $A \in \mathcal{M}$ alors $X \setminus A \in \mathcal{M}$.

(iii) Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{M}$ alors

$$\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{M}.$$

Les éléments de $\mathcal{M}$ sont appelés parties mesurables de X. On dit alors que $(X, \mathcal{M})$ est un espace mesurable.

Définition 1.17. Soit $(X, \mathcal{M})$ un espace mesurable. On appelle mesure (positive) sur $(X, \mathcal{M})$ une application $\mu : \mathcal{M} \to [0, +\infty] = [0, +\infty[\cup\{+\infty\}$ vérifiant les propriétés suivantes.

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

(ii) Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{M}$ deux à deux disjoints alors on a

$$\mu\left(\bigsqcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n).$$

On dit alors que $(X, \mathcal{M}, \mu)$ est un espace mesuré.

Les propriétés imposées pour la définition d'une mesure sont assez naturelles. Souvent, la tribu $\mathcal{M}$ est en fait l'ensemble $\mathcal{P}(X)$ des parties de X, mais il est parfois utile de pouvoir se restreindre à un sous-ensemble de $\mathcal{P}(X)$. Typiquement, pour la mesure de Lebesgue. On regroupe alors dans la notion de tribu les propriétés naturelles d'une collection de parties de X sur lesquelles on veut définir une mesure. Si on sait donner une mesure à une partie, on sait donner une mesure à son complémentaire, etc.

On a déjà évoqué la mesure de Lebesgue. Deux autres exemples importants de mesures, définies sur toutes les parties d'un ensemble X quelconque, sont les suivants.

Exemple 1.18. Pour $A \in \mathcal{P}(X)$ on pose

$$\mu(A) = \begin{cases} \mathsf{Card}(A) & \text{si } A \text{ est fini,} \\ +\infty & \text{sinon.} \end{cases}$$

Cela définit une mesure sur X, appelée *mesure de comptage*.

Exemple 1.19. On suppose que $X \neq \emptyset$ et on considère $x \in X$. On définit alors la mesure δ_x, appelée *mesure de Dirac* en x, par

$$\forall A \in \mathcal{M}, \quad \delta_x(A) = \begin{cases} 1 & \text{si } x \in A, \\ 0 & \text{si } x \notin A. \end{cases}$$

Un exemple important de tribu non triviale est la tribu borélienne. Elle fait le lien avec la topologie de l'ensemble considéré.

Définition 1.20. Soit X un espace topologique. On appelle tribu borélienne sur X la plus petite tribu[1] contenant tous les ouverts de X. On la note $\mathcal{B}(X)$.

1. il est facile de vérifier que cela a un sens

Sur $\mathbb{R}$ (muni de sa topologie usuelle), un critère utile en pratique est que la tribu $\mathcal{B}(\mathbb{R})$ est la plus petite tribu contenant tous les intervalles de la forme $]a, +\infty[$ avec $a \in \mathbb{R}$. On a d'autres caractérisations de ce type sur $\mathbb{R}$ ou sur $\mathbb{R}^d$, $d \geqslant 2$.

Revenons à la mesure de Lebesgue. La tribu borélienne $\mathcal{B}(\mathbb{R}^d)$ est contenue dans la tribu de Lebesgue. On peut donc restreindre la mesure de Lebesgue à $\mathcal{B}(\mathbb{R}^d)$. On la note toujours λ.

Théorème 1.21. *Soit $d \in \mathbb{N}^*$. Soit μ une mesure sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ invariante par translation et finie sur les bornés. Alors il existe une constante $C \geqslant 0$ telle que $\mu = C\lambda$. En particulier, la mesure de Lebesgue est l'unique mesure sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ telle que la mesure d'un pavé $P = \prod_{j=1}^d [b_j - a_j]$ de $\mathbb{R}^d$ est $\prod_{j=1}^d (b_j - a_j)$.*

Pour finir on introduit la notion de propriété vraie presque partout. Une propriété $\mathcal{P}(x)$ dépendant de $x \in X$ est presque toujours vraie (ou vraie pour presque tout $x \in X$) s'il existe une partie mesurable $E \in \mathcal{M}$ telle que $\mu(E) = 0$ et $\mathcal{P}(x)$ est vraie pour tout $x \in X \setminus E$.

1.5 Construction de l'intégrale de Lebesgue

L'enjeu pour pouvoir définir l'intégrale d'une fonction $f : X \to \mathbb{R}$ est de pouvoir donner une mesure aux images réciproques des intervalles de $\mathbb{R}$ par f.

Définition 1.22. Soient $(X, \mathcal{M})$ et $(Y, \mathcal{N})$ deux espaces mesurables et f une fonction de $\mathcal{M}$ dans $\mathcal{N}$. On dit que f est mesurable si

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

Évidemment la condition est vide si $\mathcal{M} = \mathcal{P}(X)$. Lorsque X et Y sont des espaces topologiques munis de leurs tribus boréliennes, les fonctions mesurables sont également appelés fonctions boréliennes (en particulier, les fonctions continues sont boréliennes).

La notion de mesurabilité est très souple, si bien que, contrairement à l'ensemble des fonctions continues ou des fonctions Riemann-intégrables, l'ensemble des fonctions mesurables est stable par toutes les opérations usuelles, et en particulier le passage à la limite simple (en fait, dans le cas de la mesure de Lebesgue on peut, comme pour les ensembles, faire comme si toutes les fonctions étaient mesurables).

On commence par définir l'intégrale pour une classe de fonctions pour lesquelles la définition est évidente. Le but sera ensuite d'approcher une fonction mesurable quelconque par ces fonctions simples (c'est le rôle des fonctions en escalier pour l'intégrale de Riemann). Pour la suite on fixe un espace mesuré $(X, \mathcal{M}, \mu)$.

Proposition-Définition 1.23. On dit qu'une fonction $f : X \to \mathbb{C}$ est étagée si elle est mesurable et ne prend qu'un nombre fini de valeurs. Une fonction étagée est alors de la forme

$$f = \sum_{j=1}^n \alpha_j \mathbb{1}_{A_j},$$

avec $n \in \mathbb{N}$, $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ et $A_1, \dots, A_n$ sont des parties mesurables deux à deux disjointes. Dans ce cas on pose

$$\int_X f\,d\mu = \sum_{j=1}^{n} \alpha_j \mu(A_j).$$

La deuxième étape consiste à définir l'intégrale des fonctions à valeurs positives (éventuellement infinies).

Définition 1.24. Soit $f : X \to [0, +\infty]$ une fonction mesurable. On appelle intégrale de f (sur X et pour la mesure μ) la quantité

$$\int_X f\,d\mu = \sup_{\substack{h \text{ étagée} \\ 0 \leqslant h \leqslant f}} \int_X h\,d\mu \quad \in [0, +\infty].$$

La propriété suivante est d'abord vérifiée pour les fonctions étagées puis, vue la définition, elle est alors vraie pour toute fonction à valeurs positives.

Proposition 1.25. *Soient $f, g : X \to [0, +\infty]$ mesurables et telles que $f \leqslant g$. Alors on a*

$$\int_X f\,d\mu \leqslant \int_X g\,d\mu.$$

La linéarité de l'intégrale se montre également « à la main » pour les fonctions étagées. Dans le cas général, on utilise le fait que toute fonction mesurable à valeurs positives est limite simple d'une suite de fonctions étagées, et le théorème de convergence monotone (énoncé ici dans le paragraphe suivant, mais qui peut être démontré à ce stade).

Proposition 1.26. *Soient f et g deux fonctions mesurables sur X et $\lambda \in \mathbb{R}$. Alors on a*

$$\int_X (f + \lambda g)\,d\mu = \int_X f\,d\mu + \lambda \int_X g\,d\mu.$$

Proposition 1.27. *Soient $f, g : X \to [0, +\infty]$ deux fonctions mesurables.*

(i) On a $\int_X f\,d\mu = 0$ si et seulement si $f = 0$ p.p..

(ii) Si $\int_X f\,d\mu < +\infty$ alors $f < \infty$ p.p..

(iii) Si $f = g$ presque partout, alors

$$\int_X f\,d\mu = \int_X g d\mu.$$

On peut maintenant définir les intégrales de fonctions de signe variable (dans ce cas on ne peut plus inclure les intégrales infinies).

Définition 1.28. Soit f une fonction de X dans $\mathbb{R}$. On dit que f est intégrable (sur X par rapport à μ) si elle est mesurable et

$$\int_X |f|\,d\mu < +\infty.$$

Dans ce cas on pose

$$\int_X f\,d\mu = \int_X f_+\,d\mu - \int_X f_-\,d\mu.$$

Proposition 1.29. *(i)* $\mathcal{L}^1(X, \mathcal{M}, \mu)$ *est un* $\mathbb{R}$*-espace vectoriel.*

(ii) L'application $f \mapsto \int_X f \, d\mu$ *est une forme linéaire sur* $\mathcal{L}^1(X, \mathcal{M}, \mu)$.

(iii) Pour tout $f \in \mathcal{L}^1(X, \mathcal{M}, \mu)$ *on a*

$$\left| \int_X f \, d\mu \right| \leqslant \int_X |f| \, d\mu.$$

(iv) Soient $f, g \in \mathcal{L}^1(X, \mathcal{M}, \mu)$ *telles que* $f \leqslant g$. *Alors on a*

$$\int_X f \, d\mu \leqslant \int_X g \, d\mu.$$

(v) Soient $f, g \in \mathcal{L}^1(X, \mathcal{M}, \mu)$ *telles que* $f = g$ *p.p.. Alors on a*

$$\int_X f \, d\mu = \int_X g \, d\mu.$$

On peut également définir l'intégrale de fonctions à valeurs complexes, en intégrant séparément les parties réelle et imaginaire.

Pour faire le lien avec l'intégrale de Riemann des fonctions continues par morceaux, on note qu'une fonction continue par morceaux sur un segment est intégrable au sens de Lebesgue (avec la mesure de Lebesgue !), et les deux définitions de l'intégrale coïncident. Une fonction continue par morceaux sur un intervalle quelconque de $\mathbb{R}$ est intégrable au sens de Lebesgue si et seulement si l'intégrale au sens de Riemann est absolument convergente. La terminologie de l'intégrale de Lebesgue ne s'embarasse pas des intégrales semi-convergentes. Mais, bien entendu, la limite

$$\lim_{A \to +\infty} \int_1^A \frac{\cos(x)}{x} \, dx$$

existe toujours, et sa valeur est la même, si l'intégrale sur $[1, A]$ est comprise au sens de Lebesgue.

1.6 Passage à la limite sous l'intégrale - Intégrales à paramètres

L'intérêt principal de l'intégrale de Lebesgue par rapport à l'intégrale de Riemann est qu'elle est mieux adaptée au passage à la limite sous l'intégrale. Néanmoins, il est facile de voir sur des exemples simples qu'une suite (f_n) peut converger simplement vers une fonction f sans que l'intégrale de f_n ne converge vers l'intégrale de f. Dans les trois théorèmes suivants, il y a donc toujours une concession.

Théorème 1.30 (Théorème de la convergence monotone ou Théorème de Beppo-Levi)**.** *Soient* $(f_n)_{n \in \mathbb{N}}$ *une suite croissante de fonctions mesurables de* X *dans* $[0, +\infty]$. *On note* $f : X \to [0, +\infty]$ *la limite ponctuelle de cette suite. Alors* f *est mesurable et on a*

$$\int_X f_n \, d\mu \xrightarrow[n \to \infty]{} \int_X f \, d\mu.$$

Théorème 1.31 (Lemme de Fatou)**.** *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables de X dans $[0,+\infty]$. Alors on a*

$$\int_X \liminf_{n\to\infty} f_n\,d\mu \leqslant \liminf_{n\to\infty} \int_X f_n\,d\mu.$$

Théorème 1.32 (Théorème de convergence dominée)**.** *Soient $(X,\mathcal{M},\mu)$ un espace mesuré et $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions mesurables de X dans $\mathbb{R}$ convergeant simplement vers $f : X \to \mathbb{R}$. On suppose qu'il existe une fonction intégrable $g : X \to [0,+\infty]$ telle que*

$$\forall n \in \mathbb{N}, \forall x \in X, \quad |f_n(x)| \leqslant g(x).$$

Alors f_n est intégrable pour tout $n \in \mathbb{N}$, f est intégrable, et on a

$$\int_X f_n\,d\mu \xrightarrow[n\to\infty]{} \int_X f\,d\mu.$$

Ce dernier théorème est celui qui est le plus souvent utilisé. On note tout de même que, contrairement au théorème de convergence monotone, le théorème de convergence dominée ne permet jamais de conclure qu'une suite d'intégrales tend vers l'infini.

Définition 1.33. On dit que la suite $(f_n)_{n\in\mathbb{N}}$ converge simplement presque partout vers f si $f_n(x)$ tend vers $f(x)$ pour presque tout $x \in X$.

Les théorèmes précédents sont encore valables si la convergence simple de la suite de fonctions n'a lieu que presque partout.

Après les suites définies par des intégrales, on s'intéresse aux fonctions définies par une intégrale. Soit I un intervalle non trivial de $\mathbb{R}$. On considère une fonction $f : X \times I \to \mathbb{R}$ (on peut aussi considérer f à valeurs dans $\mathbb{C}$). Lorsqu'elle est bien définie, on s'intéresse à la régularité de la fonction $F : I \to \mathbb{R}$ définie par

$$F(\lambda) = \int_X f(x,\lambda)\,d\mu(x).$$

On commence par une version du théorème de convergence dominée avec un paramètre continu. Le thoérème de convergence monotone peut également être adapté à ce cadre. On déduit ces résultats à partir des versions précédentes via le critère séquentiel des limites.

Théorème 1.34 (Théorème de convergence dominée)**.** *Soient $\overline{\lambda} \in \overline{I}$ et $\ell : X \to \mathbb{R}$. On suppose que*

(i) la fonction $x \mapsto f(x,\lambda)$ est mesurable pour tout $\lambda \in I$ et ℓ est mesurable,

(ii) pour μ-presque tout $x \in X$ on a

$$f(x,\lambda) \xrightarrow[\lambda\to\overline{\lambda}]{} \ell(x),$$

(iii) il existe $g : X \to \mathbb{R}_+$ intégrable telle que pour tout $\lambda \in I$ on a, pour μ-presque tout $x \in X$,

$$|f(x,\lambda)| \leqslant g(x). \tag{1.2}$$

Alors F est bien définie sur I, ℓ est intégrable sur X, et on a

$$F(\lambda) \xrightarrow[\lambda\to\bar\lambda]{} \int_X \ell(x)\,d\mu(x).$$

A partir du moment où on maîtrise la notion de limite, il n'est pas difficile de discuter de la continuité et de la dérivabilité.

Théorème 1.35 (Théorème de continuité sous l'intégrale)**.** *On suppose que*

(i) pour tout $\lambda \in I$ la fonction $x \mapsto f(x,\lambda)$ est mesurable,

(ii) pour μ-presque tout $x \in X$ la fonction $\lambda \mapsto f(x,\lambda)$ est continue,

(iii) il existe $g : X \to \mathbb{R}_+$ intégrable telle que pour tout $\lambda \in I$ on a, pour μ-presque tout $x \in X$,

$$|f(x,\lambda)| \leqslant g(x).$$

Alors F est bien définie et est continue sur I.

Exemple 1.36 (Fonction Gamma). Pour $s \in]0,+\infty[$, on pose

$$\Gamma(s) = \int_0^{+\infty} x^{s-1}e^{-x}\,dx.$$

Pour tout $s > 0$ la fonction $x \mapsto x^{s-1}e^{-x}$ est continue (et donc borélienne) et à valeurs positives. Soient $a,b > 0$ tels que $a < b$. Pour $x > 0$ et $s \in [a,b]$ on a

$$\left|x^{s-1}e^{-x}\right| \leqslant g(x) = \begin{cases} x^{a-1}e^{-x} & \text{si } x \leqslant 1, \\ x^{b-1}e^{-x} & \text{si } x > 1. \end{cases}$$

La fonction g est continue sur $]0,+\infty[$, on a $g(x) \underset{x\to 0}{\sim} x^{a-1}$ et, par croissance comparée, $g(x) \underset{x\to+\infty}{=} O(e^{-\frac{x}{2}})$. La fonction g est donc intégrable sur $]0,+\infty[$. Par le théorème de continuité sous l'intégrale, Γ est intégrable sur $[a,b]$. Ceci étant valable pour tout segment de $]0,+\infty[$, on obtient que Γ est continue sur $]0,+\infty[$.

Pour $x > 1$ la fonction $s \mapsto x^{s-1}e^{-x}$ est positive et croit vers $+\infty$, donc par le théorème de convergence monotone on a

$$\Gamma(s) \geqslant \int_1^{+\infty} x^{s-1}e^{-x}\,dx \xrightarrow[s\to+\infty]{} +\infty.$$

De même,

$$\Gamma(s) \geqslant \int_0^{1} x^{s-1}e^{-x}\,dx \xrightarrow[s\to 0]{} +\infty$$

(si $s_n > 0$ décroît vers 0 alors $x^{s_n-1}e^{-x}$ croît vers $+\infty$ pour tout $x \in]0,1[$).

Théorème 1.37 (Théorème de dérivation sous l'intégrale)**.** *On suppose que*

(i) pour tout $\lambda \in I$ la fonction $x \mapsto f(x,\lambda)$ est intégrable,

(ii) pour μ-presque tout $x \in X$ la fonction $\lambda \mapsto f(x,\lambda)$ est dérivable,

(iii) il existe $g : X \to \mathbb{R}_+$ intégrable telle que pour μ-presque tout $x \in X$ on a

$$\forall \lambda \in I, \quad \left|\frac{\partial f}{\partial \lambda}(x,\lambda)\right| \leqslant g(x).$$

Alors F est bien définie et est dérivable sur I de dérivée

$$F' : \lambda \mapsto \int_X \frac{\partial f}{\partial \lambda}(x, \lambda)\, d\mu(x).$$

Exemple 1.38. On revient sur l'exemple de la fonction Γ définie à l'exemple 1.36. On commence par noter que la fonction $s \mapsto x^{s-1}e^{-x}$ est dérivable pour tout $x > 0$. Soient $a, b \in]0, +\infty[$ tels que $a < b$. Pour tous $x > 0$ et $s \in]a, b[$ on a

$$\left| \frac{\partial}{\partial s} x^{s-1} e^{-x} \right| = \left| \ln(x) x^{s-1} e^{-x} \right| \leqslant g(x),$$

où on a noté

$$g(x) = \begin{cases} \ln(x)x^{a-1}e^{-x} & \text{si } x \in]0, 1], \\ \ln(x)x^{b-1}e^{-x} & \text{si } x \in [1, +\infty[. \end{cases}$$

Comme précédemment on obtient que g est intégrable sur $]0, +\infty[$. Par le théorème de dérivation sous l'intégrale, on obtient que Γ est dérivable sur $]a, b[$ et pour tout $x \in]a, b[$ on a

$$\Gamma'(s) = \int_0^{+\infty} \ln(x) x^{s-1} e^{-x}\, dx.$$

Ceci étant valable pour tous $a, b > 0$, cela prouve que γ est de classe C^1 sur $]0, +\infty[$ et l'expression de Γ' est bien valable sur ce domaine. On montre alors par récurrence que Γ est en fait dérivable à tout ordre et les dérivées successives sont obtenues en dérivant sous l'intégrale.

Tous ces résultats font intervenir deux variables qui jouent des rôles complètement différents, attention à ne pas tout mélanger dans les hypothèses (d'autant qu'en général X est $\mathbb{R}$ ou un intervalle de $\mathbb{R}$, ce qui augmente les possibilités de confusion).

Il est important de se rappeller que la continuité et la dérivabilité sont des propriétés locales. Il suffit donc de les vérifier au voisinage de chaque point. On note dans l'exemple de la fonction Γ qu'on n'est pas capable de vérifier l'hypothèse de domination uniformément pour tout $s \in]0, +\infty[$, mais qu'on travaille sur $[a, b]$ pour n'importe quels a, b tels que $0 < a < b < +\infty$. Cela suffit car tout point de $]0, +\infty[$ a un voisinage de cette forme. Par contre on ne peut pas toucher au domaine d'intégration...

1.7 Intégrales multiples - Théorème de Fubini

Le but du théorème de Fubini est d'exprimer l'intégrale d'une fonction de plusieurs variables en fonction d'intégrales d'une seule variable. En effet, tous les calculs que l'on sait faire jusqu'à présent (calcul explicite via une primitive, intégration par parties, etc.) sont des résultats pour une fonction d'une seule variable réelle. Le but du théorème de Fubini est de se ramener à ce cadre plus agréable. En fait, quasiment tous les calculs ou résultats abstraits obtenus en dimension supérieure sont déduit du cas de la dimension 1 via le théorème de Fubini.

Avant d'énoncer des théorèmes généraux, il faut introduire les structures adéquates. Si on veut ramener une intégrale sur $X \times Y$ à des intégrales simples sur X et Y, il faut que la structure d'espace mesuré sur $X \times Y$ soit compatible avec celles sur X et Y. Par exemple, la mesure de Lebesgue sur $\mathbb{R}^2$ est construite à partir des mesures des rectangles, celle sur $\mathbb{R}$ à partir des mesures des intervalles, et la mesure d'un rectangle est exactement le produit des mesures de ses deux côtés (qui sont des intervalles de $\mathbb{R}$). C'est ce schéma que l'on reproduit dans le cas général.

Définition 1.39. Soient $(X_1, \mathcal{M}_1)$ et $(X_2, \mathcal{M}_2)$ deux espaces mesurables. On appelle rectangle mesurable de $X_1 \times X_2$ un ensemble de la forme $A_1 \times A_2$ avec $A_1 \in \mathcal{M}_1$ et $A_2 \in \mathcal{M}_2$. On appelle tribu produit et on note $\mathcal{M}_1 \otimes \mathcal{M}_2$ la plus petite tribu de $X_1 \times X_2$ contenant les rectangles mesurables.

Proposition 1.40. *On a $\mathcal{B}(\mathbb{R}^2) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$. Plus généralement, on a $\mathcal{B}(\mathbb{R}^{n+m}) = \mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}(\mathbb{R}^m)$ pour tous $n, m \in \mathbb{N}^*$.*

On note $X = X_1 \times X_2$ et $\mathcal{M} = \mathcal{M}_1 \otimes \mathcal{M}_2$. Pour $A \subset X$ et $x_1 \in X_1$ on notera

$$A(x_1, \cdot) = \{x_2 \in X_2 \,|\, (x_1, x_2) \in A\} \subset A_2.$$

Pour $x_2 \in X_2$ on notera également

$$A(\cdot, x_2) = \{x_1 \in X_1 \,|\, (x_1, x_2) \in A\} \subset A_1.$$

Définition 1.41. On dit que l'espace mesuré $(X, \mathcal{M}, \mu)$ est σ-fini si X est union dénombrables de parties de mesures finies.

Ainsi la mesure de Lebesgue sur $\mathbb{R}^d$ n'est pas finie ($\lambda(\mathbb{R}^d) = +\infty$) mais elle est σ-finie ($\mathbb{R}^d$ est union des $[-n, n]^d$ qui sont tous de mesure finie).

Le résultat suivant officialise l'idée que la surface d'une partie de $\mathbb{R}^2$ doit être la « somme continue » des longueurs des « tranches » horizontales ou verticales.

Théorème 1.42. *On suppose que $(X_1, \mathcal{M}_1, \mu_1)$ et $(X_2, \mathcal{M}_2, \mu_2)$ sont σ-finis. Alors il existe une unique mesure $\mu_1 \otimes \mu_2$ sur $(X, \mathcal{M})$ telle que*

$$\forall A_1 \in \mathcal{M}_1, \forall A_2 \in \mathcal{M}_2, \quad (\mu_1 \otimes \mu_2)(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2). \tag{1.3}$$

En outre pour $A \in \mathcal{M}$ les fonctions $x_1 \mapsto \mu_2(A(x_1, \cdot))$ et $x_2 \mapsto \mu_1(A(\cdot, x_2))$ sont mesurables (de X_1 dans $[0, +\infty]$ et de X_2 dans $[0, +\infty]$, respectivement) et on a

$$(\mu_1 \otimes \mu_2)(A) = \int_{X_1} \mu_2(A(x_1, \cdot))\, d\mu_1(x_1) = \int_{X_2} \mu_1(A(\cdot, x_2))\, d\mu_2(x_2). \tag{1.4}$$

Si λ_1 désigne la mesure de Lebesgue sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, alors $\lambda_1 \otimes \lambda_1$ est la mesure de Lebesgue λ_2 sur $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$.

Par exemple, si on considère le triangle $T = \{(x, y) \in [0, 1]^2 \,|\, x + y \leqslant 1\}$, alors pour $x \in \mathbb{R}$ on a

$$\lambda_1(T(x, \cdot)) = \begin{cases} \emptyset & \text{si } x \notin [0, 1], \\ [0, 1 - x] & \text{si } x \in [0, 1], \end{cases}$$

et donc

$$\lambda_2(T) = \int_0^1 (1 - x)\, dx = \frac{1}{2}.$$

On note $\mu = \mu_1 \otimes \mu_2$. On peut maintenant donner les théorèmes de Fubini. Le premier ne s'applique qu'aux fonctions à valeurs positives. Le deuxième, qui autorise les fonctions de signe variable, nécessite de vérifier l'intégrabilité de la fonction considérée. En général, cette hypothèse d'intégrabilité de f qui apparaît dans le théorème de Fubini-Lebesgue est démontrée… en appliquant le théorème de Fubini-Tonelli à $|f|$.

Théorème 1.43 (Théorème de Fubini-Tonelli)**.** *On suppose que* $(X_1, \mathcal{M}_1, \mu_1)$ *et* $(X_2, \mathcal{M}_2, \mu_2)$ *sont* σ*-finis. Soit* f *une fonction* mesurable *de* $(X, \mathcal{M})$ *dans* $[0, +\infty]$.

(i) *Les fonctions*

$$x_1 \mapsto \int_{X_2} f(x_1, x_2)\, d\mu_2(x_2) \quad \text{et} \quad x_2 \mapsto \int_{X_1} f(x_1, x_2)\, d\mu_1(x_1)$$

sont bien définies et mesurables sur X_1 *et* X_2*, respectivement, à valeurs dans* $[0, +\infty]$.

(ii) *On a*

$$\int_X f\, d\mu = \int_{X_1} \left(\int_{X_2} f(x_1, x_2)\, d\mu_2(x_2) \right) d\mu_1(x_1) = \int_{X_2} \left(\int_{X_1} f(x_1, x_2)\, d\mu_1(x_1) \right) d\mu_2(x_2). \tag{1.5}$$

Théorème 1.44 (Théorème de Fubini-Lebesgue)**.** *On suppose que* $(X_1, \mathcal{M}_1, \mu_1)$ *et* $(X_2, \mathcal{M}_2, \mu_2)$ *sont* σ*-finis. Soit* f *une fonction* intégrable *de* $(X, \mathcal{M})$ *dans* $\mathbb{R}$.

(i) *La fonction* $f(x_1, \cdot)$ *est intégrable sur* X_2 *pour presque tout* $x_1 \in X_1$*, et la fonction* $x_1 \mapsto \int_{X_2} f(x_1, x_2)\, d\mu_2(x_2)$ *(définie* μ_1*-presque partout) est intégrable sur* X_1*. De même,* $f(\cdot, x_2)$ *est intégrable sur* X_1 *pour presque tout* $x_2 \in X_2$*, et la fonction* $x_2 \mapsto \int_{X_1} f(x_1, x_2)\, d\mu_1(x_1)$ *(définie* μ_2*-presque partout) est intégrable sur* X_2.

(ii) *On a*

$$\int_X f\, d\mu = \int_{X_1} \left(\int_{X_2} f(x_1, x_2)\, d\mu_2(x_2) \right) d\mu_1(x_1) = \int_{X_1} \left(\int_{X_2} f(x_1, x_2)\, d\mu_1(x_1) \right) d\mu_2(x_2).$$

1.8 Théorème de changement de variables

On commence par le résultat abstrait de changement de variable.

Proposition-Définition 1.45. Soit $(X, \mathcal{M}, \mu)$ un espace mesuré. Soient Y un ensemble et φ une fonction de X dans Y.

(i) La famille

$$\mathcal{N} = \left\{ B \in \mathcal{P}(Y) \,|\, \varphi^{-1}(B) \in \mathcal{M} \right\}$$

est une tribu sur Y, appelée tribu image de $\mathcal{M}$ par f.

(ii) La fonction φ est alors mesurable de $(X, \mathcal{M})$ dans $(Y, \mathcal{N})$.

(iii) Pour $B \in \mathcal{N}$ on note $\nu(B) = \mu(f^{-1}(B))$. Cela définit une mesure ν sur l'espace mesurable $(X, \mathcal{N})$, appelée mesure image de μ par f.

Théorème 1.46 (Théorème abstrait de changement de variables)**.** *Soient* $(X, \mathcal{M}, \mu)$ *et* $(Y, \mathcal{N}, \mu)$ *deux espaces mesurés et* φ *une fonction de* X *dans* Y*. On suppose que* $\mathcal{N}$ *et* ν *sont la tribu et la mesure image de* $\mathcal{M}$ *et* μ *par* φ.

(i) Soit $f : Y \to [0, +\infty]$ *une fonction mesurable. Alors* $f \circ \varphi$ *est mesurable sur* X *on a*

$$\int_Y f\, d\nu = \int_X (f \circ \varphi)\, d\mu.$$

(ii) *Soit $f : Y \to \mathbb{R}$ une fonction mesurable. Alors f est intégrable sur Y si et seulement si $(f \circ \varphi)$ est intégrable sur X, et dans ce cas on a*

$$\int_Y f\,d\nu = \int_X (f \circ \varphi)\,d\mu.$$

En identifiant la mesure image de la mesure de Lebesgue par un C^1-difféomorphisme, on obtient le résultat suivant :

Théorème 1.47 (Théorème de changement de variable dans $\mathbb{R}^d$)**.** *Soient U et V deux ouverts de $\mathbb{R}^d$ (munis de leurs tribus boréliennes et de la mesure de Lebesgue λ), et $\varphi : U \to V$ un difféomorphisme de classe C^1.*

(i) *Soit $f : V \to [0, +\infty]$ une fonction mesurable. Alors $(f \circ \varphi)\,|J\varphi|$ est une fonction mesurable sur U et on a*

$$\int_V f(y)\,d\lambda(y) = \int_U f(\varphi(x))\,|J\varphi(x)|\;d\lambda(x).$$

(ii) *Soit $f : V \to \mathbb{R}$ une fonction mesurable. Alors $(f \circ \varphi)\,|J\varphi|$ est une fonction mesurable sur U. En outre f est intégrable sur V si et seulement si $(f \circ \varphi)\,|J\varphi|$ est intégrable sur U et dans ce cas on a*

$$\int_V f(y)\,d\lambda(y) = \int_U f(\varphi(x))\,|J\varphi(x)|\;d\lambda(x).$$

Ce résultat est à comparer avec le théorème 1.12 donné en dimension 1. Dans ce résultat, on tenait compte du signe de la dérivée de φ, ce qui ne serait pas le cas en appliquant le théorème 1.47 en dimension 1. La différence est que l'intégrale définie au premier paragraphe est orientée (ce n'est pas pareil de regarder l'intégrale de a vers b ou l'intégrale de b vers a), alors qu'ici l'intégrale de Lebesgue est une intégrale sur un ensemble non orienté (par exemple le segment $[a, b]$, sans préciser de sens). On n'insiste pas sur cette différence ici, mais il ne faut pas être perturbé par le fait que le résultat qu'on utilise en dimension supérieure n'est pas une généralisation exacte du résultat connu en dimension 1 (c'est possible, mais ce n'est pas notre objectif ici). Une autre conséquence de cette différence est qu'ici on doit supposer que φ est un C^1-difféomorphisme, ce qui n'était pas le cas pour le théorème 1.12.

On introduit maintenant des changements de variables particulièrement utiles. En fonction des symétries du problème étudié, ces changements de variables peuvent permettre de considérablement simplifier l'expression des intégrales à calculer.

Proposition 1.48 (Coordonnées polaires)**.** *L'application*

$$\Phi : \begin{cases} \mathbb{R}_+^* \times]-\pi, \pi[& \to \quad \mathbb{R}^2 \setminus (\mathbb{R}_- \times \{0\}) \\ \quad (r, \theta) & \mapsto \quad (r\cos(\theta), r\sin(\theta)) \end{cases}$$

est un C^1-difféomorphisme. En outre pour tout $(r, \theta) \in \mathbb{R}_+^* \times]-\pi, \pi[$ on a

$$J\Phi(r, \theta) = r.$$

Remarque 1.49. — On a un résultat analogue en enlevant à $\mathbb{R}^2$ n'importe quelle demi-droite issue de l'origine.

— De façon générale, si φ est un C^1-difféomorphisme de U dans V, et $\tilde{U}$ est un ouvert inclus dans U, alors φ réalise un C^1-difféomorphisme de $\tilde{U}$ dans $\varphi(\tilde{U})$. En particulier, on peut passer en coordonnées polaires pour n'importe quel ouvert de $\mathbb{R}^2 \setminus (\mathbb{R}_- \times \{0\})$.

Ce changement de variables est agréable quand la frontière du domaine d'intégration s'exprime plus facilement comme courbe paramétrée en polaire et/ou que la fonction à intégrer présente une symétrie radiale.

Exemple 1.50. Soit $R > 0$. On veut calculer $\lambda(D_R)$, où D_R est le disque centré en 0 et de rayon R. On note $\tilde{D} = D \setminus (\mathbb{R}_- \times \{0\})$. Comme une demi-droite est de mesure de Lebesgue nulle dans $\mathbb{R}^2$ on a

$$\lambda(D) = \lambda(\tilde{D}) = \int_{\tilde{D}} 1\, d\lambda = \int_{[0,R]\times]-\pi,\pi[} r\, d\lambda(r,\theta) = \int_{\theta=-\pi}^{\pi} \int_{r=0}^{R} r\, dr\, d\theta = \int_{-\pi}^{\pi} \frac{R^2}{2}\, d\theta = \pi R^2.$$

Pour la troisième égalité, on a utilisé la formule de changement de variable, sachant que l'application $(r,\theta) \mapsto (r\cos(\theta), r\sin(\theta))$ réalise un C^1-difféomorphisme de $]0,R[\times]-\pi,\pi[$ dans $\tilde{D}$. Pour l'égalité suivante on a utilisé le théorème de Fubini.

Proposition 1.51 (Coordonnées cylindriques)**.** *L'application*

$$\Phi : \left\{ \begin{array}{ccc} \mathbb{R}_+^* \times]-\pi,\pi[\times\mathbb{R} & \to & \mathbb{R}^3 \setminus (\mathbb{R}_- \times \{0\} \times \mathbb{R}) \\ (r,\theta,z) & \mapsto & (r\cos(\theta), r\sin(\theta), z) \end{array} \right.$$

est un C^1-difféomorphisme. En outre pour tout $(r,\theta,z) \in \mathbb{R}_+^* \times]-\pi,\pi[\times\mathbb{R}$ on a

$$J\Phi(r,\theta,z) = r.$$

Proposition 1.52 (Coordonnées sphériques)**.** *L'application*

$$\Phi : \left\{ \begin{array}{ccc} \mathbb{R}_+^* \times]-\pi,\pi[\times]-\frac{\pi}{2},\frac{\pi}{2}[& \to & \mathbb{R}^3 \setminus (\mathbb{R}_- \times \{0\} \times \mathbb{R}) \\ (r,\theta,\varphi) & \mapsto & \begin{pmatrix} r\cos(\theta)\cos(\varphi) \\ r\sin(\theta)\cos(\varphi) \\ r\sin(\varphi) \end{pmatrix} \end{array} \right.$$

est un C^1-difféomorphisme. En outre pour tout $(r,\theta,\varphi) \in \mathbb{R}_+^* \times]-\pi,\pi[\times]-\frac{\pi}{2},\frac{\pi}{2}[$ on a

$$J\Phi(r,\theta) = r^2\cos(\varphi).$$

1.9 Espaces de Lebesgue

Soit $(X, \mathcal{M}, \mu)$ un espace mesuré.

Définition 1.53. Soit $p \in [1,+\infty[$. Étant donnée une fonction f mesurable de X dans $\mathbb{R}$, on note

$$\|f\|_p = \left(\int_X |f|^p \, d\mu \right)^{\frac{1}{p}}$$

et

$$\|f\|_\infty = \sup_{x\in X} \operatorname{ess} |f(x)| = \inf \{\rho \geqslant 0 \mid |f| \leqslant \rho \text{ presque partout}\}.$$

Pour $p \in [1,+\infty]$ on note $\mathcal{L}^p(X,\mathcal{M},\mu)$ (ou $\mathcal{L}^p(X)$ s'il n'y a pas d'ambiguïté) l'ensemble des fonctions mesurables $f : X \to \mathbb{C}$ telles que

$$\|f\|_p < +\infty.$$

$\mathcal{L}^1$ n'est rien d'autre que l'ensemble des fonctions intégrables, tandis que $\mathcal{L}^\infty$ est l'ensemble des fonctions essentiellement bornées (bornées si on oublie un ensemble de mesure nulle). On définit maintenant les exposants conjugués, qui joueront un rôle important dans l'analyse des espaces de Lebesgue. On note que 1 et $+\infty$ sont conjugués, tandis que 2 est conjugué à lui-même.

Définition 1.54. Soient $p, q \in [1, \infty]$. On dit que p et q sont des exposants conjugués si

$$\frac{1}{p} + \frac{1}{q} = 1$$

(avec la convention $1/\infty = 0$).

Théorème 1.55 (Inégalité de Hölder)**.** *Soient $p, q \in [1, +\infty]$ des exposants conjugés. Alors pour $f \in \mathcal{L}^p(X, \mathcal{M}, \mu)$ et $g \in \mathcal{L}^q(X, \mathcal{M}, \mu)$ on a $fg \in \mathcal{L}^1(X, \mathcal{M}, \mu)$ et*

$$\|fg\|_1 \leqslant \|f\|_p \|g\|_q .$$

Théorème 1.56 (Inégalité de Minkowski)**.** *Soient $p \in [1, +\infty]$ et $f, g \in \mathcal{L}^p(X, \mathcal{M}, \mu)$. Alors $f + g \in \mathcal{L}^p(X, \mathcal{M}, \mu)$ et*

$$\|f + g\|_p \leqslant \|f\|_p + \|g\|_p .$$

Avec l'inégalité de Minkiwski, on obtient que $\mathcal{L}^p(X, \mathcal{M}, \mu)$ est un espace vectoriel sur lequel l'application $\|\cdot\|_p$ est une semi-norme. Mais pas une norme, puisqu'une fonction f qui est non nulle mais presque partout nulle (ce qui est typiquement possible avec la mesure de Lebesgue) est telle que $f \neq 0$ et $\|f\|_p = 0$. Pour palier ce défaut, on va considérer comme nulles les fonctions qui sont presque partout nulles. Pour cela, on considère sur l'espace des fonctions mesurables de X dans $\mathbb{R}$ la relation d'équivalence $\mathcal{R}$ définie par

$$f \,\mathcal{R}\, g \quad \iff \quad f(x) = g(x) \text{ pour presque tout } x,$$

puis on note $L^p(X, \mathcal{M}, \mu)$ l'espace obtenu en quotientant $\mathcal{L}^p(X, \mathcal{M}, \mu)$ par la relation d'équivalence $\mathcal{R}$.

Autrement dit, un élément de L^p est une classe d'équivalence d'éléments de $\mathcal{L}^p$. Quand on parle d'une « fonction » de L^p on parle en fait de n'importe quelle fonction dans la même classe d'équivalence. En particulier, avec la mesure de Lebesgue, on ne peut pas parler de la valeur en un point d'une fonction de L^p, puisqu'on ne change pas la fonction en changeant la valeur en ce point (on reste dans la même classe d'équivalence).

En général on identifiera une fonction de $\mathcal{L}^p$ avec sa classe d'équivalence dans L^p, ce qui peut être source de maux de tête tant qu'on n'en a pas l'habitude.

L'intérêt de travailler avec des classes d'équivalence de fonctions plutôt qu'avec les fonctions elle-mêmes est d'avoir un espace vectoriel normé avec de bonnes propriétés.

Proposition 1.57. *L'espace $L^p(X, \mathcal{M}, \mu)$ hérite de la stucture de $\mathbb{R}$-espace vectoriel de $\mathcal{L}^p(X, \mathcal{M}, \mu)$ et de la semi-norme $\|\cdot\|_p$, qui est en fait une norme.*

On a dit que lorsqu'on parle d'une « fonction » f de $L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, cela n'a pas de sens de parler de la valeur de f en un point, puisque f représente toute une famille de fonctions qui n'ont pas la même valeur en ce point. Par contre on peut parler de l'intégrale de f sur $\mathbb{R}$, puisque cette quantité est commune à toutes les fonctions que f représente.

Théorème 1.58 (Inégalité de Hölder, le retour)**.** *Soient $p, q \in [1, +\infty]$ des exposants conjugés. Alors pour $f \in L^p(X, \mathcal{M}, \mu)$ et $g \in L^q(X, \mathcal{M}, \mu)$ on a $fg \in L^1(X, \mathcal{M}, \mu)$ et*

$$\|fg\|_1 \leqslant \|f\|_p \|g\|_q .$$

Les éléments des espaces $\mathcal{L}^p$ sont compliqués. Les fonctions L^p sont déjà potentiellement très désagréables, et on devra de plus gérer le passage au quotient identifiant les fonctions égales presque partout. Mais tous ces efforts ne sont pas vains, puisqu'on obtient au final des espaces de fonctions intégrables qui eux sont très agréables :

Théorème 1.59 (Théorème de Riesz-Fisher)**.** *L'espace de Lebesgue $L^p(X, \mathcal{M}, \mu)$, muni de la norme $\|\cdot\|_p$, est un espace de Banach (espace vectoriel normé complet).*

Dans la preuve du théorème de Riesz-Fisher, on montre en passant la propriété suivante.

Proposition 1.60. *Soit $p \in [1, +\infty[$ et $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy dans $L^p(X, \mathcal{M}, \mu)$. Alors il existe une sous-suite $(f_{n_k})_{k\in\mathbb{N}}$ qui converge simplement presque partout.*

Parmi tous les espaces L^p, le cas $p = 2$ jouera un rôle particulier, du fait que l'exposant 2 est conjugué à lui-même.

Corollaire 1.61. *L'espace de Lebesgue $L^2(X, \mathcal{M}, \mu)$ est un espace de Hilbert pour le produit scalaire défini par*

$$\forall f, g \in L^2(X, \mathcal{M}, \mu), \quad \langle f, g\rangle = \int_X f\overline{g}\, d\mu.$$

On note qu'en général il n'y a pas d'inclusion entre les espaces de Lebesgue. Par exemple, pour $p, q \in [1, \infty]$ tels que $p < q$ et $q \in]p, q[$ on a

$$x^{-\frac{1}{q}}\mathbb{1}_{[1,+\infty[} \in \mathcal{L}^q(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda) \setminus \mathcal{L}^p(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$$

et

$$x^{-\frac{1}{q}}\mathbb{1}_{]0,1]} \in \mathcal{L}^p(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda) \setminus \mathcal{L}^q(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda).$$

Néanmoins, pour un espace de mesure fini on a les inclusions suivantes.

Proposition 1.62. *On suppose que $\mu(X) < \infty$. Alors pour tous $p, q \in [1, +\infty]$ tels que $p < q$ on a*

$$\mathcal{L}^q(X, \mathcal{M}, \mu) \subset \mathcal{L}^p(X, \mathcal{M}, \mu),$$

et pour $f \in \mathcal{L}^q(X, \mathcal{M}, \mu)$ on a

$$\|f\|_p \leqslant \mu(X)^{\frac{1}{p}-\frac{1}{q}} \|f\|_q .$$

En particulier, sur un espace de mesure fini, les espaces $\mathcal{L}^p$ sont constitués de fonctions intégrables. Sur $\mathbb{R}^d$ on n'a pas ce type de résultat, sauf si on se restreint à un compact (qui est lui de mesure finie). Ainsi une fonction L^p est au moins localement intégrable :

Définition 1.63. On note $\mathcal{L}^1_{\text{loc}}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ l'ensemble des fonctions mesurables $f : \mathbb{R} \to \mathbb{R}$ telles que pour tout compact K de $\mathbb{R}$ on a $f\mathbb{1}_K \in \mathcal{L}^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$.

On note que sur $\mathbb{N}$, muni de la mesure de comptage (ou sur tout espace mesuré tel que tout ensemble est soit de mesure nulle soit de mesure plus grande qu'un certain $\nu > 0$), on a des inclusions inverses entre les espaces de Lebesgue par rapport à la proposition précédente.

Pour $p \in [1, \infty]$ on note $l^p(\mathbb{N})$ l'espace $\mathcal{L}^p(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$, où μ est la mesure de comptage sur $\mathbb{N}$. Pour $a = (a_n)_{n\in\mathbb{N}}$ on a alors

$$\|a\|_p = \left(\sum_{n=0}^{\infty} |a_n|^p \right)^{\frac{1}{p}}, \quad \text{pour } p \in [1, +\infty[,$$

et

$$\|a\|_\infty = \sup_{n\in\mathbb{N}} |a_n| .$$

Pour $p, q \in]1, +\infty[$ avec $p < q$ on a

$$l^p(\mathbb{N}) \subset l^q(\mathbb{N}).$$

On termine ce passage sur les inclusions entre espaces de Lebesgue avec le résultat d'interpolation suivant.

Proposition 1.64. *Soient $p, q \in [1, +\infty]$ avec $p < q$ et $f \in \mathcal{L}^p(X, \mathcal{M}, \mu) \cap \mathcal{L}^q(X, \mathcal{M}, \mu)$. Alors pour tout $r \in [p, q]$ on a $f \in \mathcal{L}^r(X, \mathcal{M}, \mu)$*

$$\|f\|_r \leqslant \|f\|_p^\theta \|f\|_q^{1-\theta}$$

où $\theta \in [0, 1]$ est tel que

$$\frac{\theta}{p} + \frac{1-\theta}{q} = \frac{1}{r}.$$

On termine ce paragraphe avec un résultat de densité dans les espaces de Lebesgue, et des exemples d'applications.

Théorème 1.65. *Soit $p \in [1, +\infty[$. L'ensemble $C_c^0(\mathbb{R}^d)$ des fonctions continues à supports compacts (c'est-à-dire nulles en dehors d'un compact) dans $\mathbb{R}^d$ est dense dans $L^p(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda)$.*

Proposition 1.66. *Soit $p \in [1, +\infty[$. Alors $L^p(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \lambda)$ est séparable (admet une partie dénombrable dense).*

Pour $f : \mathbb{R}^d \to \mathbb{R}$ et $y \in \mathbb{R}^d$, on note

$$\tau_y f : \left\{ \begin{array}{ccc} \mathbb{R}^d & \to & \mathbb{R} \\ x & \mapsto & f(x-y) \end{array} \right.$$

Proposition 1.67. *Soient $p \in [1, +\infty[$ et $f \in L^p(\mathbb{R}^d)$. Alors on a*

$$\|\tau_p f - f\|_p \xrightarrow[y\to 0]{} 0.$$

Ces derniers résultats ne sont pas valables avec $p = \infty$. En effet, si on considère $f = \mathbb{1}_{\mathbb{R}_+}$ alors pour tout $\phi \in C_c^0(\mathbb{R})$ on a

$$\|f - \phi\|_\infty \geqslant \frac{1}{2}$$

et pour tout $y \in \mathbb{R}^*$ on a

$$\|\tau_y f - f\|_\infty = 1.$$