Intégrale de Lebesgue sur $\mathbf{R}^d$

Polycopié de l'UE 2MA211

Pierre-Antoine Guihéneuf, Ayman Moussa, Frédéric Paugam

Année 2021/2022

Table des matières

Introduction

Qu'est-ce que l'aire d'une surface ? Le volume d'un objet ? En mettant de côté les formes élémentaires (rectangles, triangles...) pour lesquelles la réponse est communément connue, si l'on dessine une courbe fermée dans le plan, que signifie précisément l'*aire* délimitée par cette partie ? Et surtout : comment la calcule-t-on ? Ces questions naturelles sont au cœur de la théorie de l'intégration, qui offre un outil permettant de généraliser le concept de moyenne arithmétique pondérée au cas d'une somme « continue » (en un certain sens) où la pondération se fait par la taille (la longueur, l'aire, le volume) d'éléments infinitésimaux.

Rome ne s'est pas faite un jour... la théorie de l'intégration non plus ! Il a fallu plus de 200 ans entre la définition d'intégrale par Newton et Leibniz à la fin du XVII^e siècle et la formalisation rigoureuse en 1902 par Lebesgue de l'intégrale qu'on utilise encore aujourd'hui [1] et que l'on se propose de présenter dans ce cours.

La notion d'intégrale peut s'aborder selon des points de vues complètement différents, qui viennent eux-même des différents aspects du concept de fonction.

1. L'aspect algébrique ou calculatoire : une fonction est une expression faisant intervenir des opérations élémentaires (additions, multiplications...), des fonctions élémentaires (exponentielle, sinus, logarithme), des séries etc.
2. L'aspect géométrique ou graphique : une fonction est une courbe.
3. L'aspect ensembliste : une fonction est un objet qui à certaines quantités x associe une autre quantité y.

Si le troisième aspect est celui par lequel on *définit* de nos jours ce qu'est une fonction, ce sont les deux premiers qui, historiquement, sont apparus les premiers. Le point de vue (3), le plus général, n'est devenu prépondérant qu'à partir de la fin du XIX^e siècle avec la croissance de la formalisation et l'émergence de la construction axiomatique. Il ne faut pas s'imaginer que la recherche de la bonne définition de fonction a hanté les nuits des mathématiciens pendant tout ce temps : pour la pratique des mathématiques de l'époque, ils n'avaient simplement pas besoin d'une définition aussi abstraite.

Ces différents points de vue sur l'objet mathématique fonction amènent chacun à des définitions différentes d'intégrale (ou de primitive).

1. La primitive d'une fonction est définie algorithmiquement. Cela pose bon nombre de problèmes puisqu'il y a beaucoup de primitives de fonctions simples n'admettant pas d'expression à l'aide des fonctions usuelles.
2. L'intégrale d'une courbe est simplement l'aire (algébrique [2]) de la région délimitée par la courbe et l'axe des abscisses. On voit là le lien étroit entre les notions d'intégrale et de mesure des ensembles.
3. L'intégrale est une opération sur les fonctions vérifiant une certaine liste de propriétés. On y retrouve la définition de l'intégrale (ou la primitivation) comme opération inverse de la dérivation, mais aussi la définition axiomatique que nous adopterons dans ce cours.

1. L'histoire ne s'arrête néanmoins pas avec Lebesgue : viendront ensuite les intégrales de Kurzweil-Henstock, Daniell...

2. Cela signifie que l'aire associée à une portion de courbe située *sous* l'axe des abscisses est comptée négativement

Les différents aspects de l'intégrale ont été très vite compris, mais la première définition rigoureuse d'intégrale date du milieu du XIX$^\text{e}$ siècle avec les travaux de Cauchy puis de Riemann. Seulement, l'intégrale de Riemann ne se comporte pas très bien lors de passages à la limite sous l'intégrale, et c'est en cela l'apport fondamental de la théorie de Lebesgue : un énoncé de convergence dominée (Théorème 2.13).

La manière classique de définir ces intégrales est du type (2) (géométrique) : l'aire sous la courbe est calculée *via* la théorie de la mesure de Borel et Lebesgue, grande avancée du début du XX$^\text{e}$ siècle qui compte parmi ses autres applications la formalisation des probabilités pour des variables à densité (Définition 1.55) ainsi que la théorie ergodique, pendant probabiliste de la théorie du chaos qui vise à mesurer à quel point des systèmes déterministes ont des comportements aléatoires [3].

Malheureusement, la définition rigoureuse de l'intégration de Lebesgue demande beaucoup de travail : à la fois des notions de théorie de la mesure, et la construction “à la main” de la mesure de Lebesgue. Ces notions, habituellement vues en L3, dépassent largement le contenu de ce cours, c'est pourquoi nous adopterons une démarche inverse : à la manière du point (3), nous commencerons par définir de manière *axiomatique* l'intégrale de Lebesgue [4] pour *ensuite* en déduire les propriétés de la mesure de Lebesgue.

Notons que cette théorie de l'intégration ne permet pas de donner un sens à l'intégrale de n'importe quelle fonction : elle a comme objet l'ensemble des fonctions *mesurables*, qui contient toutes les fonctions raisonnables auxquelles on peut penser. Il existe néanmoins des fonctions non mesurables dont l'existence amène à des énoncés contre-intuitifs comme le paradoxe de Banach-Tarski.

3. Par exemple, la trajectoire d'un dé est déterministe, mais on peut raisonnablement dire que le résultat qu'il donne est aléatoire.

4. C'est-à-dire qu'on supposera l'existence d'un objet appelé intégrale qui vérifie une liste de propriétés, et qu'on ne démontrera pas l'existence de cet objet.

Préliminaires : relations d'équivalence

Les relations d'équivalence permettent de ranger les éléments d'un ensemble dans ce qu'on appelle des classes. L'exemple le plus trivial de relation d'équivalence est celle de l'égalité, mais on en rencontre dans la vie de tous les jours, avec par exemple la relation « avoir la même couleur ». L'ensemble des classes forme une *partition* de l'ensemble (Proposition 0.6) ; on a en fait une correspondance complète entre relation d'équivalence et ensemble des classes d'équivalence : il revient au même de définir la relation « avoir la même couleur » ou de définir, pour chaque couleur, la classe des objets de cette couleur.

Les relations d'équivalences se retrouvent un peu partout en mathématiques parce qu'elles amènent à la notion d'*ensemble quotient*, défini comme étant l'ensemble des classes d'équivalence d'une relation donnée. Dans l'exemple qu'on vient de voir, l'ensemble quotient est l'ensemble des couleurs. En mathématiques, « passer au quotient » ou bien « quotienter » permet de faire comme si certains objets sont identiques. Par exemple, on peut définir l'ensemble $\mathbf{Q}$ des nombres rationnels comme étant l'ensemble des couples $(p, q) \in \mathbf{Z} \times \mathbf{N}^*$, quotienté par la relation $(p, q) \sim (p', q') \iff pq' = p'q$ (ce qui correspond au critère d'égalité $p/q = p'/q'$). Dans ce cours, nous définirons les espaces de Lebesgue $L^1(\mathbf{R})$ et $L^2(\mathbf{R})$ par passage au quotient par la relation d'égalité presque partout (Chapitre 3).

Définition 0.1

Une relation d'équivalence R sur un ensemble E est une relation binaire vérifiant :

1. reflexivité : pour tout $x \in E$, on a xRx ;
2. symétrie : si xRy, alors yRx ;
3. transitivité : si xRy et yRz, alors xRz.

Exemple 0.2

Les relations suivantes sont des relations d'équivalence.

1. Pour une application $f : X \to Y$, la relation sur X définie par $xRx' \iff f(x) = f(x')$.
2. La congruence modulo n sur $\mathbf{Z}$: $xRy \iff x \equiv y \mod n$.
3. La relation de parallélisme sur l'ensemble des droites du plan.

Définition 0.3

Soit R une relation d'équivalence sur un ensemble E. La *classe d'équivalence* de $x \in E$ est

$$C_x = \{y \in E \mid yRx\}.$$

L'*ensemble quotient* de R, noté E/R, est l'ensemble des classes d'équivalence de tous les éléments de E.

Notez bien que l'ensemble quotient est bien l'*ensemble* et non la liste des classes d'équivalence. Le mieux est d'illustrer ça sur un exemple.

Exemple 0.4

Pour tout ensemble E, il y a deux relations d'équivalence très simples.

— On vérifie facilement que la relation R définie par $xRy \iff x, y \in E$ (autrement dit tout le monde est équivalent à tout le monde) est une relation d'équivalence. Dans ce cas, la classe d'équivalence de tout $x \in E$ est $C_x = E$. Ainsi, l'ensemble quotient est $E/R = \{C_x \mid x \in E\} = \{E \mid x \in E\} = \{E\}$; en particulier il ne contient qu'un seul élément.

— De même, il est facile de voir que la relation R définie par $xRy \iff x = y$ est une relation d'équivalence. Ici, pour tout $x \in E$, on a $C_x = \{x\}$, et l'ensemble quotient est $E/R = \{C_x \mid x \in E\} = \{\{x\} \mid x \in E\}$, autrement dit l'ensemble des singletons de E.

Voyons ce que ça donne sur les exemples, moins triviaux, de 0.2.

Exemple 0.5

1. Les classes d'équivalence sont les ensembles préimages de singletons par f.
2. Les classes d'équivalence sont les $C_x = \{y \in \mathbf{Z} \mid x \equiv y \mod n\}$. On peut les indexer par $\{0, \dots, n-1\}$: par division euclidienne, tout nombre entier est congruent modulo n à un et un seul nombre entre 0 et $n-1$. Ainsi, l'ensemble quotient, noté $\mathbf{Z}/n\mathbf{Z}$, est de cardinal n.
3. On voit que deux droites sont parallèles si et seulement si elles font le même angle, modulo π, avec l'axe des abscisses (il faut bien voir que l'angle entre deux droites est bien défini modulo π et non 2π !). L'ensemble quotient peut donc s'écrire

$$\{E_\theta \mid 0 \leq \theta < \pi\} \quad \text{où} \quad E_\theta = \{D \text{ droite de } \mathbf{R}^2 \mid \widehat{D, Ox} = \theta\}.$$

Proposition 0.6

L'ensemble quotient E/R forme une partition de E.

Remarquons que cette propriété est bien vérifiée par les deux exemples de 0.4.

Démonstration : Pour montrer que l'ensemble quotient forme une partition de E, il faut montrer que tout élément $x \in E$ appartient à une et une seule classe d'équivalence.

Tout d'abord, pour tout $x \in E$, on a xRx (réflexivité) et donc $x \in C_x$. En particulier, x appartient à au moins une classe d'équivalence.

Ensuite, si on suppose que $y \in E$ appartient à la fois à C_{x_1} et à C_{x_2}. On veut montrer que $C_{x_1} = C_{x_2}$ (attention, ce n'est pas vrai en général que $x_1 = x_2$!). Soit donc $z \in C_{x_1}$. Alors on a zRx_1 mais aussi x_1Ry et x_2Ry. Par symétrie et transitivité, cela implique que zRy puis que zRx_2. Ainsi, $z \in C_{x_2}$ et donc $C_{x_1} \subset C_{x_2}$. Par symétrie des rôles joués par x_1 et x_2, on a aussi $C_{x_2} \subset C_{x_1}$, et donc $C_{x_1} = C_{x_2}$, ce qu'on voulait démontrer. ■

Remarque 0.7

On peut retrouver la relation R à partir de l'ensemble E/R en définissant la relation $x \sim y \iff x$ et y sont dans la même classe d'équivalence. Plus généralement, cette idée permet de construire naturellement une relation d'équivalence à partir d'une partition d'un ensemble : se donner une relation d'équivalence ou bien se donner une partition, ça revient au même.

Chapitre 1

L'intégrale de Lebesgue

Un défaut fâcheux de l'intégrale de Riemann est sa portée : elle ne s'applique qu'à un ensemble de fonctions « suffisamment régulières ». Cela se constate anecdotiquement sur des exemples spécifiques (comme la fonction $\mathbf{1}_{\mathbf{Q}}$ valant 1 sur les rationnels et 0 autrement) et beaucoup plus sérieusement lorsque l'on souhaite *passer à la limite* dans une intégrale : en général on ne sait pas donner un sens à l'intégrale d'une fonction obtenue comme limite simple. Cette question est réglée de manière spectaculaire par l'intégrale de Lebesgue puisque celle-ci est définie pour *pratiquement*[1] toutes les fonctions. C'est cet objet mathématique que nous nous proposons d'introduire dans ce chapitre.

La construction de l'intégrale de Lebesgue vient habituellement en deux morceaux : dans un premier temps on construit la mesure éponyme pour ensuite lui associer l'objet intégrale. Afin de produire un cours digérable en un semestre, nous adoptons ici un point de vue axiomatique où toute la première étape de la théorie est admise. Plus précisément, nous prenons pour acquise une liste d'axiomes concernant l'intégrale de Lebesgue (qui sont autant de propriétés normalement démontrées lors de sa construction) pour en déduire tous les résultats importants la concernant, mais aussi en dégager des intuitions concernant la mesure de Lebesgue.

Les axiomes seront présentés dans le cadre des fonctions positives (Section 1.1). À partir de ces axiomes, nous « devinerons » la définition de la mesure de Lebesgue et présenterons quelques une de ses propriétés (Section 1.2). Nous étendrons ensuite la définition de l'intégrale aux fonctions de signe quelconque, voire complexes (Section 1.3). L'intégration de Lebesgue *recouvre* celle de Riemann : nous établirons rapidement ce lien dans la Section 1.5.

1. Nous reviendrons sur la signification de ce terme.

Résumé :

- On admet l'existence d'une classe de fonctions — appelées fonctions *mesurables* — qui contient toutes les fonctions usuelles. C'est parmi cette classe qu'on peut définir une intégrale.
- L'*intégrale de Lebesgue* se définit d'abord sur les fonctions mesurables positives (Énoncé 1.5).
- À partir de la notion d'intégrale on déduit celle de *mesure de Lebesgue*.
- On ne peut pas définir l'intégrale de n'importe quelle fonction mesurable, mais celle des fonctions *intégrables*, qui sont celles dont l'intégrale du module est finie (Définition 1.31).
- L'intégrale d'une fonction intégrable réelle est définie à l'aide des parties positive et négative (Énoncé 1.39). Pour une fonction complexe on considère les parties réelle et imaginaire.
- L'*intégrale de Riemann* est un autre procédé d'intégration défini pour toute fonction continue par morceaux (Énoncé 1.62) ; les intégrales de Riemann et Lebesgue coïncident pour les fonctions continues par morceaux sur un segment de $\mathbf{R}$ (Énoncé 1.67).

1.1 Intégrale de fonctions mesurables positives

Pour toute la suite, on fixe un entier $d \geq 1$ qui sera la dimension de l'espace ambiant. Commençons par quelques définitions préliminaires.

Définition 1.1 (***Droite achevée***)

On note $\overline{\mathbf{R}}$ la droite réelle étendue par $\overline{\mathbf{R}} = \{-\infty\} \cup \mathbf{R} \cup \{+\infty\}$ et $\overline{\mathbf{R}}_+ = \{x \in \overline{\mathbf{R}}, x \geq 0\}$ l'ensemble des nombres réels positifs, auquel on ajoute $+\infty$.

On étend de manière évidente les opérations $+$, $\times$ et $\geq$ de $\mathbf{R}$ à $\overline{\mathbf{R}}$, lorsque c'est possible. L'indétermination $0 \times \pm\infty$ est définie par convention comme étant égale à 0, et l'indétermination $\infty - \infty$ est fixée comme étant égale à 0.

Définition 1.2

Pour tout ensemble A, on note $\mathbf{1}_A$ la *fonction indicatrice* de A, définie par

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A. \end{cases}$$

Définition 1.3

Un ensemble de fonctions définies sur $\mathbf{R}^d$ est dit *universellement stable* s'il l'est par limite simple, somme, produit, composition à droite par des applications affines et par division ou passage aux bornes supérieures/inférieures dénombrables (lorsque ces opérations sont définies !).

Exemples 1.4

L'ensemble de toutes les fonctions définies sur $\mathbf{R}^d$ à valeurs dans $\overline{\mathbf{R}}$ (resp. $\overline{\mathbf{R}}_+$) est universellement stable. L'ensemble des fonctions réelles continues ne l'est pas (pourquoi ?), celui des indicatrices non plus (bis ?).

Énoncé indispensable 1.5 (***Propriétés de l'intégrale***)

Il existe un ensemble universellement stable $\mathscr{M}$ de fonctions à valeurs dans $\overline{\mathbf{R}}$ dites mesurables *contenant $\mathscr{C}^0(\mathbf{R}^d)$ et $\{\mathbf{1}_O : O \text{ ouvert}\}$ et muni d'une* unique *application appelée* l'intégrale de Lebesgue *définie sur l'ensemble $\mathscr{M}^+$ fonctions mesurables sur $\mathbf{R}^d$ à valeurs dans $\overline{\mathbf{R}}_+$*

$$\begin{aligned} \mathscr{M}^+ &\longrightarrow \overline{\mathbf{R}}_+ \\ f &\longmapsto \int f, \end{aligned}$$

que l'on peut aussi noter

$$\int_{\mathbf{R}^d} f, \text{ ou } \int_{\mathbf{R}^d} f(x)\,\mathrm{d}x,$$

obéissant aux propriétés suivantes.

1. ***Linéarité positive*** *: pour f et g mesurables positives, et pour $\lambda, \mu \in \mathbf{R}_+$, on a*
$$\int (\mu_1 f + \mu_2 g) = \mu_1 \int f + \mu_2 \int g.$$
2. ***Normalisation*** *: pour $a_1 \le b_1, \cdots, a_d \le b_d \in \mathbf{R}$, on a*
$$\int \mathbf{1}_{]a_1,b_1[\times\cdots\times]a_d,b_d[} = (b_1 - a_1) \times \cdots \times (b_d - a_d).$$
3. ***Le Théorème de Beppo-Levi ou de convergence croissante*** *: si $(f_n)_{n\in\mathbf{N}}$ est une suite croissante de fonctions mesurables positives, alors*
$$\int \lim_{n\to+\infty} f_n = \lim_{n\to+\infty} \int f_n.$$

Cette application vérifie de plus une propriété d'approximation : pour toute fonction f mesurable positive telle que $\int f < +\infty$ et tout $\varepsilon > 0$, il existe $\varphi \in \mathscr{C}^0(\mathbf{R}^d; \mathbf{R}_+)$ nulle en dehors d'un compact, telle que
$$\int |f - \varphi| < \varepsilon.$$

Remarques 1.6

- La stabilité universelle est une propriété très forte. Dans notre cadre, où l'on travaille sur $\mathbf{R}^d$ avec l'intégrale de Lebesgue, on peut sans trop de risque considérer que *toutes* les fonctions sont mesurables. En réalité, l'existence de tels objets est une question liée aux fondements de la théorie des ensembles (voir l'Annexe A.6).
- Attention ! La remarque précédente semble suggérer que la mesurabilité est une notion sans grand intérêt. Il ne s'agit ici que d'un cas (très spécifique) d'un concept omniprésent en théorie des probabilités et dans laquelle la mesurabilité de telle ou telle fonction peut à la fois être un résultat difficile à établir et un outil de modélisation très puissant. Mais cela concerne d'autres espaces que $\mathbf{R}^d$ et (donc) des intégrales et mesures bien différentes de celles de Lebesgue.
- L'ensemble $\mathscr{M}$ des fonctions mesurables n'est pas défini par cet énoncé. En pratique, on aura juste besoin de ses propriétés : il contient un ensemble de fonctions raisonnables (les fonctions continues et les indicatrices d'ouverts) et est stable par toutes les opérations courantes sur les fonctions. Ces propriétés permettent de démontrer que toutes les fonctions qu'on aura à étudier sont mesurables. Le lecteur désirant avoir une définition précise de cet ensemble $\mathscr{M}$ pourra se référer à l'Annexe A.1.
- Les fonctions constantes sont continues, et à ce titre sont dans l'ensemble $\mathscr{M}$. Celui-ci étant stable par produit, il est aussi stable par multiplication par un scalaire : multiplier par $\mu \in \mathbf{R}$, c'est la même chose que multiplier par la fonction constante égale à μ

- La propriété de normalisation exprime que le volume d'un parallélépipède rectangle est égal au produit des longueurs de ses arêtes.
- Notons que $|f - \varphi| = \max(f - \varphi, 0) + \max(\varphi - f, 0)$ est bien mesurable positive. L'intérêt de la propriété d'approximation se mesure à l'aune de la taille considérable de $\mathscr{M}$: f peut être incroyablement biscornue, on arrivera toujours à approcher le volume qu'elle délimite par une fonction continue.
- Le Théorème de Beppo-Levi est la grande nouveauté de l'intégrale de Lebesgue. Il est parfois appelé *théorème de convergence monotone*, seulement il devient faux si on suppose la suite de fonctions décroissante… Gare à cette terminologie !

La linéarité et la positivité de l'intégrale permettent d'en déduire sa croissance.

Proposition 1.7 (*Croissance de l'intégrale*)

Si $f \leq g$ sont deux fonctions mesurables positives, alors $\int f \leq \int g$.

Démonstration : Il suffit de poser $h = g - f$. Alors h est une fonction mesurable positive, par stabilité universelle. Donc $\int h$ existe par l'Énoncé 1.5 et est positif, et par linéarité $\int g = \int f + \int h \geq \int f$. ■

Cette proposition assure que la limite de droite dans l'énoncé du Théorème de Beppo-Levi existe bien, en tant que limite (dans $\overline{\mathbf{R}}_+$) d'une suite croissante de nombres positifs. Remarquons que la fonction intégrée à gauche de l'égalité (à savoir $\lim f_n$) existe bien elle aussi, en tant que limite croissante de fonctions positives.

On peut facilement adapter le Théorème de Beppo-Levi aux séries de fonctions.

Corollaire 1.8

Si $(g_k)_{k \in \mathbf{N}}$ est une suite de fonctions mesurables positives, alors

$$\int \sum_{k=0}^{+\infty} g_k = \sum_{k=0}^{+\infty} \int g_k.$$

Démonstration : Il suffit d'appliquer le théorème de Beppo-Levi à la suite de fonctions $f_n = \sum_{k=0}^{n} g_k$. ■

Cette version permet (déjà) de calculer certaines intégrales.

Exemple 1.9

$$\int_{\mathbf{R}} \mathbf{1}_{\mathbf{Q}} = 0.$$

En effet, on sait que $\mathbf{Q}$ est dénombrable (exercice : pourquoi ?), on peut donc écrire $\mathbf{Q} = \{x_0, x_1, x_2, \ldots\}$. Ainsi, $\mathbf{1}_{\mathbf{Q}} = \sum_{k=0}^{+\infty} \mathbf{1}_{\{x_k\}}$. Or par la propriété de normalisation on a $\int \mathbf{1}_{\{x_k\}} = \int \mathbf{1}_{[x_k, x_k]} = x_k - x_k = 0$. Par conséquent, par le corollaire précédent,

$$\int_{\mathbf{R}} \mathbf{1}_{\mathbf{Q}} = \sum_{k=0}^{+\infty} \int_{\mathbf{R}} \mathbf{1}_{\{x_k\}} = \sum_{k=0}^{+\infty} 0 = 0.$$

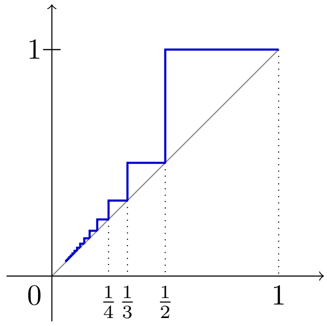

FIGURE 1.1 – La fonction f de l'exemple 1.10.

Exemple 1.10

Soit la fonction $f(x) = \mathbf{1}_{]0,1](x)} \frac{1}{\lfloor 1/x \rfloor}$ (voir la Figure 1.1). Alors

$$\int f = \frac{\pi^2}{6} - 1.$$

En effet, pour tout $n \geq 1$, on a $f(x) = 1/n$ si $1/(n+1) < x \leq 1/n$. Ainsi, $f = \sum_{n \geq 1} f_n$, avec $f_n = \frac{1}{n} \mathbf{1}_{]1/(n+1),1/n]}$. Ainsi, par le corollaire précédent, on obtient

$$\begin{aligned}
\int f &= \int \sum_{n=1}^{+\infty} f_n = \sum_{n=1}^{+\infty} \int f_n \\
&= \sum_{n=1}^{+\infty} \frac{1}{n} \left(\frac{1}{n} - \frac{1}{n+1} \right) = \sum_{n=1}^{+\infty} \frac{1}{n^2} - \sum_{n=1}^{+\infty} \frac{1}{n(n+1)} \\
&= \frac{\pi^2}{6} - \sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+1} \right) = \frac{\pi^2}{6} - 1,
\end{aligned}$$

où la dernière égalité s'obtient par téléscopage.

On a pour l'instant très peu de moyens pour *calculer* explicitement des intégrales. On verra au chapitre suivant le lien entre intégrale et primitive, qui permet de calculer une bonne quantité d'intégrales, surtout lorsqu'on lui adjoint les outils — qui en sont des conséquences — d'intégration par parties et de changement de variables.

1.2 Ensembles mesurables et mesure de Lebesgue

Dans cette section, nous déduisons de la définition que nous avons vue de l'intégrale de Lebesgue les propriétés de la mesure de Lebesgue. C'est un objet qui permet de mesurer la taille des sous-ensembles de $\mathbf{R}^d$, et qui coïncide avec les notions intuitives de longueur ($d = 1$), aire ($d = 2$), volume ($d = 3$), pour de nombreux ensembles simples (segments, disques, polygones...). De la même manière que pour ce qui concerne l'intégrale où l'on se borne à étudier des fonctions mesurables, il faut se restreindre à une sous-classe de parties $\mathbf{R}^d$ pour lesquelles la mesure est bien définie.

Définition 1.11

On dit qu'un sous-ensemble A de $\mathbf{R}^d$ est *mesurable* si sa fonction indicatrice $\mathbf{1}_A$ est mesurable.

L'ensemble des sous-ensembles mesurables de $\mathbf{R}^d$ sera noté $\mathcal{L}(\mathbf{R}^d)$.

Proposition 1.12

L'ensemble $\mathcal{L}(\mathbf{R}^d)$ des mesurables de $\mathbf{R}^d$ est :

1. *non vide ;*
2. *stable par complémentaire : si $A \in L(\mathbf{R}^d)$, alors $A^\complement \in L(\mathbf{R}^d)$;*
3. *stable par union dénombrable : si $(A_n)_{n\in\mathbf{N}}$ est une suite d'éléments de $\mathcal{L}(\mathbf{R}^d)$, alors $\bigcup_{n\in\mathbf{N}} A_n \in \mathcal{L}(\mathbf{R}^d)$.*

Il contient de plus tous les ouverts et tous les fermés de $\mathbf{R}^d$.

Démonstration : L'ensemble $\mathcal{L}(\mathbf{R}^d)$ est non vide car il contient tous les pavés (Énoncé 1.5). Il est stable par passage au complémentaire : si $A \in \mathcal{L}(\mathbf{R}^d)$, alors la fonction 1_A est mesurable, et donc la fonction $\mathbf{1}_{A^\complement} = 1 - \mathbf{1}_A$ l'est aussi, donc $A^\complement \in \mathcal{L}(\mathbf{R}^d)$.

Si $(A_n)_{n\in\mathbf{N}} \in \mathcal{A}$, alors $\mathbf{1}_{\cup_n A_n} = \sup_n \mathbf{1}_{A_n}$ est une fonction mesurable, et donc $\cup_n A_n \in \mathcal{L}(\mathbf{R}^d)$. Enfin, l'Énoncé 1.5 assure que tous les ensembles ouverts sont mesurables, et comme on vient de voir que $\mathcal{L}(\mathbf{R}^d)$ est stable par passage au complémentaire, cela implique aussi que tout ensemble fermé est mesurable. ∎

De la preuve de cette proposition on retiendra les correspondances suivantes entre les opérations sur les ensembles et celles sur leurs fonctions indicatrices :

Ensembles	Fonctions indicatrices
A mesurable	$\mathbf{1}_A$ mesurable
$A^\complement$	$\mathbf{1}_{A^\complement} = 1 - \mathbf{1}_A$
$A \cap B$	$\mathbf{1}_{A\cap B} = \mathbf{1}_A \mathbf{1}_B = \inf(\mathbf{1}_A, \mathbf{1}_B)$
$A \cup B$	$\mathbf{1}_{A\cup B} = \mathbf{1}_A + \mathbf{1}_B - \mathbf{1}_A\mathbf{1}_B = \sup(\mathbf{1}_A, \mathbf{1}_B)$
$A \subset B$	$\mathbf{1}_A \leq \mathbf{1}_B$

Exercice 1.13

Montrer que si $f : \mathbf{R}^d \to \mathbf{R}$ est mesurable, alors pour tout $a \in \mathbf{R}$, les ensembles $\{f \leq a\}$ et $\{f < a\}$ le sont également.

La proposition précédente exprime en fait que l'ensemble des mesurables de $\mathbf{R}^d$ forme *une tribu*, selon la définition suivante.

Définition 1.14 (*Tribus*)

Soit X un ensemble. Une *tribu*, ou *σ-algèbre*[2], sur X est un ensemble $\mathcal{A}$ de parties[3] de X qui vérifie les propriétés 1. à 3. de la Proposition 1.12. Si $(X, \mathcal{A})$ est $(Y, \mathcal{B})$ sont des ensembles munis de σ-algèbres, on dit qu'une fonction $f : X \to Y$ est mesurable si

$A \in \mathcal{B}$ implique $f^{-1}(A) \in \mathcal{A}$.

Maintenant que nous avons défini les ensembles mesurables, nous pouvons passer à la définition de la mesure de Lebesgue.

Énoncé indispensable 1.15 (*La mesure de Lebesgue*)

La mesure de Lebesgue sur $\mathbf{R}^d$ est l'application — notée ici λ — définie sur l'ensemble $\mathcal{L}(\mathbf{R}^d)$ des mesurables sur $\mathbf{R}^d$ par

$$\forall A \in \mathcal{A}, \quad \lambda(A) = \int_{\mathbf{R}^d} \mathbf{1}_A.$$

Elle vérifie les propriétés suivantes :

1. *l'ensemble vide est mesurable et de mesure nulle : $\lambda(\varnothing) = 0$;*
2. *si $(A_n)_{n\in\mathbf{N}}$ est une famille dénombrable d'ensembles mesurables, deux à deux disjoints[4], alors*

$$\lambda\left(\bigcup_{n\in\mathbf{N}} A_n\right) = \sum_{n\in\mathbf{N}} \lambda(A_n).$$

Remarques 1.16

- La mesure de Lebesgue dépend de la dimension : en toute rigueur il faudrait la noter λ_d ;
- C'est ce λ qui intervient dans une notation que l'on rencontre parfois pour l'intégrale de Lebesgue

$$\int f \mathrm{d}\lambda(x),$$

il indique alors que l'intégrale est prise par rapport à la mesure de Lebesgue.
- Voir l'Annexe A.1 pour une autre définition de la mesure de Lebesgue, sans l'aide de la construction de l'intégrale de Lebesgue.

Démonstration :

1. On a bien

$$\lambda(\varnothing) = \int_{\mathbf{R}^d} \mathbf{1}_\varnothing = \int_{\mathbf{R}^d} 0 = 0.$$

2. Si $(A_n)_{n\in\mathbf{N}}$ une famille dénombrable d'ensembles mesurables, deux à deux disjoints, alors $\mathbf{1}_{\cup_{n\in\mathbf{N}} A_n} = \sum_{n\in\mathbf{N}} \mathbf{1}_{A_n}$, si bien que, par le Corollaire 1.8,

$$\lambda\left(\bigcup_{n\in\mathbf{N}} A_n\right) = \int_{\mathbf{R}^d} \mathbf{1}_{\cup_{n\in\mathbf{N}} A_n} = \int_{\mathbf{R}^d} \sum_{n\in\mathbf{N}} \mathbf{1}_{A_n} = \sum_{n\in\mathbf{N}} \int_{\mathbf{R}^d} \mathbf{1}_{A_n} = \sum_{n\in\mathbf{N}} \lambda(A_n).$$

■

2. Lire "sigma-algèbre".
3. Ce qui signifie que $\mathcal{A}$ est une collection de sous-ensembles de X.
4. C'est-à-dire que si $n \neq m$, alors $A_n \cap A_m = \varnothing$.

Exercice 1.17

Que vaut la mesure de Lebesgue de $\mathbf{R}$ *? Montrer que dans* $\mathbf{R}^2$*, la mesure de Lebesgue de l'ensemble* $\mathbf{R} \times \{0\}$ *est nulle.*

Les propriétés de l'Énoncé 1.15 indiquent que la mesure de Lebesgue est un objet raisonnable pour mesurer des longueurs, des aires, des volumes (suivant la valeur qu'on prend pour d)... Le fait d'isoler ces propriétés qui font d'une application un outil pertinent de mesure amène à la définition abstraite suivante.

Définition 1.18

Soit $\mathcal{A}$ une tribu sur un ensemble X. Une *mesure* sur $\mathcal{A}$ est une application $\mu : \mathcal{A} \to \overline{\mathbf{R}}_+$ vérifiant les deux propriétés de l'Énoncé 1.15. Dans ce cas, on appelle le triplet $(X, \mathcal{A}, \mu)$ un *espace mesuré*. Dans le cas où $\mu(X) = 1$, on parle d'*espace probabilisé* et de *mesure de probabilité*.

Exercice 1.19

Montrer que sur tout ensemble X*, l'ensemble* $P(X)$ *des parties de* X *est une tribu, et que l'application* $P(X) \ni A \mapsto \operatorname{card}(A)$ *est une mesure, appelée* mesure de comptage.

Voici quelques propriétés simples vérifiées par les mesures, et donc en particulier la mesure de Lebesgue.

Proposition 1.20

Si $(X, \mathcal{A}, \mu)$ *est un espace mesuré, alors*

1. *si* $A, B \in \mathcal{A}$*, alors* $\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$ *;*
2. *si* $(A_n)_{n\in\mathbf{N}} \in \mathcal{A}^{\mathbf{N}}$*, alors*
$$\mu\left(\bigcup_{n\in\mathbf{N}} A_n\right) \leq \sum_{n\in\mathbf{N}} \mu(A_n).$$
3. *si* $A \subset B \in \mathcal{A}$*, alors* $\mu(B \setminus A) + \mu(A) = \mu(B)$*, en particulier* $\mu(A) \leq \mu(B)$ *(c'est la* croissance *de la mesure) ;*
4. *pour toute famille* $(B_n)_{n\in\mathbf{N}}$ *d'ensemble mesurables, croissante pour l'inclusion (i.e. pour tout* n*, on a* $B_n \subset B_{n+1}$*), on a*
$$\mu\left(\bigcup_{n\in\mathbf{N}} B_n\right) = \lim_{n\to+\infty} \mu(B_n).$$

Démonstration : 1. On commence par remarquer qu'en prenant $A_n = \varnothing$ à partir d'un certain rang, la seconde propriété de la définition d'une mesure est aussi vraie pour une famille *finie* de mesurables deux à deux disjoints. On pose alors $A_0 = A \setminus B$, $A_1 = B \setminus A$ et $A_2 = A \cap B$. Alors (A_0, A_1, A_2) est une famille finie de mesurables, deux à deux disjoints, et de plus $A \cup B = A_0 \cup A_1 \cup A_2$. La seconde propriété de la définition de mesure implique alors que

$$\mu(A \cup B) = \mu(A_0) + \mu(A_1) + \mu(A_2). \tag{1.1}$$

Or, $A_0 \cup A_2 = A$ et $A_1 \cup A_2 = B$. Ainsi, appliquant de nouveau la seconde propriété,

$$\mu(A) = \mu(A_0) + \mu(A_2) \quad \text{et} \quad \mu(B) = \mu(A_1) + \mu(A_2) \tag{1.2}$$

En additionnant $\mu(A_2)$ à (1.1) et en utilisant (1.2), on obtient

$$\mu(A \cup B) + \mu(A_2) = \mu(A) + \mu(B),$$

ce qu'il fallait démontrer.

2. Pour tout $n \geq 0$, on pose

$$B_n = A_n \setminus \bigcup_{i<n} A_i.$$

Alors les ensembles B_n sont deux-à-deux disjoints. En effet, si on suppose qu'il existe $x \in B_n \cap B_m$, avec $n < m$, alors $x \in B_n \subset A_n$, mais puisque $x \in B_m$, $x \notin \bigcup_{i<m} A_i$ et en particulier $x \notin A_n$, ce qui est une contradiction.
De plus, $\bigcup_{n\in\mathbf{N}} B_n = \bigcup_{n\in\mathbf{N}} A_n$. En effet, une inclusion découle immédiatement du fait que $B_n \subset A_n$; d'autre part, si $x \in \bigcup_{n\in\mathbf{N}} A_n$, alors il existe $n_0 \in \mathbf{N}$ tel que $x \in A_n$ mais $x \notin A_i$ pour tout $i < n_0$ (on considère le premier indice n tel que $x \in A_n$), et donc $x \in B_{n_0}$.
Ainsi, en utilisant la propriété 2. d'une mesure,

$$\mu\left(\bigcup_{n\in\mathbf{N}} A_n\right) = \mu\left(\bigcup_{n\in\mathbf{N}} B_n\right) = \sum_{n\in\mathbf{N}} \mu\left(B_n\right) \leq \sum_{n\in\mathbf{N}} \mu\left(A_n\right)$$

(où la dernière inégalité découle de l'inclusion $B_n \subset A_n$).

3. On procède de même que pour le point 1. en notant que B est alors l'union disjointe de A et de $B \setminus A$; la croissance se déduit de la positivité de la mesure.

4. Commençons par écarter le cas trivial où l'un des ensembles B_n est de mesure infinie. En effet, dans ce cas (en utilisant la croissance démontrée dans le point précédent), tous les ensembles suivants sont également de mesure infinie et l'union des $(B_n)_n$ également : la formule est vérifiée. Autrement, si $\mu(B_n) < \infty$ pour tout $n \in \mathbf{N}$, posons $A_n = B_n \setminus B_{n-1}$ (avec la convention $B_{-1} = \varnothing$). Alors $(A_n)_{n\in\mathbf{N}}$ est une famille dénombrable d'ensembles mesurables deux à deux disjoints ; on peut donc appliquer la seconde propriété des mesures :

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

Or

$$\bigcup_{n\in\mathbf{N}} A_n = \bigcup_{n\in\mathbf{N}} \left(B_{n+1} \setminus B_n\right) = \bigcup_{n\in\mathbf{N}} B_n,$$

et (remarquant qu'on a une somme télescopique)

$$\begin{aligned}\sum_{n\in\mathbf{N}} \mu(A_n) &= \lim_{N\to+\infty} \sum_{n=0}^{N} \mu\bigl(B_n \setminus B_{n-1}\bigr) = \lim_{N\to+\infty} \sum_{n=0}^{N} \Bigl(\mu(B_n) - \mu(B_{n-1})\Bigr) \\ &= \lim_{N\to+\infty} \Bigl(\mu(B_N) - \mu(B_{-1})\Bigr) = \lim_{n\to+\infty} \mu(B_n).\end{aligned}$$

■

Exercice 1.21

Soit $(B_n)_n$ une suite décroissante *(pour l'inclusion) d'ensembles mesurables d'un espace mesuré $(X, \mathcal{A}, \mu)$. Montrer qu'en général $\mu(\cap_n B_n) \neq \lim_n \mu(B_n)$; fournir cependant une condition suffisante qui assure l'égalité.*

Passons à une propriété spécifique à la mesure de Lebesgue : par la propriété de normalisation de l'intégrale (Énoncé indispensable 1.5), la mesure de Lebesgue d'un pavé est égale à ce qu'on définit usuellement comme son volume :

$$\lambda\Big([a_1, b_1] \times \cdots \times [a_d, b_d]\Big) = (b_1 - a_1) \times \cdots \times (b_d - a_d).$$

On a en fait un peu mieux que ça : la mesure de Lebesgue est invariante par découpage, déplacement et recollement. Autrement dit, si on a un ensemble mesurable A qu'on découpe en une quantité *dénombrable* de morceaux mesurables, et si on déplace chacun de ces morceaux à l'aide d'une isométrie, alors l'ensemble obtenu aura la même mesure que l'ensemble de départ. Cette propriété se déduit de la proposition suivante, qui est une version linéaire du théorème de changement de variables (Théorème 2.36).

Proposition 1.22 (*Lebesgue et applications affines*)

Si $A \subset \mathbf{R}^d$ est un ensemble mesurable, et $L : \mathbf{R}^d \to \mathbf{R}^d$ est une application affine de partie linéaire $M \in \mathrm{M}_d(\mathbf{R})$ i.e. *$L(x) = L(0) + Mx$. Alors*

$$\lambda(L(A)) = |\det(M)|\lambda(A).$$

Exercice 1.23

Le but de l'exercice est de démontrer la proposition précédente.

1. *En utilisant l'unicité de l'intégrale de Lebesgue (Énoncé indispensable 1.5), montrer qu'il suffit de démontrer cette égalité dans le cas où A est un pavé.*
2. *Montrer que la mesure de Lebesgue est invariante par translation : pour tout $u \in \mathbf{R}^d$ et tout mesurable $A \subset \mathbf{R}^d$, $\lambda(A+u) = \lambda(A)$. On suppose donc dorénavant $L(0) = 0$.*
3. *Montrer qu'il suffit de prouver l'identité $\lambda(MC) = |\det M|$ où $C := [0, 1]^d$ est l'hypercube unité.*
4. *Montrer que toute matrice de M s'écrit comme le produit d'une matrice diagonale avec des matrices de transposition $\mathrm{I}_d + \alpha \mathrm{E}_{ij}$ $(i \neq j)$ et conclure.*

Exemple 1.24

Si on déforme la boule unité de $\mathbf{R}^3$ par une opération linéaire diagonale valeurs propres $\lambda_1, \lambda_2, \lambda_3$ le volume de l'ellipsoïde obtenu est $|\lambda_1\lambda_2\lambda_3| \times \frac{4}{3}\pi$.

Terminons cette section par quelques considérations autour de la notion *d'ensemble négligeable.*

Définition 1.25

On dit qu'une partie mesurable de $\mathbf{R}^d$ est *négligeable* si sa mesure de Lebesgue est nulle (on peut aussi dire « de mesure nulle »).

Une propriété est vraie *presque partout* (en abrégé « p.p. ») si elle est satisfaite en dehors d'un ensemble négligeable.

Exemple 1.26

Un sous-ensemble dénombrable de $\mathbf{R}^d$ est toujours négligeable (car tout singleton est négligeable, et toute union dénombrable d'ensemble négligeables l'est aussi).

Remarque 1.27

La tribu de Borel mentionnée à la Remarque A.1 présente un défaut : il existe des ensembles n'appartenant pas à cette tribu et toutefois inclus dans un ensemble de mesure nulle (on aurait sacrément envie de pouvoir les négliger !). La *tribu de Lebesgue* $\mathcal{L}(\mathbf{R}^n) \supset \mathcal{B}(\mathbf{R}^n)$ consiste grossièrement à ajouter ces ensembles manquants à la tribu de Borel. Un ensemble Lebesgue mesurable est ainsi défini comme une réunion $B \cup N$ avec B un Borelien et N un ensemble inclus dans un Borelien de mesure de Borel-Lebesgue nulle. La collection d'ensembles alors obtenue (et les fonctions indicatrices associées) correspond aux parties mesurables que l'on utilise dans ce cours. Les fonctions $f : \mathbf{R}^d \to \overline{\mathbf{R}}$ qui sont mesurables pour les tribus de Lebesgue sont les fonctions mesurables utilisées dans ce cours. Dans ce cadre spécifique, l'obtention d'un ensemble non mesurable ne peut se faire que par un acte volontaire qui remonte aux fondements de la théorie des ensembles (voir par exemple l'annexe A.6). Comme nous l'avons déjà souligné dans les Remarques 1.6, c'est le prisme euclidien par lequel nous regardons la notion de mesurabilité qui la vide de son essence. La théorie des probabilités fournit de nombreux exemples d'ensembles dont il est naturel de questionner la mesurabilité, sans avoir recours à un axiome supplémentaire de théorie des ensembles.

Remarque 1.28

Les terminologies « négligeables » et « presque partout » sont employées pour d'autres mesures que la mesure de Lebesgue. Une petite exception demeure dans le cas particulier d'une mesure de probabilité, où l'on utilise traditionnellement l'expression « *presque sûrement* » à la place de « presque partout ».

Exemple 1.29

Étant données deux fonctions $f, g : \mathbf{R}^d \to \mathbf{C}$, si $\lambda\{x \in \mathbf{R}^d \mid f(x) \neq g(x)\} = 0$, on dira que « $f = g$ presque partout ». Ainsi, $\mathbf{1}_{\mathbf{Q}} = 0$ presque partout (eh oui !).

Il est assez clair (par axiome de linéarité) que la fonction nulle a une intégrale nulle. Pour une fonction positive, l'inverse est « presque » vrai : c'est l'objet de la proposition suivante.

Proposition 1.30

Une fonction mesurable positive est d'intégrale nulle si et seulement si elle est nulle *presque partout.*

Démonstration : Soit $f : \mathbf{R}^d \to \overline{\mathbf{R}}_+$ mesurable, et supposons que $A := \{x \in \mathbf{R}^d \ : \ f(x) > 0\}$ est de mesure nulle. Posons $g = \infty \mathbf{1}_A \geq f$, et $(g_n)_n := (n\mathbf{1}_A)_n$. Puisque g_n converge simplement en croissant vers g, le Théorème de Beppo-Levi assure que

$$\int g_n \underset{n\to+\infty}{\longrightarrow} \int g.$$

Mais par définition de g_n on a $\int g_n = n\lambda(A) = 0$, on en déduit la nullité de $\int g$, puis celle de $\int f$ qui découle des inégalités $0 \leq \int f \leq \int g$ (par la croissance de l'intégrale, Proposition 1.7).

Réciproquement si $g : \mathbf{R}^d \to \overline{\mathbf{R}}_+$ positive mesurable est d'intégrale nulle, posons

$$A_n = \big\{x \in \mathbf{R}^d \mid g(x) \geq 1/n\big\}.$$

Alors la suite d'ensembles mesurables $(A_n)_{n>0}$ est croissante, et son union est

$$A = \bigcup_{n>0} A_n = \big\{x \in \mathbf{R}^d \mid g(x) > 0\big\}.$$

Le but est donc de montrer que $\lambda(A) = 0$. On peut appliquer le point 3) de la Proposition 1.20 : la suite $\big(\lambda(A_n)\big)_n$ est croissante, de limite $\lambda(A)$. Or, puisque $\mathbf{1}_{A_n} \leq ng$, par croissance de l'intégrale,

$$\lambda(A_n) = \int \mathbf{1}_{A_n} \leq n \int g = 0.$$

Ainsi, $\lambda(A_n) = 0$, et donc $\lambda(A) = \lim_n \lambda(A_n) = 0$. ■

1.3 Extension de l'intégrale aux fonctions complexes

Maintenant que l'on a à notre disposition l'intégrale des fonctions positives, on peut s'attaquer à la définition de l'intégrale de fonctions à valeurs réelles ou complexes.

Malheureusement, on ne peut pas espérer intégrer toutes les fonctions : quel sens pourrait-on donner à $\int_{\mathbf{R}} \cos$?

La solution, dans le cas de l'intégration de Lebesgue, consiste à ne considérer que les fonctions dont le module (ou la valeur absolue dans le cas réel) est intégrable.

Définition 1.31

On dit qu'une fonction mesurable $f : \mathbf{R}^d \to \mathbf{C}$ est *intégrable* (ou sommable) si

$$\int_{\mathbf{R}^d} |f| < +\infty.$$

On note $\mathscr{L}^1(\mathbf{R}^d)$ l'ensemble des fonctions $f : \mathbf{R}^d \to \mathbf{C}$ intégrables.

Remarque 1.32

On utilise le même terme pour parler des fonctions mesurables $f : \mathbf{R}^d \to \overline{\mathbf{R}}_+$ d'intégrale finie.

Attention, la terminologie est un peu trompeuse : les fonctions mesurables positives ont une intégrale bien définie (par l'Énoncé indispensable 1.5) mais ne sont pas toujours "intégrables" car leur intégrale peut être infinie.

Ainsi, pour être intégrable, une fonction doit être mesurable. Par ailleurs, comme nous l'avons déjà remarqué, si f est à valeurs réelles, la fonction $|f| = \max(f, 0) + \max(-f, 0)$ est bien mesurable positive : on peut parler de son intégrale. Si f est à valeurs complexes, la fonction $|f|$ est mesurable comme composée de fonctions mesurables, car $|\cdot| : \mathbf{C} \to \mathbf{R}_+$ est continue, donc mesurable.

Exercice 1.33

Montrer que si $f : \mathbf{R}^d \to \overline{\mathbf{R}}_+$ est d'intégrale finie, alors elle est finie presque partout. La réciproque est-elle vraie ?

Dans la pratique, si on veut montrer qu'une fonction est intégrable, on utilisera la croissance de l'intégrale (Proposition 1.7) en majorant la valeur absolue de la fonction par une autre fonction qui elle est positive et intégrable.

Exemple 1.34

Attention, une fonction intégrable sur $\mathbf{R}$ ne tend pas forcément vers 0 en $\pm\infty$. Par exemple, la fonction

$$\sum_{n\geq 0} \mathbf{1}_{[n,n+2^{-n}]}$$

est mesurable positive et d'intégrale $\sum_{n\geq 0} 2^{-n} = 2$, mais ne tend pas vers 0 en $+\infty$.

Exemple 1.35

Si $f : \mathbf{R}^d \to \mathbf{C}$ est mesurable, on peut approcher sa valeur absolue $|f|$ par la suite croissante de fonctions mesurables positives $f_n = \mathbf{1}_{[-n,n]^d}|f|$. Le théorème de Beppo-Levi permet donc de montrer que f est intégrable si et seulement si la limite

$$\int |f| = \lim_n \int f_n$$

est finie. Par exemple, la fonction constante $1 = \mathbf{1}_{\mathbf{R}^d}$ n'est pas intégrable car dans ce cas, cette limite vaut

$$\int \mathbf{1}_{\mathbf{R}^d} = \lim_n (2n)^d = \infty.$$

Exemple 1.36

La fonction $f(x) = \sin(x)$ n'est pas intégrable sur $\mathbf{R}$. En effet, $|\sin|$ est π-périodique et on voit sur son graphe que son intégrale (aire sous le graphe) est une somme infinie de termes non nuls tous égaux, donc qu'elle est infinie. Donnons maintenant une démonstration plus précise de cet énoncé intuitif. La fonction $|\sin|$ est continue, strictement croissante sur $[0, \pi/2]$ et strictement décroissante sur $[\pi/2, \pi]$ et son graphe est symétrique par rapport à $\pi/2$ sur $[0, \pi]$. On peut donc la minorer, sur $[0, \pi]$ par la fonction $g = \sin(\pi/4)\mathbf{1}_{[\pi/4,3\pi/4]}$, qui est non nulle au voisinage de $\pi/2$ (dessiner le petit rectangle correspondant sous le graphe de $|\sin|$). Ainsi, par Beppo-Levi et périodicité, on a

$$\int |f| \geq \sum_{n\in\mathbf{Z}} \int_0^\pi g = \left(\int_0^\pi g\right) \sum_{n\in\mathbf{Z}} 1 = \infty$$

car la normalisation et la linéarité positive donnent

$$\int_0^\pi g = \sin(\pi/4)(3\pi/4 - \pi/4) = \sin(\pi/4)\pi/2 > 0.$$

La définition de l'intégrale d'une fonction intégrable repose sur la notion de partie positive/négative.

Définition 1.37

Soit $f : \mathbf{R}^d \to \mathbf{R}$. La *partie positive* (respectivement *partie négative*) de f est la fonction

$$f_+ = \max(f, 0) \quad \text{respectivement} \quad f_- = \max(-f, 0).$$

Exercice 1.38

Démontrer les égalités

$$f_+ = \frac{|f| + f}{2}, \qquad f_- = \frac{|f| - f}{2} \qquad \text{et} \qquad f = f_+ - f_-.$$

Ces propriétés permettent d'exprimer la fonction f comme la différence de deux fonctions positives. De plus, chacune de ces fonctions est intégrable. En effet, on a $f_+ \leq |f|$ et $f_- \leq |f|$, et donc

$$\int_{\mathbf{R}^d} f_+ \leq \int_{\mathbf{R}^d} |f| < \infty \qquad \text{et} \qquad \int_{\mathbf{R}^d} f_- \leq \int_{\mathbf{R}^d} |f| < \infty.$$

Pour les fonctions à valeurs réelles f, ces remarques permettent de définir l'intégrale de f à partir de celles des fonctions positives f_+ et f_-, dont l'intégrale a été définie au paragraphe précédent. Pour les fonctions à valeurs complexes f, il suffit de savoir intégrer les fonctions $\operatorname{Re} f$ et $\operatorname{Im} f$, mais celles-ci sont intégrables par croissance de l'intégrale, grâce aux inégalités

$$\int_{\mathbf{R}^d} |\operatorname{Re} f| \leq \int_{\mathbf{R}^d} |f| \qquad \text{et} \qquad \int_{\mathbf{R}^d} |\operatorname{Im} f| \leq \int_{\mathbf{R}^d} |f|.$$

Énoncé indispensable 1.39 (*Définition de l'intégrale*)

Soit $f : \mathbf{R}^d \to \mathbf{R}$ une fonction mesurable, intégrable et à valeurs réelles. L'intégrale de la fonction f est définie par

$$\int_{\mathbf{R}^d} f = \int_{\mathbf{R}^d} f_+ - \int_{\mathbf{R}^d} f_-.$$

L'intégrale d'une fonction $f : \mathbf{R}^d \to \mathbf{C}$ mesurable, intégrable et à valeurs complexes *est définie par*

$$\int_{\mathbf{R}^d} f = \int_{\mathbf{R}^d} \operatorname{Re} f + i \int_{\mathbf{R}^d} \operatorname{Im} f.$$

Remarquons qu'ici, contrairement à ce que nous avons fait pour l'intégration des fonctions positives, aussi bien l'intégrande que l'intégrale ne peuvent prendre des valeurs infinies.

Bien sûr, cette définition d'intégrale donne le même résultat que la précédente pour les fonctions à valeurs réelles positives. En effet, dans ce cas la partie imaginiare de la fonction est nulle, ainsi que sa partie négative. Cela implique en particulier la normalisation de l'intégrale à valeurs complexes. De l'Énoncé indispensable 1.5, on déduit aussi la linéarité de l'intégrale pour les fonctions à valeurs complexes.

Proposition 1.40 (*Linéarité*)

Soient f et g deux fonctions intégrables de $\mathbf{R}^d$ dans $\mathbf{C}$, et pour $\lambda, \mu \in \mathbf{C}$, la fonction $\lambda f + \mu g$ est intégrable et on a

$$\int (\lambda f + \mu g) = \lambda \int f + \mu \int g.$$

Démonstration : L'idée de la preuve est de se ramener au cas de fonctions positives, dans lequel on a la linéarité par l'Énoncé indispensable 1.5. Sans perte de généralité, il suffit bien sûr de traiter $\mu = 1$. L'intégrabilité s'obtient en notant, par inégalité triangulaire, que $|\lambda f + g| \leq |\lambda||f| + |g|$, qui est effectivement une fonction positive d'intégrale finie.

On commence par montrer que $\int (\lambda f) = \lambda \int f$ pour $\lambda \in \mathbf{R}$ et f à valeur réelles. Si $\lambda \geq 0$, cela vient du fait que $(\lambda f)+ = \lambda f_+$ et $(\lambda f)_- = \lambda f_-$, qui implique, par linéarité positive,

$$\int (\lambda f) = \int (\lambda f)_+ - \int (\lambda f)_- = \int (\lambda f_+) - \int (\lambda f_-) = \lambda \int f_+ - \lambda \int f_- = \lambda \int f.$$

Si en revanche $\lambda < 0$, on a $(\lambda f)_+ = -\lambda f_-$ et $(\lambda f)_- = -\lambda f_+$, si bien que

$$\int (\lambda f) = \int (\lambda f)_+ - \int (\lambda f)_- = \int (-\lambda f_-) - \int (-\lambda f_+) = (-\lambda) \int f_- - (-\lambda) \int f_+ = \lambda \int f.$$

Montrons maintenant que si f et g sont à valeurs réelles, alors $\int (f+g) = \int f + \int g$. Pour cela on utilise l'égalité

$$f + g = f_+ - f_- + g_+ - g_- = (f+g)_+ - (f+g)_-,$$

qui donne

$$f_+ + g_+ + (f+g)_- = (f+g)_+ + f_- + g_-,$$

et en intégrant chacun des membres on obtient (par la linéarité dans le cas des fonctions positives)

$$\int f_+ + \int g_+ + \int (f+g)_- = \int (f+g)_+ + \int f_- + \int g_-,$$

si bien que

$$\int f_+ - \int f_- + \int g_+ - \int g_- = \int (f+g)_+ - \int (f+g)_-,$$

ce qui équivaut à $\int f + \int g = \int (f+g)$.

Dans le cas général où f et λ sont complexes, on pose $\lambda = a+ib$. On a alors, en utilisant la linéarité dans le cas réel,

$$\begin{aligned}\int (\lambda f) &= \int \mathrm{Re}(\lambda f) + i \int \mathrm{Im}(\lambda f)\\ &= \int \big(a \,\mathrm{Re}\, f - b\,\mathrm{Im}\, f\big) + i \int \big(a\,\mathrm{Im}\, f + b\,\mathrm{Re}\, f\big)\\ &= a\int \mathrm{Re}\, f - b \int \mathrm{Im}\, f + ia \int \mathrm{Im}\, f + ib\int \mathrm{Re}\, f\\ &= (a+ib)\left(\int \mathrm{Re}\, f + i\int \mathrm{Im}\, f\right)\\ &= \lambda \int f.\end{aligned}$$

Enfin, si on considère deux fonctions intégrables f et g à valeurs complexes, on a

$$\begin{aligned}\int (f+g) &= \int \mathrm{Re}(f+g) + i\int \mathrm{Im}(f+g)\\ &= \int \big(\mathrm{Re}\, f + \mathrm{Re}\, g\big) + i \int \big(\mathrm{Im}\, f + \mathrm{Im}\, g\big) \quad = \int \mathrm{Re}\, f + \int \mathrm{Re}\, g + i\int \mathrm{Im}\, f + i\int \mathrm{Im}\, g\\ &= \int f + \int g.\end{aligned}$$

■

L'intégrale de Lebesgue possède aussi les propriétés de croissance et d'inégalité triangulaire.

Proposition 1.41 (*Croissance de l'intégrale, 2*)

Pour toutes fonctions $f, g \in \mathscr{L}^1(\mathbf{R}^d)$ ***à valeurs réelles****, si $f \leq g$, alors $\int f \leq \int g$.*

Démonstration : Par les définitions des parties positive et négative d'une fonction, on voit que $f_+ = \max(f,0) \leq \max(g,0) = g_+$ et $f_- = \max(-f,0) \geq \max(-g,0) = g_-$. Ainsi,

$$\int_{\mathbf{R}^d} f = \int_{\mathbf{R}^d} f_+ - \int_{\mathbf{R}^d} f_- \leq \int_{\mathbf{R}^d} g_+ - \int_{\mathbf{R}^d} g_- = \int_{\mathbf{R}^d} g.$$

■

Proposition 1.42 (*Inégalité triangulaire*)

Pour toute $f \in \mathscr{L}^1(\mathbf{R})$ (à valeurs complexes), on a

$$\left|\int_{\mathbf{R}^d} f\right| \leq \int_{\mathbf{R}^d} |f|.$$

Démonstration : Commençons par le cas où f est à valeurs réelles. Des inégalités $f \leq |f|$ et $-f \leq |f|$ on déduit, à l'aide de la croissance de l'intégrale, que $\int f \leq \int |f|$ et $-\int f \leq \int |f|$, d'où $|\int f| \leq \int |f|$.

Dans le cas où f est à valeurs complexes, on écrit $\int_{\mathbf{R}^d} f = re^{i\theta}$, avec $r = |\int_{\mathbf{R}^d} f|$; on cherche donc à majorer r. Mais

$$r = e^{-i\theta}\int_{\mathbf{R}^d} f = \int_{\mathbf{R}^d} (e^{-i\theta} f) = \int_{\mathbf{R}^d} \mathrm{Re}(e^{-i\theta} f) + i \int_{\mathbf{R}^d} \mathrm{Im}(e^{-i\theta} f),$$

si bien que $\int_{\mathbf{R}^d} \operatorname{Im}(e^{-i\theta} f) = 0$ (car $r \in \mathbf{R}$). Par conséquent, puisque $r \geq 0$,

$$r = \int_{\mathbf{R}^d} \operatorname{Re}(e^{-i\theta} f) = \left| \int_{\mathbf{R}^d} \operatorname{Re}(e^{-i\theta} f) \right|,$$

et donc, par l'inégalité triangulaire dans le cas réel, et par croissance,

$$r = \left| \int_{\mathbf{R}^d} \operatorname{Re}(e^{-i\theta} f) \right| \leq \int_{\mathbf{R}^d} \left| \operatorname{Re}(e^{-i\theta} f) \right| \leq \int_{\mathbf{R}^d} \left| e^{-i\theta} f \right| = \int_{\mathbf{R}^d} |f|.$$

■

Nous avons introduit dans le paragraphe précédent la notion d'ensemble négligeable (voir Définition 1.25). La proposition suivante justifie quelque peu cette appellation : en pratique, une fois le symbole intégral écrit, peu importe la valeur des fonctions sur les ensembles négligeables. Cela permet de renforcer (légèrement) les énoncés de nombreux théorèmes en les parsemant du vocable « presque partout ». Par exemple le Théorème de Beppo-Levi reste vrai sous l'hypothèse plus faible que la suite $(f_n)_n$ est croissante presque partout.

Proposition 1.43

Deux fonctions intégrables égales presque partout ont la même intégrale.

Démonstration : Soient f et g de telles fonctions. Alors $h := |f - g|$ est positive mesurable et nulle presque partout, donc d'intégrale nulle par la Proposition 1.30. L'inégalité triangulaire de la Proposition 1.42 assure donc l'égalité des deux intégrales. ■

Remarque 1.44

Attention ! La Proposition 1.30 n'est pas vraie dans le cas de fonctions non positives : une fonction à valeurs réelles ou complexes peut avoir une intégrale nulle sans pour autant être nulle presque partout.

1.4 Intégrale sur un sous-ensemble

Définition 1.45

Soit $A \subset \mathbf{R}^d$ un sous-ensemble mesurable et $f : A \to \mathbf{C}$ une fonction mesurable. On dit que f est *intégrable* sur A si la fonction

$$\begin{array}{rcl} \widetilde{f}_A : \mathbf{R}^d & \longrightarrow & \mathbf{C} \\ x & \longmapsto & \begin{cases} 0 & \text{si } x \notin A \\ f(x) & \text{si } x \in A \end{cases} \end{array}$$

est intégrable. On note alors

$$\int_A f = \int_A f(x)\,\mathrm{d}x \stackrel{\text{déf.}}{=} \int_{\mathbf{R}^d} \widetilde{f}_A,$$

et l'ensemble des fonctions intégrables sur A est noté $\mathscr{L}^1(A)$.

Toutes les propriétés que nous avons établies pour l'intégrale sur $\mathbf{R}^d$ se transfèrent *mutadis mutandis* au cas de l'intégration sur un sous-ensemble mesurable. L'intégrale sur de telles sous-parties est donc linéaire, croissante (pour les fonctions à valeurs réelles), satisfait l'inégalité triangulaire et est insensible au changement de valeurs sur des ensembles négligeables.

Proposition 1.46 (*Relation de Chasles*)

Si A et B sont deux ensembles mesurables tels que $\lambda(A \cap B) = 0$, alors pour toute fonction f mesurable intégrable sur $A \cup B$, on a

$$\int_{A\cup B} f = \int_A f + \int_B f.$$

Cette propriété est en particulier vraie pour des intervalles : on retrouve la relation de Chasles classique : pour des réels $a < b < c$,

$$\int_{[a,c]} f = \int_{[a,b]} f + \int_{[b,c]} f.$$

Démonstration : On voit que sur la différence symétrique $A\Delta B := (A \setminus B) \cup (B \setminus A)$, on a $f = \widetilde{f}_A + \widetilde{f}_B$. Donc f et $\widetilde{f}_A + \widetilde{f}_B$ coïncident sur $A \cup B$ en dehors de l'ensemble $A \cap B$ qui est de mesure nulle ; autrement dit ces fonctions sont égales presque partout sur $A \cup B$. Ainsi (Proposition 1.43),

$$\int_{A\cup B} f = \int_{A\cup B} (\widetilde{f}_A + \widetilde{f}_B) = \int_{A\cup B} \widetilde{f}_A + \int_{A\cup B} \widetilde{f}_B = \int_A f + \int_B f.$$ ∎

Remarquons que cette preuve n'utilise que l'existence des objets utilisés : la relation de Chasles est aussi vraie pour des fonctions à valeurs positives, sans hypothèse d'intégrabilité de celles-ci.

Bien souvent, par la suite, lorsqu'on voudra intégrer une fonction réelle à l'aide d'une de ses primitives, on voudra se ramener au cas d'un segment. L'énoncé suivant, plus général, permet de le faire lorsque la fonction est positive. Le cas des fonctions à valeurs complexes sera traité au chapitre suivant comme corollaire du théorème de convergence dominée (Corollaire 2.16).

Corollaire 1.47 (*Réduction à un compact, cas positif*)

Soit $(A_n)_n$ une suite croissante de sous-ensembles mesurables de $\mathbf{R}^d$, et $A = \bigcup_{n\in\mathbf{N}} A_n$. Soit f une fonction mesurable positive sur A. Alors

$$\int_{A_n} f \underset{n\to+\infty}{\longrightarrow} \int_A f.$$

Remarque 1.48

Ce corollaire est souvent utile lorsqu'on souhaite se ramener à des fonctions sur des ensembles de mesures finies. Il est fréquemment utilisé dans le cas $A = \mathbf{R}$ et $A_n = [-n, n]$.

Démonstration : Il suffit d'appliquer le théorème de Beppo-Levi à la suite de fonctions $f_n = \mathbf{1}_{A_n} f$ qui est croissante et converge simplement vers f. ■

Lorsque l'on intègre sur une partie de mesure finie, on dispose d'un critère très simple d'intégrabilité. Commençons par un petit rappel.

Définition 1.49

Soit A une partie de $\mathbf{R}^d$. Une fonction $f : A \to \mathbf{C}$ est dite *bornée sur* A si $\sup_A |f| < +\infty$ ou, de manière équivalente, s'il existe une constante $M > 0$ pour laquelle, pour tout élément x de A, on a $|f(x)| \leq M$. Étant donnée une suite $(f_n)_n$ de fonctions bornées sur A, on dit qu'elle *converge uniformément* vers f si

$$\sup_{x \in A} |f(x) - f_n(x)| \underset{n \to +\infty}{\longrightarrow} 0$$

Proposition 1.50

Si A est une partie mesurable de mesure finie, toute fonction mesurable et bornée sur A est intégrable sur A.

Démonstration : Il suffit d'écrire, par croissance et linéarité de l'intégrale,

$$\int_A |f| \leq \int_A \sup_A |f| = \sup_A |f| \lambda(A) < +\infty.$$

■

Corollaire 1.51

Si A est un sous-ensemble compact de $\mathbf{R}^d$, alors toute fonction continue sur A est intégrable sur A.

Démonstration : A est compact donc en particulier borné et donc inclus dans un ensemble $[-R, R]^d$ pour un certain $R > 0$ suffisamment grand. On en déduit en particulier par croissance de la mesure de Lebesgue que $\lambda(A) \leq (2R)^d < +\infty$. Ensuite, si f est une fonction continue sur A, alors $|f|$ est une fonction continue sur le compact A et à valeurs réelles : on sait qu'elle est bornée (et atteint ses bornes !). On peut donc applique la Proposition 1.50. ■

Poursuivons par un énoncé qui nous servira pour établir la coïncidence des intégrales de Lebesgue et Riemann, mais qui n'est en fait qu'un cas particulier du théorème de convergence dominée (Théorème 2.13) qui sera vu au Chapitre 2.

Proposition 1.52

Soit A un ensemble de mesure finie. Si une suite de fonctions bornées $(f_n)_{n \in \mathbf{N}}$ converge uniformément vers une fonction bornée f, alors

$$\int_A f_n \underset{n \to +\infty}{\longrightarrow} \int_A f.$$

Remarque 1.53

La preuve établit en fait, en des termes (à peine) plus sophistiqués, la continuité de l'intégration sur une partie de mesure finie, pour la norme uniforme.

Démonstration : Par inégalité triangulaire, croissance et linéarité de l'intégrale on a

$$\left|\int_A f_n - \int_A f\right| = \left|\int_A (f_n - f)\right| \leq \lambda(A) \sup_A |f_n - f| \underset{n\to+\infty}{\longrightarrow} 0,$$

où l'on a utilisé $\lambda(A) < \infty$ pour la convergence. ■

À partir de la mesure de Lebesgue, on peut construire d'autres mesures qu'on appelle *mesures à densité*, qui ont une importance capitale en théorie des probabilités.

Proposition 1.54

Si f est une fonction positive, l'application

$$\begin{aligned} \mu : \mathcal{A} &\longrightarrow \mathbf{R}_+ \\ A &\mapsto \int_A f, \end{aligned}$$

définit une mesure.

Démonstration : En reprenant la Définition 1.45, $\widetilde{f_\varnothing}$ est nulle presque partout et donc d'intégrale nulle par la Proposition 1.43. Si $(A_n)_n$ est une famille dénombrable d'ensembles mesurables deux à deux disjoints c'est le Théorème de Beppo-Levi associé à la linéarité de l'intégrale de Lebesgue qui permet d'établir que $\mu(\sqcup_n A_n) = \sum_n \mu(A_n)$. ■

Définition 1.55 (*Mesure à densité*)

Une mesure μ est dite *à densité* (par rapport à la mesure de Lebesgue) s'il existe une fonction positive mesurable f_μ pour laquelle μ s'exprime comme dans la Proposition 1.54. On dit alors que f_μ est une *densité* de μ. L'intégrale d'une fonction positive mesurable g *par rapport* à μ est alors simplement donnée par la formule

$$\int_{\mathbf{R}^d} g \, \mathrm{d}\mu := \int_{\mathbf{R}^d} g f_\mu \, .$$

La terminologie et les définitions de l'intégrale de Lebesgue se transmettent alors à l'intégration par rapport à μ.

Remarque 1.56

D'après la Proposition 1.43 la densité d'une mesure n'est pas unique : on peut la modifier sur un ensemble de mesure nulle et obtenir une autre fonction positive qui jouera le rôle de densité ; par contre la mesure est entièrement déterminée par sa densité. En pratique, on quotiente par la relation d'équivalence « être égal presque partout » pour identifier la densité. Nous aurons à pratiquer ce quotient lors de l'étude des espaces $L^1(\mathbf{R})$ et $L^2(\mathbf{R})$ dans le Chapitre 3.

Exemple 1.57

- La mesure de probabilité uniforme sur $[0,1]$ est la mesure de densité $\mathbf{1}_{[0,1]}$.
- La mesure de la loi normale (ou loi gaussienne), est la mesure de densité

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$

On vérifiera à l'exemple 2.40 que c'est bien une mesure de probabilité. Cette mesure est centrale en théorie des probabilités, puisqu'elle apparaît naturellement dans le théorème central limite.

Exercice 1.58

*Après avoir (re)démontré que le graphe d'une fonction convexe réelle admet une droite d'appui en tout point, établir l'**inégalité de Jensen** : pour toute mesure de probabilité μ à densité, toute fonction convexe $\varphi : \mathbf{R} \to \mathbf{R}_+$ et toute fonction $g : \mathbf{R}^d \to \mathbf{R}$ intégrable par rapport à μ on a*

$$\varphi\left(\int_{\mathbf{R}^d} g \,\mathrm{d}\mu\right) \leq \int_{\mathbf{R}^d} \varphi(g) \,\mathrm{d}\mu.$$

Remarque 1.59

La « vraie » inégalité de Jensen ne réclame pas que la mesure de probabilité soit à densité... mais encore faut-il définir l'intégration selon une mesure quelconque pour l'énoncer !

1.5 Sommes et intégrale de Riemann

Terminons ce chapitre en effectuant le lien avec une autre intégrale, historiquement attribuée à Riemann. Cette intégrale est largement antérieure à celle de Lebesgue et sa construction est beaucoup plus simple : prenons la peine de la présenter ici, dans le cas particulier des fonctions continues sur un segment. Cette simplicité de construction a un prix et l'objet obtenu est bien moins pratique que l'intégrale de Lebesgue ; on ne peut définir l'intégrale que d'un nombre assez restreint de fonctions, et l'intégration correspondante ne se comporte pas très bien par passage à la limite [5]. Heureusement, et c'est le message à retenir de cet ultime paragraphe, l'intégrale de Riemann coïncide complètement avec l'intégrale de Lebesgue et on peut donc se contenter de manipuler cette dernière, sans ambiguïté.

5. En tous cas, la preuve des théorèmes de passage à la limite est bien moins naturelle.

Définition 1.60 (*Sommes de Riemann*)

Soit $f:[a,b]\to\mathbf{C}$ une fonction. Une *somme de Riemann* pour f est une somme du type

$$S_N(f)=\frac{b-a}{N}\sum_{k=1}^{N} f\left(a+k\frac{b-a}{N}\right).$$

L'idée derrière cette définition est d'approcher l'intégrale d'une fonction f par l'intégrale d'une fonction « *en escalier* » (voir la Figure 1.2). Plus précisément si on note $r_{k,N}:=a+k\frac{b-a}{N}$ et $I_{k,N}:=[r_{k-1,N},r_{k,N}[$ et que l'on introduit

$$\begin{array}{rcl} E_N(f):[a,b] & \longrightarrow & \mathbf{C} \\ x & \longmapsto & \sum_{k=1}^{N}\mathbf{1}_{I_{k,N}} f\left(a+k\frac{b-a}{N}\right), \end{array}$$

nous verrons que $S_N(f)$ n'est rien d'autre que *l'intégrale de Riemann* de $E_N(f)$.

Proposition 1.61

Si f est une fonction continue sur un segment $[a,b]$, alors la suite des fonctions en escalier $(E_N(f))_{N\in\mathbf{N}}$ converge uniformément vers f.

Démonstration : Par le Théorème de Heine, puisque f est continue sur $[a,b]$, elle y est uniformément continue : pour tout $\varepsilon>0$, il existe $\delta>0$ tel que si $|x-y|\le\delta$, alors $|f(x)-f(y)|\le\varepsilon$. Soit maintenant $N\ge(b-a)/\delta$ et $x\in[a,b[$. Alors il existe k tel que $x\in[a+(k-1)(b-a)/N,a+k(b-a)/N[$. Ainsi, $|x-a+k(b-a)/N|\le(b-a)/N\le\delta$ et donc

$$\left|f(x)-E_N(f)(x)\right|=\left|f(x)-f\left(a+k\frac{b-a}{N}\right)\right|\le\varepsilon.$$

■

L'intégrale d'une fonction en escalier est définie comme l'aire (algébrique) de l'union des rectangles formés par son graphe. Cela permet de définir l'intégrale de Riemann de la fonction f comme la limite des intégrales des fonctions en escalier qui l'approchent. Bien sûr, il faut démontrer l'existence de cette limite, et c'est l'objet de l'énoncé indispensable qui suit.

Énoncé indispensable 1.62 (*Convergence des sommes de Riemann*)

Soit $f:[a,b]\to\mathbf{C}$ une fonction continue. Alors les sommes de Riemann $S_N(f)$ convergent (quand N tend vers $+\infty$) vers un nombre complexe appelé l'intégrale de Riemann de f sur $[a,b]$, et notée

$$\int_a^b f(x)\,\mathrm{d}x.$$

Remarque 1.63

Par convention, si on a $a\le b$ et $f:[a,b]\to\mathbf{C}$, on définit

$$\int_b^a f(x)\,\mathrm{d}x:=-\int_a^b f(x)\,\mathrm{d}x.$$

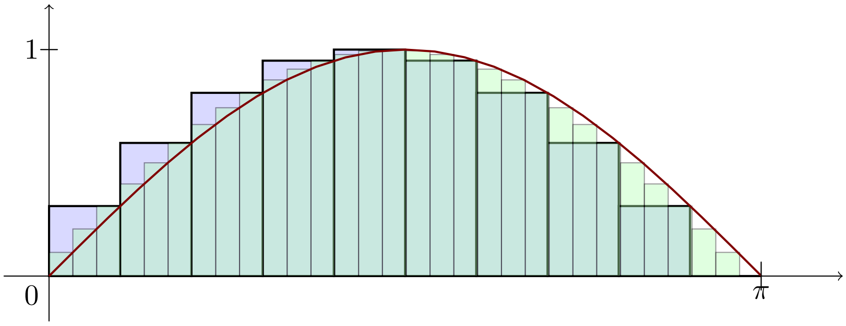

FIGURE 1.2 – Fonctions en escalier E_{10} (bleu) et E_{30} (vert) associées à la fonction sin.

Démonstration : L'idée de preuve est assez simple : on montre que les sommes de Riemann forment une suite de Cauchy. Autrement dit, on va prouver que si M et N sont assez grands, alors les sommes $S_N(f)$ et $S_M(f)$ sont très proches. Pour simplifier les calculs, une astuce consiste à comparer non pas directement $S_N(f)$ et $S_M(f)$, mais $S_N(f)$ et $S_{MN}(f)$.

La notion clef pour cette preuve est celle de continuité uniforme. Soit donc $\varepsilon > 0$. Par le théorème de Heine, puisque f est continue sur $[a, b]$, elle y est uniformément continue : il existe $\delta > 0$ tel que si $|x - y| \leq \delta$, alors $|f(x) - f(y)| \leq \varepsilon$. Prenons $M, N \in \mathbf{N}$ tels que $M, N \geq (b - a)/\delta$. Alors

$$S_N(f) - S_{MN}(f) = \frac{b-a}{N} \sum_{k=1}^{N} f\left(a + k\frac{b-a}{N}\right) - \frac{b-a}{MN} \sum_{m=1}^{MN} f\left(a + m\frac{b-a}{MN}\right).$$

L'idée est alors de découper la seconde somme en N morceaux de taille M et de montrer que chacun de ces morceaux est proche d'un des termes de la première somme ; sur la Figure 1.2, cela correspond à considérer les rectangles de la subdivision la plus fine par paquets de 3 et à observer que l'aire de ces paquets de 3 petits rectangles (verts) est proche de celle du grand rectangle (bleu) correspondant. Tout entier n compris entre 0 et $MN - 1$ se décompose de manière unique sous la forme $kM + \ell$, où $k \in [\![0, N-1]\!]$ et $\ell \in [\![0, M-1]\!]$: il s'agit simplement de la division euclidienne de n par M. En appliquant cela pour $n = m - 1$ dans la somme qui suit, nous obtenons

$$\begin{aligned}
\sum_{m=1}^{MN} f\left(a + m\frac{b-a}{MN}\right) &= \sum_{m=1}^{MN} f\left(a + \frac{b-a}{MN} + (m-1)\frac{b-a}{MN}\right) \\
&= \sum_{k=0}^{N-1} \sum_{\ell=0}^{M-1} f\left(a + \frac{b-a}{MN} + (kM+\ell)\frac{b-a}{MN}\right) \\
&= \sum_{k=0}^{N-1} \sum_{\ell=0}^{M-1} f\left(a + (k+1)\frac{b-a}{N} + (\ell+1-M)\frac{b-a}{MN}\right).
\end{aligned}$$

En réindexant selon $k \leftrightarrow k+1$ et $\ell \leftrightarrow (M-1) - \ell$, il vient finalement

$$\sum_{m=1}^{MN} f\left(a + m\frac{b-a}{MN}\right) = \sum_{k=1}^{N} \sum_{\ell=0}^{M-1} f\left(a + k\frac{b-a}{N} - \ell\frac{b-a}{MN}\right).$$

Finalement, nous obtenons :

$$S_N(f) - S_{MN}(f) = \frac{b-a}{N} \sum_{k=1}^{N} \left(f\left(a + k\frac{b-a}{N}\right) - \frac{1}{M} \sum_{\ell=0}^{M-1} f\left(a + k\frac{b-a}{N} - \ell\frac{b-a}{MN}\right) \right).$$

On utilise maintenant le fait que

$$f\left(a+k\frac{b-a}{N}\right)=\frac{M}{M}f\left(a+k\frac{b-a}{N}\right)=\frac{1}{M}\sum_{\ell=0}^{M-1}f\left(a+k\frac{b-a}{N}\right),$$

qui implique que

$$S_N(f)-S_{MN}(f)=\frac{b-a}{N}\sum_{k=1}^{N}\frac{1}{M}\sum_{\ell=0}^{M-1}\left(f\left(a+k\frac{b-a}{N}\right)-f\left(a+k\frac{b-a}{N}-\ell\frac{b-a}{MN}\right)\right),$$

et l'inégalité triangulaire donne :

$$|S_N(f)-S_{MN}(f)|\leq\frac{b-a}{N}\sum_{k=1}^{N}\frac{1}{M}\sum_{\ell=0}^{M-1}\left|f\left(a+k\frac{b-a}{N}\right)-f\left(a+k\frac{b-a}{N}-\ell\frac{b-a}{MN}\right)\right|.$$

Mais puisque $\ell(b-a)/(MN)\leq(b-a)/N\leq\delta$, on a, pour tous les indices de sommation k et ℓ,

$$\left|f\left(a+k\frac{b-a}{N}\right)-f\left(a+k\frac{b-a}{N}-\ell\frac{b-a}{MN}\right)\right|\leq\varepsilon,$$

si bien que

$$|S_N(f)-S_{MN}(f)|\leq\frac{b-a}{N}\sum_{k=1}^{N}\frac{1}{M}\sum_{\ell=0}^{M-1}\varepsilon=(b-a)\varepsilon.$$

Par symétrie des rôles, on en déduit que $|S_N(f)-S_M(f)|\leq 2(b-a)/\varepsilon$, ce qui montre bien que la suite $(S_N(f))_N$ est de Cauchy. ■

Exemple 1.64

Essayons d'obtenir le comportement limite de la somme

$$T_N=\sum_{k=1}^{N}\frac{1}{N+k}.$$

Pour faire apparaître une somme de Riemann, il nous faut un $1/N$ devant la somme, on multiplie donc par N/N :

$$T_N=\frac{1}{N}\sum_{k=1}^{N}\frac{N}{N+k}=\frac{1}{N}\sum_{k=1}^{N}\frac{1}{1+\frac{k}{N}}.$$

On reconnaît la somme de Riemann associée à la fonction $f(x)=1/x$ et à l'intervalle $[a,b]=[1,2]$. Le Théorème 1.62 assure alors que

$$T_N\underset{N\to+\infty}{\longrightarrow}\int_1^2\frac{1}{x}\,\mathrm{d}x=\ln 2.$$

La définition même de l'intégrale de Riemann fournit un moyen pratique d'approcher numériquement l'intégrale d'une fonction donnée, par le simple calcul des sommes de Riemann. On a même une majoration de l'erreur dans le cas où la fonction est lipschitzienne.

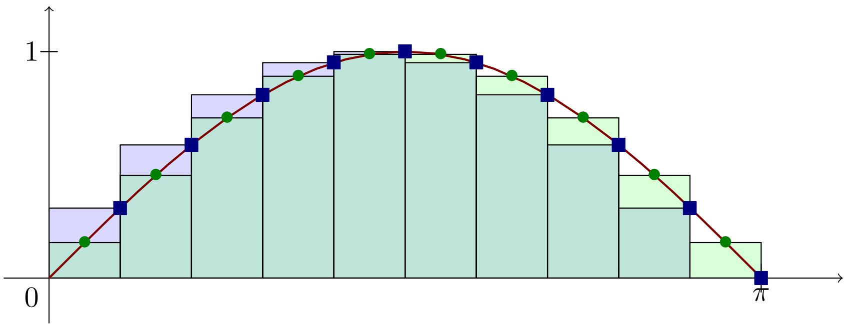

FIGURE 1.3 – Les fonctions en escalier associées à la fonction sin : celle que nous utilisons ici pour définir les sommes de Riemann (bleu, avec le point d'évaluation de f marqués par des carrés) et celle pour la méthode point-milieu (vert, avec le point d'évaluation de f marqués par des disques). On voit que dans la méthode du point-milieu il y a des compensations qu'il n'y a pas avec l'autre méthode, ce qui explique sa convergence plus rapide.

Exercice 1.65

Montrer que si $f : [a, b] \to \mathbf{C}$ est k-lipschitzienne alors

$$\left| S_N(f) - \int_a^b f(x)\,\mathrm{d}x \right| \leq k\frac{(b-a)^2}{2N}.$$

Remarque 1.66

Si au lieu de choisir le point de droite des intervalles pour évaluer la fonction sur chacun des intervalles on choisit le point du milieu (voir la figure 1.3), on obtient la méthode d'intégration dite du « point milieu », qui est bien plus rapide lorsque la fonction est de classe $\mathscr{C}^2([a, b])$: si $|f''(x)| \leq k$ pour tout $x \in [a, b]$, alors

$$\left| S_N(f) - \int_a^b f(x)\,\mathrm{d}x \right| \leq k\frac{(b-a)^3}{24N^2}.$$

Énoncé indispensable 1.67 (*Lebesgue généralise Riemann*)

Soit $f : [a, b] \to \mathbf{C}$ une fonction continue. Alors elle est intégrable (au sens de Lebesgue) et les intégrales de Riemann et de Lebesgue de f coïncident :

$$\int_{[a,b]} f = \int_a^b f(t)\,\mathrm{d}t.$$

Démonstration : Il suffit de combiner les faits suivants :

- le Corollaire 1.51 assure l'intégrabilité de f au sens de Lebesgue ;

- par définition, les intégrales de Riemann et de Lebesgue coïncident sur l'ensemble des fonctions en escalier ;
- toute fonction continue est la limite uniforme des fonctions en escalier qui lui sont associées (Proposition 1.61), et l'intégrale de Riemann d'une telle fonction est définie comme étant la limite des intégrales de ces fonctions en escalier ;
- on peut intervertir limite uniforme et intégrale de Lebesgue (Proposition 1.52). ■

Remarque 1.68

En fait, ce théorème, tout comme le Théorème 1.62, est vrai pour une classe de fonctions bien plus large : celle des fonctions dites *Riemann-intégrables.* Cet ensemble de fonction contient notamment l'ensemble des fonctions *réglées*, soit l'ensemble des fonctions que l'on peut obtenir comme limite uniforme de fonctions en escaliers. Plutôt que de définir précisément la « Riemann-intégrabilité », contentons-nous de citer un critère établi par Lebesgue qui permet de comprendre à quel point son intégrale généralise la précédente : *une fonction est Riemann-intégrable sur un segment si et seulement si elle est bornée et presque partout continue.*

5. Pour $[a, b] \subset \mathbf{R}$ un intervalle, on dit qu'une fonction $f : [a, b] \to \mathbf{C}$ est *continue par morceaux* s'il existe $a = x_0 < x_1 < \cdots < x_n = b$ tels que pour tout $1 \leq i \leq n$, il existe une fonction continue $g : [x_{i-1}, x_i]$ dont la restriction à l'intérieur de l'intervalle de définition coïncide avec $f_{|]x_{i-1}, x_i[}$.

Exemple 1.69

Attention ! Il existe des fonctions très simples pour lesquelles les sommes de Riemann convergent sans pour autant que la limite coïncide avec l'intégrale de Lebesgue de la fonction. Prenons la fonction indicatrice de $\mathbf{Q} \cap [0,1]$. On a vu à l'exemple 1.9 que son intégrale de Lebesgue est nulle. En revanche, ses sommes de Riemann valent

$$\begin{aligned} S_N(\mathbf{1}_{\mathbf{Q}\cap[0,1]}) &= \frac{1}{N}\sum_{k=1}^{N} \mathbf{1}_{\mathbf{Q}\cap[0,1]}\left(0 + \frac{k}{N}\right) \\ &= \frac{1}{N}\sum_{k=1}^{N} 1 = 1, \end{aligned}$$

elles convergent donc vers 1.

Chapitre 2

Outils de calculs

Comme mentionné au chapitre précédent, nous n'avons pour l'instant pas beaucoup de moyens de *calculer* la valeur d'une intégrale. Nous allons ici présenter plusieurs résultats fondamentaux le permettant. Les premiers de ceux-ci sont les héritiers naturels de l'intégrale de Riemann et s'expriment pour les fonctions de variables réelles ; il s'agit du théorème fondamental de l'analyse, de l'intégration par parties et du changement de variables classique que nous présentons dans la Section 2.1. L'une des manifestations les plus spectaculaires de l'avancée de Lebesgue par son intégrale est le théorème de convergence dominée présenté dans la Section 2.2. En plus d'être un outil très efficace pour calculer des limites, ce résultat éclaire l'étude des intégrales à paramètres que nous présentons dans la Section 2.3. Nous aborderons ensuite dans la Section 2.4 les théorèmes de Fubini qui permettent de ramener une intégrale double à deux intégrales d'une seule variable et terminerons le chapitre par la formule générale de changement de variables.

Résumé :

- Le théorème fondamental de l'analyse fait le lien entre intégrale et primitive (2.1). Il admet comme corollaires la formule d'intégration par parties (Théorème 2.3) et de changement de variables (Théorème 2.5).
- Le passage à la limite pour une suite d'intégrales se fait par le théorème de convergence dominée (Théorème 2.13) ; pour l'appliquer il faut que chacune des fonctions de la suite soit majorée en module par une fonction intégrable fixe.
- Ce théorème implique des résultats de continuité et de dérivabilité d'intégrales à paramètres, avec là aussi des hypothèses de domination (Théorèmes 2.19 et 2.23).
- Le Théorème 2.26 de Fubini-Tonelli (pour les fonctions positives) et le Théorème 2.28 de Fubini (pour les fonctions intégrables) permettent de faire le lien entre intégrale sur $\mathbf{R}^2$ et doubles intégrales de fonctions. Ils impliquent en particulier la possibilité d'intervertir les symboles intégrale.
- Le Théorème 2.36 donne une formule générale de changement de variables valide en toute dimension. C'est un outil de calcul précieux pour les intégrales multiples, notamment grâce aux changements de variables polaire et sphérique (à connaître ou savoir retrouver !).

2.1 Intégrale et primitives

On démontre ici le théorème fondamental de l'analyse, qui fait le lien entre intégrale et dérivée : il énonce que la fonction $x \mapsto \int_a^x f(t)dt$ est dérivable, de dérivée f.

Énoncé indispensable 2.1 (***Théorème fondamental de l'analyse***)

Soit $f : [a, b] \to \mathbf{C}$ une fonction continue. Alors la fonction

$$F : x \mapsto \int_a^x f(t)\,\mathrm{d}t$$

est dérivable sur $]a, b[$, de dérivée égale à f. On dit que c'est une primitive de f.

Démonstration : On calcule le taux d'accroissement en un point $x \in]a, b[$: pour $h > 0$, par la relation de Chasles,

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h}\left(\int_a^{x+h} f - \int_a^x f\right) = \frac{1}{h}\int_x^{x+h} f.$$

Or f est continue en x, donc pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que si $|h| \leq \eta$, alors $|f(x) - f(x+h)| \leq \varepsilon$. Ainsi, pour $|h| \leq \eta$

$$\begin{aligned}\left|\frac{F(x+h) - F(x)}{h} - f(x)\right| &= \left|\frac{1}{h}\int_x^{x+h} f(t)\,\mathrm{d}t - f(x)\right| = \left|\frac{1}{h}\int_x^{x+h} f(t)\,\mathrm{d}t - \frac{1}{h}\int_x^{x+h} f(x)\,\mathrm{d}t\right| \\ &\leq \frac{1}{h}\int_x^{x+h} |f(t) - f(x)|\,\mathrm{d}t \leq \frac{1}{h}\int_x^{x+h} \varepsilon\,\mathrm{d}t = \varepsilon.\end{aligned}$$

■

Corollaire 2.2

Si $f : [a, b] \to \mathbf{C}$ est une fonction de classe $\mathscr{C}^1$, alors

$$\int_a^b f'(t)\,\mathrm{d}t = f(b) - f(a).$$

Démonstration : Considérons l'application

$$\varphi : x \mapsto f(x) - \int_a^x f'(t)dt.$$

Par le théorème précédent, cette fonction est dérivable sur $]a, b[$, de dérivée $\varphi'(x) = f'(x) - f'(x) = 0$. Cela implique (par l'inégalité des accroissements finis) que la fonction φ est constante. Donc $\varphi(b) = \varphi(a) = f(a)$, ce qui donne le résultat. ■

Ces résultats sont très utiles pour calculer des intégrales en pratique, et invitent à connaître les primitives des fonctions usuelles (voir le tableau dans la Section A.7 de l'annexe). Ils impliquent aussi deux théorèmes précieux : l'intégration par parties et le changement de variables.

Théorème 2.3 (***Intégration par parties***)

Soit $a < b \in \mathbf{R}$, et $f, g : [a, b] \to \mathbf{C}$ deux fonctions de classe $\mathscr{C}^1$. Alors

$$\boxed{\int_a^b f(x)g'(x)\,\mathrm{d}x = [f(x)g(x)]_a^b - \int_a^b f'(x)g(x)\,\mathrm{d}x.}$$

Démonstration : Il suffit d'intégrer entre a et b l'égalité

$$(fg)' = f'g + fg',$$

et d'utiliser le Corollaire 2.2. ■

Exemple 2.4

Il est possible d'étendre la fonction factorielle aux nombres réels positifs, qu'on appelle fonction Γ. Cette fonction a des applications allant des statistiques à la théorie analytique des nombres.

Pour $t > 0$, posons

$$\Gamma(t) = \int_0^{+\infty} e^{-x}x^{t-1}\,\mathrm{d}x.$$

Cette fonction est bien définie, puisque la fonction intégrée est positive et continue.

On peut aussi vérifier que pour tout $t > 0$, on a $\Gamma(t) < +\infty$. En effet, pour t fixé, il existe $C > 0$ tel que si $x \geq C$, alors $e^{-x}x^{t-1} \leq e^{-x/2}$ (car $x^{t-1} = o(e^{x/2}$ en $+\infty$). Ainsi, par la relation de Chasles,

$$\begin{aligned}\Gamma(t) &= \int_0^C e^{-x}x^{t-1}\,\mathrm{d}x + \int_C^{+\infty} e^{-x}x^{t-1}\,\mathrm{d}x \\ &\leq \int_0^C x^{t-1}\,\mathrm{d}x + \int_C^{+\infty} e^{-x/2}\,\mathrm{d}x\end{aligned}$$

Le premier terme est une nombre fini par le critère de Riemann, voir la Proposition 2.7. Le calcul du second terme se fait par le théorème de Beppo-Levi : la suite de fonctions positives $x \mapsto e^{-x/2}\mathbf{1}_{[C,n]}$ converge simplement et en croissant vers la fonction $x \mapsto e^{-x/2}\mathbf{1}_{[C,+\infty[}$, et donc, par utilisation du Corollaire 2.2,

$$\int_C^{+\infty} e^{-x/2} = \lim_{n\to+\infty}\int_C^n e^{-x/2} = \lim_{n\to+\infty}\left[-2e^{-x/2}\right]_C^n = 2e^{-C/2}.$$

Cela démontre que $\Gamma(t)$ est fini.

Intégrons par parties cette intégrale pour trouver une équation fonctionnelle vérifiée par Γ. Soit $t > 1$. Puisque l'énoncé du théorème d'intégration par parties est valable pour un segment, intégrons entre $1/M$ et M. Posant

$$\begin{array}{rclcrcl} f(x) & = & x^{t-1} & \qquad & g'(x) & = & e^{-x} \\ f'(x) & = & (t-1)x^{t-2} & & g(x) & = & -e^{-x} \end{array},$$

on obtient

$$\int_{1/M}^{M} e^{-x}x^{t-1}\,\mathrm{d}x = \left[-x^{t-1}e^{-x}\right]_{1/M}^{M} + \int_{1/M}^{M}(t-1)x^{t-2}e^{-x}\,\mathrm{d}x$$
$$= -M^{t-1}e^{-M} + \frac{1}{M^{t-1}}e^{-1/M} + (t-1)\int_{1/M}^{M} x^{t-2}e^{-x}\,\mathrm{d}x,$$

et en faisant tendre M vers $+\infty$ on obtient (par croissance comparée, et puisque $t > 1$) :

$$\Gamma(t) = (t-1)\Gamma(t-1).$$

Or, $\Gamma(1) = 1$, si bien que par récurrence, on montre que

$$\forall n \in \mathbf{N}^*, \quad \Gamma(n) = (n-1)!.$$

Théorème 2.5 (***Changement de variables réel***)

Soient I un segment, $f : I \to \mathbf{C}$ une fonction continue, et $\varphi : [a,b] \to I$ une fonction de classe $\mathscr{C}^1$. Alors

$$\boxed{\int_{\varphi(a)}^{\varphi(b)} f(x)\,\mathrm{d}x = \int_a^b f(\varphi(t))\varphi'(t)\,\mathrm{d}t.}$$

Remarque 2.6

Voici un moyen pratique pour retrouver la formule si on l'a oubliée : on écrit « à la physicienne » que $x = \varphi(t)$ et on remplace dans l'intégrale, en utilisant la formule $\mathrm{d}x = \mathrm{d}(\varphi(t)) = \varphi'(t)\,\mathrm{d}t$, ce qui donne

$$\int f(x)\,\mathrm{d}x = \int f(\varphi(t))\varphi'(t)\,\mathrm{d}t.$$

Il reste à s'occuper des bornes. Pour celle du bas, si $t = a$, alors $x = \varphi(t) = \varphi(a)$. La borne du bas est donc a pour la variable t et $\varphi(a)$ pour la variable x. On fait de même pour la borne du haut, ce qui donne la formule.

Démonstration : On définit deux fonctions :

$$\psi : u \mapsto \int_a^u f(\varphi(t))\varphi'(t)\,\mathrm{d}t \qquad \text{et} \qquad F : v \mapsto \int_{\varphi(a)}^{v} f(x)\,\mathrm{d}x.$$

Par le théorème 2.1, ces deux fonctions sont dérivables, de dérivées

$$\psi'(u) = f(\varphi(u))\varphi'(u) \qquad \text{et} \qquad F'(v) = f(v).$$

Ainsi, on a $\psi'(u) = F'(\varphi(u))\varphi'(u) = (F\circ\varphi)'(u)$, avec de plus $\psi(a) = 0 = (F\circ\varphi)(a)$. Cela implique que $\psi = F\circ\varphi$, ce qui donne le résultat. ■

Terminons par une dernière conséquence très pratique du théorème fondamental de l'analyse : le critère de Riemann.

Proposition 2.7 (***Critère de Riemann***)

Pour $\alpha \in \mathbf{R}$, la fonction $t \mapsto \frac{1}{t^\alpha}$ est intégrable sur $]0,1]$ si et seulement si $\alpha < 1$ et intégrable sur $[1,+\infty[$ si et seulement si $\alpha > 1$.

Remarque 2.8

Les fonctions en question étant continues sur $\mathbf{R}^\star$, d'après le Corollaire 1.51 elle sont intégrables sur tout segment ne contenant pas l'origine. La difficulté levée par le critère de Riemann concerne donc soit la singularité $t = 0$, soit le comportement en l'infini. Ce critère fournit une échelle de référence qui permet souvent de vérifier qu'une fonction est intégrable, grâce à la croissance de l'intégrale (voir Exemple 2.9).

Démonstration : Grâce au Théorème de Beppo-Levi, l'intégrabilité de $t \mapsto \frac{1}{t^\alpha}\mathbf{1}_{]0,1]}$ est équivalente à l'existence d'une borne (indépendante de $n \in \mathbf{N}^\star$) pour la suite d'intégrales

$$I_n = \int_{1/n}^{1} \frac{1}{t^\alpha}\,\mathrm{d}t.$$

Remarquons que par le changement de variable $t \leftrightarrow 1/s$ on a

$$I_n = \int_1^n \frac{1}{s^{2-\alpha}}\,\mathrm{d}s.$$

À nouveau grâce au Théorème de Beppo-Levi on voit donc que $(I_n)_n$ est une suite bornée si et seulement si $s \mapsto \frac{1}{s^{2-\alpha}}\mathbf{1}_{[1,\infty[}$ est intégrable. Les deux critères de Riemann s'obtiennent donc simultanément puisque $\alpha < 1 \Leftrightarrow 2-\alpha > 1$. Pour conclure la preuve on utilise par exemple la première expression de I_n, à l'aide du Théorème fondamental de l'analyse : si $\alpha < 1$ on a

$$I_n = \left[\frac{1}{1-\alpha}t^{1-\alpha}\right]_{t=1/n}^{t=1} \underset{n\to+\infty}{\longrightarrow} \frac{1}{1-\alpha},$$

et $(I_n)_n$ est donc effectivement bornée. Inversement, si $\alpha \leq 1$ on a

$$I_n \geq \int_{1/n}^{1} \frac{1}{t}\,\mathrm{d}t = \Big[\ln(t)\Big]_{t=1/n}^{t=1} = \ln(n) \underset{n\to+\infty}{\longrightarrow} +\infty.$$

■

Exemple 2.9

Pour tout $a > 0$, la fonction $g_a : t \mapsto e^{-at^2}$ est intégrable. En effet, par croissance comparée on sait que la fonction $t \mapsto g_a(t)t^2$ est bornée (c'est une fonction continue qui tend vers 0 en l'infini) par une certaine constante $M > 0$. Puisque g_a est paire et positive on a, par relation de Chasles pour les fonctions positives,

$$\int_{\mathbf{R}} |g_a| = 2\int_0^{+\infty} g_a = 2\int_0^1 g_a + \int_1^{+\infty} g_a.$$

g_a est continue, donc intégrable sur tout segment ce qui assure que l'intégrale sur $[0,1]$ est bien finie. Pour l'intégrale portant sur $[1,+\infty[$ on peut par exemple écrire

$$g_a(t)\mathbf{1}_{[1,+\infty[}(t) = g_a(t)t^2\frac{1}{t^2}\mathbf{1}_{[1,+\infty[}(t) \leq M\frac{1}{t^2}\mathbf{1}_{[1,+\infty[},$$

où le membre de droite de l'inégalité est bien intégrable, grâce au critère de Riemann.

Exercice 2.10

Pour $\alpha, \beta \in \mathbf{R}$*, établir le* critère de Bertrand *:*
- $t \mapsto \frac{1}{t^\alpha \ln(t)^\beta} \in \mathscr{L}^1(]0, \frac{1}{e}]) \Leftrightarrow$ *(*$\alpha < 1$ *ou* $\alpha = 1$ *et* $\beta > 1$*) ;*
- $t \mapsto \frac{1}{t^\alpha \ln(t)^\beta} \in \mathscr{L}^1([e, +\infty[) \Leftrightarrow$ *(*$\alpha > 1$ *ou* $\alpha = 1$ *et* $\beta > 1$*).*

2.2 La convergence dominée

On a vu à la Proposition 1.52 un cadre permettant de permuter intégrale et limite uniforme. Cela n'est en général pas vrai pour la convergence simple (et donc encore moins pour la convergence presque partout !) et il est bon de garder à l'esprit les contre-exemples suivants.

Exemple 2.11

Les suites de fonctions $(\mathbf{1}_{[n,n+1]})_n$ et $(n\mathbf{1}_{[0,1/n]})_n$ convergent vers 0 presque partout (et même partout pour la première) alors que chaque terme de ces suites a pour intégrale 1.

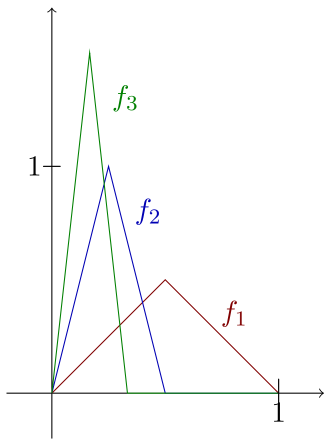

FIGURE 2.1 – Les fonctions f_n de l'exemple 2.12.

Exemple 2.12

Bien sûr, on peut également produire des contre-exemples avec des fonctions continues (voire de classe $\mathscr{C}^\infty$) : pour $n \in \mathbf{N}^\star$, on pose f_n la fonction affine par morceaux allant de $[0, 1]$ dans $\mathbf{R}_+$, et définie par

$$f_n(x) = \begin{cases} n^2 x & \text{si } 0 \leq x \leq 1/(2n) \\ n - n^2 x & \text{si } 1/(2n) \leq x \leq 1/n \\ 0 & \text{si } 1/n \leq x \end{cases}$$

(voir la figure 2.1). D'une part, l'intégrale de f_n est constante égale à $1/4$: le domaine situé sous la courbe est un triangle de base $1/n$ et de hauteur $1/(2n)$. D'autre part, les fonctions f_n convergent simplement vers la fonction nulle. En effet, soit $x \in [0, 1]$.

- Si $x = 0$, alors pour tout $n \in \mathbf{N}^*$, $f(x) = f_n(x) = 0$.
- Si $x > 0$, alors $f_n(x) = 0$ pour tout $n \geq 1/x$, en particulier la suite $(f_n(x))_n$ converge vers 0.

Ainsi, la suite (f_n) converge simplement vers la fonction nulle mais son intégrale — constante égale à $1/4$ — ne converge pas vers l'intégrale de la fonction nulle (qui vaut 0).

Par conséquent, si on veut espérer un énoncé d'interversion intégrale-limite simple, il nous faudra une hypothèse supplémentaire. La plus classique est celle de *domination*, qui donne le théorème fondamental suivant.

Énoncé indispensable 2.13 (*Théorème de convergence dominée*)

Soit $A \subset \mathbf{R}^d$ un ensemble mesurable, et $(f_n)_{n\in\mathbf{N}}$ une suite de fonctions intégrables sur A telle que :

- *f_n converge presque partout vers une fonction f ;*
- *il existe une fonction mesurable, positive et intégrable g telle que pour tout $n \in \mathbf{N}$, on ait $|f_n| \leq g$ presque partout.*

Alors f est intégrable et

$$\int_A f = \lim_{n\to+\infty} \int_A f_n.$$

De plus, $\lim_{n\to+\infty} \int_A |f_n - f| = 0$.

On dit alors que la fonction g *domine* la suite $(f_n)_n$, ou que g est un *chapeau intégrable* pour f.

Démonstration : Comme $|f_n| \leq g$ pour tout n, par passage à la limite, on en déduit que $|f| \leq g$ et que donc f est intégrable.

Remarquons que par l'inégalité triangulaire (Proposition 1.42), on a

$$\left|\int_A f_n - \int_A f\right| \leq \int_A |f_n - f|,$$

si bien qu'il suffit de montrer que $\int_A |f_n - f| \to_{n\to+\infty} 0$.

L'idée est d'utiliser le Théorème de Beppo-Levi, donc de se ramener au cas d'une suite croissante de fonctions positives. Par hypothèse la suite de fonctions positives $|f - f_n|$ tend vers 0 presque partout et on en déduit que la suite décroissante $h_n := \sup_{k\geq n} |f - f_k|$ converge vers 0 également. Par inégalité triangulaire et hypothèse de domination, nous avons $h_n \leq 2g$. Cela permet de se ramener aux hypothèses du Théorème de Beppo-Levi : $(2g - h_n)_n$ est alors une suite croissante de fonctions positives et elle converge vers $2g$. Nous avons donc que

$$\lim_{n\to+\infty} \int_A (2g - h_n) = \int_A 2g.$$

Puisque g est intégrable il en est de même de $|h_n| \leq 2g$ et il est donc licite d'écrire

$$\int_A (2g - h_n) = \int_A 2g - \int_A h_n,$$

ce qui, combiné à la convergence précédente, établit que

$$\lim_{n\to+\infty} \int_A h_n = 0.$$

Par définition $|f - f_n| \leq h_n$ et la preuve est donc terminée, par croissance de l'intégrale. ■

Remarque 2.14

Le cas de la convergence uniforme vu à la proposition 1.52 rentre dans le cadre du théorème de convergence dominée, puisque (on l'a vu lors de la preuve de la proposition 1.52) la suite est alors dominée.

Exercice 2.15

Démontrer que si $f \in \mathscr{L}^1(\mathbf{R})$ alors la fonction $x \mapsto \displaystyle\int_0^x f$ est une fonction continue.

On a déjà vu au chapitre précédent un moyen de se ramener au calcul d'intégrales sur des ensembles compacts pour des fonctions positives (Corollaire 1.47). Le cas des fonctions à valeurs complexes requiert (bien sûr) une hypothèse d'intégrabilité et est traité par l'énoncé qui suit.

Corollaire 2.16 (*Réduction à un compact, cas intégrable*)

Soit $(A_n)_n$ une suite croissante de sous-ensembles mesurables de $\mathbf{R}^d$, et $A = \bigcup_{n\in\mathbf{N}} A_n$. Soit f une fonction intégrable sur A. Alors

$$\int_{A_n} f \underset{n\to+\infty}{\longrightarrow} \int_A f.$$

Remarque 2.17

Ce corollaire est souvent utile lorsqu'on souhaite se ramener à des fonctions sur des ensembles de mesures finies. Il est fréquemment utilisé dans le cas $A = \mathbf{R}$ et $A_n = [-n, n]$.

Démonstration : Il suffit d'appliquer le théorème de convergence dominée à la suite de fonctions $f_n = \mathbf{1}_{A_n} f$ qui est dominée par $g = |f|$ et converge simplement vers f. ■

Dans le premier chapitre, on a vu une version du théorème de Beppo-Levi pour les séries de fonctions (Corollaire 1.8). Il existe aussi une version du théorème de convergence dominée pour les séries de fonctions *de signe quelconque*, qui contient lui aussi une hypothèse de domination.

Théorème 2.18 (*Interversion de $\sum$ et $\int$*)

Soit $A \subset \mathbf{R}^d$ un ensemble mesurable, et $(f_n)_n$ une suite de fonctions mesurables sur A, telles que

$$\sum_{n=0}^{+\infty} \int_A |f_n| < +\infty.$$

Alors

$$\int_A \sum_{n=0}^{+\infty} f_n = \sum_{n=0}^{+\infty} \int_A f_n.$$

L'hypothèse de convergence de la somme des intégrales est analogue à celle de domination du théorème de convergence dominée.

La conclusion de ce théorème sous-entend en fait plusieurs faits :

- la série $\sum_{n=0}^{+\infty} f_n(x)$ est convergente pour presque tout $x \in A$; en fait, la preuve du théorème montre que cette série est absolument convergente presque partout ;
- la fonction $x \mapsto \sum_{n=0}^{+\infty} f_n(x)$ est intégrable sur A ;
- la série numérique $\sum_{n=0}^{+\infty} \int_A f_n$ est convergente.

Démonstration : Soit $\varphi : A \to \mathbf{R}_+$ définie par $\varphi(x) = \sum_{n\in\mathbf{N}} |f_n(x)|$. Alors par le corollaire 1.8 du théorème de convergence croissante, on a

$$\int_A \varphi = \sum_{n\in\mathbf{N}} \int |f_n| < +\infty$$

(la dernière inégalité est juste l'hypothèse de l'énoncé). Cela implique que φ est intégrable ; en particulier on a $\varphi(x) < +\infty$ presque partout. Ainsi, la série $\sum_{n\in\mathbf{N}} f_n(x)$ est absolument convergente pour presque tout $x \in A$. On peut donc appliquer le théorème de convergence dominée à la suite $(\sum_{k=0}^{n} f_k)_n$, dominée par φ, qui donne

$$\int_A f = \lim_{n\to+\infty} \int_A \sum_{k=0}^{n} f_k = \lim_{n\to+\infty} \sum_{k=0}^{n} \int_A f_k = \sum_{k=0}^{+\infty} \int_A f_k. \qquad \blacksquare$$

2.3 Intégrales à paramètres

Maintenant qu'on a à notre disposition une théorie de l'intégration efficace, avec un théorème puissant de passage à la limite (le théorème de convergence dominée), on peut s'en servir pour définir de nouvelles fonctions. En fait, beaucoup de fonctions classiques sont (ou peuvent être) définies à partir d'une intégrale : le logarithme népérien, la fonction Γ mais aussi la transformée de Fourier (comme on le verra au chapitre 4), la transformée de Laplace...

Dans tout ce paragraphe, on considère une fonction

$$\begin{aligned} f : A \times \Lambda &\longrightarrow \mathbf{C} \\ (x,t) &\longmapsto f(x,t), \end{aligned}$$

où $A \subset \mathbf{R}^d$ est un ensemble mesurable, et $\Lambda \subset \mathbf{R}^n$ un ouvert. On cherche alors essentiellement à savoir à quelles conditions la fonction suivante

$$\begin{aligned} F : \Lambda &\longrightarrow \mathbf{C} \\ t &\longmapsto \int_A f(x,t)\,\mathrm{d}x \end{aligned} \tag{2.1}$$

est bien définie et quelles propriétés de f sont transmises à la fonction F. Les deux résultats que nous allons présenter se basent sur le théorème de convergence dominée.

Théorème 2.19 (*Continuité sous domination*)

Sous les hypothèses suivantes :

(*i*) *pour tout* $t \in \Lambda$, *la fonction* $x \mapsto f(x,t)$ *est intégrable sur* A ;

(*ii*) *pour presque tout* $x \in A$, *la fonction* $t \mapsto f(x,t)$ *est continue sur* Λ ;

(*iii*) *il existe une fonction intégrable* $g : A \to \mathbf{R}_+$ *telle que pour tout* $t \in \Lambda$

$$|f(x,t)| \leq g(x) \quad \textit{pour presque tout } x \in A;$$

la fonction F *est continue.*

Remarque 2.20

1. Dans le théorème de convergence dominée, on avait besoin d'une domination indépendante de n. Ici, la variable discrète n est remplacée par la variable continue t, il est donc naturel qu'on ait besoin d'une majoration uniforme en t.
2. Si dans (*ii*) on suppose seulement que $t \mapsto f(x,t)$ est continue en un point $t_0 \in \Lambda$, on obtient que F est continue en t_0

Démonstration : Tout d'abord, la condition (*i*) assure que la fonction F est bien définie. Ensuite, on utilise la caractérisation séquentielle de la continuité : soit $t_0 \in \Lambda$, ainsi que $(t_n)_n \in \Lambda^{\mathbf{N}}$ une suite tendant vers t_0. Alors par (*ii*), $f(x,t_n) \to_{n\to+\infty} f(x,t_0)$ pour presque tout $x \in A$. D'autre part, par (*iii*), $|f(x,t_n)| \leq g(x)$ pour tout $n \in \mathbf{N}$ et presque tout $x \in A$; plus précisément, cela veut dire que pour tout $n \in \mathbf{N}$ on dispose d'un ensemble N_n négligeable pour lequel, si $x \notin N_n$, alors l'inégalité précédente est vérifiée. Une union dénombrable d'ensembles négligeables étant négligeable, l'ensemble $N_\infty := \cup_n N_n$ est toujours négligeable. Nous sommes alors en mesure d'appliquer le théorème de convergence dominée : la suite de fonctions $x \mapsto f(x,t_n)$ converge presque partout vers $x \mapsto f(x,t_0)$, et elle est dominée (pour $x \notin N_\infty$, donc presque partout) par la fonction intégrable g. Cela assure donc que

$$F(t_n) = \int_A f(x,t_n)\,\mathrm{d}x \underset{n\to+\infty}{\longrightarrow} \int_A f(x,t_0)\,\mathrm{d}x = F(t).$$

Cela montre la continuité de F en t_0. ■

Exemple 2.21

Soit f une fonction mesurable intégrable sur $\mathbf{R}_+^*$. On définit sa *transformée de Laplace* F sur $\mathbf{R}_+$ par $F(p) = \int_0^{+\infty} e^{-px} f(x)\,\mathrm{d}x$. Alors F est continue sur $\mathbf{R}_+$. Pour voir ça, il suffit d'appliquer le théorème précédent à $f(x,p) = e^{-px}f(x)$, qui et dominée par $|f(x,p)| = |e^{-px}f(x)| \leq |f(x)|$.

Exemple 2.22

Tout comme pour le théorème de convergence dominée, l'hypothèse de domination est ici cruciale : en s'inspirant de la suite de fonctions $(f_n)_n$ définie à l'Exemple 2.12, on peut poser

$$f(x,t) = \begin{cases} \frac{x}{t^2} & \text{si } 0 \le x \le t/2 \\ \frac{1}{t} - \frac{x}{t^2} & \text{si } t/2 \le x \le t \\ 0 & \text{si } t \le x \end{cases}$$

pour tout $t > 0$, et $f(x,0) = 0$. Alors la fonction F définie par (2.1) est discontinue en 0, bien que les hypothèses (i) et (ii) soient satisfaites.

Théorème 2.23 (*Dérivabilité sous domination*)

Supposons que l'ensemble Λ soit un intervalle ouvert de $\mathbf{R}$*. Sous les hypothèses suivantes :*

(i) *pour tout $t \in \Lambda$, la fonction $x \mapsto f(x,t)$ est intégrable sur A ;*

(ii) *pour presque tout $x \in A$, la fonction $t \mapsto f(x,t)$ est dérivable sur Λ ;*

(iii) *il existe une fonction intégrable $g : A \to \mathbf{R}_+$ telle que pour presque tout $x \in A$ et tout $t \in \Lambda$,*

$$\left|\frac{\partial f}{\partial t}(x,t)\right| \le g(x);$$

la fonction F est dérivable, et

$$F'(t) = \int_A \frac{\partial f}{\partial t}(x,t)\,\mathrm{d}x.$$

Remarques 2.24

- Attention, même si on ne souhaite que la dérivabilité en un point $t_0 \in \Lambda$, il faut quand même avoir l'hypothèse (ii) sur un petit intervalle ouvert contenant t_0.
- Dans ce théorème comme dans le précédent, si Λ est par exemple donné par $\mathbf{R}$ tout entier, il est parfois impossible d'obtenir une domination indépendante de $t \in \mathbf{R}$. Il faut bien comprendre que la continuité et la dérivabilité étant des propriétés *locales* on peut *toujours* se ramener à étudier la fonction F sur un intervalle (ou une partie) bornée.
- En itérant le Théorème 2.23 (si f et ses dérivées partielles le permettent !) on peut bien sûr établir le caractère $\mathscr{C}^k$ de F pour $k \in \mathbf{N}^*$.

Démonstration : La fonction F est bien définie grâce à (i). Ensuite, comme pour la preuve du théorème précédent, on choisit $t_0 \in \Lambda$, ainsi que $(t_n)_n \in \Lambda^{\mathbf{N}}$ une suite tendant vers t_0, telle que $t_n \ne t_0$ pour tout $n \in \mathbf{N}$.

Pour montrer que F est dérivable en t_0, on calcule les taux d'accroissement

$$\frac{F(t_n) - F(t_0)}{t_n - t_0} = \int_A \frac{f(t_n,x) - f(t_0,x)}{t_n - t_0}\,\mathrm{d}x = \int_A h_n(x)\,\mathrm{d}x,$$

où

$$h_n(x) = \frac{f(t_n,x) - f(t_0,x)}{t_n - t_0}.$$

Le but est donc de pouvoir appliquer le théorème de convergence dominée (Théorème 2.13) à la suite $(h_n)_n$; vérifions point par point les hypothèses :

- par l'hypothèse (i), h_n est mesurable pour tout n ;
- par l'hypothèse (ii), pour (presque) tout $x \in A$, $h_n(x)$ converge vers $\frac{\partial f}{\partial t}(x, t_0)$, autrement dit la suite (h_n) converge presque partout vers $x \mapsto \frac{\partial f}{\partial t}(x, t_0)$; par ailleurs, le théorème des accroissements finis assure également que pour presque tout x et pour tout $n \in \mathbf{N}$, il existe $\xi_n \in]t_n, t_0[$ tel que $h_n(x) = \frac{\partial f}{\partial t}(x, \xi_n)$.
- en combinant l'égalité précédente avec l'hypothèse (iii) (l'union de deux ensembles négligeables l'est toujours) on en déduit que l'on a presque partout

$$\left|h_n(x)\right| \leq \sup_{t\in[t_0,t_n]} \left|\frac{\partial f}{\partial t}(x,t)\right| \leq g(x),$$

où g est une fonction positive intégrable.

Le théorème de convergence dominée s'applique donc ici et établit que

$$\lim_{n\to+\infty} \frac{F(t_n) - F(t_0)}{t_n - t_0} = \int_A \frac{\partial f}{\partial t}(x, t_0)\,\mathrm{d}x.$$

■

Exemple 2.25

Reprenons la fonction Γ vue à l'Exemple 2.4, et montrons qu'elle est dérivable. Rappelons que

$$\Gamma(t) = \int_0^{+\infty} f(x,t)\,\mathrm{d}x,$$

avec $f(x,t) = e^{-x}x^{t-1} = e^{-x+(t-1)\ln x}$. La fonction f admet une dérivée partielle par rapport à t et on a

$$\frac{\partial f}{\partial t}(x,t) = e^{-x}x^{t-1}\ln x.$$

Pour appliquer le Théorème 2.23, il faut montrer une domination sur cette dérivée. On va le faire sur tout intervalle $[a,b] \subset \mathbf{R}_+^*$, ce qui est suffisant car la dérivabilité est une propriété *locale*. Soit donc $t \in [a,b]$. On a

$$\left|\frac{\partial f}{\partial t}(x,t)\right| \leq \begin{cases} e^{-x}|\ln x|x^{a-1} & \text{si } x \in]0,1] \\ e^{-x}|\ln x|x^{b-1} & \text{si } x \in [1,+\infty[. \end{cases}$$

Définissons $g : \mathbf{R}_+^* \to \mathbf{R}_+$ par

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

et vérifions que g est bien intégrable sur $\mathbf{R}_+$. On a

$$\int_0^{+\infty} g(x)\,\mathrm{d}x = \int_0^1 e^{-x}|\ln x|x^{a-1}\,\mathrm{d}x + \int_1^{+\infty} e^{-x}|\ln x|x^{a-1}\,\mathrm{d}x.$$

Le premier terme est fini puisque la fonction $x \mapsto |\ln x|x^{a-1}$ est intégrable au voisinage de 0 (c'est le critère de Bertrand, voir l'exercice 2.10) ; le second l'est aussi puisque la fonction $x \mapsto e^{-x}|\ln x|x^{b-1}$ est majorée au voisinage de $+\infty$ par $e^{-x/2}$ (on refait le raisonnement de l'Exemple 2.4).

On peut donc appliquer le Théorème 2.23, qui nous dit que la fonction Γ est dérivable sur tout segment $[a,b] \subset \mathbf{R}_+^*$, donc en fait sur tout $\mathbf{R}_+^*$, et que de plus

$$\Gamma'(t) = \int_0^{+\infty} e^{-x} x^{t-1} \ln x \, \mathrm{d}x.$$

Avec les mêmes arguments, on montre que

$$\Gamma''(t) = \int_0^{+\infty} e^{-x} (\ln x)^2 \, x^{t-1} \, \mathrm{d}x,$$

qui est un nombre toujours positif ; cela montre que Γ est une fonction convexe.

2.4 Théorèmes de Fubini

Les théorèmes sont énoncés dans le cas de deux variables (intégrales sur $\mathbf{R}^2$), seulement pour des raisons de simplicité des notations. Dans le cas général d'intégrales sur $\mathbf{R}^d$ où $d = m + n$, on peut intervertir intégration sur $\mathbf{R}^m$ et $\mathbf{R}^n$ sous des hypothèses analogues.

Énoncé indispensable 2.26 (*Fubini-Tonelli pour les fonctions positives*)

Soit $f : \mathbf{R}^2 \to \overline{\mathbf{R}}_+$ une fonction mesurable positive. Alors les fonctions

$$x \mapsto \int_{\mathbf{R}} f(x,y) \, \mathrm{d}y, \;\text{ et } y \mapsto \int_{\mathbf{R}} f(x,y) \, \mathrm{d}x,$$

sont mesurables et on a l'égalité suivante, dans $\overline{\mathbf{R}}_+$,

$$\int_{\mathbf{R}^2} f = \int_{\mathbf{R}} \left\{ \int_{\mathbf{R}} f(x,y) \, \mathrm{d}y \right\} \mathrm{d}x = \int_{\mathbf{R}} \left\{ \int_{\mathbf{R}} f(x,y) \, \mathrm{d}x \right\} \mathrm{d}y.$$

Les preuves des théorèmes de Fubini-Tonelli et Fubini font appel à des résultats fins de théorie de la mesure, ils seront donc admis.

Remarque 2.27

Il faut bien comprendre que le Théorème de Fubini-Tonelli n'est pas juste un énoncé d'interversion d'intégrales : il énonce surtout l'égalité entre l'intégrale double (sur tout le plan $\mathbf{R}^2$) et l'intégrale des intégrales selon une variable (voir la Figure 2.2), autrement dit l'égalité entre un volume et une intégrale de surfaces. Ce théorème donne aussi un moyen de *calculer* des intégrales doubles : il suffit d'intégrer d'abord par rapport à une variable et ensuite par rapport à l'autre, et l'intégration par rapport à une seule variable réelle est parfois beaucoup plus simple notamment grâce aux résultats de la Section 2.1. La possibilité d'intervertir les signes $\int$ est alors un corollaire de ce résultat : si on a une intégrale d'intégrales, alors on a une intégrale double et on peut intervertir les intégrales : ça ne change rien si on intègre d'abord par rapport à x puis à y ou bien d'abord par rapport à y puis à x.

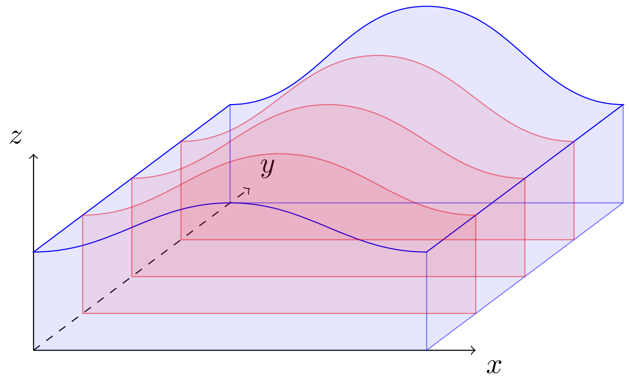

FIGURE 2.2 – Le théorème de Fubini-Tonelli : le volume sous la surface $z = f(x, y)$ (en bleu) est égal à l'intégrale sur y des aires en rouge, obtenues en découpant le volume par des plans $y = cte$ parallèles à Oxz. Bien sûr, on peut échanger les rôles de x et de y et découper selon des tranches parallèles à Oyz.

Le théorème de Fubini-Tonelli ne concerne que les fonctions positives. Dans le cas général, il faut au moins que les fonctions soient intégrables pour qu'il soit possible de définir l'intégrale ; cette condition est en fait suffisante pour avoir un théorème similaire pour les fonctions à valeurs complexes.

Énoncé indispensable 2.28 (*Fubini pour les fonctions intégrables*)

Soit $f : \mathbf{R}^2 \to \mathbf{C}$ une fonction intégrable (sur $\mathbf{R}^2$). Alors les fonctions

$$x \mapsto \int_{\mathbf{R}} f(x, y)\,\mathrm{d}y, \text{ et } y \mapsto \int_{\mathbf{R}} f(x, y)\,\mathrm{d}x,$$

(définies presque partout) sont intégrables (sur $\mathbf{R}$) et on a l'égalité suivante,

$$\int_{\mathbf{R}^2} f = \int_{\mathbf{R}} \left\{ \int_{\mathbf{R}} f(x, y)\,\mathrm{d}y \right\} \mathrm{d}x = \int_{\mathbf{R}} \left\{ \int_{\mathbf{R}} f(x, y)\,\mathrm{d}x \right\} \mathrm{d}y.$$

Remarque 2.29

En pratique, comment vérifier qu'une fonction $f : \mathbf{R}^2 \to \mathbf{C}$ est intégrable sur $\mathbf{R}^2$? On étudie la fonction positive $|f| : \mathbf{R}^2 \to \mathbf{R}_+$ et on se base sur le Théorème 2.26 de Fubini-Tonelli pour effectuer le calcul dans un sens ou l'autre : si on obtient une quantité finie, la fonction est bien intégrable.

Remarque 2.30

Certains ou certaines seront peut-être confus quant à la signification d'une intégrale double : est-ce que ça représente une aire ou un volume ? Réponse : les deux ! En fait, tout est une question d'interprétation :

— Si $f : \mathbf{R}^2 \to \mathbf{R}_+$, alors $\int_{\mathbf{R}^2} f$ représente le volume situé sous la surface définie par $z = f(x, y)$. Dans ce cas, on interprète f comme la hauteur sous la courbe ; les physiciens diront que f est homogène à une distance. Et l'intégrale double d'une distance donne un volume.

— Si en revanche on considère $A \subset \mathbf{R}^2$, alors $\int_{\mathbf{R}^2} \mathbf{1}_A$ représente l'aire de A. Dans ce cas, on interprète la fonction indicatrice comme une variable binaire, sans dimension. Et l'intégrale double d'une variable adimensionnée donne une surface.

Exemple 2.31

Retrouvons la formule exprimant l'aire d'un disque de $\mathbf{R}^2$. Grâce à la Proposition 1.22, moyennant une homothétie et une translation, il nous suffit de traiter le cas du disque unité, soit l'ensemble $\mathbf{D} := \{(x, y) \in \mathbf{R}^2 \,:\, x^2 + y^2 \leq 1\}$. Par définition de la mesure de Lebesgue, celle de $\mathbf{D}$ vaut

$$\int_{\mathbf{R}^2} \mathbf{1}_{\mathbf{D}}.$$

Comme on intègre une fonction positive, le Théorème de Fubini-Tonelli s'applique et on a ainsi

$$\int_{\mathbf{R}} \left\{ \int_{\mathbf{R}} \mathbf{1}_{\mathbf{D}}(x, y) \, \mathrm{d}x \right\} \mathrm{d}y.$$

Notons que si $|x| > 1$ ou $|y| > 1$, $(x, y) \notin \mathbf{D}$. Par ailleurs, si $|y| \leq 1$, l'appartenance $(x, y) \in \mathbf{D}$ équivaut à l'inégalité $|x| \leq \sqrt{1 - y^2}$. L'intégrale double précédente se réécrit donc

$$\int_{-1}^{1} \left\{ \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} 1 \, \mathrm{d}x \right\} \mathrm{d}y = 2 \int_{-1}^{1} \sqrt{1 - y^2} \, \mathrm{d}y = 4 \int_{0}^{1} \sqrt{1 - y^2} \, \mathrm{d}y,$$

où la dernière égalité est obtenue par parité. Finalement en appliquant le Théorème 2.5 de changement de variables avec la fonction $\sin : [0, \pi/2] \to [0, 1]$ en posant $y = \sin(z)$, on a finalement

$$\lambda(\mathbf{D}) = 4 \int_{0}^{\pi/2} \sqrt{1 - \sin(z)^2} \cos(z) \, \mathrm{d}z = 4 \int_{0}^{\pi/2} \cos(z)^2 \, \mathrm{d}z.$$

On se demande parfois dans les changements de variables trigonométriques pourquoi on choisit une fonction plutôt que l'autre. En posant ici $y = \cos(z)$ on obtient

$$\lambda(\mathbf{D}) = 4 \int_{0}^{\pi/2} \sin(z)^2 \, \mathrm{d}z,$$

et en faisant la moyenne des deux résultats il vient

$$\lambda(\mathbf{D}) = \pi.$$

On peut également utiliser le Théorème de Fubini pour calculer les coordonnées du centre de gravité d'un sous-ensemble de $\mathbf{R}^d$, une notion qui généralise le concept d'isoba-

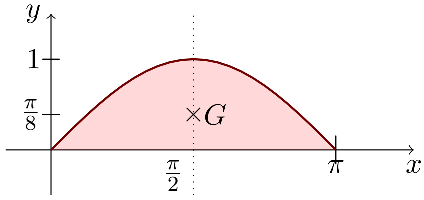

FIGURE 2.3 – L'ensemble A de l'exemple 2.33

rycentre, mais avec une « sommation continue ».

Définition 2.32

Soit A un sous-ensemble mesurable borné de $\mathbf{R}^d$, de mesure positive. Le *centre de gravité* de A est le point $G = (g_1, \ldots, g_d) \in \mathbf{R}^d$ sont les coordonnées sont définies par

$$g_i = \frac{1}{\lambda(A)} \int_A x_i \, \mathrm{d}x_1 \ldots \mathrm{d}x_d.$$

Exemple 2.33

Soit $A \subset \mathbf{R}^2$ le domaine délimité par l'axe des abscisses et par le graphe $\{(x, \sin x) \mid 0 \leq x \leq \pi\}$ (voir la Figure 2.3). On veut calculer les coordonnées de son centre de gravité.

On commence par calculer la mesure de A, autrement dit son aire. Elle est donnée par

$$\int_0^\pi \sin(x) \, \mathrm{d}x = 2.$$

Ensuite, on voit que le graphe est symétrique par rapport à l'axe $x = \pi/2$, ce qui implique que l'abscisse du centre de gravité est $g_1 = \pi/2$. Ensuite, il faut calculer l'ordonnée, ce qu'on fait en appliquant la formule :

$$g_2 = \frac{1}{\lambda(A)} \int_A x_2 \, \mathrm{d}x_1 \, \mathrm{d}x_2.$$

Par le théorème de Fubini-Tonelli, on peut séparer cette intégrale :

$$\begin{aligned} g_2 &= \frac{1}{2} \int_{x_1=0}^{\pi} \left\{ \int_0^{\sin(x_1)} x_2 \, \mathrm{d}x_2 \right\} \mathrm{d}x_1 \\ &= \frac{1}{2} \int_{x_1=0}^{\pi} \left[\frac{x_2^2}{2} \right]_{x_2=0}^{x_2=\sin(x_1)} \mathrm{d}x_1 \\ &= \frac{1}{4} \int_{x_1=0}^{\pi} \sin(x_1)^2 \, \mathrm{d}x_1 \\ &= \frac{\pi}{8}. \end{aligned}$$

Au final, $G = (\pi/2, \pi/8)$.

Terminons par un exemple où la permutation des intégrales échoue à cause du fait que la fonction de deux variables n'est pas intégrable.

Exemple 2.34

Soit la fonction

$$\begin{aligned} f : \mathbf{R}^2 &\longrightarrow \mathbf{R} \\ (x,y) &\longmapsto \frac{x^2-y^2}{(x^2+y^2)^2}\mathbf{1}_{]0,1[}(x)\mathbf{1}_{]0,1[}(y). \end{aligned}$$

Pour tout $x \in]0,1[$, on a

$$\int_{\mathbf{R}} f(x,y)\,\mathrm{d}y = \int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2}\,\mathrm{d}y = \left[\frac{y}{x^2+y^2}\right]_{y=0}^{y=1} = \frac{1}{x^2+1},$$

et donc

$$\int_{\mathbf{R}} \left\{\int_{\mathbf{R}} f(x,y)\,\mathrm{d}y\right\}\mathrm{d}x = \int_0^1 \frac{1}{x^2+1}\,\mathrm{d}x = \left[\arctan x\right]_{x=0}^{x=1} = \frac{\pi}{4}.$$

Mais puisque $f(x,y) = -f(y,x)$, on voit que

$$\int_{\mathbf{R}} \left\{\int_{\mathbf{R}} f(x,y)\,\mathrm{d}x\right\}\mathrm{d}y = \int_{\mathbf{R}} \left\{\int_{\mathbf{R}} f(y,x)\,\mathrm{d}y\right\}\mathrm{d}x = -\int_{\mathbf{R}} \left\{\int_{\mathbf{R}} f(x,y)\,\mathrm{d}y\right\}\mathrm{d}x = -\frac{\pi}{4},$$

et donc

$$\int_{\mathbf{R}} \left\{\int_{\mathbf{R}} f(x,y)\,\mathrm{d}x\right\}\mathrm{d}y \neq \int_{\mathbf{R}} \left\{\int_{\mathbf{R}} f(x,y)\,ddy\right\}\mathrm{d}x.$$

On peut en effet vérifier que la fonction n'est pas intégrable : par le Théorème 2.26 de Fubini-Tonelli,

$$\begin{aligned} \int_{\mathbf{R}^2} |f| &\geq \int_{\mathbf{R}^2} \mathbf{1}_{x>y}|f(x,y)|\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_0^1 \left\{\int_0^x \frac{x^2-y^2}{(x^2+y^2)^2}\,\mathrm{d}y\right\}\mathrm{d}x \\ &= \int_0^1 \left[\frac{y}{x^2+y^2}\right]_{y=0}^{y=x}\mathrm{d}x \\ &= \int_0^1 \frac{1}{2x}\,\mathrm{d}x = \infty. \end{aligned}$$

2.5 Formule générale de changement de variables

Définition 2.35 ($\mathscr{C}^1$-*difféomorphisme*)

Soient U et V deux ouverts de $\mathbf{R}^d$. On dit que $\varphi : U \to V$ est un $\mathscr{C}^1$-*difféomorphisme de U sur V* si :

(i) φ est bijective de U sur V ;
(ii) les dérivées partielles de φ existent et sont continues sur U ;
(iii) pour tout $x \in U$, la *matrice jacobienne* J_φ de $\varphi := (\varphi_1, \cdots, \varphi_d)$ définie par

$$\mathrm{J}_\varphi(x) = \begin{pmatrix} \frac{\partial \varphi_1}{\partial x_1}(x) & \cdots & \frac{\partial \varphi_1}{\partial x_d}(x) \\ \vdots & & \vdots \\ \frac{\partial \varphi_d}{\partial x_1}(x) & \cdots & \frac{\partial \varphi_s}{\partial x_d}(x) \end{pmatrix}$$

est inversible en tout point $x \in U$, *i.e.* si $\det \mathrm{J}_\varphi(x) \neq 0$.

Énoncé indispensable 2.36 (*Formule générale de changement de variables*)

Soient U et V deux ouverts de $\mathbf{R}^d$ et $\varphi : U \to V$ un $\mathscr{C}^1$-difféomorphisme de U sur V. Une fonction f est intégrable sur V si et seulement si la fonction $(f \circ \varphi)|\det \mathrm{J}_\varphi|$ l'est sur U, et dans ce cas on a l'égalité

$$\int_V f(x)\,\mathrm{d}x = \int_U f(\varphi(y))\,|\det \mathrm{J}_\varphi(y)|\,\mathrm{d}y.$$

Remarques 2.37

- On dit alors que cette égalité est obtenue par le changement de variable « $x = \varphi(y)$ » ; l'élément différentiel $\mathrm{d}x$ est transformé en $|\det \mathrm{J}_\varphi(y)|\,\mathrm{d}y$.
- Lorsque U et V sont des intervalles de $\mathbf{R}$ on retrouve le cas réel donné par le Théorème 2.5 *modulo* la remarque suivante : dans le Théorème 2.5, le changement de variables peut ne pas être bijectif, le terme différentiel est signé et l'ordre des bornes peut être échangé ; alors que dans la formule générale de changement de variable, le terme différentiel est positif (grâce à la valeur absolue) et on ne s'intéresse pas à l'ordre des bornes.

L'idée de la preuve est qu'un $\mathscr{C}^1$-difféomorphisme est localement bien approché par son développement de Taylor à l'ordre 1, qui est affine, et une application affine multiplie la mesure par le déterminant de sa partie linéaire (Proposition 1.22). La mise en forme rigoureuse de cette idée (en particulier la gestion des termes de reste) est bien trop technique pour ce cours : nous en admettrons la preuve.

Exemple 2.38

Changement de variables affine. C'est le plus simple des changements de variables : si $f : \mathbf{R}^d \to \mathbf{C}$ est intégrable, si $\lambda \in \mathbf{R}^\star$ et $a \in \mathbf{R}^d$, alors posant $x = \lambda y + a$ on obtient

$$\int_{\mathbf{R}^d} f(x)\,\mathrm{d}x = |\lambda^d| \int_{\mathbf{R}^d} f(\lambda y + a)\,\mathrm{d}y.$$

Coordonnées polaires La configuration des coordonnées polaires est représentée sur la Figure 2.4.

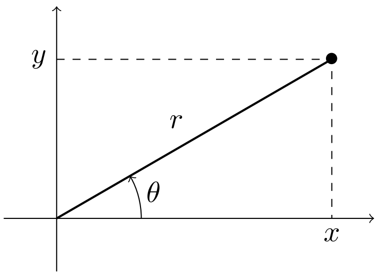

FIGURE 2.4 – Les coordonnées polaires

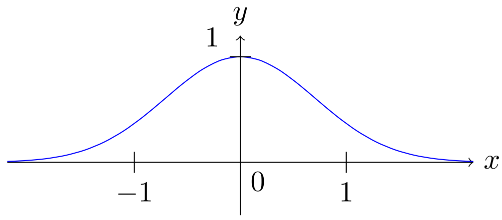

FIGURE 2.5 – Le graphe de la gaussienne $x \mapsto e^{-x^2}$.

Soit $D \subset \mathbf{R}^2$ la demi-droite $\mathbf{R}_-^* \times \{0\}$ et

$$\begin{array}{rcl} \varphi : \mathbf{R}_+^\star \times]-\pi,\pi[& \longrightarrow & \mathbf{R}^2 \setminus D \\ (r,\theta) & \longmapsto & (r\cos\theta, r\sin\theta). \end{array}$$

On enlève la demi-droite D pour que φ soit bien un difféomorphisme, mais on pourrait enlever n'importe quelle demi-droite passant par l'origine (auquel cas il faut adapter l'intervalle $]-\pi,\pi[$ à l'angle de la droite).

Vérifions que φ est bien un difféomorphisme. Son aspect bijectif provient de la paramétrisation du cercle unité épointé de $(-1,0)$ par l'application $\theta \mapsto (\cos\theta, \sin\theta)$ sur $]-\pi,\pi[$. L'application φ est différentiable, et sa matrice jacobienne vérifie

$$\det \mathrm{J}_\varphi(r,\theta) = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r\cos^2\theta + r\sin^2\theta = r.$$

Nous avons finalement obtenu que φ est un $\mathscr{C}^1$-difféomorphisme de $\mathbf{R}_+^\star \times]-\pi,\pi[$ sur $\mathbf{R}^2 \setminus D$ le changement de variable $z = \varphi(p)$ permet donc d'écrire, pour toute fonction $f : \mathbf{R}^2 \to \mathbf{C}$ intégrable,

$$\int_{\mathbf{R}^2\setminus D} f(z)\,\mathrm{d}z = \int_{\mathbf{R}_+^\star \times]-\pi,\pi[} f(\varphi(p))\,\left|\det \mathrm{J}_\varphi(p)\right| \mathrm{d}p.$$

Puisque D est de mesure nulle dans le plan (voir Exercice 1.17), en écrivant $p = (r,\theta)$ dans l'égalité précédente, on a finalement grâce au Théorème de Fubini pour les fonctions intégrables

$$\boxed{\int_{\mathbf{R}^2} f = \int_0^\infty \left\{ \int_{-\pi}^{\pi} f(r\cos\theta, r\sin\theta)\,\mathrm{d}\theta \right\} r\,\mathrm{d}r.}$$

Remarque 2.39

Lorsque dans l'espace $\mathbf{R}^3$, on choisit deux des coordonnées (typiquement x et y) et on opère un changement de variables en polaires sur ces coordonnées en laissant la dernière fixe (dans ce cas, z), l'opération est traditionnellement appelée *changement de variables cylindrique.*

Exemple 2.40

Soit $a > 0$. Utilisons le changement de variables en coordonnées polaires pour calculer l'intégrale de la gaussienne (représentée sur la Figure 2.5 dans le cas $a = 1$) :

$$I_a = \int_{\mathbf{R}} e^{-ax^2}\,\mathrm{d}x.$$

On remarque que

$$\begin{aligned} I_a^2 &= \left(\int_{\mathbf{R}} e^{-ax^2}\,\mathrm{d}x\right)\left(\int_{\mathbf{R}} e^{-ay^2}\,\mathrm{d}y\right) \\ &= \int_{\mathbf{R}} e^{-ax^2}\left\{\int_{\mathbf{R}} e^{-ay^2}\,\mathrm{d}y\right\}\mathrm{d}x \\ &= \int_{\mathbf{R}^2} e^{-a(x^2+y^2)}\mathrm{d}x\,\mathrm{d}y. \end{aligned}$$

Ainsi, grâce au Théorème 2.26 de Fubini-Tonelli, I_a^2 n'est rien d'autre que l'intégrale de la fonction $z \mapsto e^{-a\|z\|^2}$ sur $\mathbf{R}^2$. La fonction intégrée est intégrable : on voit que I_a est fini en appliquant le critère de Riemann, ainsi la fonction est intégrable car elle est positive et son intégrale I_a^2 est finie. On peut donc appliquer le changement de variables polaire, qui permet d'écrire

$$\begin{aligned} I_a^2 &= \int_0^{+\infty}\left\{\int_{-\pi}^{\pi}\mathrm{d}\theta\right\}e^{-ar^2}r\mathrm{d}r \\ &= \frac{\pi}{a}\int_0^{+\infty} e^{-ar^2}2ar\,\mathrm{d}r. \end{aligned}$$

Puisque $r \mapsto 2are^{-ar^2}$ est la dérivée de $r \mapsto -e^{-ar^2}$, en associant le théorème fondamental de l'analyse au Corollaire 2.16 du théorème de Beppo-Levi, on peut finalement écrire

$$\begin{aligned} I_a^2 &= \frac{\pi}{a}\lim_{n\to+\infty}\int_0^n e^{-ar^2}2ar\,\mathrm{d}r \\ &= \frac{\pi}{a}\lim_{n\to+\infty}\Big[-e^{-ar^2}\Big]_{r=0}^{r=n} \\ &= \frac{\pi}{a}, \end{aligned}$$

nous avons alors établi que $I_a = \sqrt{\frac{\pi}{a}}$.

Coordonnées sphériques Commençons par retrouver les formules de passage en coordonnées sphériques (voir la Figure 2.6). Soit P un point ayant pour coordonnées cartésiennes (x, y, z). On définit le point M comme étant le projeté de P sur le plan Oxy. L'angle orienté $\widehat{xOM}$ est noté θ, et l'angle orienté $\widehat{zOP}$ noté ϕ. Enfin, la distance OP est notée r.

On a $z = r\cos\phi$. D'autre part, en se plaçant dans le plan xOy, on voit que $x = OM\cos\theta$ et $y = OM\sin\theta$. Or, en se plaçant dans le plan POM, on voit que $OM = r\sin\phi$.

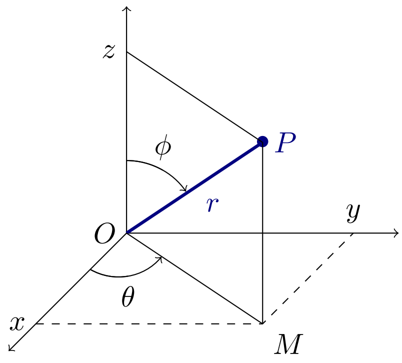

FIGURE 2.6 – Les coordonnées sphériques

On en déduit les formules :

$$\begin{cases} x = r\cos\theta\sin\phi, \\ y = r\sin\theta\sin\phi, \\ z = r\cos\phi. \end{cases}$$

Si on pose

$$\begin{array}{rcl} \varphi : \mathbf{R}_+^\star \times]-\pi,\pi[\times]0,\pi[& \longrightarrow & \mathbf{R}^3 \\ (r,\theta,\phi) & \longmapsto & (r\cos\theta\sin\phi,\ r\sin\theta\sin\phi,\ r\cos\phi), \end{array}$$

et Q le demi-plan $\{y = 0, x \leq 0\}$, alors on voit que φ est un difféomorphisme sur son image $\mathbf{R}^3 \setminus Q$. Le déterminant de sa jacobienne vaut

$$\det J\varphi(r,\theta,\phi) = \begin{vmatrix} \cos\theta\sin\phi & -r\sin\theta\sin\phi & r\cos\theta\cos\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\phi & 0 & -r\sin\phi \end{vmatrix}.$$

Le développement par rapport à la dernière ligne donne :

$$\begin{aligned} \det J\varphi(r,\theta,\phi) &= \cos\phi \begin{vmatrix} -r\sin\theta\sin\phi & r\cos\theta\cos\phi \\ r\cos\theta\sin\phi & r\sin\theta\cos\phi \end{vmatrix} - r\sin\phi \begin{vmatrix} \cos\theta\sin\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi \end{vmatrix} \\ &= -r^2\cos\phi\big(\sin^2\theta\sin\phi\cos\phi + \cos^2\theta\sin\phi\cos\phi\big) \\ &\quad - r^2\sin\phi\big(\cos^2\theta\sin^2\phi + \sin^2\theta\sin^2\phi\big) \\ &= -r^2\cos^2\phi\sin\phi - r^2\sin^3\phi \\ &= -r^2\sin\phi. \end{aligned}$$

Et la formule de changement de variables implique que, pour une fonction mesurable intégrable de $\mathbf{R}^3$:

$$\boxed{\int_{\mathbf{R}^3} f(x,y,z)dxdydz = \int_{\mathbf{R}_+^\star\times]-\pi,\pi[\times]0,\pi[} f\big(r\cos\theta\sin\phi,\ r\sin\theta\sin\phi,\ r\cos\phi\big)\, r^2\sin\phi\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\phi.}$$

Attention ! La définition des angles θ et ϕ dépend des auteurs, parfois ϕ est pris par rapport au segment $[0M]$ et non l'axe des z. Méfiance !

Exemple 2.41

Calculons le volume d'une boule B de rayon R dans l'espace de dimension 3. Ce volume est donné par

$$\mathrm{Vol}(B) = \int_B 1 \,\mathrm{d}x \,\mathrm{d}y \,\mathrm{d}z.$$

Par un changement de variables en coordonnées sphériques, et le théorème de Fubini, on obtient :

$$\begin{aligned}\mathrm{Vol}(B) &= \int_0^{\pi} \left\{ \int_{-\pi}^{\pi} \left\{ \int_0^R r^2 \sin\phi \,\mathrm{d}r \right\} \mathrm{d}\theta \right\} \mathrm{d}\phi \\ &= \left(\int_0^{\pi} \sin\phi \,\mathrm{d}\phi \right) \left(\int_{-\pi}^{\pi} \mathrm{d}\theta \right) \left(\int_0^R r^2 \,\mathrm{d}r \right) \\ &= \big[-\cos\phi \big]_{\phi=0}^{\phi=\pi} \times 2\pi \times \left[\frac{r^3}{3} \right]_{r=0}^{r=R} = \frac{4}{3}\pi R^3.\end{aligned}$$

Chapitre 3

Convolution

La *convolution* est une opération qui à deux fonctions en associe une troisième pouvant être vue comme un « mélange » des deux premières. L'étymologie [1] du terme est éclairante puisque la convolution est « l'acte de s'enrouler autour de quelque chose ». Étant données deux fonctions f et g, la convolution pourra ainsi être vue en tout point comme une moyenne des valeurs de f avec des pondérations fournies par la fonction g (ou l'inverse). Nous fournirons dans la Section 3.1 un cadre fonctionnel précis dans lequel nous pourrons définir cette opération. Cette définition se base sur l'intégrale de Lebesgue et le théorème de Fubini que nous avons présenté dans le chapitre précédent. Ensuite nous verrons rapidement en quoi cette « moyennisation » peut être un outil de régularisation (ou lissage) des fonctions. Mieux : raffiner la zone de pondération la convolution permet de fournir un procédé d'approximation extrêmement efficace que nous présenterons à travers le concept d'*approximation de l'unité* dans la Section 3.3. Nous en déduirons l'un des résultats les plus élaborés de ce cours qui est l'approximation de toute fonction intégrable par une suite de fonctions extrêmement régulières. Tous les résultats de ce chapitre sont abordés en dimension 1 afin d'en simplifier la présentation mais se généralisent sans difficulté (autre que des artefacts techniques) aux dimensions supérieures.

1. Le verbe à employer est *convoluer* (ou plus rarement : *convoluter*). Même si dans l'esprit on s'en rapproche, on ne *convole* pas les fonctions !

Résumé :

- On note $\mathscr{L}^1(\mathbf{R})$ l'ensemble des fonctions intégrables et $\mathscr{L}^2(\mathbf{R})$ l'ensemble des fonctions de carré intégrable.
- Sur ces ensembles on introduit respectivement
$$\|f\|_1 := \int_{\mathbf{R}} |f|, \quad \|f\|_2 := \left(\int_{\mathbf{R}} |f|^2\right)^{1/2}.$$
- Ces expressions définissent des normes sur les espaces $L^1(\mathbf{R})$ et $L^2(\mathbf{R})$ obtenus par quotient de sur $\mathscr{L}^1(\mathbf{R})$ et $\mathscr{L}^2(\mathbf{R})$ par la relation « presque partout égal » (Définition 3.6).
- La convolution de deux fonctions est définie par la formule (Énoncé indispensable 3.20)
$$(f \star g)(x) = \int_{\mathbf{R}} f(x-y)g(y)\,\mathrm{d}y.$$
- Pour $f \in L^1(\mathbf{R})$, cette opération est bien définie si $g \in L^1(\mathbf{R})$ (Proposition 3.23), ou $g \in L^2(\mathbf{R})$ (Corollaire 3.24) ou $g \in \mathscr{C}_b^0(\mathbf{R})$ (Proposition 3.25) ; dans ce cernier cas $f \star g$ a la même régularité que g.
- À l'aide du concept d'approximation de l'unité (Définition 3.28) on peut montrer que $\mathscr{C}_c^\infty(\mathbf{R})$ est dense dans $L^1(\mathbf{R})$ et dans $L^2(\mathbf{R})$ (Corollaire 3.34).

3.1 Les espaces $L^1(\mathbf{R})$, $L^2(\mathbf{R})$ et $\mathscr{C}_b^0(\mathbf{R})$

Définition 3.1

On introduit les ensembles suivants

$$\begin{aligned}\mathscr{L}^1(\mathbf{R}) &:= \{f : \mathbf{R} \to \mathbf{C} \,:\, f \text{ intégrable}\},\\ \mathscr{L}^2(\mathbf{R}) &:= \{f : \mathbf{R} \to \mathbf{C} \,:\, |f|^2 \text{ intégrable}\},\\ \mathscr{C}_b^0(\mathbf{R}) &:= \{f : \mathbf{R} \to \mathbf{C} \,:\, f \text{ continue et bornée}\}.\end{aligned}$$

Proposition 3.2

$\mathscr{L}^1(\mathbf{R})$, $\mathscr{L}^2(\mathbf{R})$ et $\mathscr{C}_b^0(\mathbf{R})$ sont des espaces vectoriels.

Démonstration : Pour $\mathscr{L}^1(\mathbf{R})$ et $\mathscr{C}_b^0(\mathbf{R})$, c'est l'inégalité triangulaire pour le module : si f, g sont deux éléments de l'un ou autre de ces ensembles et $\lambda \in \mathbf{C}$, alors en écrivant $|f + \lambda g| \leq |f| + |\lambda||g|$ on obtient l'intégrabilité ou le caractère borné de $f + \lambda g$. Si maintenant $f, g \in \mathscr{L}^2(\mathbf{R})$, alors par la même inégalité on a $|f + \lambda g|^2 \leq |f|^2 + |\lambda|^2|g|^2 + 2|\lambda||f||g|$ et une identité remarquable nous fournit que $2|f||g| \leq |f|^2 + |g|^2$: l'intégrabilité de $|f + \lambda g|^2$ est établie. ■

On rappelle qu'une *norme* sur un $\mathbf{C}$-espace vectoriel E est une application $N : E \to \mathbf{R}_+$ satisfaisant $N(\lambda x) = |\lambda| N(x)$ pour $\lambda \in \mathbf{C}$ (homogénéité), $N(x + y) \leq N(x) + N(y)$ (inégalité triangulaire) et $N(x) = 0 \Leftrightarrow x = 0$ (séparation).

Définition 3.3

Pour $f \in \mathscr{L}^1(\mathbf{R})$, $g \in \mathscr{L}^2(\mathbf{R})$ et $h \in \mathscr{C}_b^0(\mathbf{R})$ on pose

$$\begin{aligned}\|f\|_1 &:= \int_{\mathbf{R}} |f|,\\ \|g\|_2 &:= \left(\int_{\mathbf{R}} |g|^2\right)^{1/2},\\ \|h\|_\infty &:= \sup_{\mathbf{R}} |h|.\end{aligned}$$

Proposition 3.4

$\|\cdot\|_\infty$ est une norme sur $\mathscr{C}_b^0(\mathbf{R})$: on parle de « norme infinie » ou « norme uniforme ».

Remarque 3.5

La convergence uniforme sur un segment $[a, b]$ que nous avons manipulée au premier chapitre provient de la norme uniforme sur ce segment, soit $\sup_{[a,b]} |\cdot|$: c'est une norme sur $\mathscr{C}^0([a, b])$.

Démonstration : Pour $h \in \mathscr{C}_b^0(\mathbf{R})$, $\lambda \in \mathbf{C}$ et $x \in \mathbf{R}$ on a $|\lambda h(x)| \leq |\lambda|\|h\|_\infty$ ce qui implique $\|\lambda h\|_\infty \leq |\lambda|\|h\|_\infty$ et en dehors du cas trivial $\lambda = 0$ on peut invoquer cette même inégalité sur la fonction λh avec λ^{-1} pour obtenir $\|\lambda h\|_\infty = |\lambda|\|h\|_\infty$. On a par ailleurs, pour

$h_1, h_2 \in \mathscr{C}_b^0(\mathbf{R})$, par inégalité triangulaire du module $|h_1(x) + h_2(x)| \leq |h_1(x)| + |h_2(x)|$ pour tout $x \in \mathbf{R}$ ce qui implique $\|h_1 + h_2\|_\infty \leq \|h_1\|_\infty + \|h_2\|_\infty$ après passage à la borne supérieure. Enfin $\|h\|_\infty = 0 \Leftrightarrow h(x) = 0$ pour tout $x \in \mathbf{R}$, ce qui est précisément la séparation. ■

Dans le cas des espaces $\mathscr{L}^1(\mathbf{R})$ et $\mathscr{L}^2(\mathbf{R})$ l'axiome de séparation n'est pas vérifié pour $\|\cdot\|_1$ et $\|\cdot\|_2$ à cause de la Proposition 1.30 : on se souvient en effet que $\mathbf{1_Q}$ est un élément de $\mathscr{L}^1(\mathbf{R})$ d'intégrale nulle ! Cela conduit à la définition suivante.

Définition 3.6 (*Espaces de Lebesgue*)

En notant $\sim$ la relation d'équivalence sur l'ensemble des fonctions de $\mathbf{R}$ dans $\mathbf{C}$ par « $f \sim g$ si $f = g$ presque partout », on note $L^1(\mathbf{R})$ et $L^2(\mathbf{R})$ les ensembles quotients $\mathscr{L}^1(\mathbf{R})/\sim$ et $\mathscr{L}^2(\mathbf{R})/\sim$.

Remarque 3.7

En pratique, puisque l'intégration n'est pas sensible aux modifications sur les ensembles négligeables, on confondra de manière un peu abusive les représentants et la classe elle-même.

Proposition 3.8

$\|\cdot\|_1$ est une norme sur $L^1(\mathbf{R})$: on parle de « norme 1 ».

Démonstration : L'homogénéité vient de la linéarité de l'intégrale, la séparation de la définition de $L^1(\mathbf{R})$ et enfin l'inégalité triangulaire découle de la même propriété pour le module, associée à la croissance de l'intégrale. ■

De la même manière, $\|\cdot\|_2$ est une norme sur $L^2(\mathbf{R})$, mais pour le prouver nous aurons besoin de l'inégalité suivante.

Proposition 3.9 (*Cauchy-Schwarz*)

Pour $f, g \in L^2(\mathbf{R})$ on a $fg \in L^1(\mathbf{R})$ et

$$\left| \int_{\mathbf{R}} f\overline{g} \right| \leq \|f\|_2 \|g\|_2,$$

avec égalité si et seulement si f et g sont liées.

Remarque 3.10

On fait comme si de rien n'était, mais ce que nous entrevoyons est la structure hilbertienne de $L^2(\mathbf{R})$: cet espace peut être muni d'un produit scalaire dont la norme associée est précisément $\|\cdot\|_2$!

Démonstration : Sans perte de généralité, on peut supposer que $g \neq 0$ puisque dans le cas contraire l'inégalité est immédiate (et les deux fonctions sont bien liées). Toujours sans perte de généralité, on peut également supposer que $\|g\|_2 = 1$, et il s'agit donc de démontrer, pour une telle fonction normalisée, que

$$\left| \int_{\mathbf{R}} f\overline{g} \right| \leq \|f\|_2,$$

avec égalité si et seulement si f et g sont liées. Pour cela, on introduit la fonction

$$h := \left(\int_{\mathbf{R}} f\overline{g}\right) g.$$

Remarquons que l'intégrale apparaissant dans cette fonction est bien définie, puisqu'une identité remarquable implique que $|f\overline{g}| \leq \frac{1}{2}(|f|^2 + |g|^2)$.

Pour la fonction h, on calcule

$$\begin{aligned} \|f-h\|_2^2 = \int_{\mathbf{R}} |f-h|^2 &= \int_{\mathbf{R}} (f-h)(\overline{f-h}) \\ &= \|f\|_2^2 + \|h\|_2^2 - \int_{\mathbf{R}} f\overline{h} - \int_{\mathbf{R}} \overline{f}h. \end{aligned}$$

Par définition de h et linéarité de l'intégrale on a

$$\int_{\mathbf{R}} f\overline{h} = \overline{\left(\int_{\mathbf{R}} f\overline{g}\right)} \left(\int_{\mathbf{R}} f\overline{g}\right) = \left|\int_{\mathbf{R}} f\overline{g}\right|^2,$$

et la normalisation de g montre que ce nombre n'est rien d'autre que $\|h\|_2^2$. Puisque la conjugaison commute avec l'intégrale on a aussi

$$\int_{\mathbf{R}} \overline{f}h = \|h\|_2^2,$$

et on a en réalité démontré

$$\|f-h\|_2^2 = \|f\|_2^2 - \left|\int_{\mathbf{R}} f\overline{g}\right|^2,$$

qui est effectivement un nombre positif, nul uniquement lorsque $f = h$, soit une liaison entre f et g. ■

Corollaire 3.11

$\|\cdot\|_2$ est une norme sur $L^2(\mathbf{R})$: on parle de « norme 2 ».

Démonstration : L'homogénéité et la séparation se démontre comme pour la norme $\|\cdot\|_1$ sur $L^1(\mathbf{R})$. L'inégalité triangulaire provient de l'identité,

$$\|f+g\|_2^2 = \int_{\mathbf{R}} |f+g|^2 = \|f\|_2^2 + \|g\|_2^2 + 2\mathrm{Re}\left(\int_{\mathbf{R}} f\overline{g}\right),$$

de laquelle on déduit (grâce à l'inégalité de Cauchy-Schwarz pour la deuxième ligne)

$$\begin{aligned} \|f+g\|_2^2 &\leq \|f\|_2^2 + \|g\|_2^2 + 2\|f\overline{g}\|_1 \\ &\leq \|f\|_2^2 + \|g\|_2^2 + 2\|f\|_2\|g\|_2 = \big(\|f\|_2 + \|g\|_2\big)^2. \end{aligned}$$

■

Terminons ce premier paragraphe par deux résultats d'approximation que l'on peut exprimer à travers le concept de densité.

Définition 3.12

Soit $(E, \|\cdot\|)$ espace vectoriel normé et $A \subset B$ deux parties de E. On dit que A est *dense* dans B si tout élément de B peut être approché par une suite constituée d'éléments de A : pour tout $b \in B$ il existe $(a_n)_n \in A^{\mathbf{N}}$ telle que $(\|b - a_n\|)_n \to 0$.

Exemple 3.13

L'ensemble $\mathbf{Q}$ est dense dans $\mathbf{R}$.

Proposition 3.14

Soit $(E, \|\cdot\|)$ un espace vectoriel normé et A, B, C trois parties de E. Si A est dense dans B et B est dense dans C, alors A est dense dans C.

Démonstration : Soit $c \in C$. Pour tout $n \in \mathbf{N}^\star$, par hypothèse, on peut trouver $b_n \in B$ tel que $\|c - b_n\| \leq 1/n$ et $a_n \in A$ tel que $\|a_n - b_n\| \leq 1/n$. Par inégalité triangulaire on a alors $\|c - a_n\| \leq 2/n$ et $(a_n)_n \in A^{\mathbf{N}}$ est bien convergente vers c. ■

Définition 3.15

On dit qu'une fonction $f : \mathbf{R} \to \mathbf{C}$ est à *support compact* s'il existe un segment $[-R, R]$ en dehors duquel elle est nulle. Pour un élément de $L^1(\mathbf{R})$ ou $L^2(\mathbf{R})$, on dit qu'il est à support compact s'il admet un représentant à support compact.

Cette définition laissera sans doute le lecteur sur sa faim. Existe-t-il des fonctions dont le support n'est pas compact ? D'ailleurs, c'est quoi le support d'une fonction ? Nous n'aurons pas à utiliser cette notion en dehors du cas « compact » de la Définition 3.15, mais nous donnons tout de même ici la signification précise (à omettre en première lecture) de cette notion dans le cadre général. Il faut pour la comprendre savoir ce qu'est l'*adhérence* d'une partie de $\mathbf{R}$: il s'agit du plus petit ensemble fermé la contenant.

Définition 3.16

Le *support* d'une fonction $f : \mathbf{R} \to \mathbf{C}$ est l'adhérence de l'ensemble des points où elle ne s'annule pas, soit $\overline{\{x \in \mathbf{R} \ : \ f(x) \neq 0\}}$.

Exercice 3.17

Montrer que le support d'une fonction (au sens de la Définition 3.16) est un ensemble compact si et seulement si cette fonction est à support compact (au sens de la définition 3.15).

Le premier résultat de densité que nous allons prouver est une conséquence immédiate de la convergence dominée. Le second, beaucoup plus fin, sera lui hérité de la propriété d'approximation de l'Énoncé indispensable 1.5.

Lemme 3.18

L'ensemble $B_c(\mathbf{R})$ des fonctions mesurables bornées et à support compact est dense dans $L^1(\mathbf{R})$ et dense dans $L^2(\mathbf{R})$.

Démonstration : Tout d'abord, si f est mesurable, nulle en dehors de $[-R, R]$ et majorée en module par M, alors f est bien intégrable (et également de carré intégrable) : l'ensemble des fonctions mesurables bornées à support compact est inclus dans $L^1(\mathbf{R})$ et $L^2(\mathbf{R})$. Ensuite, si f appartient à $L^1(\mathbf{R})$ ou $L^2(\mathbf{R})$, $f_n := \mathbf{1}_{|f|\leq n}\mathbf{1}_{[-n,n]}f$ converge presque partout vers f et $|f - f_n| \leq 2|f|$, ce qui fournit une domination (en élevant au carré cette inégalité dans le cas quadratique). On a donc, par théorème de convergence dominée, $(f_n)_n \to f$ dans $L^1(\mathbf{R})$ ou $L^2(\mathbf{R})$ selon l'espace d'appartenance de f, ce qui est la caractérisation séquentielle de la densité. ∎

Lemme 3.19

L'espace $\mathscr{C}_c^0(\mathbf{R})$ des fonctions continues à support compact est dense dans $L^1(\mathbf{R})$ et dense dans $L^2(\mathbf{R})$.

Démonstration : Considérons d'abord le cas d'une fonction $f \in L^1(\mathbf{R})$ positive. Par la propriété d'approximation de l'Énoncé indispensable 1.5 on a pour tout $\varepsilon > 0$ l'existence de $\varphi_\varepsilon \in \mathscr{C}_c^0(\mathbf{R})$ telle que $\|f - \varphi_\varepsilon\|_1 < \varepsilon$. En particulier, si l'on prend $\varepsilon := 1/n$ avec $n \in \mathbf{N}^\star$ on construit ainsi une suite $(\varphi_n)_n$ approchant f dans $L^1(\mathbf{R})$. Si maintenant f est une fonction intégrable à valeurs complexes, elle se décompose ainsi

$$f = \mathrm{Re}(f)_+ - \mathrm{Re}(f)_- + i(\mathrm{Im}(f)_+ - \mathrm{Im}(g)_-),$$

où les fonctions $\mathrm{Re}(f)^\pm$ et $\mathrm{Im}(f)^\pm$ sont toutes les quatre positives et intégrables. L'argument précédent fournit pour chacune de ces quatre fonctions une approximation par une suite d'éléments de $\mathscr{C}_c^0(\mathbf{R})$ et l'inégalité triangulaire permet d'aboutir à la même propriété pour f, en utilisant le fait que $\mathscr{C}_c^0(\mathbf{R})$ est un espace vectoriel.

Il nous reste à traiter le cas quadratique, que nous allons ramener au précédent. D'après le Lemme 3.18 on sait que $B_c(\mathbf{R})$ est dense dans $L^2(\mathbf{R})$. D'après, la Proposition 3.14, il nous suffit donc de montrer que $\mathscr{C}_c^0(\mathbf{R})$ est dense dans $\mathscr{C}_c^0(\mathbf{R}) \cup B_c(\mathbf{R})$, pour la norme $\|\cdot\|_2$. Cela revient à montrer que pour tout élément $f \in B_c(\mathbf{R})$ il existe une suite $(\varphi_n)_n$ d'éléments de $\mathscr{C}_c^0(\mathbf{R})$ telle que $(\varphi_n)_n \to f$ dans $L^2(\mathbf{R})$.

Comme lors du cas précédent, il suffit de traiter le cas des fonctions à valeurs réelles, le cas à valeurs complexes s'en déduisant par décomposition selon les parties réelle et imaginaire, puis à l'aide de l'inégalité triangulaire.

Si $f \in B_c(\mathbf{R})$, c'est en particulier une fonction intégrable, et le premier paragraphe nous fournit donc une suite $(\psi_n)_n$ d'éléments de $\mathscr{C}_c^0(\mathbf{R})$ telle que $(\psi_n)_n \to f$, mais cette convergence a lieu dans $L^1(\mathbf{R})$ et il nous faut une convergence quadratique. Puisque f est dans $B_c(\mathbf{R})$, elle est majorée en module par une constante $M > 0$. Soit alors $\gamma(x)$ la fonction continue valant l'identité sur $[-M, M]$ et constante pour $|x| > M$. Il s'agit d'une fonction 1-lipschitzienne, c'est-à-dire vérifiant $|\gamma(x) - \gamma(y)| \leq |x - y|$. On a donc $\|\gamma(f) - \gamma(\psi_n)\|_1 \leq \|f - \psi_n\|_1 \to 0$. Puisque $|f| \leq M$ on a $\gamma(f) = f$ et les fonctions $\varphi_n := \gamma(\psi_n)$ sont continues, à support compact, et satisfont

$$\|f - \varphi_n\|_2 = \|\gamma(f) - \gamma(\psi_n)\|_2 \leq \sqrt{2M}\|\gamma(f) - \gamma(\psi_n)\|_1^{1/2} \to 0,$$

où l'inégalité provient de $|\gamma(f) - \gamma(\psi_n)|^2 \leq 2M|\gamma(f) - \gamma(\psi_n)|$ (puisque $|\gamma| \leq M$). ∎

Il est possible de considérablement renforcer le résultat précédent, en réduisant (énormément) la classe de fonctions formant une partie dense. Pour y parvenir, nous utiliserons la notion centrale de ce chapitre : la convolution.

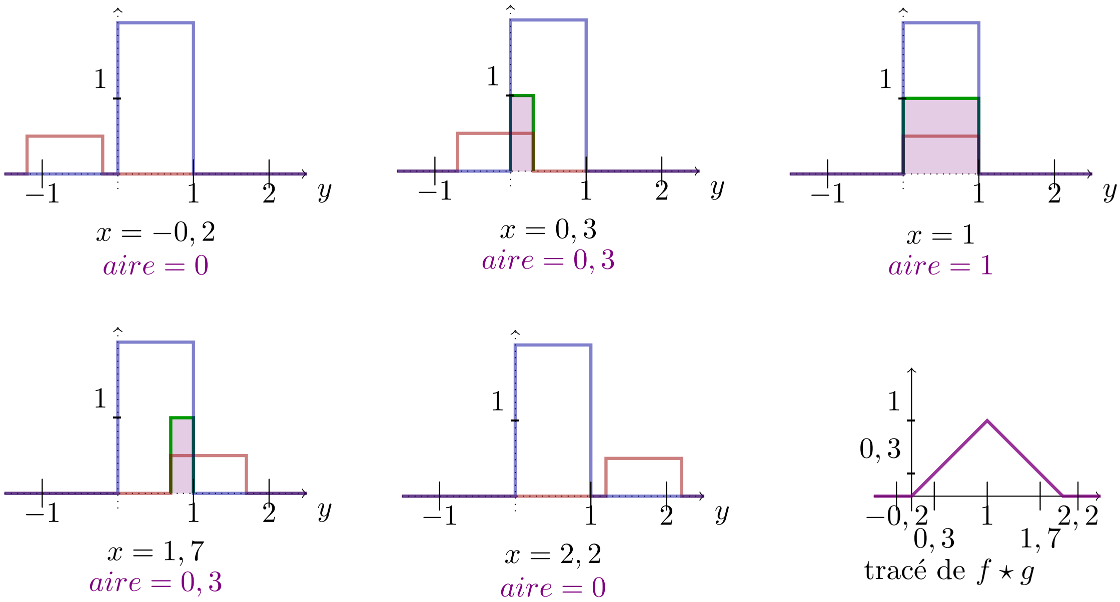

FIGURE 3.1 – La convolution dans l'exemple 3.22. Pour les 5 premiers graphiques, le nombre x est différent. La fonction $y \mapsto f(x-y)$ est représentée en rouge, la fonction $y \mapsto g(y)$ en bleu, et leur produit en vert. Le dernier graphique représente la convolution $f \star g$, obtenue comme l'intégrale (violette) de la fonction verte (qui dépend de x) en fonction de x.

3.2 Le produit de convolution

Énoncé indispensable 3.20

Soient $f, g : \mathbf{R} \to \mathbf{C}$ deux fonctions mesurables. Le produit de convolution *de f par g est l'application $f \star g : \mathbf{R} \to \mathbf{C}$ définie par (lorsque l'intégrale a un sens)*

$$\boxed{(f \star g)(x) = \int_{\mathbf{R}} f(x-y) g(y) \, \mathrm{d}y.}$$

Remarquons tout de suite que par le changement de variables $y' = x - y$, si l'intégrale est bien définie, alors

$$\boxed{f \star g = g \star f}.$$

De plus, par linéarité de l'intégrale, le produit de convolution est bilinéaire.

En un certain sens, $(f \star g)(x)$ représente la moyenne des valeurs de la fonction f autour de x, « pondérée » par la fonction $y \mapsto g(y)$. On peut donc voir la convolution $(f \star g)(x)$ comme une version continue du barycentre de tous les nombres $f(x-y)$, pondérés par les nombres $g(y)$ (à un facteur $\int g$ près, de la même manière qu'on divise par la somme des coefficients pour avoir un vrai barycentre).

Exemple 3.21

La convolution par une fonction indicatrice d'un segment permet de saisir l'une des propriétés fondamentales de la convolution : cette opération a tendance à « lisser » ses arguments, précisément parce qu'elle en réalise une « moyenne » (en un certain sens). Prenons par exemple le cas de $g = \mathbf{1}_{[-1,1]}$. Alors le produit de convolution $f \star g$ a un sens dès lors que f est intégrable et dans ce cas, la valeur de $f \star g$ en x est donnée par (on effectue le changement de variable $z = x - y$)

$$\int_{-1}^{1} f(x-y)\,\mathrm{d}y = \int_{x-1}^{x+1} f(z)\,\mathrm{d}z,$$

qui n'est rien d'autre que la moyenne des valeurs (au sens intégral) des valeurs de f sur l'intervalle $[x-1, x+1]$. Le « lissage » se constate par la remarque suivante : on n'a supposé aucune forme de régularité sur la fonction f et pourtant le produit $f \star g$ est une fonction continue (voir Exercice 2.15) ! Nous verrons un peu plus loin comme exploiter cette propriété de régularisation à un cran bien plus élevé.

Exemple 3.22

Voici un exemple de visualisation dans le cas $f = \frac{1}{2}\mathbf{1}_{[0,1]}$ et $g = 2 \cdot \mathbf{1}_{[0,1]}$. On se référera à la Figure 3.1 pour une représentation graphique.

On voit par un petit calcul que pour tout $x \in \mathbf{R}$, la fonction $y \mapsto f(x-y)$ est nulle en dehors de $[x-1, x]$, tandis que $y \mapsto g(y)$ est nulle en dehors $[0, 1]$. Donc si $x \leq 0$ ou $x \geq 2$, les fonctions $y \mapsto f(x-y)$ et $y \mapsto g(y)$ ne sont jamais simultanément non nulles, si bien que le produit de ces fonctions est nul et donc l'intégrale de ce produit est nulle.

Si $0 < x \leq 1$, on forme alors le produit de ces fonctions (en vert sur les figures) et calcule l'intégrale de ce produit, qu'on note $(f \star g)(x)$. Ici, ce nombre est l'intégrale de g sur $[x-1, x]$, pondérée par un facteur $1/2$ (qui est le facteur devant l'indicatrice dans f), il vaut donc $\frac{2}{2}x = x$.

De la même manière, on voit que si $1/2 \leq x < 1$, alors $(f \star g)(x) = 2 - x$.

Pour l'instant, on n'a donné aucun critère d'existence du produit de convolution. On va en donner trois dans la suite de ce paragraphe, mais sachez que ce ne sont pas les seuls ! Le premier d'entre eux concerne les fonctions intégrables, et se déduit du théorème de Fubini.

Proposition 3.23

Si $f, g \in L^1(\mathbf{R})$, alors $f \star g$ est bien définie presque partout, $f \star g \in L^1(\mathbf{R})$ et $\|f \star g\|_1 \leq \|f\|_1 \|g\|_1$.

Démonstration : On considère l'intégrale double

$$I = \int_{\mathbf{R}} \left\{ \int_{\mathbf{R}} |f(x-y)g(y)|\,\mathrm{d}y \right\} \mathrm{d}x.$$

Alors, en utilisant le Théorème 2.26 de Fubini-Tonelli et le changement de variable $t = x - y$, on a

$$\begin{aligned}
I &= \int_{\mathbf{R}} \left\{ \int_{\mathbf{R}} |f(x-y)g(y)| \, \mathrm{d}x \right\} \mathrm{d}y \\
&= \int_{\mathbf{R}} \left\{ \int_{\mathbf{R}} |f(x-y)| \, \mathrm{d}x \right\} |g(y)| \, \mathrm{d}y \\
&= \int_{\mathbf{R}} \left\{ \int_{\mathbf{R}} |f(t)| \, \mathrm{d}t \right\} |g(y)| \, \mathrm{d}y \\
&= \left(\int_{\mathbf{R}} |f(t)| \, \mathrm{d}t \right) \left(\int_{\mathbf{R}} |g(y)| \, \mathrm{d}y \right).
\end{aligned}$$

Puisque $f, g \in L^1(\mathbf{R})$, on en déduit que I est fini, et en particulier (voir Exercice 1.33) que pour presque tout $x \in \mathbf{R}$,

$$\int_{\mathbf{R}} |f(x-y)g(y)| \, \mathrm{d}y. < \infty$$

Autrement dit, pour presque tout x, la fonction $y \mapsto f(x-y)g(y)$ est intégrable, et donc le produit de convolution $(f \star g)(x)$ est bien défini et intégrable. De plus, par l'inégalité triangulaire (intégrale),

$$\int_{\mathbf{R}} |(f \star g)(x)| \, \mathrm{d}x \leq \int_{\mathbf{R}} \left\{ \int_{\mathbf{R}} |f(x-y)g(y)| \, \mathrm{d}y \right\} \mathrm{d}x \leq \left(\int_{\mathbf{R}} |f(t)| \, \mathrm{d}t \right) \left(\int_{\mathbf{R}} |g(y)| \, \mathrm{d}y \right),$$

ce qui est exactement $\|f \star g\|_1 \leq \|f\|_1 \|g\|_1$. ■

Corollaire 3.24

Si $f \in L^1(\mathbf{R})$ et $g \in L^2(\mathbf{R})$, alors $f \star g$ est bien définie presque partout, $f \star g \in L^2(\mathbf{R})$ et $\|f \star g\|_2 \leq \|f\|_1 \|g\|_2$.

Démonstration : On a par l'inégalité de Cauchy-Schwarz

$$\begin{aligned}
\int_{\mathbf{R}} |f(x-y)g(y)| \, \mathrm{d}y &= \int_{\mathbf{R}} |f(x-y)|^{1/2} |f(x-y)|^{1/2} |g(y)| \, \mathrm{d}y \\
&\leq \left(\int_{\mathbf{R}} |f(x-y)| \, \mathrm{d}y \right)^{1/2} \left(\int_{\mathbf{R}} |f(x-y)| |g(y)|^2 \, \mathrm{d}y \right)^{1/2} \\
&= \sqrt{\|f\|_1} \sqrt{|f| \star |g|^2}(x),
\end{aligned}$$

où l'on a utilisé le changement de variable $t = x - y$ à la dernière ligne. Le cadre de convolution $L^1(\mathbf{R}) \times L^1(\mathbf{R})$ que nous avons obtenu dans la Proposition 3.23 nous assure que le membre de droite est bien fini presque partout (puisque $f, |g|^2 \in L^1(\mathbf{R})$). En élevant cette inégalité au carré et en intégrant en la variable x, on obtient d'après la même proposition

$$\int_{\mathbf{R}} \left\{ \int_{\mathbf{R}} |f(x-y)g(y)| \, \mathrm{d}y \right\}^2 \mathrm{d}x \leq \|f\|_1^2 \|g^2\|_1,$$

ce qui fournit l'estimation annoncée puisque $|f \star g| \leq |f| \star |g|$ et $\|g^2\|_1 = \|g\|_2^2$. ■

Les deux propositions précédentes nous montrent que la convolution avec un élément de $L^1(\mathbf{R})$ est une opération qui semble conserver les propriétés d'*intégrabilité* : partant d'une fonction intégrable on obtient une fonction du même type, puis partant d'une fonction de carré intégrable on obtient à nouveau le même phénomène. La convolution se comporte de la même manière pour la *régularité* : c'est l'objet de la proposition suivante.

Proposition 3.25

Si $f \in L^1(\mathbf{R})$ et $g \in \mathscr{C}_b^0(\mathbf{R})$, alors $f \star g$ est bien définie, continue, bornée, et vérifie $\|f \star g\|_\infty \leq \|f\|_1 \|g\|_\infty$. Si de plus $g \in \mathscr{C}^1(\mathbf{R})$ de dérivée bornée alors il en est de même de $f \star g \in \mathscr{C}^1(\mathbf{R})$ et celle-ci vérifie

$$(f \star g)' = f \star (g'),$$

ainsi que $\|(f \star g)'\|_\infty \leq \|f\|_1 \|g'\|_\infty$.

Démonstration : Rappelons que

$$(f \star g)(x) = \int_{\mathbf{R}} f(y) g(x - y) \, \mathrm{d}y.$$

Pour tout $x \in \mathbf{R}$ et presque tout $y \in \mathbf{R}$, on a $|f(y)g(x-y)| \leq |f(y)| \|g\|_\infty$, si bien que $y \mapsto f(y)g(x-y)$ est intégrable et que $(f \star g)(x)$ est bien définie. Puisque g est continue, on peut appliquer le Théorème 2.19 de continuité à la fonction $h(x,t) = f(x)g(t-x)$, avec la domination $x \mapsto |f(x)| \|g\|_\infty$, ce qui implique la continuité du produit de convolution et par la même occasion, après intégration de l'inégalité $|h(x,t)| \leq |f(x)| \|g\|_\infty$, on récupère l'estimation annoncée.

Enfin, si maintenant $g \in \mathscr{C}^1(\mathbf{R})$ de dérivée bornée, on applique le Théorème 2.23 de dérivation sous l'intégrale à $h(x,t) = f(x)g(t-x)$. Pour (presque) tout $x \in \mathbf{R}$, la fonction $t \mapsto h(x,t)$ est dérivable, et on a

$$\frac{\partial h}{\partial t}(x,t) = f(x) g'(t - x).$$

Comme par hypothèse, g' est bornée, on a

$$\left| \frac{\partial h}{\partial t}(x,t) \right| \leq \|g'\|_\infty |f(x)|,$$

qui est une fonction intégrable en x et indépendante de t ; on en fait notre domination. On peut donc appliquer le Théorème 2.23 de dérivation sous l'intégrale, qui nous assure que $f \star g$ est dérivable, de dérivée $f \star g'$. La dernière inégalité s'obtient comme précédemment. ■

Proposition 3.26

Soient f et g deux fonctions à support compacts. Si le produit de convolution $f \star g$ est défini[2] alors $f \star g$ est également à support compact.

Démonstration : Par hypothèse on dispose de $R > 0$ tel que f et g sont presque partout nulles en dehors de $[-R, R]$. Mais alors, si $x \notin [-2R, 2R]$, on en déduit $y \in [-R, R] \Rightarrow x - y \notin [-R, R]$ et donc $f(x-y)g(y) = 0$ pour presque tout $y \in \mathbf{R}$. Puisque $x \mapsto f \star g(x)$ est presque partout défini par l'intégrale de la fonction précédente, la conclusion en découle. ■

3.3 Approximation de l'unité

Nous avons démontré (c'est la Proposition 3.23) que la convolution est une loi de composition interne sur $L^1(\mathbf{R})$. Il est naturel de se demander si une telle loi admet un élément neutre. Malheureusement, ce n'est pas le cas :

2. Au sens des Proposition 3.23, Corollaire 3.24 ou Proposition 3.25.

Exercice 3.27

Démontrer qu'il n'existe pas d'élément $f \in L^1(\mathbf{R})$ pour lequel, pour tout $g \in L^1(\mathbf{R})$, on ait $f \star g = g$.

On peut remédier à ce défaut par le concept d'*approximation de l'unité*.

Définition 3.28

Une *approximation de l'unité* est une suite de fonctions mesurables positives $\varphi_k : \mathbf{R} \to \mathbf{R}_+$ telle que

$$\int_{\mathbf{R}} \varphi_k = 1 \qquad \text{et} \qquad \forall \delta > 0, \quad \lim_{k \to +\infty} \int_{[-\delta,\delta]} \varphi_k = 1.$$

Remarque 3.29

Grâce à la relation de Chasles et la première condition, la deuxième admet une formulation parfaitement équivalente :

$$\forall \delta > 0, \quad \lim_{k \to +\infty} \int_{\mathbf{R} \setminus [-\delta,\delta]} \varphi_k = 0.$$

Commençons par donner une simple recette pour fabriquer une telle suite à partir d'une fonction intégrable positive donnée.

Lemme 3.30

Si $\varphi : \mathbf{R} \to \mathbf{R}_+$ est une fonction mesurable d'intégrale 1, alors la suite définie par $\varphi_k(x) = k\varphi(kx)$ est une approximation de l'unité.

Démonstration : Par le changement de variables $u = kx$, on voit que

$$\int_{\mathbf{R}} \varphi_k(x)\,\mathrm{d}x = \int_{\mathbf{R}} \varphi(u)\,\mathrm{d}u = 1.$$

De même, pour tout $\delta > 0$,

$$\int_{[-\delta,\delta]} \varphi_k(x)\,\mathrm{d}x = \int_{[-k\delta,k\delta]} \varphi(u)\,\mathrm{d}u = \int_{\mathbf{R}} \varphi(u)\mathbf{1}_{[-k\delta,k\delta]}(u)\,\mathrm{d}u.$$

On applique alors le théorème de convergence dominée à la suite de fonctions $u \mapsto \varphi(u)\mathbf{1}_{[-k\delta,k\delta]}(u)$, qui converge simplement vers la fonction φ, qui la domine. Par conséquent, la limite des intégrales de ces fonctions converge vers l'intégrale de φ, qui vaut 1. ■

L'appellation « approximation de l'unité » est justifiée par le résultat fondamental suivant.

Théorème 3.31

Soit $(\varphi_k)_k$ une approximation de l'unité.

(i) Si $f \in L^1(\mathbf{R})$, la suite $(f \star \varphi_k)_k$ converge vers f dans $L^1(\mathbf{R})$;

(ii) Si $f \in L^2(\mathbf{R})$, la suite $(f \star \varphi_k)_k$ converge vers f dans $L^2(\mathbf{R})$;

(iii) Si f est bornée et uniformément continue, la suite $(f \star \varphi_k)_k$ converge vers f uniformément sur $\mathbf{R}$.

Démonstration : Tout d'abord : les produits de convolution mentionnés sont bien définis grâce aux Propositions 3.23 et 3.25 ainsi que le Corollaire 3.24, ces résultats justifiant également l'appartenance à $L^1(\mathbf{R})$, $L^2(\mathbf{R})$ ou $\mathscr{C}_b^0(\mathbf{R})$, selon les cas. Commençons par traiter (iii). Puisque les fonctions φ_k sont d'intégrale 1 nous avons

$$(f \star \varphi_k)(x) - f(x) = \int_{\mathbf{R}} \big(f(x-y) - f(x)\big)\varphi_k(y)\,\mathrm{d}y.$$

On en déduit, pour $\delta > 0$ arbitraire

$$\big|f \star \varphi_k - f\big|(x) \leq \int_{\mathbf{R}} \big|f(x-y) - f(x)\big|\varphi_k(y)\mathbf{1}_{|y|\leq\delta}\,\mathrm{d}y + \int_{\mathbf{R}} \big|f(x-y) - f(x)\big|\varphi_k(y)\mathbf{1}_{|y|>\delta}\,\mathrm{d}y.$$

En particulier si $\tau_y f$ désigne la fonction $x \mapsto f(x-y)$, l'inégalité précédente implique, puisque encore une fois φ_k est d'intégrale 1,

$$\|f \star \varphi_k - f\|_\infty \leq \sup_{|y|\leq\delta} \|\tau_y f - f\|_\infty + 2\|f\|_\infty \int_{\mathbf{R}} \varphi_k(y)\mathbf{1}_{|y|>\delta}\,\mathrm{d}y.$$

Ainsi, pour tout $\varepsilon > 0$, la continuité uniforme de f fournit l'existence de δ pour lequel

$$\|f \star \varphi_k - f\|_\infty \leq \frac{1}{2}\varepsilon + 2\|f\|_\infty \int_{\mathbf{R}} \varphi_k(y)\mathbf{1}_{|y|>\delta}\,\mathrm{d}y.$$

Ensuite, δ étant fixé, c'est la deuxième propriété des approximations de l'unité (voir Définition 3.28) qui assure l'existence d'un rang k_0 au-delà duquel

$$\|f \star \varphi_k - f\|_\infty \leq \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon.$$

Le point (iii) est donc établi. Notons que cela couvre en particulier le cas d'une fonction $f \in \mathscr{C}_c^0(\mathbf{R})$: une telle fonction est effectivement bornée et uniformément continue par le théorème de Heine.

Les deux autres en découlent, par un raisonnement de densité. Nous traitons uniquement le cas (i), le second étant totalement similaire. Nous venons de remarquer que l'on a la convergence

$$\forall f \in \mathscr{C}_c^0(\mathbf{R}), \qquad \|f \star \varphi_k - f\|_\infty \underset{k\to+\infty}{\longrightarrow} 0. \tag{3.1}$$

Montrons que nous avons également

$$\forall f \in \mathscr{C}_c^0(\mathbf{R}), \qquad \|f \star \varphi_k - f\|_1 \underset{k\to+\infty}{\longrightarrow} 0, \tag{3.2}$$

où l'utilisation de la norme $\|\cdot\|_1$ est ici justifiée par le fait que $f \in L^1(\mathbf{R})$ et donc $f \star \varphi_k$ également, pour tout $k \in \mathbf{N}$. Puisque $f \in \mathscr{C}_c^0(\mathbf{R})$, nous disposons de $R > 0$ tel que $f(x) = 0$ pour $|x| > R$. En particulier, on peut écrire

$$\begin{aligned}\|f \star \varphi_k - f\|_1 &= \int_{|x|\leq R+1} |f - \varphi_k \star f| + \int_{|x|>R+1} |\varphi_k \star f| \\ &\leq 2(R+1)\,\|f - \varphi_k \star f\|_\infty + \int_{|x|>R+1} |\varphi_k \star f|, \end{aligned} \tag{3.3}$$

où l'inégalité provient de la croissance de l'intégrale. Pour le second terme, on sépare l'intégrale sur $|x| > R+1$ en $x < -R-1$ et $x > R+1$. On traite le cas de $x > R+1$, l'autre étant similaire : on écrit, par inégalité triangulaire,

$$\int_{x>R+1} |\varphi_k \star f| = \int_{x>R+1} \left| \int_{y\in\mathbf{R}} \varphi_k(y) f(x-y)\mathrm{d}y \right| \mathrm{d}x \leq \int_{x>R+1} \int_{y\in\mathbf{R}} \varphi_k(y)|f(x-y)|\mathrm{d}y\mathrm{d}x.$$

Mais si $x > R+1$ et $y < 1$, alors $f(x-y) = 0$, si bien que

$$\int_{x>R+1} |\varphi_k \star f| \leq \int_{x>R+1} \int_{y\geq 1} \varphi_k(y)|f(x-y)|\mathrm{d}y\mathrm{d}x.$$

On applique le théorème de Fubini-Tonelli, qui donne

$$\begin{aligned}\int_{x>R+1} |\varphi_k \star f| &\leq \int_{y\geq 1} \varphi_k(y) \int_{x>R+1} |f(x-y)|\mathrm{d}x\mathrm{d}y \\ &\leq \int_{y\geq 1} \varphi_k(y)\|f\|_1 \mathrm{d}y \\ &\leq \|f\|_1 \int_{|y|\geq 1} \varphi_k(y)\mathrm{d}y,\end{aligned}$$

et donc

$$\int_{|x|>R+1} |\varphi_k \star f| \leq 2\|f\|_1 \int_{|y|\geq 1} \varphi_k(y)\mathrm{d}y.$$

En combinant cette dernière majoration à (3.3) on aboutit à

$$\|f \star \varphi_k - f\|_1 \leq 2(R+1)\|f - \varphi_k \star f\|_\infty + 2\|f\|_1 \int_{\mathbf{R}\setminus[-1,1]} \varphi_k,$$

où le membre de droite tend vers 0 grâce à (3.1) et la Remarque 3.29, ce qui permet d'établir (3.2).

Maintenant, revenons à (i) en considérant $f \in L^1(\mathbf{R})$. Le Lemme 3.19 nous fournit, pour tout $\varepsilon > 0$, l'existence de $\psi_\varepsilon \in \mathscr{C}_c^0(\mathbf{R})$ telle que

$$\|f - \psi_\varepsilon\|_1 < \varepsilon.$$

En particulier, l'inégalité de la Proposition 3.23 montre pour tout $k \in \mathbf{N}$ que (on utilise $\|\varphi_k\|_1 = 1$)

$$\|(f - \psi_\varepsilon) \star \varphi_k\|_1 < \varepsilon,$$

et on déduit par inégalité triangulaire

$$\begin{aligned}\|f - f \star \varphi_k\|_1 &\leq \|f - \psi_\varepsilon\|_1 + \|\psi_\varepsilon - \psi_\varepsilon \star \varphi_k\|_1 + \|\psi_\varepsilon \star \varphi_k - f \star \varphi_k\|_1 \\ &< 2\varepsilon + \|\psi_\varepsilon - \psi_\varepsilon \star \varphi_k\|_1.\end{aligned}$$

Maintenant, grâce à (3.2), puisque $\psi_\varepsilon \in \mathscr{C}_c^0(\mathbf{R})$, on dispose de k_0 au-delà duquel $\|\psi_\varepsilon - \psi_\varepsilon \star \varphi_k\|_1 < \varepsilon$ et on a donc pour $k \geq k_0$

$$\|f - f \star \varphi_k\|_1 < 3\varepsilon,$$

ce qui termine la preuve de (i), celle de (ii) étant identique (en utilisant l'inégalité du Corollaire 3.24). ■

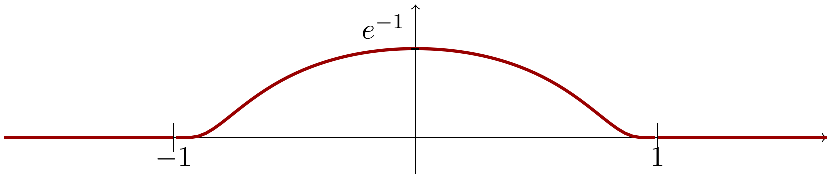

FIGURE 3.2 – La fonction g de la preuve du Lemme 3.33, qui est de classe $\mathscr{C}^\infty$ et à support compact.

3.4 Les fonctions lisses à support compact

Comme nous l'avions annoncé après sa preuve, le Lemme 3.19 admet une généralisation beaucoup plus forte. Celle-ci exprime la densité, dans $L^1(\mathbf{R})$ et $L^2(\mathbf{R})$ d'une classe très réduite de fonctions.

Définition 3.32

On note $\mathscr{C}_c^\infty(\mathbf{R})$ l'ensemble des fonctions de $\mathbf{R}$ dans $\mathbf{C}$, de classe $\mathscr{C}^\infty$ et à support compact.

Avant d'espérer établir ce résultat de densité, il faut déjà se convaincre que l'espace vectoriel $\mathscr{C}_c^\infty(\mathbf{R})$ n'est pas trivial.

Lemme 3.33

Il existe une fonction réelle positive mais non identiquement nulle dans $\mathscr{C}_c^\infty(\mathbf{R})$.

Démonstration : Posons

$$\begin{aligned} g : \mathbf{R} &\longrightarrow \mathbf{R} \\ x &\longmapsto \begin{cases} 0 & \text{si } |x| \geq 1 \\ \exp\left(\frac{1}{(x-1)(x+1)}\right) & \text{si } |x| < 1 \end{cases} \end{aligned}$$

(voir la Figure 3.2). Remarquons tout de suite que g est à support compact et non nulle (car $g(0) = e^{-1}$). Remarquons aussi que si $-1 < x < 1$, alors $\frac{1}{(x-1)(x+1)} < 0$, donc $g(x)$ est inférieure à 1, et donc $0 \leq g(x) \leq 1$ pour tout x. De plus, $\lim_{x \to \pm 1} \frac{1}{(x-1)(x+1)} = -\infty$, donc $g(x)$ tend vers 0 en ± 1 et donc g est continue sur $\mathbf{R}$, puisque continue sur $\mathbf{R} \setminus \{-1, 1\}$ et continue en ± 1.

On montre alors par récurrence que g est de classe $\mathscr{C}^\infty$ sur $]-1, 1[$, et que ses dérivées successives sont de la forme $P_k(x)g(x)$, où P_k est une fraction rationnelle. Les théorèmes de comparaison assurent alors que les dérivées k-ièmes de g tendent vers 0 en ± 1, ce qui permet de dire, par le théorème de la limite dérivée, que le prolongement par continuité de g en -1 et 1 est une fonction de classe $\mathscr{C}^\infty$, dont toutes les dérivées k-ièmes sont nulles en ± 1. Puisque la même propriété est vraie pour la restriction de g à $[-1, 1]^\complement$, on en déduit que g est de classe $\mathscr{C}^\infty$. ∎

Corollaire 3.34

L'espace $\mathscr{C}_c^\infty(\mathbf{R})$ est dense dans $L^1(\mathbf{R})$ et dense dans $L^2(\mathbf{R})$.

Démonstration : Les preuves étant rigoureusement identiques, nous ne traitons que le cas de $L^1(\mathbf{R})$. Le Lemme 3.33 nous permet de fabriquer une approximation de l'unité dont les éléments sont à support compact, en suivant le procédé suggéré par le Lemme 3.30. En effet, si $\varphi \in \mathscr{C}_c^\infty(\mathbf{R})$ est positive non identiquement nulle, alors $\|\varphi\|_1 \neq 0$, $\varphi/\|\varphi\|_1$ a une intégrale valant 1 et la suite $\varphi_k(x) := k\varphi(kx)$ est une approximation de l'unité. Pour tout $k \in \mathbf{N}$, le produit de convolution $f \star \varphi_k$ est bien défini, grâce à la Proposition 3.25. Mieux : en itérant cette proposition on constate que pour tout $k \in \mathbf{N}$, $f \star \varphi_k \in \mathscr{C}^\infty(\mathbf{R})$. En se remémorant le point (i) du Théorème 3.31, on obtient ainsi la convergence d'une suite $(f \star \varphi_k)_k$ de fonctions $\mathscr{C}^\infty(\mathbf{R})$ vers f, dans $L^1(\mathbf{R})$. En l'état le résultat n'est démontré qu'à moitié puisqu'il manque l'hypothèse de compacité sur les supports. On utilise alors la Proposition 3.14 et le Lemme 3.18 : il suffit de démontrer que $\mathscr{C}_c^\infty(\mathbf{R})$ est dense dans l'ensemble des éléments à supports compacts de $L^1(\mathbf{R})$. Ceci étant noté, notre raisonnement précédent nous montre que pour f intégrable à support compact, on a $(f \star \varphi_k)_k \to f$ dans $L^1(\mathbf{R})$. Mais cette fois les fonctions $f \star \varphi_k$ sont bien des éléments de $\mathscr{C}_c^\infty(\mathbf{R})$ grâce aux Propositions 3.26 et 3.26. ■

Chapitre 4

Transformation de Fourier

L'acte fondateur de l'analyse de Fourier est le *Mémoire sur la propagation de la chaleur dans les corps solides* publié en 1807 par Joseph Fourier. Si, dans des travaux antérieurs, on avait pu voir certaines fonctions décomposées en séries trigonométriques, l'idée géniale de Fourier est que cette opération peut être effectuée pour *n'importe quelle fonction* ; il l'applique avec succès à l'étude de l'équation de la chaleur. Le XIX[e] siècle verra des efforts pour donner des bases rigoureuses à l'analyse ; en particulier, avec la découverte du phénomène de Gibbs, on voit apparaître un contre exemple frappant à la convergence ponctuelle des séries de Fourier pour les fonctions discontinues, convergence qui était pourtant conjecturée (ou plutôt, considérée comme évidente) par Fourier.

La transformée de Fourier apparaît comme une généralisation des séries de Fourier pour des fonctions qui ne sont pas périodiques. Le prix à payer est qu'on ne peut pas espérer exprimer la fonction étudiée sous la forme d'une série trigonométrique, mais seulement comme une intégrale de fonctions trigonométriques, autrement dit comme "une somme indexée par $\mathbf{R}$". Pour que cela fonctionne, autrement dit pour avoir une formule d'inversion, nous verrons au Théorème 4.7 qu'il faut des hypothèses d'intégrabilité à la fois sur la fonction et sa transformée de Fourier.

La transformée de Fourier peut être vue comme un filtre qui, à partir d'un signal (une fonction f intégrable sur $\mathbf{R}$) et un nombre $\xi \in \mathbf{R}$, va détecter la contribution de la fréquence ξ à la fonction f, à l'image de l'oreille humaine qui peut détecter la hauteur des sons dans un bruit (vos oreilles font de la transformée de Fourier !). D'un point de vue mathématique, à partir d'une fonction f dépendant du temps x, on obtient une nouvelle fonction $\hat{f}$ dépendant de la fréquence ξ. On peut alors établir un dictionnaire bilingue entre les domaines temporel et fréquentiel ; par exemple une dérivation dans le domaine temporel (pour f) correspond à une multiplication par $i\xi$ dans le domaine fréquentiel (pour $\hat{f}$, Proposition 4.2), et la transformée de Fourier d'une convolution correspond au produit ponctuel des transformées de Fourier (Théorème 4.6).

Cette transformation de la dérivation en multiplication par $i\xi$ est cruciale, c'est cette propriété qui donne à la transformée de Fourier toute sa force dans l'étude des équations différentielles. Elle vient du fait que les fonctions exponentielles $x \mapsto e^{ax}$ sont les valeurs propres de la dérivation sur $\mathbf{R}$. En termes algébriques : la transformée de Fourier d'une fonction, c'est sa décomposition dans une base formée de vecteurs propres de la dérivation. Cette phrase peut être rendue (plus ou moins) rigoureuse dans l'espace de Hilbert $L^2(\mathbf{R})$, dans lequel la transformée de Fourier peut effectivement se voir comme une projection orthogonale dans la base “orthogonale” des exponentielles complexes. C’est un espace bien plus pratique que $L^1(\mathbf{R})$ pour Fourier, dont nous ne ferons qu’en entrevoir la théorie au travers la formule de Parseval (Théorème 4.10).

La puissance de la transformée de Fourier fait que ses applications, tant théoriques que pratiques, sont innombrables ; c’est sans doute une des découvertes mathématiques les plus importantes. Bien sûr, elle permet la résolution d’équations différentielles et est à la base de la théorie du filtrage, centrale tant en physique qu’en électricité, mécanique ou théorie de l’information. L’existence d’un algorithme extrêmement rapide pour son calcul (la FFT, pour *Fast Fourier Transform*) la rend omniprésente dans nos ordinateurs, par exemple pour calculer une convolution (on passe dans le domaine fréquentiel, fait une multiplication simple, puis applique la transformée de Fourier inverse), un produit de polynômes ou même le produit de deux grands nombres. Cela en fait l’outil central lors du traitement des données audio, photo ou vidéo (l’algorithme JPEG est basé sur une transformée de Fourier).

Résumé :

- La transformée de Fourier d'une fonction intégrable f est définie par la formule
$$(\mathcal{F}f)(\xi) = \widehat{f}(\xi) = \int_{\mathbf{R}} f(x)e^{-i\xi x}\,\mathrm{d}x.$$
Elle vérifie les propriétés (Proposition 4.2)
$$\widehat{f'}(\xi) = i\xi\widehat{f}(\xi) \qquad \text{et} \qquad \mathcal{F}(f)' = -i\mathcal{F}(x \mapsto xf(x)).$$
- La transformée de Fourier transforme une convolution en produit ponctuel (Théorème 4.6) : si $f, g \in L^1(\mathbf{R})$, alors
$$\widehat{f \star g}(\xi) = \widehat{f}(\xi)\widehat{g}(\xi).$$
- Le Théorème 4.7 ou *formule d'inversion* : si $f \in L^1(\mathbf{R})$ et $\hat{f} \in L^1(\mathbf{R})$, alors
$$\mathcal{F}\mathcal{F}f = 2\pi\check{f}.$$
- La formule de Parseval (Théorème 4.10) : si $\varphi \in \mathscr{C}_c^\infty(\mathbf{R})$, alors $\widehat{\varphi} \in L^2(\mathbf{R})$ et
$$\|\hat{\varphi}\|_2 = \sqrt{2\pi}\|\varphi\|_2.$$

4.1 Définition et premiers exemples

Énoncé indispensable 4.1

Si $f : \mathbf{R} \to \mathbf{C}$ est une fonction intégrable, alors sa transformée de Fourier *est la fonction notée $\mathcal{F}(f)$ ou $\hat{f}$, allant de $\mathbf{R}$ vers $\mathbf{C}$, et définie par*

$$\widehat{f}(\xi) = \int_{\mathbf{R}} f(x) e^{-i\xi x} \, \mathrm{d}x.$$

Cette définition a bien un sens : on vérifie très facilement que pour tout $\xi \in \mathbf{R}$, l'intégrale est bien définie, puisque $x \mapsto |f(x)e^{-i\xi x}| = |f(x)|$ est intégrable.

Attention ! Il existe plusieurs conventions pour la définition de transformée de Fourier, chacune adaptée à des problèmes spécifiques : certains choisissent de remplacer le $e^{-i\xi x}$ de l'intégrale par un $e^{-2i\pi\xi x}$, d'autres encore multiplient l'intégrale par un facteur $1/\sqrt{2\pi}$... Bref, soyez attentifs à la définition de la transformée de Fourier quand vous allez chercher des informations dans d'autres références : elle a des chances d'être différente !

Voici quelques unes des propriétés fondamentales de la transformée de Fourier. En particulier, le comportement vis-à-vis de la dérivation (les points 3. et 4.) est remarquable.

Proposition 4.2

Pour une fonction $f : \mathbf{R} \to \mathbf{C}$ et un réel a non nul on note $e_a : \xi \mapsto e^{ia\xi}$ et on introduit les opérations suivantes : $\tau_a f : x \mapsto f(x-a)$ et $\sigma_a f : x \mapsto f(x/a)$.

1. *L'application $\mathcal{F}$ est linéaire et continue de $L^1(\mathbf{R})$ dans $\mathscr{C}_b^0(\mathbf{R})$, et vérifie plus précisément*
$$\|\hat{f}\|_\infty \leq \|f\|_1.$$
2. *Si $f \in L^1(\mathbf{R})$ et $a \in \mathbf{R}^\star$, alors $\widehat{\tau_a f} = e_{-a}\widehat{f}$ et $\widehat{\sigma_a f} = |a|\sigma_{1/a}\widehat{f}$.*
3. *Si $f \in L^1(\mathbf{R}) \cap \mathscr{C}^1(\mathbf{R})$ et $f' \in L^1(\mathbf{R})$, alors*
$$\boxed{(\mathcal{F} f')(\xi) = i\xi(\mathcal{F} f)(\xi).}$$
4. *Si $x \mapsto (1+|x|)f(x) \in L^1(\mathbf{R})$, alors $\widehat{f} \in \mathscr{C}^1(\mathbf{R})$ et*
$$\boxed{(\mathcal{F} f)' = -i\mathcal{F}(x \mapsto xf(x)).}$$
5. *Si $f, g \in L^1(\mathbf{R})$, alors*
$$\int_{\mathbf{R}} \hat{f} g = \int_{\mathbf{R}} f \hat{g}.$$

Démonstration : 1. La linéarité provient de celle de l'intégrale de Lebesgue. La transformée de Fourier d'une fonction intégrable est bien une fonction continue grâce au Théorème 2.19 (la domination est donnée par $|f|$). C'est même une fonction bornée grâce à l'inégalité suivante, qui assure aussi la continuité annoncée de la transformée de

Fourier :

$$\left|\hat{f}(\xi)\right| = \left|\int_{\mathbf{R}} e^{-i\xi x} f(x)\,\mathrm{d}x\right| \leq \int_{\mathbf{R}} \left|e^{-i\xi x} f(x)\right|\,\mathrm{d}x = \int_{\mathbf{R}} |f(x)|\,\mathrm{d}x = \|f\|_1.$$

2. Calculons

$$\widehat{\tau_a f}(\xi) = \int_{\mathbf{R}} e^{-i\xi x} \tau_a f(x)\,\mathrm{d}x = \int_{\mathbf{R}} e^{-i\xi x} f(x-a)\,\mathrm{d}x,$$

et le changement de variables $y = x - a$ donne

$$\widehat{\tau_a f}(\xi) = \int_{\mathbf{R}} e^{-i\xi(y+a)} f(y)\,\mathrm{d}y = e^{-i\xi a} \int_{\mathbf{R}} e^{-i\xi y} f(y)\,\mathrm{d}y = e^{-ia\xi} \widehat{f}.$$

D'autre part,

$$\widehat{\sigma_a f}(\xi) = \int_{\mathbf{R}} e^{-i\xi x} f(x/a)\,\mathrm{d}x.$$

Ici, le changement de variable $y = x/a$ donne

$$\widehat{\sigma_a f}(\xi) = |a| \int_{\mathbf{R}} e^{-i\xi y a} f(y)\,\mathrm{d}y = |a| \sigma_{1/a} \widehat{f}(\xi).$$

3. Par définition,

$$\mathcal{F}(f')(\xi) = \int_{\mathbf{R}} e^{-i\xi x} f'(x)\,\mathrm{d}x.$$

L'idée est de faire une intégration par parties. Pour cela, on se ramène à une intégrale sur un segment en appliquant le Corollaire 2.16 aux segments $[-R, R]$:

$$\mathcal{F}(f')(\xi) = \lim_{R\to+\infty} \int_{-R}^{R} e^{-i\xi x} f'(x)\,\mathrm{d}x.$$

On fait alors une intégration par parties, qui donne :

$$\mathcal{F}(f')(\xi) = \lim_{R\to+\infty} \left(\left[e^{-i\xi x} f(x)\right]_{x=-R}^{x=R} + i\xi \int_{-R}^{R} e^{-i\xi x} f(x)\,\mathrm{d}x \right).$$

En appliquant de nouveau le Corollaire 2.16, on voit que le terme de l'intégrale tend vers $i\xi\mathcal{F}(f)(\xi)$; il s'agit donc de montrer que le premier terme tend vers 0. Attention, le fait que la fonction $x \mapsto e^{-i\xi x} f(x)$ soit intégrable n'implique pas que la fonction tende vers 0 en $\pm\infty$ (voir l'exemple 1.34) ; il faut donc utiliser le fait que f est de dérivée intégrable.
Puisque f est de classe $\mathscr{C}^1$, par le Corollaire 2.2, on a

$$f(R) = f(0) + \int_0^R f'(t)\,\mathrm{d}t,$$

qui admet le nombre $f(0) + \int_0^{+\infty} f'(t)\,\mathrm{d}t$ comme limite en $+\infty$. Ainsi, f admet une limite en $+\infty$, qui ne peut être que 0 puisque f est intégrable. La même propriété est vraie en $-\infty$, et donc le premier terme admet comme limite 0 lorsque R tend vers $+\infty$.

4. On applique le théorème 2.23 de dérivation sous le signe somme à la fonction
$$\mathcal{F}(f) : \xi \mapsto \int_{\mathbf{R}} e^{-i\xi x} f(x)\,\mathrm{d}x.$$
Pour tout x, la fonction $\xi \mapsto e^{-i\xi x} f(x)$ est bien dérivable de dérivée $\xi \mapsto -ixe^{-i\xi x} f(x)$ qui est dominée par la fonction $x \mapsto |xf(x)|$, d'intégrale finie par hypothèse. Le théorème 2.23 de dérivation sous le signe somme s'applique donc, et on obtient
$$(\mathcal{F}f)'(\xi) = -i \int_{\mathbf{R}} e^{-i\xi x} x f(x)\,\mathrm{d}x.$$

5. Premièrement, on vérifie que l'intégrale a bien un sens : par le premier point, on a $\hat{f} \in \mathscr{C}^0_b(\mathbf{R})$ et donc $\hat{f}g \in L^1(\mathbf{R})$. On peut donc calculer
$$\int \hat{f}(\xi)g(\xi)\,\mathrm{d}\xi = \int_{\mathbf{R}} \left\{\int_{\mathbf{R}} e^{-i\xi x} f(x)\,\mathrm{d}x\right\} g(\xi)\,\mathrm{d}\xi = \int_{\mathbf{R}} \left\{\int_{\mathbf{R}} e^{-i\xi x} f(x)g(\xi)\,\mathrm{d}x\right\} \mathrm{d}\xi.$$
Par le Théorème 2.26 de Fubini-Tonelli on a
$$\int_{\mathbf{R}} \left\{\int_{\mathbf{R}} \left|e^{-i\xi x} f(x)g(\xi)\right| \mathrm{d}x\right\} \mathrm{d}\xi = \left(\int_{\mathbf{R}} |f(x)|\,\mathrm{d}x\right)\left(\int_{\mathbf{R}} |g(\xi)|\,\mathrm{d}\xi\right) = \|f\|_1 \|g\|_1 < +\infty,$$
et on peut donc appliquer le Théorème 2.28 de Fubini, qui donne
$$\begin{aligned}\int \hat{f}(\xi)g(\xi)\,\mathrm{d}\xi &= \int_{\mathbf{R}} \left\{\int_{\mathbf{R}} e^{-i\xi x} f(x)g(\xi)\,\mathrm{d}x\right\} \mathrm{d}\xi \\ &= \int_{\mathbf{R}} \left\{\int_{\mathbf{R}} e^{-i\xi x} f(x)g(\xi)\,\mathrm{d}\xi\right\} \mathrm{d}x \\ &= \int_{\mathbf{R}} \left\{\int_{\mathbf{R}} e^{-i\xi x} g(\xi)\mathrm{d}\xi\right\} f(x)\,\mathrm{d}x \\ &= \int_{\mathbf{R}} \hat{g}(x)f(x)\,\mathrm{d}x.\end{aligned}$$
■

Donnons l'un des exemples les plus simples de calcul de transformée de Fourier.

Exemple 4.3

Soit $a \in \mathbf{R}^\star_+$, et $f = \mathbf{1}_{[-a,a]}$. La fonction f est bornée et à support compact, sa transformée de Fourier est donc bien définie, et on a, pour tout $\xi \neq 0$ (on trouve $\hat{f}(0) = 2a$),
$$\hat{f}(\xi) = \int_{\mathbf{R}} f(x)e^{-i\xi x}\,\mathrm{d}x = \int_{-a}^{a} e^{-i\xi x}\,\mathrm{d}x = \left[\frac{e^{-i\xi x}}{-i\xi}\right]_{x=-a}^{x=a} = \frac{e^{-i\xi a}}{-i\xi} - \frac{e^{i\xi a}}{-i\xi} = \frac{2}{\xi}\sin(a\xi).$$
On voit apparaître la fonction *sinus cardinal* $\operatorname{sinc} x = \frac{\sin x}{x}$ (prolongée par continuité en 0 en la valeur 1).

Cet exemple montre que l'appartenance $f \in L^1(\mathbf{R})$ n'implique pas nécessairement $\hat{f} \in L^1(\mathbf{R})$. En revanche, elle tend vers 0 en l'infini. On pouvait s'y attendre : si on considère l'intégrale de $e^{i\xi x}$ pour x entre $-a$ et a et pour ξ grand, des compensations vont

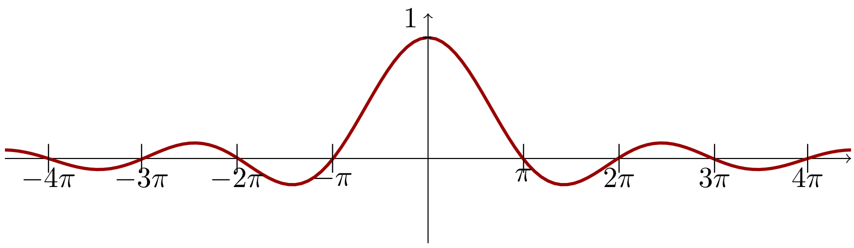

FIGURE 4.1 – Le sinus cardinal.

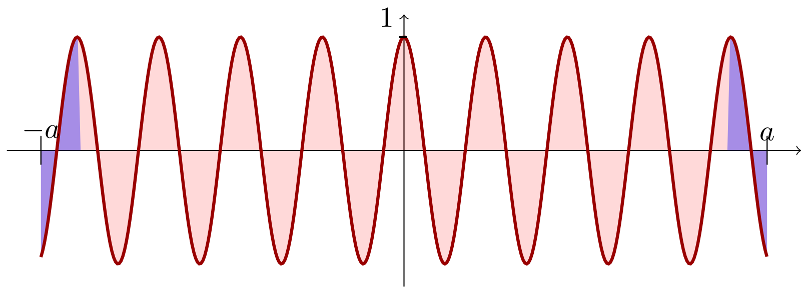

FIGURE 4.2 – Les compensations du cosinus : l'intégrale sur $[-a, a]$ est égale à celle sur la région bleue, vu que celle sur la région rouge est nulle.

s'effectuer étant donné que l'application $x \mapsto e^{i\xi x}$ est $2\pi/\xi$-périodique (voir la Figure 4.2). Cette remarque implique que l'intégrale coïncide avec celle de $e^{i\xi x}$ sur un intervalle de longueur au plus $2\pi/\xi$; comme la fonction intégrée est de module inférieur à un, cette intégrale est inférieure (en module) à $2\pi/\xi$.

Un argument par densité permet de généraliser cette convergence vers 0 en l'infini à toute fonction de $L^1(\mathbf{R})$.

Théorème 4.4 (*Lemme de Lebesgue*)

Si $f \in L^1(\mathbf{R})$, alors

$$\hat{f}(\xi) \underset{\xi \to \pm\infty}{\longrightarrow} 0.$$

Démonstration : Considérons d'abord le cas d'une fonction $\varphi \in \mathscr{C}_c^\infty(\mathbf{R}) \subset L^1(\mathbf{R})$. Commençons par remarquer que φ' étant continue à support compact, elle est intégrable, ce qui permet d'appliquer le point 3. de la Proposition 4.2 et d'écrire $\widehat{\varphi'}(\xi) = i\xi\widehat{\varphi}(\xi)$. En particulier, $\widehat{\varphi}(\xi) \leq \|\widehat{\varphi'}\|_\infty/\xi$, si bien que par le point 1. de cette même proposition, $\widehat{\varphi}(\xi) \leq \|\varphi'\|_1/\xi$. Ceci montre que $\lim_{\xi \to \pm\infty} \widehat{\varphi(\xi)} = 0$.

Retournons au cas général en considérant $f \in L^1(\mathbf{R})$. Grâce au Corollaire 3.34, pour tout $\varepsilon > 0$, on dispose de $\varphi_\varepsilon \in \mathscr{C}_c^\infty(\mathbf{R})$ telle que $\|f - \varphi_\varepsilon\|_1 \leq \varepsilon$. Par inégalité triangulaire et en utilisant le premier point de la Proposition 4.2, on a ainsi

$$|\widehat{f}(\xi)| \leq |\widehat{\varphi}_\varepsilon(\xi)| + |\widehat{f - \varphi_\varepsilon}(\xi)| \leq |\widehat{\varphi}_\varepsilon(\xi)| + \|\widehat{f - \varphi_\varepsilon}\|_\infty \leq |\widehat{\varphi}_\varepsilon(\xi)| + \|f - \varphi_\varepsilon\|_1 \leq |\widehat{\varphi}_\varepsilon(\xi)| + \varepsilon.$$

Mais on a déjà établi $|\widehat{\varphi}_\varepsilon(\xi)| \to_{\xi\to\pm\infty} 0$, donc il existe $A > 0$ tel que $|\xi| \geq A \Rightarrow |\widehat{\varphi}_\varepsilon(\xi)| < \varepsilon$; cela montre que pour $\varepsilon > 0$, il existe $A > 0$ tel que $|\xi| \geq A \Rightarrow |\widehat{f}(\xi)| \leq 2\varepsilon$. ■

Un second exemple classique est celui de la transformée de Fourier de la gaussienne, qui s'avère être aussi une gaussienne. Ce fait nous servira pour la preuve de la formule d'inversion.

Proposition 4.5

Soit $a > 0$ et g_a la gaussienne associée : $g_a(x) = e^{-ax^2}$. Alors

$$\boxed{\widehat{g_a}(\xi) = \sqrt{\frac{\pi}{a}} g_{1/(4a)}(\xi).}$$

Démonstration : En tout premier lieu, les courbes gaussiennes mentionnées sont bien toutes intégrables, par croissance comparée et critère de Riemann par exemple. Ensuite, notons que la formule se symétrise dans le cas $a = 1/2$ (Gaussienne centrée réduite) : la fonction $g_{1/2}$ est alors un vecteur propre de la transformée de Fourier, de valeur propre $\sqrt{2\pi}$. Mieux : il nous suffit de prouver l'égalité $\widehat{g_{1/2}} = \sqrt{2\pi} g_{1/2}$ pour déduire toutes les autres car on a $g_a(x) = g_{1/2}(\sqrt{2a}x)$ et grâce au deuxième point de la Proposition 4.2, on en déduirait donc que $\widehat{g_a}(x) = (2a)^{-1/2}\,\widehat{g_{1/2}}(x/(2a)^{1/2}) = (\pi/a)^{1/2} g_{1/2}(x/(2a)^{1/2})$ qui est exactement la formule annoncée.

On va établir une équation différentielle vérifiée par

$$\widehat{g_{1/2}} : \xi \mapsto \int_{\mathbf{R}} e^{-i\xi x} e^{-x^2/2}\,\mathrm{d}x;$$

que l'on pourra résoudre explicitement.

On commence par appliquer le point 4. de la Proposition 4.2 à la fonction $g_{1/2}$ (on montre facilement, comme au-dessus, que $x \mapsto (1 + |x|)g_{1/2}(x) \in L^1(\mathbf{R})$). Ainsi, $\widehat{g_{1/2}} \in \mathscr{C}^1(\mathbf{R})$ et

$$\begin{aligned}
\widehat{g_{1/2}}'(\xi) &= -i \int_{\mathbf{R}} x e^{-i\xi x} e^{-x^2/2}\,\mathrm{d}x \\
&= i \int_{\mathbf{R}} e^{-i\xi x}(-x) e^{-x^2/2}\,\mathrm{d}x \\
&= i \lim_{n\to+\infty} \int_{[-n,n]} e^{-i\xi x}(-x) e^{-x^2/2}\,\mathrm{d}x,
\end{aligned}$$

où la dernière ligne s'obtient par application du Théorème 2.13 de convergence dominée. On fait une intégration par parties (où on intègre $x \mapsto -xe^{-x^2/2}$) pour obtenir

$$\widehat{g_{1/2}}'(\xi) = i \lim_{n\to+\infty} \left(\left[e^{-i\xi x} e^{-x^2/2} \right]_{x=-n}^{x=n} + \int_{-n}^{n} i\xi e^{-i\xi x} e^{-x^2/2}\,\mathrm{d}x \right).$$

Une seconde application du théorème de convergence dominée permet de passer à la limite, et donc

$$\widehat{g_{1/2}}'(\xi) = -\xi \int_{\mathbf{R}} e^{-i\xi x} e^{-x^2/2}\,\mathrm{d}x = -\xi \widehat{g_{1/2}}(\xi).$$

Ainsi, la fonction $\widehat{g_{1/2}}$ vérifie l'équation différentielle $y'(\xi) = -\xi y(\xi)$. Multipliant par $e^{\xi^2/2}$ cette identité on voit que $\xi \mapsto y(\xi)e^{\xi^2/2}$ a une dérivée nulle : on récupère ainsi $y = y(0)g_{1/2}$ et la preuve se termine en remarquant que

$$y(0) = \int_{\mathbf{R}} e^{-x^2/2}\, \mathrm{d}x = \sqrt{2\pi},$$

grâce au calcul effectué dans l'Exemple 2.40. ■

Théorème 4.6 (***Transformée de Fourier d'une convolution***)

Si $f, g \in L^1(\mathbf{R})$, alors

$$\boxed{\widehat{f \star g}(\xi) = \widehat{f}(\xi)\widehat{g}(\xi).}$$

Démonstration : Par la Proposition 3.23, la convolution $f \star g$ est bien définie et est intégrable ; sa transformée de Fourier est donc bien définie. Par définition,

$$\widehat{f \star g}(\xi) = \int_{\mathbf{R}} e^{-i\xi x} \left\{ \int_{\mathbf{R}} f(x-y)g(y)\, \mathrm{d}y \right\} \mathrm{d}x.$$

Avant d'appliquer le Théorème 2.28 de Fubini il faut vérifier l'intégrabilité de la fonction en question. Par le Théorème 2.26 de Fubini-Tonelli, puis par un changement de variable affine, on a

$$\begin{aligned}
\int_{\mathbf{R}^2} \left| e^{-i\xi x} f(x-y)g(y) \right| \mathrm{d}y\, \mathrm{d}x &= \int_{\mathbf{R}} \left\{ \int_{\mathbf{R}} |f(x-y)||g(y)|\, \mathrm{d}x \right\} \mathrm{d}y \\
&= \int_{\mathbf{R}} \left\{ \int_{\mathbf{R}} |f(x-y)|\, \mathrm{d}x \right\} |g(y)| \mathrm{d}y \\
&= \int_{\mathbf{R}} \left\{ \int_{\mathbf{R}} |f(t)|\, \mathrm{d}t \right\} |g(y)| \mathrm{d}y \\
&= \left(\int_{\mathbf{R}} |f(t)| \mathrm{d}t \right) \left(\int_{\mathbf{R}} |g(y)| \mathrm{d}y \right) \\
&= \|f\|_1 \|g\|_1 < \infty.
\end{aligned}$$

La fonction $(x, y) \mapsto e^{-i\xi x} f(x-y)g(y)$ est donc intégrable pour tout ξ, et on peut appliquer le Théorème 2.28 de Fubini :

$$\widehat{f \star g}(\xi) = \int_{\mathbf{R}^2} e^{-i\xi x} f(x-y)g(y)\, \mathrm{d}x\, \mathrm{d}y.$$

Le changement de variables affine $(x, y) = (u + v, v)$ est de déterminant jacobien égal à 1 (la matrice est triangulaire avec des 1 sur la diagonale), et donne donc

$$\widehat{f \star g}(\xi) = \int_{\mathbf{R}^2} e^{-i\xi(u+v)} f(u)g(v)\, \mathrm{d}u\, \mathrm{d}v.$$

Une nouvelle application du théorème de Fubini donne alors

$$\widehat{f \star g}(\xi) = \left(\int_{\mathbf{R}} e^{-i\xi u} f(u)\, \mathrm{d}u \right) \left(\int_{\mathbf{R}} e^{-i\xi v} g(v)\, \mathrm{d}v \right),$$

qui n'est rien d'autre que $\widehat{f}(\xi)\widehat{g}(\xi)$. ■

4.2 Formules d'inversion et de Parseval

Nous terminons ce chapitre par deux résultats fondamentaux liés à la transformée de Fourier.

Le premier indique ce qu'il se passe lorsqu'on applique deux fois cette transformation : sous réserve que la transformée de Fourier de $f \in L^1(\mathbf{R})$ soit elle-même intégrable, la double transformée de Fourier $\mathcal{F}\mathcal{F}f$ est égale à (2π fois) $\check{f}$, l'application définie par $\check{f}(x) = f(-x)$. Cette formule permet d'exprimer la fonction f comme une « combinaison » d'exponentielles $\xi \mapsto e^{i\xi x}$, à une infinité de termes.

Énoncé indispensable 4.7 (*Inversion de la transformée de Fourier*)

Si $f \in L^1(\mathbf{R})$ et $\widehat{f} \in L^1(\mathbf{R})$, alors

$$\boxed{\mathcal{F}\mathcal{F}f = 2\pi\check{f}.}$$

La preuve de cette formule se trouve dans les compléments (Section A.4).

Un corollaire important de ce résultat est l'injectivité de la transformée de Fourier.

Corollaire 4.8

L'application $\mathcal{F} : L^1(\mathbf{R}) \to \mathscr{C}_b^0(\mathbf{R})$ est injective.

Démonstration: Il s'agit d'une application linéaire (Proposition 4.2) et un élément f de son noyau vérifiant (trivialement) $\widehat{f} \in L^1(\mathbf{R})$, la formule d'inversion s'applique pour aboutir à l'égalité $f = 0$. ∎

Il est naturel de se demander à quelles fonctions s'applique la formule d'inversion. Sans décrire précisément l'espace des fonctions intégrables dont la transformée de Fourier l'est aussi, nous pouvons déjà noter que les éléments de $\mathscr{C}_c^\infty(\mathbf{R})$ vérifient cette propriété.

Proposition 4.9

Pour toute fonction $\varphi \in \mathscr{C}_c^\infty(\mathbf{R})$, on a $\widehat{\varphi} \in L^1(\mathbf{R})$. En particulier, la formule d'inversion est vérifiée pour les éléments de $\mathscr{C}_c^\infty(\mathbf{R})$.

Démonstration: Puisque $\varphi \in \mathscr{C}_c^\infty(\mathbf{R})$ on a bien sûr $\varphi \in L^1(\mathbf{R})$ (ce qui permet de parler de $\widehat{\varphi}$) mais également $\varphi' \in L^1(\mathbf{R})$ et $\varphi'' \in L^1(\mathbf{R})$ (en fait : toutes les dérivées de φ). Cela nous permet d'invoquer deux fois le troisième point de la Proposition 4.2, ce qui conduit à $\widehat{\varphi''} = -\xi^2\widehat{\varphi}$. En particulier, puisque la transformée de Fourier d'une fonction intégrable est bornée, nous venons d'établir que $\xi \mapsto (1+\xi^2)\widehat{\varphi}(\xi)$ est bornée ; par le critère de Riemann cela implique bien l'intégrabilité de $\widehat{\varphi}$. ∎

À l'aide de la formule d'inversion, il est possible d'établir un autre résultat fondamental de l'analyse de Fourier.

Théorème 4.10 (*Formule de Parseval*)

Soit $\varphi \in \mathscr{C}_c^\infty(\mathbf{R})$. Alors $\widehat{\varphi} \in L^2(\mathbf{R})$ et

$$\boxed{\|\widehat{\varphi}\|_2 = \sqrt{2\pi}\|\varphi\|_2.}$$

Remarque 4.11

Par un raisonnement de densité (du type de ceux effectués dans ce cours), il est possible de prolonger cette formule aux éléments de $L^1(\mathbf{R}) \cap L^2(\mathbf{R})$ (et on utilisera cette propriété dans les exercices). La preuve de se fait se trouve en annexe (Section A.5).

La « vraie » formule de Parseval s'applique en réalité à n'importe quel élément de $L^2(\mathbf{R})$, mais il faut pour cela avoir déjà étendu la définition de la transformation de Fourier à cet espace.

Démonstration : Grâce à la Proposition 4.9, on sait que $\widehat{\varphi}$ est intégrable. Par ailleurs, grâce à la Proposition 4.2, $\widehat{\varphi}$ est également une fonction bornée. Cela nous donne déjà l'appartenance $\widehat{\varphi} \in L^2(\mathbf{R})$, puisque $|\widehat{\varphi}(\xi)|^2 \leq \|\widehat{\varphi}\|_\infty |\widehat{\varphi}(\xi)| \in L^1(\mathbf{R})$. Ensuite, l'idée est d'appliquer le point 5. de 4.2 et la formule d'inversion de Fourier. Plus précisément, écrivons que

$$\|\widehat{\varphi}\|_2^2 = \int_{\mathbf{R}} |\widehat{\varphi}|^2 = \int_{\mathbf{R}} \widehat{\varphi}\,\overline{\widehat{\varphi}}. \tag{4.1}$$

On a par ailleurs

$$\begin{aligned}
\overline{\widehat{\varphi}(\xi)} &= \overline{\int_{\mathbf{R}} e^{-ix\xi}\varphi(x)\,\mathrm{d}x} \\
&= \int_{\mathbf{R}} e^{ix\xi}\overline{\varphi}(x)\,\mathrm{d}x \\
&= \int_{\mathbf{R}} e^{-ix(-\xi)}\overline{\varphi}(x)\,\mathrm{d}x,
\end{aligned}$$

ce que l'on peut synthétiser en $\overline{\widehat{\varphi}} = \check{\overline{\varphi}}$. Retournant à (4.1), puisque φ et $\check{\overline{\varphi}}$ sont toutes les deux intégrables, le point 5. de la proposition 4.2 fournit

$$\|\widehat{\varphi}\|_2^2 = \int_{\mathbf{R}} \widehat{\varphi}\,\check{\overline{\varphi}} = \int_{\mathbf{R}} \varphi\,\widehat{\check{\overline{\varphi}}}.$$

Le deuxième point de la Proposition 4.2 permet d'intervertir les opérations $f \mapsto \widehat{f}$ et $f \mapsto \check{f}$, si bien que nous avons finalement obtenu

$$\|\widehat{\varphi}\|_2^2 = \int_{\mathbf{R}} \varphi\,\check{\widehat{\overline{\varphi}}},$$

là où la formule d'inversion (qui s'applique aux éléments de $\mathscr{C}_c^\infty(\mathbf{R})$ par la proposition 4.9) nous assure justement que

$$\check{\widehat{\overline{\varphi}}} = 2\pi\overline{\varphi},$$

ce qui permet de conclure. ∎

Annexe A

Compléments

A.1 Quelques précisions sur la mesurabilité et la mesure de Lebesgue

On donne ici quelques précisions sur la notion de mesurabilité des fonctions et des ensembles : l'ensemble $\mathscr{M}$ des fonctions mesurables, défini dans l'Énoncé indispensable 1.5, ne l'est pas d'une manière explicite. Les quelques définitions qui suivent montrent comment cet ensemble est construit. Pour ce faire, on commence par voir comment on définit l'ensemble des *sous-ensembles* mesurables de $\mathbf{R}^d$. Tous les détails de cette anexe sont du programme de L3 et, à ce titre, seront vues en détail l'année prochaine.

Proposition A.1

Si X est un ensemble et $\mathcal{T} \subset \mathcal{P}(X)$ est un ensemble de parties de X, il existe une plus petite[1] *tribu*[2] *de parties de X contenant $\mathcal{T}$, appelée* tribu engendrée par $\mathcal{T}$.

Si on prend pour $\mathcal{T}$ les ensembles ouverts de $\mathbf{R}^d$, on obtient une tribu $\mathcal{B}(\mathbf{R}^d)$ appelée *la tribu de Borel de* $\mathbf{R}^d$. Si on prend pour $\mathcal{T}$ les ouverts de $\overline{\mathbf{R}}$ (qui sont des unions d'ensembles choisis parmi des intervalles de la forme $[-\infty, b[$, $]a, b[$ et $]b, +\infty[$ pour a et b finis arbitraires), on obtient la tribu $\mathcal{B}(\overline{\mathbf{R}})$ des Boréliens de $\overline{\mathbf{R}}$.

Attention ! La tribu $\mathcal{B}(\mathbf{R}^d)$ ne coïncide pas avec la tribu des ensembles mesurables $\mathcal{L}(\mathbf{R}^d)$ de $\mathbf{R}^d$ telle que définie dans la Définition 1.11 : elle est incluse dans cette tribu. Pour obtenir la tribu des ensembles mesurables de $\mathbf{R}^d$, il faut lui ajouter les ensembles de mesure nulle, comme nous le verrons très bientôt.

Le théorème principal de la théorie de la mesure de Lebesgue, dont la démonstration, trop technique, ne sera pas donnée ici, est la construction de la mesure de Lebesgue.

2. Au sens de l'inclusion.
2. La notion de tribu est défini dans la Définition 1.14.

Théorème A.2

On munit $\mathbf{R}^d$ de sa tribu des Boréliens. Il existe une unique mesure $\lambda : \mathcal{B}(\mathbf{R}^d) \to \overline{\mathbf{R}}_+$, appelée mesure de Lebesgue, qui :

1. *est invariante par translation, i.e., vérifie $\lambda(a+B) = \lambda(B)$ pour tout Borelien B et*
2. *vérifie la normalisation*

$$\lambda(]a_1, b_1[\times \cdots \times]a_n, b_n[) = \prod_{i=1}^{n} (b_i - a_i).$$

L'introduction de cette mesure définie sur $\mathcal{B}(\mathbf{R}^d)$ permet de définir la notion d'ensemble négligeable : un sous-ensemble de $\mathbf{R}^d$ est dit *négligeable* s'il est *inclus* dans un ensemble de $\mathcal{B}(\mathbf{R}^d)$ de mesure nulle.

Cela permet de définir la tribu des *ensembles mesurables*[3] de $\mathbf{R}^d$ comme étant la tribu engendrée par les ouverts de $\mathbf{R}^d$ et les ensembles négligeables. C'est cette tribu $\mathcal{L}(\mathbf{R}^d)$ qu'on retrouve dans la Définition 1.11. Ainsi, un ensemble borélien est mesurable, mais un ensemble mesurable peut ne pas être borélien.

Un procédé simple permet d'étendre la mesure du Théorème A.2 en une mesure sur les mesurables de $\mathbf{R}^d$; la mesure obtenue est celle de l'Énoncé indispensable 1.15 (et la restriction de cette mesure aux boréliens est celle du Théorème A.2).

Passons à la notion de mesurabilité pour les fonctions.

Définition A.3

Soient $(X, \mathcal{T}_X)$ et $(Y, \mathcal{T}_Y)$ deux ensembles munis chacun d'une tribu. Une fonction $f : X \to Y$ est dite *mesurable* si pour tout $E \in \mathcal{T}_Y$, on a $f^{-1}(E) \in \mathcal{T}_X$.

L'ensemble $\mathscr{M}$ des fonctions mesurables, défini dans l'Énoncé indispensable 1.5 est alors défini comme étant l'ensemble des fonctions mesurables de $\mathbf{R}^d$ muni de la tribu des mesurables dans $\overline{\mathbf{R}}$ muni de la tribu des boréliens.

On vérifie alors que cet ensemble est bien universellement stable, et contient toutes les fonctions continues et les indicatrices d'ensembles ouverts.

A.2 Construction de l'intégrale à partir de la mesure

Comme on l'a déjà dit, la construction classique de l'intégrale de Lebesgue se fait à partir de celle de la mesure de Lebesgue (et non pas le contraire comme c'est fait ici). Nous donnons ici quelques éléments concernant cette construction : dans toute cette section on ne garde en tête que les propriétés de la mesure de Lebesgue, à partir desquelles on va définir une intégrale.

3. Lorsque cela prête à confusion, on appelle ces ensembles *Lebesgue-mesurables*.

On appelle *fonction étagée* une fonction mesurable qui ne prend qu'un nombre fini de valeurs, autrement dit une fonction du type

$$f = \sum_{i=1}^{n} \alpha_i \mathbf{1}_{A_i},$$

où les α_i sont des nombres complexes et les A_i des parties mesurables de $\mathbf{R}^d$. Par exemple, la fonction $\mathbf{1_Q}$ est une fonction étagée.

On peut alors montrer que toute fonction mesurable est approchée par des fonctions étagées. Ces dernières fonctions jouent quelque part le rôle des fonctions en escaliers dans l'intégration de Riemann : on commence par définir l'intégrale d'une fonction étagée de la manière suivante. Si $f = \sum_{i=1}^{n} \alpha_i \mathbf{1}_{A_i}$, alors on pose $\int f = \sum_{i=1}^{n} \alpha_i \lambda(A_i)$. Il faut ensuite vérifier que cette quantité est indépendante de l'écriture de f comme une somme (par exemple la fonction $\mathbf{1_Q}$ peut aussi s'écrire $\mathbf{1}_{\mathbf{Q}\cap\mathbf{R}_+^*} + \mathbf{1}_{\mathbf{Q}\cap\mathbf{R}_-}$).

Cela permet de définir l'*intégrale* d'une fonction mesurable positive par :

$$\boxed{\int f = \sup\left\{\int g \mid g \text{ étagée positive et } g \leq f\right\}}.$$

C'est à partir de cette définition qu'il faut ensuite récupérer les propriétés de l'intégrale des fonctions positives énoncées dans le Théorème 1.5 (en dehors de la la propriété d'approximation, qui elle est directement héritée de la mesure).

A.3 Espaces $L^p(\mathbf{R})$

Plusieurs énoncés du Chapitre 3 énoncent des propriétés similaires pour les espaces $L^1(\mathbf{R})$ et $L^2(\mathbf{R})$. En réalité ces deux espaces font partie de la plus large famille des espaces $L^p(\mathbf{R})$, pour $p \in [1, \infty]$. Pour $p \in [1, \infty[$, sans trop de surprise, on définit $L^p(\mathbf{R})$ comme l'ensemble des (classes de) fonctions $f : \mathbf{R} \to \mathbf{C}$ mesurables pour lesquelles

$$\int_{\mathbf{R}} |f|^p < \infty.$$

On montre qu'il s'agit d'un espace vectoriel et que l'expression précédente, lorsqu'elle est finie et élevée à la puissance $1/p$, définit une norme : $\|f\|_p$. En comparaison des cas $p = 1, 2$ que nous avons traité, l'inégalité triangulaire pour les normes $\|\cdot\|_p$ demande un peu plus de travail et est connue sous le nom d'*inégalité de Minkowski*. Une manière de la prouver repose sur une autre inégalité fondamentale : l'*inégalité de Hölder*. Ces inégalités seront traités en TD.

Terminons par le cas $p = \infty$, que nous avons laissé à part. L'étude des normes $\|\cdot\|_p$ d'un élément de $f \in \mathscr{C}_c^0(\mathbf{R})$ fixé à mesure que $p \to +\infty$ est instructive : on montre en effet que $\|f\|_p \to \|f\|_\infty$, où la norme infinie a été définie dans ce cours comme le *supremum* de $|f|$ sur $\mathbf{R}$. Attention cependant, dans le cas d'une fonction continue, il n'existe qu'un seul élément dans sa classe d'équivalence *modulo* l'égalité presque partout. Dans le cas général, on introduit donc plutôt le *supremum essentiel*, soit :

$$\|f\|_\infty := \inf\{\mathrm{C} > 0 \,:\, |f| \leq C \text{ presque partout }\}.$$

Cette expression est bien définie dès lors que f admet un majorant presque partout (ou bien, une fonction bornée dans sa classe d'équivalence) (on dit qu'une telle fonction est *essentiellement bornée*). Comme pour les autres espaces $L^p(\mathbf{R})$, l'espace $L^\infty(\mathbf{R})$ est alors défini comme quotient des fonctions essentiellement bornées par la relation d'équivalence d'égalité presque partout.

Les espaces $L^p(\mathbf{R})$ jouent un rôle central dans le domaine des mathématiques appelé l'*analyse fonctionnelle*. L'une des nombreuses forces de l'intégrale de Lebesgue par rapport à celles de Riemann est d'avoir pu produire une hiérarchie d'espaces d'intégrabilité complets. L'étude générale des espaces de Banach s'applique agréablement aux espaces $L^p(\mathbf{R})$ grâce au théorème suivant.

Théorème A.4 (*Riesz-Fisher*)

Pour tout $p \in [1, \infty]$, l'espace $(L^p(\mathbf{R}), \|\cdot\|_p)$ est complet.

A.4 Preuve de la formule d'inversion

Pour établir la preuve du Théorème 4.7 nous utiliserons deux petits lemmes.

Lemme A.5

Soit $f \in L^1(\mathbf{R})$, et $g \in L^1(\mathbf{R})$. Alors pour tout $x \in \mathbf{R}$, on a

$$\mathcal{F}(\widehat{f}g) = \check{f} \star \widehat{g}.$$

Démonstration : Notons déjà que les appartenances $f, g \in L^1(\mathbf{R})$ et $\hat{f}, \hat{g} \in \mathscr{C}_b^0(\mathbf{R})$ assurent que les deux membres sont bien définis (en utilisant notamment la Proposition 3.25). Par la définition de la transformée de Fourier et successivement les deuxième et cinquième points de la Proposition 4.2, il vient

$$\begin{aligned}\mathcal{F}(\widehat{f}g)(\xi) &= \int_{\mathbf{R}} e^{-i\xi x}\widehat{f}(x)g(x)\,\mathrm{d}x\\ &= \int_{\mathbf{R}} \widehat{\tau_\xi f}g\\ &= \int_{\mathbf{R}} (\tau_\xi f)\widehat{g},\end{aligned}$$

et par définition de la convolution, le dernier terme n'est rien d'autre que $\check{f} \star \widehat{g}$. ■

Lemme A.6

Pour $n \in \mathbf{N}^\star$ posons $\gamma_n(x) := \gamma(\frac{x}{n})$, où γ est la gaussienne centrée réduite $x \mapsto e^{-x^2/2}$. Alors

(i) $(\gamma_n(x))_n$ converge simplement vers 1 par valeurs inférieures ;

(ii) $(\frac{1}{2\pi}\widehat{\gamma_n})_n$ est une approximation de l'unité.

Démonstration : Pour (i), il suffit de remarquer que par continuité de la fonction γ, pour tout $x \in \mathbf{R}$, la suite $\gamma_n(x) := \gamma(\frac{x}{n})$ tend vers $\gamma(0) = 1$.

Pour (ii) le deuxième point la Proposition 4.2 montre que $\widehat{\gamma_n}(x) = n\widehat{\gamma}(nx)$. Par ailleurs, on sait par la Proposition 4.5 que $\widehat{\gamma} = \sqrt{2\pi}\gamma$ puisque γ n'est rien d'autre que $g_{1/2}$ suivant les notations de cette proposition. Finalement, en utilisant le Lemme 3.30 on voit que $\frac{1}{2\pi}\widehat{\gamma_n}(x) = \frac{1}{\sqrt{2\pi}}n\gamma(nx)$ définit bien une approximation de l'unité puisque l'intégrale de γ vaut $\sqrt{2\pi}$ par l'Exemple 2.40. ■

Démonstration du Théorème 4.7 : On applique le Lemme A.5 à la fonction $g = \gamma_n$ du Lemme A.6 :

$$\mathcal{F}(\widehat{f}\gamma_n) = \check{f} \star \widehat{\gamma_n}. \tag{A.1}$$

Lorsque n tend vers $+\infty$, le membre de droite converge dans $L^1(\mathbf{R})$ vers $2\pi\check{f}$ grâce au Lemme A.6 combiné au Théorème 3.31. Pour le membre de gauche, on a $|\gamma_n| \leq 1$ et la convergence simple $(\gamma_n)_n \to 1$: on récupère ainsi par convergence dominée (puisque $\widehat{f}$ est supposée intégrable) que $(\widehat{f}\gamma_n)_n$ converge vers $\widehat{f}$ dans $L^1(\mathbf{R})$. D'après le premier point de la Proposition 4.2, la transformée de Fourier envoie continûment $L^1(\mathbf{R})$ dans $\mathscr{C}_b^0(\mathbf{R})$: on a donc convergence uniforme du membre de gauche de (A.1) vers $\mathcal{F}(\widehat{f})$. Pour établir l'identité annoncée $\mathcal{F}(\widehat{f}) = 2\pi\check{f}$ presque partout on peut par exemple remarquer que les convergences obtenues de part et d'autre ($L^1(\mathbf{R})$ à droite et uniforme à gauche) impliquent toutes les deux la convergence $L^1([a,b])$ sur tout segment $[a,b]$ de $\mathbf{R}$. Par unicité de la limite dans ce dernier espace, on obtient ainsi l'égalité. ■

A.5 Extension de l'identité de Parseval

Le but de cette section est de démontrer la formule de Parseval (quasi-isométrie quadratique de la transformée de Fourier) sur $L^1 \cap L^2(\mathbf{R})$ à partir du théorème 4.10 (à savoir cette même formule mais uniquement sur $\mathscr{C}_c^\infty(\mathbf{R})$). Si on *sait* que la transformée de Fourier est continue sur $L^1 \cap L^2(\mathbf{R})$ pour la norme quadratique, alors on peut raisonner par densité. Tout se ramène donc à démontrer l'inégalité de Bessel [4] $\|\hat{f}\|_2 \lesssim \|f\|_2$, pour toute fonction $f \in L^1 \cap L^2(\mathbf{R})$.

On commence par démontrer un résultat intermédiaire.

Lemme A.7

Si $f \in L^1 \cap L^2(\mathbf{R})$ est telle que $\hat{f} \in L^2(\mathbf{R})$, alors $\|\hat{f}\|_2 \lesssim \|f\|_2$.

Bien sûr la deuxième condition est en fait superflue.

Démonstration : Pour démontrer ce lemme, on utilise $(g_n)_n \in \mathscr{C}_c^\infty(\mathbf{R})$ approchant $\overline{\hat{f}}$ dans $L^2(\mathbf{R})$. On a par ailleurs la formule (point 5. de la Proposition 4.2)

$$\int_{\mathbf{R}} \hat{f} g_n = \int_{\mathbf{R}} f \hat{g}_n,$$

on en déduit par l'inégalité de Cauchy-Schwarz

$$\left| \int_{\mathbf{R}} \hat{f} g_n \right| \leq \|f\|_2 \|\hat{g}_n\|_2 \lesssim \|f\|_2 \|g_n\|_2,$$

4. Dans la suite, $f \lesssim g$ signifie qu'il existe une constante $C > 0$ telle que $f \leq Cg$.

où la deuxième inégalité (à une constante près) provient de la formule de Plancherel satisfaite sur les fonctions tests. On peut passer à la limite (encore une fois pas Cauchy-Schwarz) dans le membre d'extrême gauche et bien sûr dans le celui d'extrême droite pour finalement aboutir à $\|\hat{f}\|_2^2 \lesssim \|f\|_2\|\hat{f}\|_2$ et le résultat s'en déduit. ■

Il reste à montrer que la deuxième condition du lemme est effectivement superflue *i.e.* que pour tout élément $f \in L^1 \cap L^2(\mathbf{R})$, $\hat{f}$ est de carré intégrable. Si on parvient à construire une approximation de l'unité $(\theta_n)_n$ dont la transformée de Fourier $(\hat{\theta}_n)_n \in (\mathscr{C}_c^\infty(\mathbf{R}))^{\mathbf{N}}$ est à valeurs positives et converge vers 1 en croissant, alors on aura gagné : le fait précédent s'applique à chaque fonction $f \star \theta_n$ (c'est un élément de $L^1 \cap L^2(\mathbf{R})$ et sa transformée de Fourier $\hat{f}\hat{\theta}_n$ est bien de carré intégrable) montre que

$$\|\hat{f}\hat{\theta}_n\|_2 = \|\widehat{f \star \theta_n}\|_2 \lesssim \|f \star \theta_n\|_2.$$

Cette fois-ci le membre tout à gauche converge vers $\|\hat{f}\|_2$ (par convergence monotone, puisque $(\hat{\theta}_n)_n$ est positive et converge en croissant vers 1, donc son carré aussi) et celui tout à droite converge vers $\|f\|_2$ (par le théorème 3.31).

Pour conclure il reste à construire une approximation de l'unité $(\theta_n)_n$ répondant aux critères demandés ; on va faire ça en utilisant la formule d'inversion (il y a peut-être plus immédiat). On prend une fonction $\varphi \in \mathscr{C}_c^\infty(\mathbf{R})$ positive, paire, décroisante sur $\mathbf{R}_+$ et valant 1 en 0 ; en particulier $\varphi_n : x \mapsto \varphi(x/n)$ est alors une suite de fonctions positives croissant vers 1. La formule d'inversion peut s'utiliser pour φ (encore un résultat du cours) et elle nous donne donc que $\varphi = \frac{1}{2\pi}\hat{\hat{\varphi}}$, en particulier la fonction $\psi := \frac{1}{2\pi}\hat{\varphi}$ a pour intégrale 1. Pour tout n on pose $\theta_n(x) := n\psi(nx)$ de sorte que $(\theta_n)_n$ est une approximation de l'unité. Par ailleurs un changement de variables et la formule d'inversion précédemment citée montrent que $\widehat{\theta_n} = \varphi_n$. Finalement, la suite $(\theta_n)_n$ réunit les conditions demandées.

Remarque : *En vrai, dans le cours, les approximations de l'unité sont supposées positives, on peut également rejoindre ce cadre en considérant $\varphi\star\varphi$ et en normalisant sa valeur (non nulle) à l'origine : en faisant cela on a bien la positivité de $\widehat{\varphi\star\varphi} = \hat{\varphi}^2$, puisque $\hat{\varphi}$ est à valeurs réelles (parité).*

A.6 Un exemple d'ensemble non mesurable

Les ensembles et fonctions *mesurables* que nous avons manipulé dans ce cours sont un peu fantomatiques : on prétend que la « plupart » des ensembles (ou fonctions) vérifient ce critère sans même en avoir donné une définition précise ! Une construction due à Vitali permet, en utilisant l'axiome du choix, d'exhiber un ensemble non mesurable. En réalité, la preuve ne manipule pas la définition de ces ensembles directement : elle n'utilise que les conséquences que avons utilisées dans ce cours (notamment les propriétés de la mesure de Lebesgue). Voici un énoncé qui permet de formuler ce contre-exemple sans même faire référence à la notion de mesurabilité. Nous précisons par les initiales « AC » l'utilisation de l'axiome du choix dans sa preuve.

Proposition A.8 (AC)

Il n'existe pas de mesure μ sur l'ensemble de toutes les parties de $\mathbf{R}$ qui soit invariante *par translation et satisfasse* $\mu([a,b]) = b - a$ *pour tous réels* $a \leq b$.

Remarque A.9

La mesure de Lebesgue est invariante par translation par la Proposition 1.22 et l'axiome de normalisation assure bien que $\lambda([a,b]) = b - a$. Mais elle n'est *a priori* définie que sur les mesurables. C'est en ce sens que cette proposition prouve, moyennant l'axiome du choix, l'existence d'ensembles « non mesurables » : l'ensemble $\mathcal{V}$ manipulé dans la preuve n'est pas mesurable, ce qui « contamine » tous ses translatés et invalide donc les manipulations que l'on opère avec la mesure de Lebesgue sur ceux-ci.

Démonstration : On commence par définir une relation d'équivalence sur les nombres de $[0,1]$ par $x\mathcal{R}y$ si $x - y \in \mathbf{Q}$. On *choisit* — et c'est ici que l'axiome du choix intervient ! — alors un élément de $[0,1]$ dans chaque classe d'équivalence de cette relation, et on note $\mathcal{V}$ l'union de tous ces éléments.

Puisque $\mathbf{Q} \cap [-1,1]$ est dénombrable, on peut énumérer ses éléments : $\mathbf{Q} \cap [-1,1] = \{r_1, r_2, \ldots\}$. On vérifie que

(*i*) $[0,1] \subset \bigcup_{n\in\mathbf{N}}(r_n + \mathcal{V})$: tout élément $x \in [0,1]$ est dans la classe d'équivalence d'un certain $v \in \mathcal{V}$, donc $x - v \in \mathbf{Q} \cap [-1,1]$ et donc $x_v = r_n$ pour un certain $n \in \mathbf{N}$;

(*ii*) $\bigcup_{n\in\mathbf{N}}(r_n + \mathcal{V}) \subset [-1,2]$ (car pour tout $n \in \mathbf{N}$ et tout $v \in \mathcal{V}$, on a $r_n, v \in [0,1]$) ;

(*iii*) les ensembles $r_n + \mathcal{V}$ sont deux à deux disjoints, simplement parce que ce sont des classes d'équivalences différentes.

Si une telle mesure μ existait sur l'ensemble de toutes les parties de $\mathbf{R}$, alors l'invariance par translation imposerait que tous les ensembles $r_n + \mathcal{V}$ aient la même mesure, soit $\mu(\mathcal{V})$. De plus, des points (*i*) et (*ii*) donnés ci-dessus et de la croissance de la mesure, on en déduirait que

$$1 = \mu([0,1]) \leq \mu\left(\bigcup_{n\in\mathbf{N}}(r_n + \mathcal{V})\right) \leq \mu([-1,2]) = 3.$$

Enfin, par (*iii*), on aurait que

$$\mu\left(\bigcup_{n\in\mathbf{N}}(r_n + \mathcal{V})\right) = \sum_{n\in\mathbf{N}} \mu(r_n + \mathcal{V}) = \sum_{n\in\mathbf{N}} \mu(\mathcal{V}),$$

qui ne peut être entre 1 et 3 pour aucune valeur de $\mu(\mathcal{V})$: c'est une contradiction. ■

A.7 Primitives usuelles

Fonction	Primitive	Conditions
$x \mapsto x^n$	$\frac{1}{n+1}x^{n+1}$	$n \neq -1$
$x \mapsto x^{-1}$	$x \mapsto \ln\lvert x\rvert$	$x \neq 0$
$x \mapsto \frac{1}{(x+1)^2}$	$x \mapsto -\frac{1}{x+1}$	$x \neq -1$
$x \mapsto \frac{1}{1+x^2}$	$x \mapsto \arctan x$	
$x \mapsto \frac{1}{1-x^2}$	$x \mapsto \frac{1}{2}\ln\left\lvert\frac{x+1}{x-1}\right\rvert$	$x \neq \pm 1$
$x \mapsto \ln x$	$x \mapsto x\ln x - x$	$x > 0$
$x \mapsto a^x$	$x \mapsto \frac{a^x}{\ln a}$	$0 < a \neq 1$
$x \mapsto \sin x$	$x \mapsto -\cos x$	
$x \mapsto \cos x$	$x \mapsto \sin x$	
$x \mapsto \tan x$	$x \mapsto \ln\frac{1}{\lvert\cos x\rvert}$	$x \not\equiv \frac{\pi}{2}[2\pi]$
$x \mapsto \frac{1}{\sqrt{1-x^2}}$	$x \mapsto \arcsin x$	$x \in]-1,1[$

Bibliographie

[1] Jean-Michel Bony, “**Cours d’analyse**”, Ellipses, 2001.
C’est du premier chapitre de ce livre qu’est inspiré le plan de ces notes. L’auteur donne le b.a.-ba de l’intégration de Lebesgue pour pouvoir au plus vite définir les distributions.

[2] Bernard Candelpergher, “**Calcul intégral**”, Cassini, 2009.
Un livre de niveau licence, qui unifie les présentations des analyses réelles et complexes. Un bon complément à ce poly (les chapitres 2,6 et 8 concernent des notions en dehors de notre programme).

[3] Thierry Gallay, “**Théorie de la mesure et de l’intégration**”, notes de cours, 2009, https://www-fourier.ujf-grenoble.fr/~edumas/integration.pdf
Un poly très bien fait qui va un peu plus loin dans les notions abordées, en particulier concernant la théorie de la mesure.

[4] Amaury Lambert, “**Théorie de la Mesure et Intégration**”, notes de cours, 2012, https://www.lpsm.paris/pageperso/levy/LM364_Integration-Lambert.pdf
L’ancien poly du cours de L3 de l’UPMC, très bien écrit.

[5] Marc Briane, Gilles Pages, “**Analyse : Théorie de l’intégration**”, De Boeck Sup, 2018.
Une référence complète et auto-contenue qui présente l’intégrale de Lebesgue pour la mesure éponyme mais également d’autres mesures. Le livre commence par une petite introduction historique à l’intégrale de Riemann et explique les avantages de la construction de Lebesgue. C’est aussi une source fournie d’exercices en lien avec ce polycopié.

[6] Michel Willem, “**Analyse fonctionnelle élémentaire**”, Cassini, 2003.
Un bon livre pour aller un peu plus loin sur l’analyse fonctionnelle ; en particulier les premiers chapitres font un bon complément à ce cours. La présentation de la mesure de Lebesgue y est un peu différente. L’ouvrage contient plein d’exercices, mais qui sont généralement assez durs !