Théorie de l'intégration
Licence de Mathématiques (Parcours Math. Fond.)
2007-2011 : UE MHT512
2011-2014 : UE MA5012 [1]

Alain Yger

30 mai 2012

1. Les notes de ce cours doivent énormément aux notes du cours dispensé à l'université Paris XI dans les années 1970 par Jacques Deny ; elles en sont très largement inspirées. Le cours oral ne pouvant cependant refléter toute la teneur de ces notes, celles ci se veulent utiles pour une lecture et un approfondissement du cours « à tête reposée ». Ces notes sont accompagnées de guides hebdomadaires d'activité mis en ligne sur la plateforme Ulysse (Annales de l'UE MHT512). Ces notes de cours peuvent être utilisées dans le cadre de l'UE MA5012 prenant la suite de MHT512 dans la nouvelle habilitation 2011-2014.

Table des matières

Chapitre 1

Tribus et mesures positives ; les exemples de Borel et Lebesgue

1.1 Introduction

On connaît depuis le lycée et les cours de L1 et L2 l'approche classique du calcul d'*aire*, basée sur les idées développées par le mathématicien allemand Bernhard Riemann[1](1826-1866). Par exemple, si f est une fonction réelle positive et majorée sur le segment $[a, b]$, on peut définir sans ambiguïté l'aire du « sous-graphe »

$$\{(x, y) \in \mathbb{R}^2\,;\, x \in [a, b] \text{ et } 0 \leq y \leq f(x)\} \tag{1.1}$$

pourvu que la condition suivante soit satisfaite :

Critère d'intégrabilité de Riemann (pour une fonction réelle positive f sur $[a, b]$) : *pour tout $\epsilon > 0$, il existe deux fonctions positives « en escalier » (c'est-à-dire dont le « sous-graphe » se présente sous forme d'histogramme) φ_ϵ et ψ_ϵ telles que $\varphi_\epsilon \leq f \leq \psi_\epsilon$ sur $[a, b]$ et que*

$$0 \leq \int_a^b (\psi_\epsilon(t) - \varphi_\epsilon(t))\, dt \leq \epsilon\,.$$

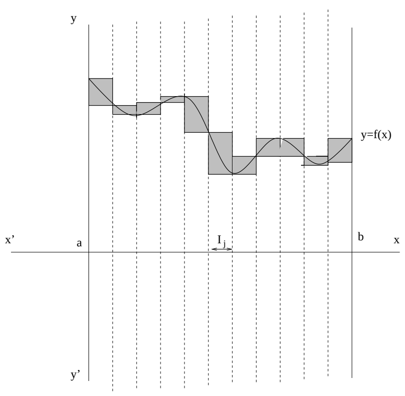

FIGURE 1.1 – Calcul approché d'aire suivant Riemann

1. Si l'on lui doit des travaux intéressants sur le concept d'intégrale (qu'avec Cauchy il contribua pour ses besoins à clarifier), ce sont ses travaux en géométrie différentielle (dont il posa dans sa dissertation en 1854 les premiers jalons) qui certainement constituent l'une des contributions majeures de Riemann aux mathématiques et à la physique.

L'aire du « sous-graphe » (1.1) (ou ici, ce qui revient au même, l'intégrale de f sur $[a,b]$) est définie alternativement comme la borne inférieure des intégrales[2] des fonctions en escalier ψ qui majorent f sur $[a,b]$, ou (ce qui revient au même pourvu que le critère d'intégrabilité soit satisfait) la borne supérieure des intégrales des fonctions en escalier positives φ qui minorent f sur $[a,b]$.

Si f est continue, ou bien monotone, ou encore s'exprimant comme différence de deux fonctions monotones (par exemple si f est « *à variations bornées*[3] »), plus généralement si f est *réglée* (c'est-à-dire admet des limites à gauche et à droite en tout point de $[a,b]$), alors le critère d'intégrabilité de Riemann est satisfait et on peut calculer l'aire du « sous-graphe » (1.1) et la noter d'ailleurs

$$\int_a^b f(t)\,dt \quad \text{ou encore} \quad \int_{[a,b]} f(t)\,dt$$

(c'est la définition dans ce cas de l'intégrale de f sur $[a,b]$ puisque f est supposée positive).

Cependant, suivant cette approche, il n'est pas possible de calculer l'aire de l'ensemble

$$\{(x,y)\in\mathbb{R}^2\,;\, x\in[0,1] \text{ et } 0\le y\le \chi_{\mathbb{Q}}(x)\}\,, \tag{1.2}$$

$\chi_{\mathbb{Q}}$ désignant la fonction définie sur $\mathbb{R}$ et valant 1 sur $\mathbb{Q}$ et 0 sur $\mathbb{R}\setminus\mathbb{Q}$, ou, ce qui revient au même, de donner un sens à l'intégrale

$$\int_0^1 \chi_{\mathbb{Q}}(t)\,dt\,.$$

Pourtant, on constate que la fonction $\chi_{\mathbb{Q}}$ que l'on prétend intégrer sur $[0,1]$ ne vaut 1 que sur $[0,1]\cap\mathbb{Q}$, c'est-à-dire sur un ensemble dénombrable, autant dire pratiquement nulle part dans l'ensemble de définition $[0,1]$, qui lui, on le sait, n'est pas un ensemble dénombrable ! 0n aurait donc envie de dire que la surface de l'ensemble défini en (1.2) vaut 0 mais la théorie inspirée de l'approche de Riemann nous en empêche : la fonction $\chi_{\mathbb{Q}}$ ne satisfait pas le critère d'intégrabilité de Riemann sur $[0,1]$ car la seule fonction en escalier positive φ qui la minore est la fonction nulle (sauf peut-être en un nombre fini de points rationnels ou elle peut prendre des valeurs entre 0 et 1) puisque $[0,1]\setminus\mathbb{Q}$ est dense dans $[0,1]$, tandis que toute fonction en escalier positive ψ qui la majore est minorée par 1 sur $[0,1]$ (sauf éventuellement en un nombre fini de points isolés) puisque $[0,1]\cap\mathbb{Q}$ est dense dans $[0,1]$; l'intégrale de $\psi-\varphi$ sur $[0,1]$ est toujours au moins égale à 1, donc ne saurait être rendue arbitrairement petite.

Nous allons voir dans cette section comment rectifier le tir, à savoir élargir la classe des sous-ensembles de $\mathbb{R}^n$ que l'on saura « mesurer » et en même temps donner un sens à la notion de mesure n-dimensionnelle dans $\mathbb{R}^n$.

2. Le calcul de l'intégrale d'une fonction en escalier positive sur $[a,b]$ revient juste au calcul de la surface d'un histogramme, donc au calcul de la somme de surfaces d'un nombre fini de rectangles à côtés parallèles aux axes, ce qui ne pose aucun problème.

3. *i.e* il existe une constante $C>0$ telle que pour toute subdivision $x_0=a<x_1<\ldots<x_N=b$ ($N\ge 1$ quelconque), on ait

$$\sum_{k=1}^N |f(x_k)-f(x_{k-1})|<C\,.$$

Rappelons cependant ici que la méthode la plus naïve de tenter de calculer une aire est inspirée des probabilités et est d'origine très ancienne : imaginons par exemple que nous souhaitions calculer la surface du domaine fractal A proposé sur la figure 1.2 (en blanc sur la figure) et que le cadre rectangulaire dans lequel cette figure est inscrite soit supposé d'aire normalisée égale à 1.

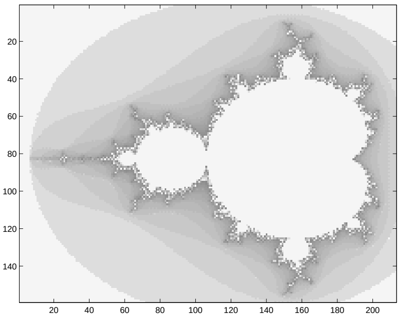

FIGURE 1.2 – Un domaine plan « fractal »

On « jette » des points sur ce cadre, tous les points de chute étant équiprobables. On fait cela un nombre N (très très grand de fois), puis on divise par N le nombre de jets (N_{fav}) où le point est tombé sur la cible (le domaine A dont on veut calculer l'aire). Dans les bons cas (c'est-à-dire lorsque parler de cette aire a un sens), la loi des grands nombres[4] (outil clef du principe sur lequel se fonde le raisonnement statistique, comme dans les sondages d'opinion par exemple) permet d'assurer que lorsque N (le nombre de jets, ceux ci étant supposés indépendants) tend vers $+\infty$, alors le quotient N_{fav}/N tend vers l'aire de A. Cette méthode très intuitive est connue comme la *méthode de Monte Carlo* et permet, lentement, c'est vrai, de calculer l'aire d'un domaine (si toutefois cela est possible, c'est-à-dire si l'on sait associer sans ambiguïté une « mesure » à ce domaine plan A en tentant de le mesurer (suivant une idée inspirée de la démarche de Riemann pour le « sous-graphe » (1.1)) de l'extérieur, puis de l'intérieur et de vérifier ensuite que les deux approches coïncident.

Il ne faut pas être trop optimiste et penser que l'on puisse tout « mesurer » ! La possibilité de faire effectivement un calcul de mesure est intimement liée à la notion profonde de *dénombrabilité* (on le verra constamment dans ce chapitre). Sortir de ce cadre, c'est se priver de procédés algorithmiques pour « calculer » et « mesurer ». En 1923 par exemple, le mathématicien, logicien et philosophe polonais Alfred Tarski (1903-1983) a proposé le célèbre paradoxe suivant, dit paradoxe de Banach[5]-Tarski : on peut faire un puzzle avec une sphère pleine S de $\mathbb{R}^3$ de rayon 1 (*i.e* réaliser une partition de S en un nombre fini de parties disjointes) et reconstituer avec les morceaux du puzzle deux sphères de même volume que la sphère initiale ! Il est clair ici que, si un tel paradoxe surgit, c'est qu'il existe des sous-ensembles de la sphère pleine auxquels on est incapable d'affecter un volume 3-dimensionnel (il en figure au moins un dans notre partition). Construire une classe de sous-ensembles

4. Il s'agit en fait ici d'une approche empirique à la notion d'*espérance*, voir l'UE MHT601 à venir.

5. Stefan Banach (1892-1945), mathématicien polonais, encore plus connu qu'Alfred Tarski (avec qui il mis en lumière ce paradoxe dit aujourd'hui de Banach-Tarski), fut l'un des fondateurs de l'*analyse fonctionnelle* ; le concept d'*espace de Banach* (espace vectoriel *normé complet*) est omniprésent dans toutes les branches des mathématiques.

de $\mathbb{R}^n$ « mesurables » est donc une première nécessité avant d'étudier le problème de l'intégration [6] des fonctions sur des sous-ensembles « mesurables ».

Les physiciens envisageant aussi le concept de répartition de masse non nécessairement homogène, il sera aussi utile de développer dans un cadre plus général le concept de mesure positive : une famille de parties d'un ensemble Ω étant donnée avec certaines propriétés, que signifie « l'équiper d'une mesure positive » ? Si Ω est un ensemble abstrait et la « mesure » de masse totale 1, ces constructions seront les ingrédients essentiels de la théorie des probabilités (dissociée de ce cours d'intégration, elle fera l'objet de l'UE MHT601 au second semestre).

Voilà donc les divers objectifs que nous nous assignons dans ce premier chapitre.

1.2 Algèbres de Boole et tribus sur un ensemble abstrait

Si Ω désigne un ensemble (pour l'instant abstrait), alors, suivant l'axiomatique de la théorie des ensembles, on peut définir un nouvel ensemble $\mathcal{P}(\Omega)$ (dit des « parties de Ω ») dont les éléments sont les sous-ensembles de Ω. Cet ensemble $\mathcal{P}(\Omega)$ peut être équipé naturellement de deux opérations internes :

$$\begin{array}{rcl} (A,B) \in \mathcal{P}(\Omega) \times \mathcal{P}(\Omega) & \longmapsto & A \cup B \text{ (union de A et B)} \\ (A,B) \in \mathcal{P}(\Omega) \times \mathcal{P}(\Omega) & \longmapsto & A \cap B \text{ (intersection de A et B)} \end{array}$$

correspondant respectivement aux opérations logiques de *disjonction* et de *conjonction.*

Définition 1.1 *Une algèbre de Boole* [7] *sur Ω est par définition un sous-ensemble de $\mathcal{E}$ de l'ensemble $\mathcal{P}(\Omega)$ des parties de Ω tel que*
- *$\mathcal{E}$ est stable par union finie, i.e, pour tout A_1, A_2 éléments de $\mathcal{E}$, $A_1 \cup A_2$ est encore un élément de $\mathcal{E}$;*
- *$\mathcal{E}$ est stable par prise de complémentaire, i.e, pour tout A élément de $\mathcal{E}$, $\Omega \setminus A$ est encore un élément de $\mathcal{E}$;*
- *l'ensemble vide $\emptyset$ est un élément de $\mathcal{E}$.*

Remarque 1.1. Une algèbre de Boole $\mathcal{E}$ sur Ω est stable à la fois sous les opérations d'union et d'intersection finies (entre parties de Ω) et contient toujours l'élément Ω de $\mathcal{P}(\Omega)$ (la partie « pleine ») ; si $\mathcal{E}$ est une algèbre de Boole, ceci veut dire que toute union (*resp.* intersection) finie d'éléments de $\mathcal{E}$ est encore un élément de $\mathcal{E}$. On a d'ailleurs la règle de distributivité classique :

$$\forall A_1, A_2, A_3 \in \mathcal{E}, \quad A_1 \cap (A_2 \cup A_3) = (A_1 \cap A_2) \cup (A_1 \cap A_3).$$

Une autre opération interne importante d'une algèbre de Boole $\mathcal{E}$ est la *différence symétrique* définie par

$$A_1 \Delta A_2 := (A_1 \setminus A_2) \cup (A_2 \setminus A_1), \quad A_1, A_2 \in \mathcal{E}.$$

6. Notons tout de même que chercher à mesurer $A \subset \mathbb{R}^n$, c'est chercher à calculer l'intégrale sur $\mathbb{R}^n$ de la fonction caractéristique χ_A de A (valant 1 sur A et 0 ailleurs) ; intégration des fonctions et théorie de la mesure sont deux questions étroitement liées ; le titre de la thèse d'Henri Lebesgue en 1901, « *Intégrale, longueur, aire* » ne le résume t-il pas ?

7. George Boole, logicien, mathématicien et philosophe anglais (1815-1864) ; c'est à lui que l'on doit les bases de la logique classique, mais ce fut aussi un analyste et, inévitablement bien sûr, un des premiers probabilistes.

Elle jouera un rôle important dans le concept de mesure que nous allons introduire ultérieurement.

Remarque 1.2. Pour vérifier qu'une partie $\mathcal{E}$ de $\mathcal{P}(\Omega)$ est une algèbre de Boole sur Ω, il peut s'avérer plus judicieux de vérifier qu'elle contient l'élément Ω, est stable par passage au complémentaire, et est telle que l'intersection de deux éléments de $\mathcal{E}$ est encore un élément de $\mathcal{E}$. En effet comme

$$\begin{aligned} A_1 \cup A_2 &= \Big(\Omega \setminus (\Omega \setminus A_1)\Big) \cup \Big(\Omega \setminus (\Omega \setminus A_2)\Big) \\ &= \Omega \setminus \Big((\Omega \setminus A_1) \cap (\Omega \setminus A_2)\Big) \end{aligned}$$

pour $A_1, A_2 \in \mathcal{E}$, $\mathcal{E}$ est aussi stable dans ce cas par union finie (c'est cette idée que l'on peut utiliser par exemple pour traiter les exemples 1.2 ci-dessous).

Exemple 1.1 (exemples basiques). Si Ω est un ensemble quelconque, $\{\emptyset, \Omega\}$ est une algèbre de Boole sur Ω (la « plus petite possible » du point de vue de l'inclusion entre parties de $\mathcal{P}(\Omega)$). L'ensemble $\mathcal{P}(\Omega)$ de toutes les parties de Ω est aussi une algèbre de Boole sur Ω (la « plus grosse possible » cette fois, toujours du point de vue de l'inclusion entre parties de $\mathcal{P}(\Omega)$).

Exemple 1.2 (sur $\mathbb{R}$ ou $\mathbb{R}^n$). Considérons $\Omega = \mathbb{R}$ et $\mathcal{E}^{(1)}$ l'ensemble des unions finies d'intervalles du type $]a, b]$ avec $-\infty < a \leq b < +\infty$ ou du type $]a, +\infty[$ ou $]-\infty, b]$ avec $a, b \in \mathbb{R}$ (ces intervalles sont dits *semi-fermés* à droite). On ajoute à cette liste l'ensemble vide $\emptyset$. On vérifie sans peine que $\mathcal{E}^{(1)}$ est bien une algèbre de Boole sur $\mathbb{R}$ (faites l'exercice). Si $\Omega = \mathbb{R}^n$, l'ensemble $\mathcal{E}^{(n)}$ des parties de $\mathbb{R}^n$ s'écrivant comme union finie de « pavés élémentaires » du type

$$I_1 \times \cdots \times I_n\,,$$

où $I_1, ..., I_n$ sont des intervalles semi-fermés à droite de $\mathbb{R}$, est aussi une algèbre de Boole, cette fois sur $\mathbb{R}^n$ (se forcer à faire encore l'exercice). On remarque que $\mathcal{E}^{(n)}$ est aussi l'ensemble des unions finies de toutes les parties de $\mathbb{R}^n$ de la forme

$$A_1 \times \cdots \times A_n\,,$$

où $A_1, ..., A_n$ sont des parties de $\mathbb{R}$ éléments de l'algèbre de Boole $\mathcal{E}^{(1)}$ sur $\mathbb{R}$ (notons qu'un tel ensemble n'est nullement en général un pavé élémentaire du type $I_1 \times \cdots \times I_n$, $I_1, ..., I_n$ étant des intervalles semi-fermés à droite de $\mathbb{R}$!).

Exemple 1.3 (algèbre de Boole sur un produit). Si $\mathcal{E}_1$ est une algèbre de Boole sur un ensemble Ω_1 et $\mathcal{E}_2$ une algèbre de Boole sur un ensemble Ω_2, la famille des unions finies de parties

$$A_1 \times A_2\,,\ A_1 \in \mathcal{E}_1\,,\ A_2 \in \mathcal{E}_2\,,$$

de l'ensemble produit $\Omega_1 \times \Omega_2$ est une algèbre de Boole sur cet ensemble produit ; c'est d'ailleurs sur ce modèle que l'on peut construire l'algèbre de Boole $\mathcal{E}^{(2)}$ sur $\mathbb{R}^2$ à partir de l'algèbre de Boole $\mathcal{E}^{(1)}$ sur $\mathbb{R}$. L'algèbre réalisée ainsi est notée $\mathcal{E}_1 \otimes \mathcal{E}_2$ et appelée *produit* des algèbres de Boole $\mathcal{E}_1$ et $\mathcal{E}_2$. On pourra vérifier en exercice la stabilité par prise de complémentaire et intersection finie de $\mathcal{E}_1 \otimes \mathcal{E}_2$.

Profitons ici de faire quelques remarques importantes autour de la notion (cruciale en théorie de la mesure, on le verra tout au long de ce cours) de *dénombrabilité*.

On rappelle qu'un ensemble *dénombrable* (infini) est un ensemble en bijection avec $\mathbb{N}$; c'est le cas de $\mathbb{Z}$, de $\mathbb{Q}$, de $\mathbb{Z}^n$, de $\mathbb{Q}^n$; dire qu'un ensemble est dénombrable signifie que l'on peut en « numéroter » les éléments sous la forme d'une suite exhaustive $x_0, x_1, ..., ...$ d'éléments deux-à-deux distincts.

On sait que $\mathbb{R}$ (qui vérifie le critère des segments emboîtés) n'est pas un ensemble dénombrable. Si Ω est dénombrable infini, l'ensemble $\mathcal{P}(\Omega)$ des parties de Ω (qui est toujours en bijection[8] avec l'ensemble des applications de Ω dans $\{0, 1\}$), est

8. Il suffit d'associer à une partie A sa *fonction caractéristique* χ_A, c'est-à-dire la fonction valant 1 sur A, 0 sur $\Omega \setminus A$.

(dans ce cas particulier où Ω est en bijection avec $\mathbb{N}$ car supposé dénombrable infini) en bijection avec l'ensemble des applications de $\mathbb{N}$ dans $\{0,1\}$, c'est-à-dire avec l'ensemble $\{0,1\}^{\mathbb{N}}$ des suites $(x_k)_{k\in\mathbb{N}}$ dont le terme général vaut 0 ou 1, lui même en bijection avec l'ensemble $\{0,1\}^{\mathbb{N}^*}$ des suites $(x_k)_{k\geq 1}$ indexées par les entiers positifs non nuls. Or on a le

Lemme 1.1 *L'application*

$$(x_k)_{k\geq 1} \longmapsto \sum_{k=1}^{\infty} \frac{x_k}{3^k}$$

est une application injective de $\{0,1\}^{\mathbb{N}^*}$ *dans* $\mathbb{R}$*, tandis que l'application qui à un élément* x *de* $[0,1[$ *associe la suite* $(x_k)_{k\geq 1}$ *correspondant à l'unique développement dyadique propre*

$$x = \sum_{k=1}^{\infty} \frac{x_k}{2^k}$$

(tel que les coefficients x_k *ne sont pas tous égaux à* 1 *au delà d'un certain rang) est une application injective de* $[0,1[$ *dans* $\{0,1\}^{\mathbb{N}^*}$.

Preuve. Si l'on a

$$\sum_{k=1}^{\infty} \frac{x_k}{3^k} = \sum_{k=1}^{\infty} \frac{x'_k}{3^k}$$

et si k_0 désigne le plus petit entier k tel que $x_k \neq x'_k$ (s'il tant est qu'il en existe un !), on a

$$\frac{x_{k_0} - x'_{k_0}}{3^{k_0}} = \sum_{k_0+1}^{\infty} \frac{x'_k - x_k}{3^k}$$

et, en prenant les valeurs absolues et en utilisant l'inégalité triangulaire :

$$3^{-k_0} \leq \sum_{k_0+1}^{\infty} 3^{-k} = \frac{3^{-k_0}}{2}\,,$$

ce qui est absurde et prouve la première assertion. La seconde est immédiate. ◇

Ce lemme, combiné avec le célèbre théorème de Cantor-Bernstein [9] que l'on admettra (on pourra le traiter en exercice), nous permet d'affirmer que $\mathbb{R}$ et l'ensemble des parties d'un ensemble dénombrable infini sont deux ensembles en bijection, donc en particulier que, dès que Ω est infini, $\mathcal{P}(\Omega)$ n'est pas dénombrable (et est même en bijection avec $\mathbb{R}$ si Ω est infini dénombrable).

Le fait qu'il soit possible, étant donné une collection $(A_\iota)_{\iota\in\mathcal{I}}$ (indexée par un ensemble d'indices $\mathcal{I}$) de parties non vides d'un ensemble Ω, de construire une application $F\ :\mathcal{I}\to\Omega$ telle que

$$\forall \iota \in \mathcal{I}\,,\ F(\iota) \in A_\iota$$

est immédiat si $\mathcal{I}$ est fini et se prouve aisément grâce aux axiomes de Peano (en particulier l'axiome fondant le raisonnement par récurrence) lorsque $\mathcal{I}$ est infini

9. S'il existe une injection d'un ensemble dans un autre et *vice versa*, alors les deux ensembles sont en bijection. Ceux qui viennent de Classes Préparatoires ont sans doute croisé cet important résultat, la preuve n'étant pas très difficile mais subtile car basée sur une bonne compréhension de l'*hérédité*.

dénombrable. En revanche, si ce n'est plus le cas, il convient de faire de ce principe un axiome (dit axiome du choix). Sortir du cadre dénombrable nous oblige bien souvent à invoquer cet axiome dont vous verrez plus tard des incarnations autrement moins naïves (par exemple le lemme de Zorn).

Les éléments de l'algèbre de Boole $\mathcal{E}^{(n)}$ sont des parties de $\mathbb{R}^n$ très particulières, au sens où l'on sait immédiatement leur attacher un *volume* : le volume d'un pavé élémentaire $]a_1, b_1] \times \cdots \times]a_n, b_n]$ (supposé non vide, sinon on décide que ce volume est nul) est naturellement

$$(b_1 - a_1) \times \cdots \times (b_n - a_n)$$

lorsque tous les nombres a_k et b_k sont réels et vaut $+\infty$ si l'un des $]a_i, b_i]$ est un intervalle non borné. Il est aisé ensuite, par un jeu de manipulations ensemblistes (voir la figure 1.3 ci-dessous), de calculer le volume d'une union finie de tels ensembles, donc d'un élément de $\mathcal{E}^{(n)}$.

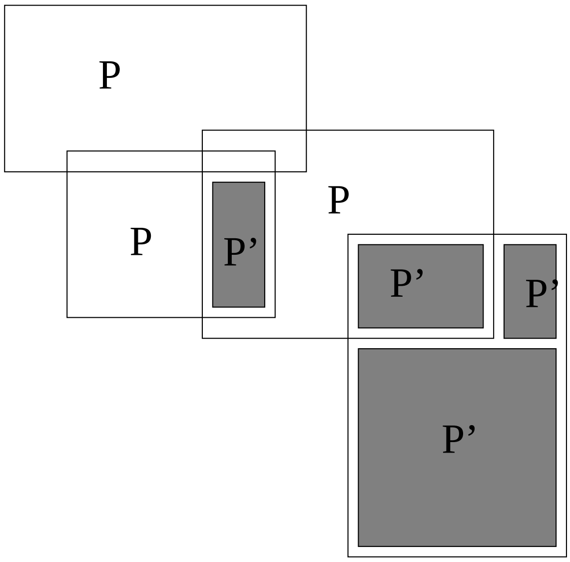

FIGURE 1.3 – « Découpage » d'un élément de $\mathcal{E}^{(2)}$

On aimerait bien sûr ultérieurement élargir la classe d'ensembles auxquels on souhaiterait pouvoir attacher un volume, donc travailler avec des familles de parties moins restreintes que les algèbres de Boole. On voudrait en fait que la famille, outre le fait d'être une algèbre de Boole, ait en plus la propriété d'être stable non plus seulement par union finie, mais par union dénombrable (pour ne pas avoir à affronter l'axiome du choix, on prendra soin de ne pas demander plus). Cela va nous donner, on le verra, un moyen de « grossir » de manière très significative des algèbres de Boole telles par exemples que notre algèbre de référence $\mathcal{E}^{(n)}$ sur $\mathbb{R}^n$.

Définition 1.2 *On appelle tribu (ou encore σ-algèbre) toute algèbre de Boole $\mathcal{T}$ sur Ω ayant de plus la propriété d'être stable par union dénombrable, ce qui signifie que si $(A_k)_{k\in\mathbb{N}}$ est une suite de parties de Ω appartenant à $\mathcal{T}$, l'ensemble*

$$\bigcup_{k\in\mathbb{N}} A_k := \{x \in \Omega\,;\, \exists k = k(x) \in \mathbb{N}\,,\ x \in A_k\}$$

est encore un élément de $\mathcal{T}$.

Si $(\mathcal{T}_\iota)_{\iota\in\mathcal{I}}$ est une famille de tribus (indexée par un ensemble arbitraire d'indices I) sur le même ensemble Ω, le sous ensemble

$$\bigcap_{\iota\in\mathcal{I}} \mathcal{T}_\iota$$

de $\mathcal{P}(\Omega)$ constitué des parties de Ω qui appartiennent à <u>toutes</u> les σ-algèbres $\mathcal{T}_\iota$ est encore une tribu ; c'est d'ailleurs la plus grosse tribu (toujours au sens de l'inclusion dans l'ensemble $\mathcal{P}(\Omega)$) incluse dans toutes les tribus $\mathcal{T}_\iota$, $\iota \in \mathcal{I}$.

En revanche, si $\mathcal{T}_1$ et $\mathcal{T}_2$ sont deux tribus sur un même ensemble Ω, la famille obtenue en concaténant les familles $\mathcal{T}_1$ et $\mathcal{T}_2$, c'est-à-dire le sous ensemble $\mathcal{T}_1 \cup \mathcal{T}_2$ de $\mathcal{P}(\Omega)$, peut fort bien ne pas être une tribu ! Il suffit de prendre sur $\mathbb{R}^2$ les deux tribus

$$\begin{aligned}\mathcal{T}_1 &:= \{A\times\mathbb{R}\,;\, A\subset\mathbb{R}\}\\ \mathcal{T}_2 &:= \{\mathbb{R}\times B\,;\, B\in\mathbb{R}\}\end{aligned}$$

et de constater que l'intersection de $[0,1]\times\mathbb{R}$ et de $\mathbb{R}\times[0,1]$, égale à $[0,1]\times[0,1]$ n'est ni un élément de $\mathcal{T}_1$, ni un élément de $\mathcal{T}_2$.

L'ensemble $\mathcal{P}(\Omega)$ est par contre toujours une tribu, comme d'ailleurs l'ensemble $\{\emptyset,\Omega\}$ (constituant la plus petite de toutes les tribus possibles, dite *tribu grossière*).

Si $(\mathcal{T}_\iota)_{\iota\in\mathcal{I}}$ est une collection de tribus sur Ω, l'intersection de toutes les tribus contenant

$$\bigcup_{\iota\in\mathcal{I}} \mathcal{T}_\iota := \{A\in\mathcal{P}(\Omega)\,;\, \exists \iota = \iota(A)\,,\ A\in\mathcal{T}_\iota\}$$

($\mathcal{P}(\Omega)$ en est une) est, d'après ce qui précède, une tribu, dite *tribu engendrée par les* $\mathcal{T}_\iota$, $\iota\in\mathcal{I}$. Cette tribu est la plus petite tribu (toujours au sens de l'inclusion dans $\mathcal{P}(\Omega)$) contenant toutes tribus $\mathcal{T}_\iota$, $\iota\in\mathcal{I}$. Plus généralement, on a la

Définition 1.3 *Si $\mathcal{G}\subset\mathcal{P}(\Omega)$, on appelle tribu engendrée par $\mathcal{G}$ et on note $\mathcal{T}(\mathcal{G})$ l'intersection de toutes les tribus sur Ω contenant $\mathcal{G}$.*

Si E est une partie spécifique de Ω et $\mathcal{T}$ une tribu sur Ω, on sait construire une tribu $\mathcal{T}_E$ sur E en définissant

$$\mathcal{T}_E := \{A\cap E\,;\, A\in\mathcal{T}\}\,.$$

Il est en effet immédiat de vérifier que $\mathcal{T}_E$ est bien une tribu sur E : l'intersection étant distributive par rapport à l'union, on a stabilité de $\mathcal{T}_E$ par union dénombrable, la stabilité par prise de complémentaire venant du fait que $E\cap(\Omega\setminus A) = E\setminus(E\cap A)$ pour tout $A\in\mathcal{T}$. Cette tribu « restreinte » $\mathcal{T}_E$ est appelée *tribu trace* de $\mathcal{T}$ sur E. On la retrouvera souvent.

1.3 La tribu borélienne sur $\mathbb{R}^n$ et $\overline{\mathbb{R}}$

L'algèbre de Boole $\mathcal{E}^{(1)}$ ne saurait être une tribu sur $\mathbb{R}$: si tel était le cas en effet, tout intervalle ouvert $]a,b[$ avec $-\infty\leq a<b\leq+\infty$ appartiendrait à $\mathcal{E}^{(1)}$ puisque l'on a

$$]a,b[= \bigcup_{n=N}^{\infty} \Big]a, b-\frac{1}{n}\Big]\,, \tag{1.3}$$

dès que $N > 1/(b-a)$; or, si $b < +\infty$, un tel intervalle $]a,b[$ ne saurait être union finie d'intervalles $]a_k, b_k]$ puisqu'une telle union finie se doit de contenir sa borne supérieure (ce qui n'est pas le cas pour $]a,b[$). *A fortiori*, l'algèbre de Boole $\mathcal{E}^{(n)}$ n'est pas une tribu sur $\mathbb{R}^n$.

Néanmoins, une tribu sur $\mathbb{R}^n$ intimement liée à l'algèbre de Boole $\mathcal{E}^{(n)}$ sera appelée à jouer un rôle majeur dans la suite du cours ; la terminologie la qualifiant fait référence à l'un des fondateurs de la théorie de la mesure moderne (autour des années 1900-1910), le mathématicien français Emile Borel (1871-1956).

Définition 1.4 *La tribu borélienne sur $\mathbb{R}^n$ est par définition la tribu sur $\mathbb{R}^n$ engendrée par l'algèbre de Boole $\mathcal{E}^{(n)}$; c'est aussi la plus petite tribu sur $\mathbb{R}^n$ contenant tous les ouverts (et par conséquent tous les fermés), ou encore la plus petite tribu contenant tous les pavés ouverts $]a_1,b_1[\times\cdots\times]a_n,b_n[$, avec $-\infty \leq a_i \leq b_i \leq +\infty$, $i=1,...,n$.*

Preuve des affirmations incluses dans la définition. Il est clair, du fait de (1.3), que la tribu borélienne contient tous les pavés ouverts $]a_1,b_1[\times\cdots\times]a_n,b_n[$, avec $-\infty \leq a_i \leq b_i \leq +\infty$, $i=1,...,n$. D'autre part, si U est un ouvert de $\mathbb{R}^n$, on sait que pour tout $x \in U$, il existe un pavé ouvert P_x tel que $x \in P_x \subset U$, ce qui nous permet d'écrire

$$U = \bigcup_{x\in U} P_x \,;$$

comme $\mathbb{Q}^n$ est dense dans P_x et que $\mathbb{Q}^+$ est dense dans $]0,+\infty[$, il est aussi possible de construire, pour chaque $x \in U$, un pavé ouvert $\widetilde{P}_x$ inclus dans P_x, contenant x, et de la forme $]\alpha_{x,1},\beta_{x,1}[\times\cdots\times]\alpha_{x,n},\beta_{x,n}[$, où les nombres $\alpha_{x,i},\beta_{x,i}$, $i=1,...,n$, sont tous rationnels (faites un dessin par exemple si $n=2$) ; U s'écrit donc aussi

$$U = \bigcup_{x\in U} P_x = \bigcup_{x\in U} \widetilde{P}_x \,,$$

et, puisque $\mathbb{Q}$ (donc aussi $\mathbb{Q}^{2n}$) est dénombrable, on peut affirmer que cette écriture de U sous forme d'union de pavés ouverts $\widetilde{P}_x$ peut se réduire à une écriture de U comme union <u>dénombrable</u> de pavés ouverts. Ce que nous venons de faire nous montre aussi que toute tribu contenant l'algèbre de Boole $\mathcal{E}^{(n)}$ doit automatiquement contenir tous les pavés ouverts, donc aussi tous les ouverts et, par conséquent, tous les complémentaires d'ouverts, donc tous les fermés. Les affirmations incluses dans la définition sont donc bien prouvées. $\diamondsuit$

On admettra ici que la tribu borélienne sur $\mathbb{R}^n$, bien que riche (elle contient toutes les unions dénombrables d'ensembles ouverts ou fermés, toutes les intersections dénombrables de telles unions, *etc.*), n'est cependant pas la tribu $\mathcal{P}(\mathbb{R}^n)$ de toutes les parties de $\mathbb{R}^n$: si l'on invoque l'axiome du choix (déjà mentionné), il est possible [10] d'affirmer l'existence de sous-ensembles de $\mathbb{R}$ qui ne soient pas dans la tribu borélienne sur $\mathbb{R}$; ceci vaut donc aussi pour $\mathbb{R}^n$. Dans la suite du cours, on notera $\mathcal{B}(\mathbb{R}^n)$ la tribu borélienne sur $\mathbb{R}^n$.

Une des propriétés importantes de la tribu borélienne $\mathcal{B}(\mathbb{R}^n)$ (qui nous sera utile lors de l'énoncé du théorème de Lebesgue 1.1 à venir, Section 1.5) est son *invariance par translation* : si A est un borélien et $x \in \mathbb{R}^n$, l'ensemble $A + x := \{a + x \,;\, a \in A\}$

10. La démarche n'est cependant nullement « effective ».

est encore un borélien. Ceci résulte du fait que cette propriété est vraie pour les éléments de $\mathcal{E}^{(n)}(\mathbb{R})$, combiné avec le fait que $\mathcal{B}(\mathbb{R}^n)$ est la tribu engendrée par cette algèbre de Boole (voir aussi l'exercice 7 dans le guide d'activités « Séance 2 » sous Ulysse).

On aura aussi besoin dans ce cours de la notion de tribu borélienne sur la droite réelle achevée $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$.

Définition 1.5 *La tribu borélienne $\mathcal{B}(\overline{\mathbb{R}})$ sur la droite réelle achevée est la tribu engendrée par les intervalles $]a, b]$ avec $-\infty \leq a \leq b \leq +\infty$ (toujours semi-fermés à droite, la borne supérieure étant incluse, même si elle vaut $+\infty$).*

1.4 Notion de mesure positive sur $(\Omega, \mathcal{T})$

Considérons un ensemble Ω équipé d'une tribu $\mathcal{T}$. Définir une mesure positive sur Ω revient à définir une « répartition de masse » sur Ω, les seules parties de Ω prises en compte étant les éléments de la tribu $\mathcal{T}$.

Définition 1.6 *Soit Ω un ensemble et $\mathcal{T}$ une tribu sur Ω. On appelle mesure positive sur $(\Omega, \mathcal{T})$ toute application*

$$\mu \ : \mathcal{T} \longrightarrow [0, +\infty]$$

telle que $\mu(\emptyset) = 0$ et que, pour toute suite $(A_k)_k$ finie ou infinie dénombrable d'éléments de $\mathcal{T}$ deux-à-deux disjoints, on ait

$$\mu\Big(\bigcup_k A_k\Big) = \sum_k \mu(A_k) \in [0, \infty]\,. \tag{1.4}$$

Exemple 1.4. Si Ω est un ensemble quelconque et $\mathcal{T} = \mathcal{P}(\Omega)$, on définit une mesure sur $(\Omega, \mathcal{P}(\Omega))$, dite *mesure de décompte* sur Ω, en posant $\mu(A) = \operatorname{card} A$ si A est une partie finie de Ω, $\mu(A) = +\infty$ si A est une partie infinie de Ω.

Exemple 1.5. Soit Ω un ensemble et $\mathcal{T}$ une tribu sur Ω ayant la propriété de séparer un point donné a des autres points de Ω (pour tout $b \neq a$, il existe des parties A, B dans $\mathcal{T}$, disjointes et telles que $a \in A$ et $b \in B$). Une mesure alors intéressante sur $(\Omega, \mathcal{T})$ est la mesure δ_a de Dirac [11] au point a définie par $\delta_a(A) = 1$ si A est un élément de $\mathcal{T}$ contenant a, $\delta_a(A) = 0$ sinon.

Voici une proposition listant les propriétés majeures d'une mesure positive.

Proposition 1.1 *Soit μ une mesure positive sur $(\Omega, \mathcal{T})$, $\mathcal{T}$ désignant une tribu sur un ensemble Ω.*

- *si A et B sont deux éléments de $\mathcal{T}$ tels que $A \subset B$, on a $\mu(A) \leq \mu(B)$.*
- *si $(A_k)_{k \geq 0}$ est une suite finie ou infinie dénombrable d'éléments de $\mathcal{T}$, on a*

$$\mu\Big(\bigcup_k A_k\Big) \leq \sum_k \mu(A_k) \tag{1.5}$$

(propriété de σ-sous-additivité) ;

11. Physicien et mathématicien britannique (1902-1984), Paul Adrien Maurice Dirac fut amené à introduire ce concept d'impulsion ou de « fonction » de support concentré en un point et de « masse » l'unité, concept auquel seule la théorie de la mesure peut donner une signification mathématique précise.

– *si* $(A_k)_{k\geq 0}$ *est une suite croissante* ($A_k \subset A_{k+1}$ *pour tout entier* k) *d'éléments de* $\mathcal{T}$, *on a*

$$\lim_{k\to+\infty} \mu(A_k) = \mu\Big(\bigcup_{k\in\mathbb{N}} A_k\Big)\,; \tag{1.6}$$

– *si* $(A_k)_{k\geq 0}$ *est une suite décroissante* ($A_{k+1} \subset A_k$ *pour tout entier* k) *d'éléments de* $\mathcal{T}$, *on a, pourvu que l'un au moins des* A_k *soit de mesure finie*[12],

$$\lim_{k\to+\infty} \mu(A_k) = \mu\Big(\bigcap_{k\in\mathbb{N}} A_k\Big)\,. \tag{1.7}$$

Preuve. Le premier point résulte de ce que

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

puisque A et $B\setminus A$ sont deux éléments de $\mathcal{T}$ distincts. Pour le second point, on peut supposer que tous les A_n sont de mesure finie (sinon, l'inégalité (1.5) est immédiate). Si l'on pose $B_0 = A_0$ et

$$B_k = A_k \setminus \bigcup_{l=0}^{k-1} A_l$$

pour $k \geq 1$, on remarque que les B_k sont deux-à-deux disjoints, tous de mesure finie, et que

$$\bigcup_k A_k = \bigcup_k B_k\,;$$

d'après la σ-additivité (propriété (1.4)), on a

$$\mu\Big(\bigcup_k A_k\Big) = \sum_k \mu(B_k) \leq \sum_k \mu(A_k)\,.$$

Pour le troisième point, on pose $B_0 = A_0$ et $B_k = A_k \setminus A_{k-1}$ pour $k \geq 1$; on a, pour tout k

$$A_k = \bigcup_{l=0}^{k} B_l$$

et, toujours par σ-additivité,

$$\mu(A_k) = \sum_{l=0}^{k} \mu(B_l)$$

puisque les B_l sont deux-à-deux disjoints. On a donc

$$\lim_{k\to+\infty} \mu(A_k) = \sum_l \mu(B_l) = \mu\Big(\bigcup_l B_l\Big) = \mu\Big(\bigcup_l A_l\Big) = \mu\Big(\bigcup_k A_k\Big)\,.$$

Concernant la dernière propriété mentionnée, on se ramène au cas où A_0 est de mesure finie et on considère la suite croissante $(A_0 \setminus A_k)_{k\geq 0}$; on a

$$\lim_{k\to+\infty} \mu(A_0\setminus A_k) = \mu\Big(A_0 \setminus \bigcap_k A_k\Big)$$

12. Cette restriction est imparable, on en donnera un exemple plus loin.

d'après la propriété (1.6), soit

$$\lim_{k\to+\infty} (\mu(A_0) - \mu(A_k)) = \mu(A_0) - \mu\Big(\bigcap_k A_k\Big)\,,$$

d'où la propriété (1.7). $\diamondsuit$

Remarque 1.3. Si A_1 et A_2 sont deux éléments de $\mathcal{T}$ tous les deux de mesure μ finie, la formule exacte (précisant l'inégalité (1.5)) est

$$\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2) - \mu(A_1 \cap A_2)\,;$$

si maintenant A_1, A_2, A_3 sont trois éléments de $\mathcal{T}$ de mesure μ finie, on trouve facilement que

$$\begin{aligned}
\mu(A_1 \cup A_2 \cup A_3) &= \mu\Big((A_1 \cup A_2) \cup A_3\Big) = \mu(A_1 \cup A_2) + \mu(A_3) - \mu\Big(A_1 \cup A_2) \cap A_3\Big) \\
&= \mu(A_1) + \mu(A_2) + \mu(A_3) - \mu(A_1 \cap A_2) - \mu\Big((A_1 \cap A_3) \cup (A_2 \cap A_3)\Big) \\
&= \sum_{j=1}^{3} \mu(A_j) - \Big(\sum_{1\leq i<j\leq 3} \mu(A_i \cap A_j)\Big) + \mu(A_1 \cap A_2 \cap A_3).
\end{aligned}$$

Il est aisé de poursuivre ces calculs (par exemple par récurrence sur n) pour vérifier que, si $A_1, ..., A_n$ sont n éléments de $\mathcal{T}$ tous de mesure μ finie, on a

$$\mu(A_1 \cup \ldots \cup A_n) = \sum_{k=1}^{n} (-1)^{k-1} \mu_k(A_1, ..., A_n)\,, \tag{1.8}$$

avec

$$\mu_k(A_1, ..., A_n) := \sum_{1\leq i_1<\cdots<i_k\leq n} \mu(A_{i_1} \cap \ldots \cap A_{i_k})\,.$$

C'est un exercice qui constitue un bon entraînement aux règles de calcul inhérentes au concept de mesure. On retrouve ce type de somme alternée (1.8) dans nombre de formules mathématiques issues, c'est naturel, de la combinatoire.

1.5 La mesure de Lebesgue sur $\mathbb{R}^n$

Dans cette importante section, nous allons nous atteler à définir une mesure positive d'un intérêt capital sur $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$; l'idée qui a présidé à la construction de cette mesure a été formalisée en 1901 [13] par le mathématicien français Henri Léon Lebesgue (1875-1941), sur la base des travaux déjà initiés par E. Borel.

Voici l'énoncé majeur de cette section (et certainement un des énoncés majeurs du cours), étant entendu que la construction effective qui l'accompagne est tout aussi importante (sinon plus) que l'énoncé lui-même :

Théorème 1.1 *Il existe une unique mesure positive μ sur $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ ayant les deux propriétés suivantes :*

- *la mesure du pavé fermé borné $[0,1]^n$ est égale à 1 ;*
- *si A est un borélien de $\mathbb{R}^n$ et x un élément de $\mathbb{R}^n$, les ensembles A et*

$$A + x := \{a + x\,;\, a \in A\}$$

sont tels que $\mu(A) = \mu(A + x)$ (on dit que la mesure μ a la propriété d'invariance par translation).

Cette section (un peu longue) va être entièrement dévolue à la construction de la mesure de Lebesgue sur $\mathbb{R}^n$.

13. C'est dans sa célèbre thèse : « *Intégrale, longueur, aire* » publiée à Milan en 1902 dans les *Annali di Mathematica* que Lebesgue développa complètement sa théorie de l'intégration.

1.5.1 La notion de mesure « extérieure » d'un sous-ensemble E de $\mathbb{R}^n$

L'idée force de Lebesgue est, dans un premier temps, de tenter de mesurer « de l'extérieur » n'importe quel sous-ensemble E de $\mathbb{R}^n$. Pour ce faire, on considère tous les recouvrements $\mathcal{R}$ possibles de E par des ensembles du type

$$\bigcup_{\iota \in \mathcal{J}_{\mathcal{R}}} P_{\mathcal{R},\,\iota}$$

où l'ensemble d'indices $\mathcal{J}_{\mathcal{R}}$ est fini ou dénombrable et les $P_{\mathcal{R},\iota}$ sont des pavés ouverts de $\mathbb{R}^n$ du type $]a_1, b_1[\times \ldots \times]a_n, b_n[$, avec $-\infty \leq a_i < b_i \leq +\infty$ pour $i = 1, ..., n$. L'idée de ne pas se contenter des recouvrements de E par des unions finies de pavés ouverts, mais d'envisager les recouvrements par des unions finies ou dénombrables de tels pavés est l'une des idées capitales de Lebesgue [14].

Définition 1.7 *Si E est un sous-ensemble de $\mathbb{R}^n$, on appelle mesure extérieure de E (et on note $\mu^*(E)$) la borne inférieure, pour tous les recouvrements $\mathcal{R}$ possibles de E du type*

$$E \subset \bigcup_{\iota \in \mathcal{J}_{\mathcal{R}}} P_{\mathcal{R},\,\iota}$$

où les $P_{\mathcal{R},\,\iota}$ sont des pavés ouverts indexés par un ensemble $\mathcal{J}_{\mathcal{R}}$ fini ou dénombrable, de l'ensemble des nombres

$$\sum_{\iota \in \mathcal{J}_{\mathcal{R}}} \mu(P_{\mathcal{R},\,\iota})\,,$$

la mesure du pavé ouvert $P =]a_1, b_1[\times \ldots]a_n, b_n[$ valant $\mu(P) = +\infty$ si ce pavé est non borné et $\mu(P) = (b_1 - a_1) \times \ldots \times (b_n - a_n)$ si P est borné. Si $E = \emptyset$, on convient de ce que $\mu^*(E) = 0$.

Exemple 1.6. Soit $n = 1$. Si E est un intervalle borné de $\mathbb{R}$ (de quelque type soit-il), la mesure extérieure de E coïncide avec la longueur usuelle de E car il suffit de recouvrir E par des intervalles ouverts le contenant (on prend dans ce cas des recouvrements par un seul pavé ouvert). Toujours si $n = 1$, on constate, comme tout ouvert de $\mathbb{R}$ est une union finie ou dénombrable d'intervalles ouverts disjoints, que l'on a

$$\mu^*(E) = \inf_{E \subset U} \mu^*(U)\,, \tag{1.9}$$

U décrivant la famille d'ouverts contenant E.

Exemple 1.7. Si n est quelconque, la mesure extérieure d'un élément E de l'algèbre de Boole $\mathcal{E}^{(n)}$ est égale à son volume euclidien et vaut $+\infty$ si E est non borné (on recouvre ici par un nombre fini de pavés ouverts).

Exemple 1.8. Si E est un sous-ensemble borné de $\mathbb{R}^n$ inclus dans un hyperplan affine $\langle a, x \rangle = C$, la mesure extérieure de E est nulle car on peut le recouvrir par des pavés « s'écrasant » sur cet hyperplan. En particulier, la mesure extérieure d'un point, plus généralement d'un sous-ensemble E fini ou dénombrable de $\mathbb{R}^n$, d'une face de pavé fermé $[a_1, b_1] \times \cdots \times [a_n, b_n]$, est toujours nulle.

14. Reprendre ici la relecture de l'introduction du chapitre, car c'est ici un des points majeurs où les points de vue de Riemann et Lebesgue se différencient : chez Riemann, la subdvision associée à une fonction étagée ψ majorant f lorsque l'on veut calculer la surface du « sous-graphe » (1.1) est toujours une subdivision finie (si l'on travaille comme dans l'introduction au dessus d'un segment $[a, b]$).

Si $E_1 \subset E_2$, on a évidemment $\mu^*(E_1) \le \mu^*(E_2)$ car tout recouvrement de E_2 est aussi un recouvrement de E_1. Si $(E_k)_{k\in\mathbb{N}}$ est une suite finie ou dénombrable de sous-ensembles de $\mathbb{R}^n$, on a [15]

$$\mu^*\Big(\bigcup_k E_k\Big) \le \sum_k \mu^*(E_k)\,. \tag{1.10}$$

Le résultat est immédiat si l'un des nombres $\mu^*(E_k)$ est infini et l'on peut donc supposer que tous ces nombres sont finis. Prouver (1.10) dans ce cas est un bon moyen de rappeler la définition de la *borne inférieure* [16] d'un sous-ensemble de $[0,+\infty]$: rappelons que

$$\mu^*(E_k) := \inf_{\mathcal{R}_k} \Big(\sum_{\iota\in\mathcal{J}_{\mathcal{R}_k}} \mu(P_{\mathcal{R}_k,\iota})\Big)\,,$$

où l'inf est pris sur tous les recouvrements $\mathcal{R}_k$

$$E_k \subset \bigcup_{\iota\in\mathcal{J}_{\mathcal{R}_k}} P_{\mathcal{R}_k,\iota}$$

de E_k par des unions finies ou dénombrables de pavés ouverts ; ceci implique que, pour tout $\epsilon > 0$, on puisse trouver un recouvrement $\mathcal{R}_{k,\epsilon}$ de E_k tel que

$$\mu^*(E_k) \le \sum_{\iota\in\mathcal{J}_{\mathcal{R}_{k,\epsilon}}} \mu(P_{\mathcal{R}_{k,\epsilon},\iota}) \le \mu^*(E_k) + \frac{\epsilon}{2^k}\,.$$

En concaténant les recouvrements $\mathcal{R}_{k,\epsilon}$, $k\in\mathbb{N}$, on réalise un recouvrement $\mathcal{R}_\epsilon$ de $E = \bigcup_k E_k$ et l'on a donc

$$\begin{aligned}\mu^*(E) \le \sum_{\iota\in\mathcal{J}_{\mathcal{R}_\epsilon}} \mu(P_{\mathcal{R}_\epsilon,\iota}) \ &= \ \sum_k \Big(\sum_{\iota\in\mathcal{J}_{\mathcal{R}_{k,\epsilon}}} \mu(P_{\mathcal{R}_{k,\epsilon},\iota})\Big)\\ &\le \ \sum_k (\mu^*(E_k) + \epsilon/2^k) \le \sum_k \mu^*(E_k) + 2\epsilon\,;\end{aligned}$$

comme ϵ est arbitraire, on a bien l'inégalité (1.10).

Ces propriétés de la mesure extérieure induisent la proposition utile suivante :

Proposition 1.2 *Pour tout $E \subset \mathbb{R}^n$, la mesure extérieure de E s'atteint comme*

$$\mu^*(E) = \inf_{E\subset U} \mu^*(U)\,,$$

l'inf étant pris sur tous les ouverts U de $\mathbb{R}^n$ contenant E.

Preuve. Il est immédiat que $\mu^*(E) \le \mu^*(U)$ si U est un ouvert contenant E. Si $\epsilon > 0$, on peut trouver (toujours grâce au fait que $\mu^*(E)$ que l'on peut supposer ici

15. On ne peut pas espérer en général mieux, en particulier par exemple avoir l'égalité dans (1.10) lorsque que de plus les E_k sont deux-à-deux disjoints.

16. La *borne inférieure* α d'un sous ensemble non vide et minoré T de $\mathbb{R}$ est par définition le plus grand des minorants de T : d'une part, c'est un minorant de T (pour tout $t\in T$, on a $t\ge\alpha$), d'autre part, c'est le plus grand des minorants de T, ce qui signifie que pour tout $\epsilon > 0$, il doit exister au moins un élément de T dans $[\alpha, \alpha+\epsilon[$ ($\alpha+\epsilon$ n'étant plus minorant de T).

fini est défini comme une borne inférieure) un recouvrement $\mathcal{R}_\epsilon$ de E par des pavés ouverts

$$E \subset \bigcup_{\iota \in \mathcal{J}_{\mathcal{R}_\epsilon}} P_{\mathcal{R}_\epsilon, \iota}$$

de telle manière que

$$\begin{aligned} \mu^*(E) &\geq \sum_{\iota \in \mathcal{J}_{\mathcal{R}_\epsilon}} \mu(P_{\mathcal{R}_\epsilon, \iota}) - \epsilon \\ &\geq \sum_{\iota \in \mathcal{J}_{\mathcal{R}_\epsilon}} \mu^*(P_{\mathcal{R}_\epsilon, \iota}) - \epsilon \\ &\geq \mu^*(U_\epsilon) - \epsilon \,, \end{aligned}$$

où

$$U_\epsilon = \bigcup_{\iota \in \mathcal{J}_{\mathcal{R}_\epsilon}} P_{\mathcal{R}_\epsilon, \iota}$$

(ceci résulte de la propriété de σ sous-addtivité (1.10)). Comme U_ϵ est ouvert (union d'ouverts) et que ϵ est arbitraire, on a bien

$$\mu^*(E) \geq \inf_{E \subset U} \mu^*(U) \,,$$

ce qui achève de prouver le lemme. $\diamondsuit$

Autre proposition importante, concernant le cas très particulier de sous-ensembles compacts de $\mathbb{R}^n$:

Proposition 1.3 *Si K_1 et K_2 sont deux compacts disjoints, on a*

$$\mu^*(K_1 \cup K_2) = \mu^*(K_1) + \mu^*(K_2) \,.$$

Preuve. On sait déjà que $\mu^*(K_1 \cup K_2) \leq \mu^*(K_1) + \mu^*(K_2)$. On sait aussi, que, pour tout $\epsilon > 0$, on peut trouver un recouvrement $\mathcal{R}_\epsilon$ de $K_1 \cup K_2$,

$$K_1 \cup K_2 \subset \bigcup_{\iota \in \mathcal{J}_{\mathcal{R}_\epsilon}} P_{\mathcal{R}_\epsilon, \iota}$$

tel que

$$\mu^*(K_1 \cup K_2) \geq \sum_{\iota \in \mathcal{J}_{\mathcal{R}_\epsilon}} \mu(P_{\mathcal{R}_\epsilon, \iota}) - \epsilon \,.$$

Comme K_1 et K_2 sont compacts, on peut extraire du recouvrement $\mathcal{R}_\epsilon$ un recouvrement fini [17]. On travaille avec ce recouvrement fini et, quitte à redécouper les pavés en jeu, on peut supposer qu'aucun d'eux n'intersecte à la fois K_1 et K_2 (ces deux compacts sont à une distance strictement positive l'un de l'autre). On peut trier ainsi pavés $P_{\mathcal{R}_\epsilon, \iota}$ en deux catégories, ceux intersectant K_1 (et dont l'union recouvre K_1) et ceux intersectant K_2 (et dont l'union recouvre K_2). Les deux classes étant disjointes (elles correspondent aux ensembles d'indices $\mathcal{J}_{1,\mathcal{R}_\epsilon}$ et $\mathcal{J}_{2,\mathcal{R}_\epsilon}$), on a

$$\begin{aligned} \mu^*(K_1 \cup K_2) &\geq \sum_{\iota \in \mathcal{J}_{1,\mathcal{R}_\epsilon}} \mu(P_{\mathcal{R}_\epsilon, \iota}) + \sum_{\iota \in \mathcal{J}_{2,\mathcal{R}_\epsilon}} \mu(P_{\mathcal{R}_\epsilon, \iota}) - \epsilon \\ &\geq \mu^*(K_1) + \mu^*(K_2) - \epsilon \,, \end{aligned}$$

ce qui induit, ϵ étant arbitraire, l'inégalité

$$\mu^*(K_1 \cup K_2) \geq \mu^*(K_1) + \mu^*(K_2)$$

et achève donc la preuve de la proposition. $\diamondsuit$

17. C'est la classique caractérisation de la compacité, dite justement *propriété de Borel-Lebesgue.*

1.5.2 Comment mesurer un ensemble « de l'intérieur » ?

Après avoir tenté de « mesurer » un sous-ensemble E « de l'extérieur », nous essayons de le mesurer « de l'intérieur » en définissant sa *mesure intérieure* :

Définition 1.8 *La mesure intérieure $\mu_*(E)$ d'un sous-ensemble E de $\mathbb{R}^n$ est par définition la borne supérieure des nombres $\mu^*(K)$, K décrivant l'ensemble des compacts contenus dans E. On convient encore que $\mu_*(\emptyset) = 0$.*

Si $E_1 \subset E_2$, on a évidemment $\mu_*(E_1) \leq \mu_*(E_2)$ car tout compact inclus dans E_1 est aussi inclus dans E_2. Si $(E_k)_{k\in\mathbb{N}}$ est une suite finie ou dénombrable de sous-ensembles de $\mathbb{R}^n$ cette fois supposés deux-à-deux disjoints, on a [18]

$$\mu_*\Big(\bigcup_k E_k\Big) \geq \sum_k \mu_*(E_k)\,. \tag{1.11}$$

Prouver ceci est un bon moyen de rappeler cette fois la définition de la borne supérieure [19] : rappelons que

$$\mu_*(E_k) := \sup_{K\subset E_k} \mu^*(K)\,,$$

où le sup est pris sur tous les compacts inclus dans E_k ; ceci implique que, pour tout $\epsilon > 0$, on puisse trouver un compact $K_{k,\epsilon}$ de E_k tel que

$$\mu^*(K_{k,\epsilon}) \geq \mu_*(E_k) - \frac{\epsilon}{2^k}\,.$$

Soit $N \in \mathbb{N}$. Comme les $K_{k,\epsilon}$, $k = 0, ..., N$ sont deux à deux disjoints (comme les E_k), on a, d'après la proposition 1.3,

$$\mu^*\Big(\bigcup_{k=0}^N K_{k,\epsilon}\Big) \geq \sum_{k=0}^N \mu_*(E_k) - \epsilon\sum_{k=0}^N \frac{1}{2^k} \geq \sum_{k=0}^N \mu_*(E_k) - 2\epsilon\,.$$

On a donc, en revenant à la définition de la mesure intérieure de l'union des E_k, $k \in \mathbb{N}$,

$$\mu_*\Big(\bigcup_k E_k\Big) \geq \mu^*\Big(\bigcup_k K_{k,\epsilon}\Big) \geq \sum_k \mu_*(E_k) - 2\epsilon\,,$$

d'où l'inégalité (1.11) puisque ϵ est arbitraire.

1.5.3 La tribu des ensembles mesurables

Si l'on souhaite attribuer une valeur numérique finie à ce qui devrait être la mesure (ou ici le volume n-dimensionnel) d'un sous-ensemble E de $\mathbb{R}^n$, il est naturel d'exiger que la mesure évaluée « de l'extérieur » (c'est-à-dire $\mu^*(E)$) coïncide avec la mesure évaluée « de l'intérieur » (en l'occurrence $\mu_*(E)$) et soit de plus finie. Cela nous conduit à la définition suivante :

---

18. On ne peut toujours pas espérer en général mieux, en particulier par exemple avoir l'égalité dans (1.11), quand bien même les E_k sont cette fois supposés deux-à-deux disjoints.

19. La *borne supérieure* β d'un sous ensemble non vide et majoré T de $\mathbb{R}$ est par définition le plus petit des majorants de T : d'une part, c'est un majorant de T (pour tout $t \in T$, on a $t \leq \beta$), d'autre part, c'est le plus petit des majorants de T, ce qui signifie que pour tout $\epsilon > 0$, il doit exister au moins un élément de T dans $]\beta - \epsilon, \beta]$ ($\beta - \epsilon$ n'étant plus majorant de T).

Définition 1.9 *Une partie intégrable (au sens de Lebesgue) de $\mathbb{R}^n$ est un sous-ensemble A de $\mathbb{R}^n$ tel que $\mu^*(A) = \mu_*(A) < +\infty$. La valeur commune de ces deux nombres est notée $\mu(A)$ et appelée mesure de Lebesgue (ou encore volume euclidien n-dimensionnel) de A. Un partie mesurable (au sens de Lebesgue) de $\mathbb{R}^n$ est un sous-ensemble A de $\mathbb{R}^n$ tel que l'intersection de A avec tout pavé $P = (a_1, b_1)\times \ldots \times (a_n, b_n)$ borné soit un sous-ensemble intégrable*[20].

Un premier indice justifiant l'intérêt ultérieur de cette définition est le lemme suivant :

Lemme 1.2 *Soit $(A_k)_k$ une collection finie ou dénombrable de sous-ensembles de $\mathbb{R}^n$ deux-à-deux disjoints tels que $\mu^*(A_k) = \mu_*(A_k)$ pour tout k. Alors*

$$\mu^*\Big(\bigcup_k A_k\Big) = \mu_*\Big(\bigcup_k A_k\Big) = \sum_k \mu^*(A_k) = \sum_k \mu_*(A_k)\,. \tag{1.12}$$

Preuve. Le résultat est immédiat si l'un des nombres $\mu^*(A_k) = \mu_*(A_k)$ est infini et l'on peut donc supposer que tous ces nombres sont finis. On sait d'après (1.11) que

$$\mu_*\Big(\bigcup A_k\Big) \geq \sum_k \mu_*(A_k)$$

et, d'après (1.10), que

$$\mu^*\Big(\bigcup A_k\Big) \leq \sum_k \mu^*(A_k)\,.$$

Le fait que les deux nombres $\sum_k \mu^*(A_k)$ et $\sum_k \mu_*(A_k)$ soient ici égaux implique la chaîne d'égalités (1.12) voulue. $\diamondsuit$

Il résulte de ce lemme que tout ouvert U de $\mathbb{R}^n$ vérifie $\mu^*(U) = \mu_*(U)$ puisqu'on peut l'écrire comme union dénombrable de pavés ouverts et, en redécoupant ces pavés, comme union dénombrable de pavés (non plus nécessairement ouverts cette fois) disjoints P pour lesquels $\mu^*(P) = \mu_*(P)$ (pour un pavé, ouvert ou non, ceci est clair) ; pour un ouvert, il n'y a donc pas d'ambiguïté à appeler $\mu(U)$ le nombre (éventuellement infini) $\mu^*(U) = \mu_*(U) \in [0, \infty]$. Si K est un compact, la définition même de la mesure intérieure fait que l'on a aussi dans ce cas $\mu^*(K) = \mu_*(K)$ et que l'on peut dénommer cette quantité $\mu(K)$ (elle est d'ailleurs ici finie, ce qui signifie que tout compact est intégrable).

Armés de ce lemme et de la remarque qui précède, nous pouvons énoncer un critère bien commode d'intégrabilité :

Proposition 1.4 *Une partie A de $\mathbb{R}^n$ est intégrable si et seulement si, pour tout $\epsilon > 0$, il existe un compact K_ϵ inclus dans A et un ouvert U_ϵ contenant A tels que $\mu(U_\epsilon \setminus K_\epsilon) < \epsilon$.*

Preuve. On sait, si A est intégrable, qu'il existe, pour tout $\epsilon > 0$ un compact $K \subset A$ tel que $\mu(A) > \mu^*(K) - \epsilon/2 = \mu(K) - \epsilon/2$ et, d'après la proposition 1.2, un ouvert U_ϵ contenant A tel que $\mu^*(U_\epsilon) = \mu(U_\epsilon) < \mu(A) + \epsilon/2$. Mais $U_\epsilon \setminus K_\epsilon$ et K_ϵ sont deux

20. Le fait de mettre des parenthèses ici pour les (a_i, b_i), $i = 1, ..., n$, $a_i, b_i \in \mathbb{R}$, signifie juste que les bornes inférieure ou supérieure peuvent être incluses ou non.

ensembles disjoints dont les mesures intérieures et extérieures coïncident (le premier est ouvert, le second compact) et le lemme 1.2 permet d'affirmer que

$$\mu(U_\epsilon) = \mu(K_\epsilon) + \mu(U_\epsilon \setminus K_\epsilon)\,;$$

les inégalités qui précèdent nous assurent donc $\mu(U_\epsilon \setminus K_\epsilon) < \epsilon$, ce que l'on voulait. Si inversement A satisfait au critère et que l'on puisse, pour tout ϵ, trouver une telle paire (U_ϵ, K_ϵ) avec $\mu(U_\epsilon \setminus K_\epsilon) < \epsilon$, on a

$$\mu^*(A) \leq \mu(U_\epsilon) = \mu(K_\epsilon) + \mu(U_\epsilon \setminus K_\epsilon) \leq \mu_*(A) + \epsilon\,,$$

ce qui induit l'inégalité $\mu^*(A) < +\infty$ car $\mu(K_\epsilon) + \epsilon < +\infty$ et même $\mu^*(A) \leq \mu_*(A)$, donc en fait $\mu_*(A) = \mu^*(A)$ puisque l'autre inégalité est triviale ; la partie A est donc intégrable et le critère est démontré. $\diamondsuit$

On a un corollaire important de ce critère :

Corollaire 1.1 *L'intersection et la différence de deux sous-ensembles intégrables de $\mathbb{R}^n$ le sont encore.*

Preuve. Soit A_1 et A_2 deux sous-ensembles intégrables et $\epsilon > 0$; avec le critère donné par la proposition 1.4, on trouve les ouverts $U_{1,\epsilon}, U_{2,\epsilon}$ et les compacts $K_{1,\epsilon}, K_{2,\epsilon}$ tels que

$$K_{j,\epsilon} \subset A_j \subset U_{j,\epsilon}$$

et $\mu(U_{j,\epsilon} \setminus K_{j,\epsilon}) < \epsilon$ pour $j = 1, 2$. On a

$$K_{1,\epsilon} \cap K_{2,\epsilon} \subset A_1 \cap A_2 \subset U_{1,\epsilon} \cap U_{2,\epsilon}$$

et

$$(U_{1,\epsilon} \cap U_{2,\epsilon}) \setminus (K_{1,\epsilon} \cap K_{2,\epsilon}) \subset (U_{1,\epsilon} \setminus K_{1,\epsilon}) \cup (U_{2,\epsilon} \setminus K_{2,\epsilon})\,.$$

Par sous-additivité de la mesure extérieure, on a

$$\mu^*\Big((U_{1,\epsilon} \cap U_{2,\epsilon}) \setminus (K_{1,\epsilon} \cap K_{2,\epsilon})\Big) = \mu\Big((U_{1,\epsilon} \cap U_{2,\epsilon}) \setminus (K_{1,\epsilon} \cap K_{2,\epsilon})\Big) \leq 2\epsilon\,,$$

ce qui montre que $A_1 \cap A_2$ satisfait au critère et est donc intégrable. Pour la différence $A_2 \setminus A_1$, on peut supposer A_1 inclus dans A_2. On construit, pour $\epsilon > 0$, les couples $(U_{1,\epsilon}, K_{1,\epsilon})$ tels que

$$K_{1,\epsilon} \subset A_1 \subset U_{1,\epsilon}$$

avec $\mu(U_{1,\epsilon} \setminus K_{1,\epsilon}) < \epsilon$ et $(U_{2,\epsilon}, K_{2,\epsilon})$ tels que

$$K_{2,\epsilon} \subset A_2 \subset U_{2,\epsilon}$$

et $\mu(U_{2,\epsilon} \setminus K_{2,\epsilon}) < \epsilon$. On suppose $U_{1,\epsilon}$ inclus dans $U_{2,\epsilon}$ (on s'y ramène en prenant l'intersection des deux ouverts) et $K_{1,\epsilon} \subset K_{2,\epsilon}$ (on s'y ramène en prenant l'intersection des deux compacts). Le compact $K_{2,\epsilon} \setminus U_{1,\epsilon}$ est contenu dans $A_2 \setminus A_1$ (faire un dessin) tandis que l'ouvert $U_{2,\epsilon} \setminus K_{1,\epsilon}$ contient, lui, $A_2 \setminus A_1$ et l'on a

$$(U_{2,\epsilon} \setminus K_{1,\epsilon}) \setminus (K_{2,\epsilon} \setminus U_{1,\epsilon}) = (U_{2,\epsilon} \setminus K_{2,\epsilon}) \cup (U_{1,\epsilon} \setminus K_{1,\epsilon})\,,$$

donc, par conséquent

$$\mu\Big((U_{2,\epsilon} \setminus K_{1,\epsilon}) \setminus (K_{2,\epsilon} \setminus U_{1,\epsilon})\Big) < 2\epsilon\,,$$

ce qui prouve que $A_2 \setminus A_1$ se plie au critère d'intégrabilité et est donc intégrable dès que A_1 et A_2 le sont. Le corollaire est ainsi prouvé. $\diamondsuit$

Les définitions d'intégrabilité et de mesurabilité posées et le critère d'intégrabilité établi avec ses conséquences, voici pourquoi l'on peut équiper $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ d'une mesure positive, dite *mesure de Lebesgue.*

Proposition 1.5 *Pour tout sous-ensemble mesurable de $\mathbb{R}^n$, on a $\mu^*(A) = \mu_*(A)$. La famille $\mathcal{M}(\mathbb{R}^n)$ des sous-ensembles mesurables (au sens de Lebesgue) de $\mathbb{R}^n$ est une tribu contenant la tribu borélienne $\mathcal{B}(\mathbb{R}^n)$. L'application*

$$A \in \mathcal{M}(\mathbb{R}^n) \longmapsto \mu(A) := \mu^*(A) = \mu_*(A) \in [0, +\infty]$$

est une mesure positive sur $(\mathbb{R}^n, \mathcal{M}(\mathbb{R}^n))$, dite mesure de Lebesgue. La tribu $\mathcal{M}(\mathbb{R}^n)$ est dite tribu de Lebesgue.

Preuve. Pour prouver la première affirmation, il suffit juste d'introduire une partition de $\mathbb{R}^n$ avec des pavés bornés (bien sûr non ouverts) deux-à-deux disjoints P_k, $k \in \mathbb{N}$, et de remarquer que l'hypothèse sur A implique que les ensembles deux-à-deux disjoints

$$A_k := A \cap P_k \,, \ k \in \mathbb{N} \,,$$

sont tous intégrables. D'après le lemme 1.2, leur union A vérifie encore $\mu_*(A) = \mu^*(A)$.

On sait que $\emptyset$ est mesurable (car $\emptyset \cap P = \emptyset$ pour tout pavé borné et que $\emptyset$ est supposé intégrable puisque $\mu^*(\emptyset) = \mu_*(\emptyset) = 0$).

Vérifions que $\mathcal{M}(\mathbb{R}^n)$ est bien stable par prise de complémentaire : si A est mesurable et si P est un pavé borné,

$$(\mathbb{R}^n \setminus A) \cap P = P \setminus A = P \setminus (P \cap A)$$

est intégrable comme différence de deux ensembles intégrables (c'est le second volet du corollaire 1.1). Le complémentaire d'un ensemble mesurable est mesurable.

Il reste à vérifier la stabilité par union dénombrable. Si les $(A_k)_k$ sont des ensembles mesurables deux-à-deux disjoints, les $P \cap A_k$ sont des ensembles intégrables deux-à-deux disjoints pour tout pavé borné P de $\mathbb{R}^n$ et, du fait du lemme 1.2, l'ensemble

$$\bigcup_k (A_k \cap P) = P \cap \bigcup_k A_k$$

est intégrable ; l'union des A_k est donc mesurable.

L'intersection et la différence de deux ensembles mesurables étant mesurable (cela résulte du corollaire 1.1), on montre en la découpant comme une union disjointe d'ensembles mesurables que toute réunion finie d'ensembles mesurables est mesurable. Si $(A_k)_k$ est une collection, $k = 0, 1, \ldots$, dénombrable infinie d'ensembles mesurables, les ensembles B_k définis par $B_0 := A_0$ et

$$B_k := A_k \setminus \bigcup_{l=0}^{k-1} A_l \,, \ k \geq 1 \,,$$

sont mesurables (différences de mesurables) et leur union (qui est aussi l'union des A_k) est mesurable. Toute union finie dénombrable d'ensembles mesurables est mesurable et cela conclut au fait que $\mathcal{M}(\mathbb{R}^n)$ est bien une tribu.

Comme les ouverts sont, on l'a vu, mesurables, la tribu $\mathcal{M}(\mathbb{R}^n)$ contient la tribu borélienne [21]. Si les $(A_k)_k$ sont des ensembles mesurables deux à deux disjoints non tous intégrables, l'union des A_k est mesurable non intégrable et l'on a immédiatement

$$\mu\Big(\bigcup_k A_k\Big) = +\infty = \sum_k \mu(A_k)\,.$$

Si tous les A_k sont intégrables, on sait d'après le lemme 1.2 que $\bigcup_k A_k$ (qui est encore intégrable si la série $[\mu(A_k)]_k$ est convergente) est tel que

$$\mu\Big(\bigcup_k A_k\Big) = \sum_k \mu(A_k)\,;$$

Tout ceci montre que

$$\mu\ : A \in \mathcal{M}(\mathbb{R}^n) \longmapsto \mu(A) \in [0,+\infty]$$

est bien une mesure positive et conclut donc la preuve de la proposition. $\diamondsuit$

1.5.4 La preuve du théorème de Lebesgue (théorème 1.1)

Si P est un pavé borné de $\mathbb{R}^n$ (non nécessairement ouvert), P est bien sûr intégrable et l'on a $\mu(x+P) = \mu(P)$ puisque la mesure de $P = (a_1,b_1)\times\ldots\times(a_n,b_n)$ est le produit des $b_i - a_i$, $i = 1,...,n$, donc aussi le produit des $(b_i + x_i) - (a_i + x_i)$, $i = 1,...,n$, si $x = (x_1,...,x_n)$. Si U est un ouvert (s'écrivant donc comme union finie ou dénombrable de pavés disjoints P_k), U et $U+x$ ont même mesure (éventuellement $+\infty$). Si A est mesurable, $A = A+x$ ont donc même mesure extérieure, donc même mesure, puisque

$$\mu^*(A) = \inf_{A\subset U} \mu(U)$$

et que

$$\mu^*(A+x) = \inf_{A+x\subset U} \mu(U) = \inf_{A\subset U-x} \mu(U-x) = \mu^*(A)$$

d'après la proposition 1.2. La mesure de Lebesgue est donc bien invariante par translation. De plus, on a $\mu([0,1]^n) = 1$ par construction même.

Pour achever la preuve du théorème de Lebesgue (théorème 1.1), il reste à vérifier que la mesure de Lebesgue construite dans la proposition 1.5 et restreinte à la tribu borélienne $\mathcal{B}(\mathbb{R}^n)$ est l'unique mesure positive sur la tribu borélienne $\mathcal{B}(\mathbb{R}^n)$ possédant les deux propriétés exigées, à savoir l'invariance par translation et la « normalisation » par la contrainte $\mu([0,1]^n) = 1$.

Si ν est une mesure positive sur la tribu borélienne et invariante par translation, on voit aisément que tout pavé $P = (0,1)\times\ldots\times(0,1)$ est tel que $\nu(P) = 1$. En effet, la ν-mesure de toute face $(n-1)$-dimensionnelle de $[0,1]^n$ est nulle : si elle ne l'était pas, le fait que ν soit invariante par translation contredirait le fait que $\nu([0,1]^n) = 1 < +\infty$ car il suffit de constater que $[0,1]^n$ s'écrit comme union d'une infinité de translatés de cette face.

21. Nous verrons en fait plus loin (section 1.6) que la tribu de Lebesgue $\mathcal{M}(\mathbb{R}^n)$ est la plus petite tribu contenant tous les ensembles du type $B \cup N$, où B est un borélien de $\mathbb{R}^n$ et N un sous-ensemble quelconque d'un borélien de mesure de Lebesgue nulle.

Si P est un pavé fermé borné $(a_1, b_1) \times \cdots \times (a_n, b_n)$, où les nombres a_i, b_i, $i = 1, ..., n$, sont tous rationnels, alors $\nu(P) = \mu(P)$: pour voir cela, on prend un dénominateur commun D pour tous les nombres a_i, b_i et l'on vérifie que

$$\nu([0, 1/D]^n) = \mu([0, 1/D]^n) = 1/D^n$$

du fait de l'invariance par translation de ν et μ et des conditions de normalisation. Une fois ceci acquis, on écrit P comme union disjointe de translatés de pavés du type $(0, 1/D) \times \ldots \times (0, 1/D)$ (de tous types) pour conclure à $\mu(P) = \nu(P)$. Du fait de la densité de $\mathbb{Q}^{2n}$ dans $\mathbb{R}^{2n}$ et des propriétés partagées par les mesures positives μ et ν et listées dans la proposition 1.1 (en fait ici (1.6) ou (1.7)), on voit que $\mu(P) = \nu(P)$ reste vraie pour tout pavé borné.

Les deux mesures μ et ν coïncident donc pour tout ouvert U (union dénombrable de pavés disjoints) et (en utilisant la propriété (1.7) et le fait que tout compact K de $\mathbb{R}^n$ s'écrive comme intersection d'une suite décroissante d'ouverts bornés) pour tout compact K de $\mathbb{R}^n$.

Mais on a, si A est un borélien, et puisque μ et ν obéissent à la première règle de la proposition 1.1,

$$\begin{aligned} \mu(A) = \mu_*(A) &= \sup_{K \subset A} \mu(K) = \sup_{K \subset A} \nu(K) \leq \nu(A) \\ &\leq \inf_{A \subset U} \nu(U) = \inf_{A \subset U} \mu(U) = \inf_{A \subset U} \mu^*(U) = \mu^*(A) = \mu(A). \end{aligned} \tag{1.13}$$

si l'on invoque l'égalité de μ et ν à la fois sur la classe des ouverts et sur celle des compacts et la proposition 1.2 (pour la dernière égalité). Les mesures μ et ν coïncident donc sur la tribu borélienne et le théorème 1.1 est complètement démontré. $\diamondsuit$

Avant de clore cette longue section dévolue à la construction de Lebesgue, signalons que nous avons en fait prouvé le résultat suivant, traduisant ce que l'on appelle la *régularité* de la mesure de Lebesgue (définie sur la tribu de Lebesgue) :

Proposition 1.6 *Si A est un élément de $\mathcal{M}(\mathbb{R}^n)$ et si μ désigne la mesure de Lebesgue, le nombre $\mu(A)$ peut se calculer*

- *soit comme la borne inférieure des nombres $\mu(U)$, ou U est un ouvert contenant l'ensemble mesurable A ;*
- *soit comme la borne supérieure des nombres $\mu(K)$, où K est un compact inclus dans l'ensemble mesurable A.*

Preuve. Cela résulte de la chaîne d'égalités-inégalités (1.13) concluant la preuve du théorème de Lebesgue. Le critère d'intégrabilité de la proposition 1.4 éclaire d'ailleurs ce résultat. $\diamondsuit$

Il faut prendre garde au fait que l'égalité $\mu_*(A) = \mu^*(A) = +\infty$ n'implique pas nécessairement la mesurabilité de A. Cependant, on peut proposer un critère de mesurabilité dans la même veine que le critère d'intégrabilité de la proposition 1.4 (voir la remarque 1.4 en fin de la section suivante).

1.6 Les notions de « μ-presque partout » et de tribu complétée

Définition 1.10 *Soit Ω un ensemble, $\mathcal{T}$ une tribu sur Ω, et μ une mesure positive $\mu : \mathcal{T} \longmapsto [0, +\infty]$. On dit qu'un sous-ensemble N de Ω est μ-négligeable si et* *seulement si N est inclus dans un sous-ensemble B appartenant à $\mathcal{T}$ et tel que B vérifie $\mu(B) = 0$.*

Tout sous-ensemble d'un sous-ensemble μ-négligeable est encore μ-négligeable et l'on dira d'une propriété $P(x)$ satisfaite pour tout x dans Ω, excepté lorsque x appartient à un sous-ensemble négligeable, qu'elle est satisfaite *μ-presque partout.*

Proposition 1.7 *Soit Ω un ensemble, $\mathcal{T}$ une tribu sur Ω et $\widetilde{\mathcal{T}}$ la collection de toutes les parties de Ω de la forme $A \cup N$, où $A \in \mathcal{T}$ et N est un sous-ensemble μ-négligeable. Alors $\widetilde{\mathcal{T}}$ est encore une tribu sur Ω, dite tribu complétée de $\mathcal{T}$ relativement à la mesure μ.*

Preuve. La partie vide étant dans $\mathcal{T}$, elle est dans $\widetilde{\mathcal{T}}$.
Si $(N_k)_k$ est une collection finie ou dénombrable de parties μ-négligeables ($N_k \subset B_k$ avec $B_k \in \mathcal{T}$ et $\mu(B_k) = 0$), l'union des N_k l'est aussi car l'union des B_k est un élément B de $\mathcal{T}$ tel que $\mu(B) = 0$ par σ-sous-additivité de la mesure μ.
Montrons enfin que $\widetilde{\mathcal{T}}$ est stable par prise de complémentaire : si $A \in \mathcal{T}$ et N est μ-négligeable, on peut écrire $A \cup N$ sous la forme $A \cup (N \setminus A)$, donc sous la forme $A \cup \widetilde{N}$ avec $\widetilde{N}$ négligeable et disjointe de A ; N est inclus dans une partie $B \in \mathcal{T}$ telle que $\mu(B) = 0$ que l'on peut (quitte à remplacer B par $B \setminus A$) supposer disjointe de A ; comme $A \cup N = (A \cup B) \setminus (B \setminus N)$, le complémentaire de $A \cup N$ s'écrit est l'union du complémentaire de $A \cup B$ et de l'ensemble négligeable $B \setminus N$; comme $A \cup B$ est dans $\mathcal{T}$, $\Omega \setminus (A \cup B)$ l'est aussi et le complémentaire de $A \cup N$ s'écrit bien comme l'union d'une partie élément de $\mathcal{T}$ et d'un sous-ensemble μ-négligeable. Ceci montre bien que $\widetilde{\mathcal{T}}$ est une tribu sur Ω. $\diamondsuit$

Si $\mathcal{T}$ est une tribu sur Ω, μ une mesure positive sur $\mathcal{T}$, et $\widetilde{\mathcal{T}}$ la tribu μ-complétée, on peut définir une mesure positive $\tilde{\mu}$ sur $\widetilde{\mathcal{T}}$ en posant

$$\tilde{\mu}(A \cup N) := \mu(A) ;$$

cette définition n'est en effet pas ambiguë car, si $A_1 \cup N_1 = A_2 \cup N_2$, l'ensemble $A_1 \setminus A_2$ (*resp.* $A_2 \setminus A_1$) est de μ-mesure nulle car inclus dans N_2 (*resp.* dans N_1) et l'on a donc $\mu(A_1) = \mu(A_2) = \mu(A_1 \cap A_2)$. Il est immédiat de vérifier la σ-additivité de $\tilde{\mu}$.
De plus, la mesure $\tilde{\mu}$ sur $\widetilde{\mathcal{T}}$ définie ainsi est la seule mesure positive sur $\widetilde{\mathcal{T}}$ qui coïncide avec μ sur $\mathcal{T}$. On appelle l'unique mesure sur $\widetilde{\mathcal{T}}$ « prolongeant » ainsi la mesure μ la *mesure complétée* de μ.

Concernant les tribus de Borel et de Lebesgue sur $\mathbb{R}^n$, on a l'important résultat suivant avec lequel nous conclurons ce chapitre :

Proposition 1.8 *La tribu de Lebesgue sur $\mathbb{R}^n$ est la complétée de la tribu borélienne relativement à la restriction à $\mathcal{B}(\mathbb{R}^n)$ de la mesure de Lebesgue ; la mesure de Lebesgue sur $\mathcal{M}(\mathbb{R}^n)$ est donc la complétée de sa restriction à la tribu borélienne.*

Preuve. Soit $A \in \mathcal{M}(\mathbb{R}^n)$ un sous-ensemble intégrable de $\mathbb{R}^n$; pour tout $k \in \mathbb{N}^*$, il existe, d'après le critère d'intégrabilité (proposition 1.4), pour tout entier $k \in \mathbb{N}^*$, un compact K_k inclus dans A, un ouvert U_k contenant A, tels que $\mu(U_k \setminus K_k)$ soit inférieur ou égal à $1/k$. On peut prendre la suite de compacts $(K_k)_k$ croissante et la suite d'ouverts $(U_k)_k$ décroissante (au sens de la relation d'ordre correspondant à l'inclusion entre ensembles). Soit K l'union des K_k pour $k \in \mathbb{N}^*$ et U l'intersection des U_k pour $k \in \mathbb{N}^*$. Or

$$U \setminus K \subset \bigcap_k (U_k \setminus K_k)\,;$$

comme la suite $(U_k \setminus K_k)_k$ est une suite décroissante et que $\mu(U_1) < +\infty$, la quatrième propriété (1.8) de la proposition 1.1 vérifiée par toute mesure positive sur une tribu (ici la mesure de Lebesgue sur la tribu des boréliens) assure que

$$\mu(U \setminus K) = \lim_{k \to +\infty} \mu(U_k \setminus K_k) = 0\,. \tag{1.14}$$

Comme $K \subset A \subset U$, (1.14) implique $\mu(A \setminus K) \leq \mu(U \setminus K) = 0$, donc que $A \setminus K$ est μ-négligeable. L'ensemble $A = K \cup (A \setminus K)$ est donc dans la tribu complétée $\widetilde{\mathcal{B}}(\mathbb{R}^n)$ relativement à la mesure de Lebesgue restreinte à $\mathcal{B}(\mathbb{R}^n)$ (puisque K est un borélien). Comme tout ensemble mesurable est union dénombrable d'ensembles intégrables, donc dans la tribu $\widetilde{\mathcal{B}}(\mathbb{R}^n)$, la tribu $\mathcal{M}(\mathbb{R}^n)$ est bien incluse dans cette tribu $\widetilde{\mathcal{B}}(\mathbb{R}^n)$.

Si N est μ-négligeable, N est inclus dans un borélien B de mesure nulle et l'on peut donc, pour tout $\epsilon > 0$, trouver un compact K_ϵ inclus dans N (on prend juste $K_\epsilon = \emptyset$) et un ouvert U_ϵ contenant B, donc N, tels que $\mu(U_\epsilon \setminus K_\epsilon) = \mu(U_\epsilon) \leq \epsilon$ (d'après la proposition 1.6 suivant laquelle la mesure de Lebesgue est régulière). Le critère d'intégrabilité (proposition 1.4) nous assure alors que N est intégrable, donc mesurable (et bien sûr de mesure nulle). Comme les boréliens sont aussi mesurables, la tribu $\mathcal{M}(\mathbb{R}^n)$ contient la tribu $\widetilde{\mathcal{B}}(\mathbb{R}^n)$ et, par conséquent, compte tenu du résultat obtenu au paragraphe précédent, lui est égale ; la proposition est établie. $\diamondsuit$

Remarque 1.4. On peut montrer (l'exercice est détaillé dans une des séquences du guide d'activités 2 proposé sous le serveur Ulysse) qu'on a le critère de mesurabilité suivant : une partie A de $\mathbb{R}^n$ est mesurable si et seulement si, pour tout $\epsilon > 0$, on peut trouver un fermé $F_\epsilon \subset A$, un ouvert U_ϵ tel que $A \subset U_\epsilon$, avec $\mu(U_\epsilon \setminus F_\epsilon) \leq \epsilon$. Il en résulte (si l'on travaille avec la suite des fermés $F_{1/k}$, $k \geq 1$ et la suite des ouverts $U_{1/k}$, $k \geq 1$, la première supposée croissante, la seconde décroissante, ce qui est toujours réalisable) que toute partie mesurable de $\mathbb{R}^n$ est telle que l'on puisse l'encadrer entre un borélien F (union dénombrable de fermés) et un autre borélien U (intersection dénombrable d'ouverts) tels que le borélien $U \setminus F$ soit de mesure nulle. À une partie μ-négligeable près, une partie mesurable est donc une union dénombrable de fermés ou aussi (toujours à ensemble μ-négligeable près) une intersection dénombrable d'ouverts. Ceci précise la proposition 1.8.

Chapitre 2

Mesurabilité et intégrabilité des fonctions

2.1 Introduction

Comme on l'a déjà fait remarquer dans l'introduction du chapitre 1, savoir « mesurer » un sous-ensemble A de $\mathbb{R}^n$ et être capable de lui assigner une mesure de Lebesgue $\mu(A)$ finie revient à savoir « intégrer » sur $\mathbb{R}^n$ la fonction caractéristique χ_A de A valant 1 sur A et 0 sur $\mathbb{R}^n \setminus A$. On retrouve ici la raison pour laquelle on a parlé, lors de la définition 1.9, de sous-ensemble « intégrable » de $\mathbb{R}^n$ pour qualifier une partie A élément de la tribu $\mathcal{M}(\mathbb{R}^n)$ des sous-ensembles mesurables (au sens de Lebesgue) de $\mathbb{R}^n$.

La seconde distinction majeure entre les points de vue de Riemann et de Lebesgue (la première étant, on la vu au chapitre 1, le fait que l'on utilise le concept « fini ou dénombrable » de manière à élargir le concept « fini » lorsqu'il s'agit de mesurer « de l'extérieur » un ensemble (section 1.5.1), est que le calcul de l'intégrale d'une fonction positive $f\ : \mathbb{R}^n \longrightarrow [0, +\infty]$ s'effectue, lorsque bien sûr il a un sens (et l'on verra sous quelles conditions dans ce chapitre), par « tranches horizontales »(avec la mesure préalable des strates de niveaux successives) et non plus, comme c'était le cas chez Riemann (voir l'introduction du chapitre 1) par découpage « vertical » (au prix d'une subdivision préalable du <u>domaine de départ</u>). Chez Lebesgue, c'est au contraire le <u>domaine d'arrivée</u> (en l'occurrence ici $[0, +\infty]$ pour une fonction positive) qui est subdivisé. Voici l'idée heuristique (elle sera précisée tout au long de ce chapitre) sous-tendant le calcul de l'intégrale d'une fonction positive (définie ici sur un sous-ensemble intégrable A de $\mathbb{R}^n$) à valeurs dans $[m, M[$: On commence par examiner le graphe de f au dessus de A et l'on découpe l'intervalle $[m, M[$ en utilisant une subdivision

$$y_0 = m < y_1 < \cdots < y_{M-1} < y_M = M\ ;$$

les sous-ensembles

$$A_j := \{(x_1, ..., x_n) \in A\,;\, y_j \leq f(x_1, ..., x_n) < y_{j+1}\}\,,\ j = 0, ..., M-1\,,$$

seront toujours ici des sous-ensembles intégrables de A partitionnant A (ceci nécessitera une hypothèse cruciale sur f, dite de *mesurabilité*) ; chacun d'eux aura donc un volume n-dimensionnel fini $\mathrm{vol}_n\,(A_j)$ (ici, sa mesure de Lebesgue dans $\mathbb{R}^n$) et l'on pourra associer à la subdivision

$$\mathcal{P} := \{y_0, y_1, ..., y_M\}$$

la quantité

$$I_*(f\,;\mathcal{P}) := \sum_{j=0}^{M-1} \mathrm{vol}_n\,(A_j) \times y_j\ ;$$

ensuite, on définira naturellement *l'intégrale de f sur A* comme la borne supérieure (finie ou infinie) de l'ensemble des nombres $I_*(f\,;\mathcal{P})$ pour toutes les subdivisions de $[m, M[$ possibles (contentons nous ici de subdivisions finies). Sur la figure 2.1, on a schématisé cette démarche en représentant le découpage en « tranches » du graphe de f, le partitionnement de A qui lui correspond, et en indiquant les correspondances.

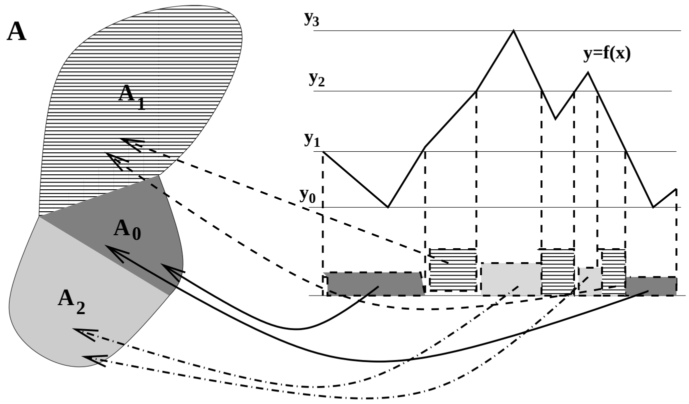

FIGURE 2.1 – Intégration « par tranches horizontales »

L'un des avantages majeurs de ce point de vue est que l'intégration des fonctions positives s'étend au cadre des fonctions $f\ : \Omega \longrightarrow [0, +\infty]$, où Ω est un ensemble abstrait, équipé d'une tribu $\mathcal{T}$ à laquelle se trouve attachée une mesure positive $\mu\ : \mathcal{T} \longrightarrow [0, +\infty]$. La théorie de l'intégration se trouve ainsi considérablement élargie et englobe, lorsque μ est une mesure de masse totale 1 ($\mu(\Omega) = 1$), la notion d'*espérance* (ou *moyenne*) inhérente à la théorie des probabilités.

Terminons cette introduction par une dernière remarque capitale pour la réalisation d'une théorie « efficace » et surtout sans pièges de l'intégration. Le point de vue de Riemann induit le concept d'*intégrale semi-convergente*, certes très intéressant et très riche, mais délicat car de fait plus proche de la notion de *primitive* que de la notion de *calcul d'aire* : une fonction définie et continue sur $[a, b[$ induit une *intégrale semi-convergente*

$$\int_a^b f(t)\,dt$$

sur $[a, b[$ si et seulement si la primitive

$$x \in [a, b[\longmapsto \int_a^x f(t)\,dt \tag{2.1}$$

de f sur $[a, b[$ admet une limite finie L lorsque x tend vers b, la valeur de l'intégrale semi-convergente (2.1) étant alors définie comme cette limite L [1]. Par exemple

$$\int_0^\infty \frac{\sin t}{t}\,dt := \lim_{x\to+\infty} \int_0^x \frac{\sin t}{t}\,dt$$

1. Pour parler l'intégrale semi-convergente sur $]a, b[$, il faut à la fois avoir semi-convergence sur $]a, c]$ et $[c, b[$ lorsque c est un point arbitraire du domaine de définition $]a, b[$ de la fonction.

existe et vaut $\pi/2$ tandis que la divergence de la série harmonique $[1/k]_{k\geq 1}$ et la π-périodicité de la fonction $t \longmapsto |\sin t|$ impliquent

$$\lim_{x\to+\infty} \int_0^x \left|\frac{\sin t}{t}\right| dt = +\infty\,.$$

Dès que les fonctions ne sont plus positives, le concept de « *semi-convergence* » se marie mal avec celui (pré-découpage en strates de niveaux « horizontales » et non plus en tranches « verticales » au dessus du domaine de définition) que nous évoquions et où la mesure des ensembles précède le calcul de l'intégrale des fonctions : par exemple

$$\lim_{x\to+\infty} \int_{-x}^x \frac{2t}{1+t^2}\, dt = \lim_{x\to+\infty} 0 = 0$$

tandis que

$$\lim_{x\to+\infty} \int_{-x}^{x^2} \frac{2t}{1+t^2}\, dt = \lim_{x\to+\infty} \log \frac{1+x^4}{1+x^2} = +\infty\,;$$

il n'est donc absolument pas question de parler de l'intégrale sur $\mathbb{R}$ de la fonction $t \longmapsto 2t/(1+t^2)$! L'extension de la théorie de l'intégration au cadre des fonctions de signe quelconque (ou des fonctions à valeurs complexes) devra prendre en compte ce type de difficulté. La notion de *semi-convergence* inhérente au point de vue Riemann ou au théorème fondamental de l'analyse ne saurait être une bonne notion dans la théorie que nous envisageons de décrire. Il n'y aura, on le verra, de fonctions intégrables que les fonctions de module intégrable, ce qui en fait nous facilitera grandement la vie [2]. Mais, bien sûr, la théorie de l'intégrale que nous présenterons dans cette section n'ôtera rien à la richesse de concepts autrement plus subtils et délicats à manier, tel celui de *semi-convergence* d'une intégrale, notions qui restent plus en phase avec l'approche de Riemann.

2.2 Une notion de fonction « simple » : celle de fonction étagée réelle sur $(\Omega, \mathcal{T})$

Définition 2.1 *Soit Ω un ensemble et $\mathcal{T}$ une tribu sur Ω. On appelle fonction étagée réelle sur $(\Omega, \mathcal{T})$ toute fonction de Ω dans $\mathbb{R}$ s'écrivant comme une combinaison linéaire finie et à coefficients réels*

$$f = \sum_{j=1}^{M} \lambda_j\, \chi_{A_j} \tag{2.2}$$

de fonctions caractéristiques χ_A de sous-ensembles non vides $A \in \mathcal{T}$, en prenant soin, c'est important pour l'écriture ici, que les λ_j sont distincts et donc que $A_j = \{f = \lambda_j\}$ [3] . Une telle fonction est dite positive si de plus tous les coefficients λ_j sont positifs ou nuls.

2. Comme s'il n'y avait de séries convergentes que les séries absolument convergentes, ce qui malheureusement n'est pas le cas bien sûr, sinon la théorie des séries perdrait tout son sel !

3. Il est pour cela parfois plus commode de penser les fonctions étagées sur Ω comme les fonctions ne prenant qu'un nombre fini de valeurs distinctes $\lambda_1, ..., \lambda_N$ (parmi elles éventuellement la valeur 0) ; notons que l'écriture de f sous la forme 2.2 est unique si les valeurs λ_j sont supposés distinctes et si N désigne exactement le nombre de valeurs réelles prises par f.

Remarque 2.1. La somme, le produit, le sup, l'inf, de deux fonctions étagées positives sont encore des fonction étagées positives. C'est un exercice immédiat que l'on laisse au lecteur ; il faut juste penser, si

$$\begin{aligned} f &= \sum_{j=1}^{M} \lambda_j \, \chi_{A_j} \\ g &= \sum_{k=1}^{N} \mu_k \, \chi_{B_k} \,, \end{aligned}$$

(les λ_j étant distincts, les μ_k aussi, les A_j partitionnant Ω, tout comme les B_k) à écrire $f+g$, fg, $\sup(f,g)$, $\inf(f,g)$ sous la forme

$$h = \sum_{\{(j,k)\,;\, A_j\cap B_k\neq\emptyset\}} \nu_{j,k} \, \chi_{A_j\cap B_k}$$

avec des coefficients positifs $\nu_{j,k}$ convenables. On utilisera souvent cette idée plus loin.

Si l'on dispose d'une mesure positive sur la tribu $\mathcal{T}$ conditionnant la définition des fonctions étagées, l'intégrale d'une fonction étagée positive se définit de la manière la plus naturelle possible :

Définition 2.2 *Soit $f = \sum_{j=1}^{M} \lambda_j \, \chi_{A_j}$ une fonction étagée positive sur un ensemble Ω équipé d'une tribu $\mathcal{T}$ et $\mu\ : \mathcal{T} \longrightarrow [0,+\infty]$ une mesure positive. L'intégrale de la fonction f relativement à la mesure positive μ est par définition la quantité*

$$\int_\Omega f(\omega) d\mu(\omega) := \sum_{j=1}^{M} \lambda_j \, \mu(A_j) \in [0,+\infty] \,. \tag{2.3}$$

Remarque 2.2. Cette définition présuppose que l'on ait adopté la convention $0\times(+\infty)=0$, ce que l'on fera toujours dans ce contexte de théorie de l'intégration (sinon, la somme dans (2.3) doit être restreinte aux j tels que $\lambda_j>0$, ce qui alourdirait l'écriture). En revanche, il n'y a pas en général possibilité de définir l'intégrale d'une fonction étagée quelconque, du fait de l'indétermination $\infty-\infty$, sauf toutefois lorsque tous les sous-ensembles A_j, $j=1,...,M$ tels que $\lambda_j\neq 0$ sont tels que $\mu(A_j)<+\infty$, auquel cas on définit l'intégrale de f (relativement à la mesure μ) par

$$\int_\Omega f(\omega) d\mu(\omega) := \sum_{j=1}^{M} \lambda_j \, \mu(A_j) \in \mathbb{R} \tag{2.4}$$

et l'on dit alors (seulement dans ce cas) que la fonction étagée f est une fonction étagée *intégrable* sur $(\Omega,\mathcal{T})$ relativement à la mesure positive μ.

La définition de l'intégrale d'une fonction étagée positive ne dépend pas de l'écriture de cette fonction sous la forme

$$f = \sum_{j=1}^{M} \lambda_j \chi_{A_j} \,.$$

En effet, si l'on dispose de deux écritures

$$f = \sum_{j=1}^{M} \lambda_j \, \chi_{A_j} \equiv \sum_{k=1}^{N} \mu_k \, \chi_{B_k} \,,$$

on peut aussi écrire

$$f = \sum_{j=1}^{M} \sum_{k=1}^{N} \nu_{j,k} \, \chi_{A_j\cap B_k} \,,$$

où $\nu_{j,k}$ est par définition la valeur commune $\lambda_j = \mu_k$ (pourvu que l'on ne garde dans la somme ci-dessus que les couples (j,k) tels que $A_j \cap B_k \neq \emptyset$). On a alors immédiatement

$$\sum_{j=1}^{M} \lambda_j \, \mu(A_j) = \sum_{k=1}^{N} \mu_k \, \mu(B_k) = \sum_{\{(j,k)\,;\, A_j \cap B_k \neq \emptyset\}} \nu_{j,k} \, \mu(A_j \cap B_k)\,.$$

En utilisant cette même idée, on voit que l'intégrale de la somme des deux fonctions étagées positives $f = \sum_{j=1}^{M} \lambda_j \, \chi_{A_j}$ et $g = \sum_{k=1}^{N} \mu_k \, \chi_{B_k}$ vaut

$$\begin{aligned}\sum_{\{(j,k)\,;\, A_j \cap B_k \neq \emptyset\}} (\lambda_j + \mu_k) \, \mu(A_j \cap B_k) &= \sum_{j=1}^{M} \lambda_j \, \mu(A_j) + \sum_{k=1}^{N} \mu_k \, \mu(B_k) \\ &= \int_\Omega f(\omega)\, d\mu(\omega) + \int_\Omega g(\omega) d\mu(\omega)\end{aligned}$$

puisque

$$\Omega = \{g = 0\} \cup \bigcup_{\{k\,;\, \mu_k > 0\}} B_k = \{f = 0\} \cup \bigcup_{\{j\,;\, \lambda_j > 0\}} A_j \tag{2.5}$$

et que l'on peut supposer les λ_j, $j = 1, ..., M$, deux-à-deux distincts (*idem* pour les μ_k, $k = 1, ..., N$), ce qui implique que l'on réalise avec (2.5) deux partitions de Ω. On a aussi

$$\int_\Omega (\lambda f)(\omega)\, d\mu(\omega) = \lambda \int_\Omega f(\omega)\, d\mu(\omega)$$

pour tout $\lambda \geq 0$, pour toute fonction étagée positive (intégrable ou non, pourvu que l'on ait pris comme règle $0 \times (+\infty) = 0$).

Enfin, si f et g sont deux fonctions étagées sur Ω (équipé de la tribu $\mathcal{T}$) telles que $f \leq g$ partout, on a

$$\int_\Omega f(\omega)\, d\mu(\omega) \leq \int_\Omega g(\omega)\, d\mu(\omega) \tag{2.6}$$

(il faut encore utiliser le découpage des A_j suivant les $A_j \cap B_k$ lorsque f et g s'expriment sous la forme $f = \sum_{j=1}^{N} \lambda_j \, \chi_{A_j}$, $g = \sum_{k=1}^{N} \mu_k \, \chi_{B_k}$).

Si E est un élément de $\mathcal{T}$ et f une fonction étagée positive

$$f = \sum_{j=1}^{M} \lambda_j \, \chi_{A_j}\,,$$

la fonction $f \times \chi_E$ est encore une fonction étagée positive, d'intégrale (relativement à la mesure positive μ)

$$\int_\Omega (f\chi_E)(\omega)\, d\mu(\omega) = \sum_{j=1}^{M} \lambda_j \, \mu(A_j \cap E) := \int_E f(\omega)\, d\mu(\omega)\,.$$

Si $(E_k)_{k\geq 0}$ est donc une suite croissante d'éléments de $\mathcal{T}$ réalisant une exhaustion de Ω (c'est-à-dire telle que l'union des E_k soit l'ensemble Ω tout entier), la propriété (1.7) de la proposition 1.1 implique

$$\int_\Omega f(\omega)\, d\mu(\omega) = \lim_{k\to+\infty} \int_{E_k} f(\omega)\, d\mu(\omega)\,. \tag{2.7}$$

2.3 $(\Omega_1,\mathcal{T}_1)$-$(\Omega_2,\mathcal{T}_2)$- mesurabilité ; le cas $(\Omega_2,\mathcal{T}_2) = ([0,\infty],\mathcal{B})$

Soient Ω_1 et Ω_2 deux ensembles, respectivement équipés de tribus $\mathcal{T}_1$ et $\mathcal{T}_2$ et F une application de Ω_1 dans Ω_2. La collection

$$\{F^{-1}(B)\,;\, B\in\mathcal{T}_2\}$$

des images inverses par F des éléments de la tribu $\mathcal{T}_2$ est une nouvelle tribu sur Ω_1, dite *tribu image réciproque* (où encore *obtenue par transfert*) de la tribu $\mathcal{T}_2$. Ceci se vérifie immédiatement (car la prise d'image réciproque au niveau ensembliste commute avec à la fois l'union quelconque, l'intersection quelconque et la prise de complémentaire [4]) et l'on appelle cette tribu la *tribu image réciproque de $\mathcal{T}_2$ par F* (on la note $F^{-1}(\mathcal{T}_2)$).

Définition 2.3 *On dit que F est $(\Omega_1,\mathcal{T}_1)$-$(\Omega_2,\mathcal{T}_2)$ mesurable si et seulement si la tribu image réciproque $F^{-1}(\mathcal{T}_2)$ est incluse dans la tribu $\mathcal{T}_1$.*

Remarque 2.3. Pour qu'une application F soit $(\Omega_1,\mathcal{T}_1)$-$(\Omega_2,\mathcal{T}_2)$ mesurable, il faut et il suffit, lorsque la tribu $\mathcal{T}_2$ est supposée engendrée par une famille $\mathcal{E}_2$ de parties de Ω_2, que l'image réciproque par F de tout élément de la famille $\mathcal{E}_2$ soit un élément de $\mathcal{T}_1$: ainsi, dans le cas particulier où $\Omega_2=\mathbb{R}^m$ et $\mathcal{T}_2=\mathcal{B}(\mathbb{R}^m)$, il suffit, pour que F soit $(\Omega_1,\mathcal{T}_1)$-$(\mathbb{R}^m,\mathcal{B}(\mathbb{R}^m))$ mesurable, que l'image réciproque par F de tout ouvert V de $\mathbb{R}^m$ soit un élément de la tribu $\mathcal{T}_1$ sur Ω_1.

Remarque 2.4. Si Ω_1 est un ensemble équipé d'une topologie [5] et que $\mathcal{T}_1=\mathcal{B}(\Omega_1)$ est la plus petite tribu (dite *tribu borélienne* sur l'espace topologique Ω_1) contenant tous les ouverts de Ω_1, on dit qu'une application $F\ :\Omega_1\longrightarrow\mathbb{R}^m$ mesurable relativement aux tribus boréliennes à la source et au but est *borélienne*. Puisque l'image réciproque par une application continue d'un ouvert est encore un ouvert, toute application <u>continue</u> de Ω_1 dans $\mathbb{R}^m$ est nécessairement <u>borélienne</u>. Le cas où Ω_1 est un sous-ensemble A appartenant à la tribu de Lebesgue $\mathcal{M}(\mathbb{R}^n)$, équipé de la topologie restreinte de la topologie usuelle de $\mathbb{R}^n$, sera pour nous particulièrement important par la suite : dans ce cas en effet, si F est une application continue de A dans $\mathbb{R}^m$, l'image réciproque par F d'un ouvert V de $\mathbb{R}^m$ est un ouvert de A, c'est-à-dire l'intersection avec A d'un ouvert de $\mathbb{R}^n$, donc en particulier un sous-ensemble de $\mathbb{R}^m$ élément de la tribu de Lebesgue sur $\mathbb{R}^n$.

Les exemples $\Omega_2=\mathbb{R}$, $\Omega_2=[0,+\infty]$, $\Omega_2=\mathbb{C}\simeq\mathbb{R}^2$, $\Omega_2=\mathbb{R}^m$, équipés chaque fois de leur tribu borélienne (que pour simplifier on notera toujours $\mathcal{B}$ sans préciser l'ensemble) seront pour nous des exemples capitaux.

Pour vérifier par exemple qu'une application $f\ :\Omega\longrightarrow\mathbb{R}$ ou $f\ :\Omega\longrightarrow[0,+\infty]$ est $(\Omega,\mathcal{T})$-($\mathbb{R}$ ou $[0,\infty],\mathcal{B}$)-mesurable, il suffit de s'assurer que, pour tout $\beta\in\mathbb{R}$,

$$f^{-1}(]\beta,\infty]) = \{\omega\in\Omega\,;\, f(\omega)>\beta\}$$

4. Ce qui n'est, attention, en général pas le cas pour la prise d'image directe au niveau ensembliste !

5. On dit aussi un *espace topologique*, comme le sont naturellement $\mathbb{R}^n$ ou une partie de $\mathbb{R}^n$.

est un élément de la tribu $\mathcal{T}$. Une fonction $f\ : \Omega \longrightarrow \mathbb{R}^m$ est $(\Omega, \mathcal{T})$-$(\mathbb{R}^m, \mathcal{B})$ mesurable si et seulement si ses m applications coordonnées $f_1, ..., f_m$ sont $(\Omega, \mathcal{T}) - (\mathbb{R}, \mathcal{B})$ mesurables.

La $(\Omega, \mathcal{T})$-$(\Omega', \mathcal{T}')$-mesurabilité des fonctions se propage naturellement par composition car $(G \circ F)^{-1}(B) = F^{-1}[G^{-1}(B)]$ lorsque F et G sont deux applications qui s'enchaînent. Ainsi la somme $f + g$ et le produit $f \times g$ de deux applications de Ω dans $\mathbb{R}$ ou $\mathbb{C}$ est $(\Omega, \mathcal{T}) - (\mathbb{R}$ ou $\mathbb{C}, \mathcal{B})$ mesurable dès que f et g le sont ; *idem* pour $1/f$ si f est une fonction mesurable de Ω dans $\mathbb{R}$ ou $\mathbb{C}$ ne s'annulant pas. Si f est une fonction $(\Omega, \mathcal{T})$-$(\mathbb{R}, \mathcal{B})$ (*resp.* $(\Omega, \mathcal{T})$-$(\mathbb{C}, \mathcal{B})$) mesurable, la fonction $|f|$ l'est aussi. Si f et g sont à valeurs dans $[0, \infty]$ et toutes les deux $(\Omega, \mathcal{T})$-$([0, \infty], \mathcal{B})$ mesurables, leur produit l'est encore (toujours une fois admise la convention $0 \times (+\infty) = 0$).

On retiendra en particulier la propriété importante suivante :

Proposition 2.1 *Si $(f_k)_k$ est une suite de fonctions de Ω dans $[0, +\infty]$, toutes mesurables (si Ω est équipé de la tribu $\mathcal{T}$ et $[0, \infty]$ de sa tribu borélienne $\mathcal{B}$), alors les fonctions*

$$\begin{aligned}
\sup_k f_k \quad &, \quad \inf_k f_k \\
\limsup_k f_k &:= \lim_{k \to +\infty} (\sup_{l \geq k} f_l) \\
\liminf_k f_k &:= \lim_{k \to +\infty} (\inf_{l \geq k} f_l)
\end{aligned}$$

sont toutes les quatre $(\Omega, \mathcal{T})$-$([0, \infty], \mathcal{B})$ mesurables[6].

Preuve. On a, si $\beta \in [0, +\infty[$,

$$\{\sup_k f_k > \beta\} = \bigcup_k \{f_k > \beta\}\,,$$

d'où la mesurabilité de $\sup_k f_k$ puisque la mesurabilité de chaque f_k implique que tous les ensembles $\{f_k > \beta\}$, donc leur union, est dans la tribu $\mathcal{T}$ (les ensembles $]\beta, +\infty]$ engendrent la tribu borélienne sur $[0, \infty]$). De même

$$\{\inf_k f_k < \beta\} = \bigcup_k \{f_k < \beta\}\,,$$

d'où la mesurabilité de $\inf_k f_k$ puisque la mesurabilité de chaque f_k implique que tous les ensembles $\{f_k < \beta\}$, donc leur union, est dans la tribu $\mathcal{T}$ (les ensembles $[0, \beta]$ engendrent aussi la tribu borélienne sur $[0, \infty]$). La fonction $\limsup_k f_k$ est la limite (décroissante), donc aussi l'inf, de la suite de fonctions $(\sup_{l \geq k} f_l)_k$ qui, d'après ce qui précède, est une suite de fonctions mesurables (chacune d'elles est un sup d'une famille de fonctions mesurables). La fonction $\liminf_k f_k$ est la limite

6. On rappelle que si $(u_k)_k$ est une suite d'éléments de la droite numérique achevée $[-\infty, +\infty]$, $\limsup_k u_k$ est par définition <u>la plus grande valeur d'adhérence</u> de la suite $(u_k)_k$ tandis que $\liminf_k u_k$ est, elle, <u>la plus petite valeur d'adhérence</u> de cette suite. Les notations se justifient par le fait que $\limsup_k u_k$ est vraiment la limite de la suite (décroissante) $(\sup_{l \geq k} u_l)_k$ tandis que $\liminf_k u_k$ est la limite de la suite (croissante) $(\inf_{l \geq k} u_l)_k$. Vous avez par exemple rencontré ces notions lors de l'énoncé des critères de Cauchy ou d'Alembert de convergence des séries numériques à termes positifs (comparaison avec les séries géométriques) et, *a fortiori*, dans l'expression $R = 1/\limsup_k |a_k|^{1/k}$ du rayon de convergence de la série entière $[a_k z^k]_{k \geq 1}$, voir le cours de MHT401 par exemple.

(croissante), donc aussi le sup, de la suite de fonctions $(\inf_{l\geq k} f_l)_k$ qui, d'après ce qui précède, est une suite de fonctions mesurables (chacune d'elles est un inf d'une famille de fonctions mesurables). La proposition est ainsi démontrée. $\diamondsuit$

En particulier, la limite supérieure d'une suite croissante de fonctions étagées positives sur Ω (équipé d'une tribu $\mathcal{T}$) est, d'après la proposition 2.1, une fonction $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ mesurable. Ce qui est intéressant et sera pour nous très important par la suite est que toute fonction de Ω dans $[0,\infty]$ qui est $(\Omega,\mathcal{T})$- $([0,\infty],\mathcal{B})$ mesurable s'exprime comme la limite d'une suite croissante de fonctions étagées positives ; on a en effet la proposition :

Proposition 2.2 *Soit Ω un ensemble, $\mathcal{T}$ une tribu sur Ω et $f\ :\Omega\longrightarrow[0,\infty]$ une fonction $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ mesurable. La fonction f est limite d'une suite croissante de fonctions étagées positives (relativement à la tribu $\mathcal{T}$).*

Preuve. Il suffit de démontrer le résultat pour une fonction mesurable bornée. En effet, toute fonction mesurable de Ω dans $[0,\infty]$ est limite croissante (ponctuellement partout) de la suite $(\inf(f,k))_{k\geq 1}$ (qui est une suite de fonctions mesurables bornées puisque l'inf de deux fonctions mesurables l'est aussi d'après la proposition 2.1). Supposons que l'on sache que chaque fonction $\inf(f,k)$ est limite partout d'une suite de fonctions étagées positives $(f_{k,l})_{l\geq 1}$. La suite $(f_{k,k})_k$ est alors une suite de fonctions étagées positives convergeant partout vers f sur Ω [7]. La suite $(h_k)_k$, où

$$h_k := \sup_{1\leq l\leq k} f_{l,l}$$

est alors une suite croissante de fonctions étagées positives convergeant vers f.

Si f est à valeurs dans $[0,M[$, on introduit, pour tout $k\in\mathbb{N}^*$, les 2^k éléments de $\mathcal{T}$

$$A_{k,l} := \Big\{\omega\,;\ \frac{lM}{2^k}\leq f(\omega)<\frac{(l+1)M}{2^k}\Big\}\,,\ \ l=0,...,2^k-1$$

et l'on pose, pour $k\in\mathbb{N}^*$,

$$h_k := \sum_{l=0}^{2^k-1}\frac{lM}{2^k}\,\chi_{A_{k,l}}\,.$$

La suite $(h_k)_k$ est une suite croissante de fonctions étagées positives convergeant (d'ailleurs en fait uniformément) vers f sur Ω. La proposition est démontrée. $\diamondsuit$

2.4 L'intégrale des fonctions $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ mesurables

Soit Ω un ensemble, $\mathcal{T}$ une tribu sur Ω, $\mu\ :\mathcal{T}\longrightarrow[0,\infty]$ une mesure positive.

Si f est une fonction $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ mesurable, on sait, d'après la proposition 2.2, que f est limite croissante d'une suite de fonctions étagées positives $(f_k)_k$. On sait aussi définir l'intégrale (relativement à la mesure positive μ) d'une fonction étagée

7. Le procédé utilisé ici est un procédé bien classique en analyse, dit *procédé diagonal*, les deux indices k,l sont pris égaux (ou encore (k,l) sur la « diagonale » du tableau d'indices).

positive (définition 2.2), opération qui a le mérite d'être « monotone » , puisque d'après (2.6),

$$\Big(f \leq g\Big) \Longrightarrow \Big(\int_\Omega f(\omega)\, d\mu(\omega) \leq \int_\Omega g(\omega)\, d\mu(\omega)\Big).$$

Ceci nous incite à la définition suivante :

Définition 2.4 *Soit Ω un ensemble, $\mathcal{T}$ une tribu sur Ω, $\mu \,: \mathcal{T} \longrightarrow [0,\infty]$ une mesure positive. L'intégrale (relativement à la mesure positive μ) d'une fonction $(\Omega, \mathcal{T})$- $([0,\infty], \mathcal{B})$ mesurable est par définition la borne supérieure (finie ou infinie) de toutes les intégrales*

$$\int_\Omega \varphi(\omega)\, d\mu(\omega)$$

lorsque φ décrit l'ensemble de toutes les fonctions étagées positives inférieures ou égales à f sur Ω. On note encore cette intégrale

$$\int_\Omega f(\omega)\, d\mu(\omega) \in [0,\infty]$$

ou encore, en abrégé,

$$\int_\Omega f\, d\mu \in [0,\infty]\,.$$

La fonction f est de plus dite intégrable (relativement à la mesure μ) si et seulement si

$$\int_\Omega f\, d\mu < +\infty\,.$$

Remarque 2.5. Si f est étagée positive, la définition proposée ici et la définition proposée dans la définition 2.2 coïncident ; notre nouvelle définition étend bien l'ancienne à une classe de fonctions plus large (celle des fonctions $(\Omega, \mathcal{T})$-$([0,\infty], \mathcal{B})$ mesurables) en restant cohérente avec la définition 2.2 lorsque l'on se limite à la classe des fonctions étagées positives.

Une inégalité (pourtant d'une extrême simplicité) est appelée à jouer un rôle très important en théorie de l'intégration : c'est *l'inégalité dite de Markov*[8] :

Proposition 2.3 *Soit Ω un ensemble équipé d'une tribu $\mathcal{T}$, $\mu \,: \mathcal{T} \longrightarrow [0,+\infty]$ une mesure positive. Si $f \,: \Omega \longrightarrow [0,\infty]$ est une fonction $(\Omega, \mathcal{T})$-$([0,\infty], \mathcal{B})$ mesurable qui de plus est intégrable relativement à la mesure μ, alors, pour tout nombre $\beta > 0$, l'ensemble $\{\omega\,;\, f(\omega) \geq \beta\}$ est un élément de $\mathcal{T}$ de mesure finie et l'on a*

$$\mu(\{f \geq \beta\}) \leq \frac{1}{\beta} \int_\Omega f\, d\mu\,. \tag{2.8}$$

En conséquence, l'ensemble $\{f = \infty\}$, intersection décroissante des sous-ensembles $\{f \geq k\}$, $k \in \mathbb{N}^*$, est un élément de $\mathcal{T}$ de μ-mesure nulle.

8. C'est sous l'impulsion de son maître Pafnuty Chebyshev que le mathématicien russe Andrei Markov (1856-1922) s'intéressera dès 1900 à la théorie des probabilités et en particulier aux théorèmes limites (la célèbre inégalité dite de Bienaymé-Chebyshev, que vous verrez dans l'UE MHT601, est une conséquence immédiate de l'inégalité dite de Markov) ; Markov développa aussi le concept probabiliste de *chaîne de Markov* appelé à jouer un rôle essentiel dans la modélisation stochastique.

Preuve. La fonction étagée positive $\beta \times \chi_{\{f \geq \beta\}}$ étant majorée par la fonction mesurable f, on a, par définition de l'intégrale de f comme borne supérieure des intégrales (relativement à μ) des fonctions étagées positives minorant f

$$\int_\Omega \beta\, \chi_{\{f\geq\beta\}}\, d\mu = \beta\, \mu(\{f \geq \beta\} \leq \int_\Omega f\, d\mu < +\infty\,,$$

d'où l'inégalité (2.8). L'assertion finale résulte de la propriété (1.8) dans la liste des propriétés partagées par toute mesure positive (ici μ) et énoncées dans la proposition 1.1. $\diamondsuit$

Voici une conséquence importante de l'inégalité de Markov :

Proposition 2.4 *Soit Ω un ensemble équipé d'une tribu $\mathcal{T}$, $\mu\ : \mathcal{T} \longrightarrow [0,+\infty]$ une mesure positive. Si $f\ : \Omega \longrightarrow [0,\infty]$ est une fonction $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ mesurable, la fonction f est d'intégrale nulle (relativement à la mesure μ) si et seulement si l'ensemble*

$$\{\omega\,;\, f(\omega) > 0\}$$

est de μ-mesure nulle (négligeable relativement à la mesure μ).

Preuve. Si f est nulle hors d'un sous-ensemble μ-négligeable de Ω, l'intégrale de toute fonction étagée positive minorant f est nulle, ce qui implique que l'intégrale de f relativement à la mesure μ est nulle. Réciproquement, si tel est le cas, l'inégalité (2.8) implique que chaque ensemble $\{f \geq 1/k\}$, $k \in \mathbb{N}^*$, est de μ-mesure nulle. L'union (dénombrable) de ces ensembles, qui est exactement $\{f > 0\}$, est donc de μ-mesure nulle. $\diamondsuit$

2.5 Le théorème de convergence monotone et ses conséquences

Soient toujours dans cette section Ω un ensemble (toujours abstrait), $\mathcal{T}$ une tribu sur Ω et $\mu\ : \mathcal{T} \longrightarrow [0,\infty]$ une mesure positive. Comme on l'a vu à la proposition 2.2, toute fonction $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ mesurable est ponctuellement limite croissante d'une suite $(\varphi_k)_k$ de fonctions étagées positives ; il serait judicieux que l'on puisse calculer l'intégrale de f sur Ω (relativement à la mesure μ), telle qu'elle a été définie à la définition 2.4, précisément en s'aidant de telles suites « approximantes » $(\varphi_k)_k$ pour la fonction f. Si tel est le cas, les propriétés auxquelles se plie l'intégrale des fonctions étagées positives (additivité, positive homogénéité, monotonie) pourront évidemment s'étendre à l'intégrale des fonctions mesurables positives. De fait, le résultat crucial ici est le *théorème de convergence croissante*, ou encore *théorème de Beppo Levi*[9] :

9. Le mathématicien italien Beppo Levi (1875-1961) prouva ce résultat, qui peut aujourd'hui sembler élémentaire, en 1906 ; mais l'on doit aussi à Beppo Levi, qui émigra à Rosario en Argentine en 1930, y créa un journal et y fonda une importante école, une quantité de travaux mathématiques (logique et axiome du choix, premières approches dès 1897 à la résolution des singularités des surfaces algébriques préludant aux travaux d'H. Hironaka 60 ans plus tard, courbes elliptiques,...) ainsi qu'en histoire des sciences. Le théorème de 1906 est donc loin d'être le point saillant du riche apport de Beppo Levi à tous les pans des mathématiques.

Théorème 2.1 (de Beppo Levi ou de convergence croissante) *Soit* $(f_k)_{k\geq 1}$ *une suite croissante*[10] *de fonctions* $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ *mesurables sur un ensemble* Ω *équipé d'une tribu* $\mathcal{T}$ *et* f *la limite de la suite* f_k *lorsque* k *tend vers* $+\infty$*. Alors* f *est aussi* $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ *mesurable et, si* $\mu\ :\mathcal{T}\longrightarrow[0,\infty]$ *est une mesure positive, on a*

$$\int_\Omega f\,d\mu=\lim_{k\to+\infty}\Big(\int_\Omega f_k\,d\mu\Big)\in[0,\infty]\,. \tag{2.9}$$

Preuve. Que f soit $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ mesurable résulte de la proposition 2.2. Par définition de l'intégrale comme borne supérieure, le fait que $f_k\leq f$ pour tout $k\in\mathbb{N}^*$ implique

$$\int_\Omega f_k\,d\mu\leq\int_\Omega f\,d\mu\qquad\forall k\in\mathbb{N}^*\,.$$

La suite (croissante)

$$\Big(\int_\Omega f_k\,d\mu\Big)_{k\geq 1}$$

converge donc dans $[0,\infty]$ vers une limite au plus égale à $\int_\Omega f\,d\mu$. Il faut montrer que cette limite est en fait minorée aussi par cette intégrale. Pour cela, on choisit un nombre $\epsilon\in]0,1[$. Si φ est une fonction étagée positive telle que $\varphi\leq f$, posons, pour $k\in\mathbb{N}^*$,

$$A_{\epsilon,\varphi,k}=\{\omega\,;\ f_k(\omega)\geq(1-\epsilon)\varphi\}\,.$$

On a

$$\int_\Omega f_k\,d\mu\geq\int_{A_{\epsilon,\varphi,k}}(1-\epsilon)\varphi\,d\mu=(1-\epsilon)\int_{A_{\epsilon,\varphi,k}}\varphi\,d\mu\,. \tag{2.10}$$

Mais la suite $(A_{\epsilon,\varphi,k})_k$ est une suite croissante d'ensembles qui exhauste Ω : en effet si $\varphi(\omega)>0$, alors, comme $f(\omega)\geq\varphi(\omega)>(1-\epsilon)\varphi(\omega)$ et que $f(\omega)=\lim_{k\to+\infty}f_k(\omega)$, on a

$$f_k(\omega)>(1-\epsilon)\varphi(\omega)$$

pour k assez grand, donc $\omega\in\bigcup_k A_{\epsilon,\varphi,k}$, tandis que, si $\varphi(\omega)=0$, ω est bien sûr dans tous les ensembles $A_{\epsilon,\varphi,k}$. En utilisant la propriété (2.7) établie à la fin de la section 2.2 (on prend $E_k=A_{\epsilon,\varphi,k}$), on a

$$\lim_{k\to+\infty}\int_{A_{\epsilon,\varphi,k}}\varphi\,d\mu=\int_\Omega\varphi\,d\mu\,.$$

En faisant tendre k vers l'infini dans (2.10), il vient donc

$$\lim_{k\to+\infty}\Big(\int_\Omega f_k\,d\mu\Big)\geq(1-\epsilon)\int_\Omega\varphi\,d\mu\,.$$

Comme $\epsilon\in]0,1[$ est arbitraire, on a, en le faisant tendre vers 0,

$$\lim_{k\to+\infty}\Big(\int_\Omega f_k\,d\mu\Big)\geq\int_\Omega\varphi\,d\mu\,;$$

10. Ceci signifie $f_k\leq f_{k+1}$ sur Ω pour tout $k\in\mathbb{N}^*$.

mais, ceci étant vrai pour toute fonction étagée positive φ minorant f, on en déduit

$$\lim_{k\to+\infty}\Big(\int_\Omega f_k\,d\mu\Big) \geq \int_\Omega f\,d\mu$$

par définition même de l'intégrale de f relativement à la mesure μ comme borne supérieure des intégrales (relativement à cette mesure) des fonctions étagées positives φ qui la minorent. Le théorème est donc bien démontré. $\diamondsuit$

Si f et g sont deux fonctions $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ mesurables, on voit, en approchant f et g respectivement par les suites croissantes de fonctions étagées $(\varphi_k)_k$ et $(\psi_k)_k$ (donc $f+g$ par la suite croissante $(\varphi_k+\psi_k)_k$), que, suivant le théorème de Beppo Levi et la propriété d'additivité de l'intégrale sur la classe des fonctions étagées positives (voir la section 2.2),

$$\begin{aligned}\int_\Omega (f+g)\,d\mu &= \lim_{k\to+\infty}\Big(\int_\Omega(\varphi_k+\psi_k)\,d\mu\Big)\\ &= \lim_{k\to+\infty}\Big(\int_\Omega \varphi_k\,d\mu + \int_\Omega \psi_k\,d\mu\Big)\\ &= \int_\Omega f\,d\mu + \int_\Omega g\,d\mu \qquad (2.11)\end{aligned}$$

et que, si $\lambda \geq 0$, on a (modulo toujours la convention $0\times(+\infty)=0$),

$$\begin{aligned}\int_\Omega (\lambda f)\,d\mu &= \lim_{k\to+\infty}\int_\Omega(\lambda\varphi_k)\,d\mu\\ &= \lambda\lim_{k\to+\infty}\int_\Omega \varphi_k\,d\mu = \lambda\int_\Omega f\,d\mu\,. \qquad (2.12)\end{aligned}$$

Ainsi donc, l'intégrale (relativement à la mesure positive μ) des fonctions $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ mesurables se plie aux règles d'additivité (2.11), de positive homogénéité (2.12), et bien sûr aussi de monotonie (2.6).

Autre application importante du théorème de Beppo Levi : si $(E_k)_k$ est une suite croissante d'éléments de $\mathcal{T}$ réalisant une exhaustion de Ω (l'union des E_k est égale à Ω tout entier) et si f est une fonction $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ mesurable, on a

$$\lim_{k\to+\infty}\int_\Omega f\chi_{E_k}\,d\mu = \int_\Omega f\,d\mu\,. \qquad (2.13)$$

Exemple 2.1. Si $(u_k)_k$ est une suite de fonctions $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ mesurables et $\mu\,:\mathcal{T}\longrightarrow[0,\infty]$ une mesure positive, le théorème de Beppo Levi implique en particulier

$$\int_\Omega\Big(\sum_k u_k\Big)\,d\mu = \sum_k\int_\Omega u_k\,d\mu \in [0,\infty]\,, \qquad (2.14)$$

résultat constituant une formulation du théorème d'intégration terme à terme des séries de fonctions mesurables positives. Il est important de souligner que ce résultat peut aussi se lire ainsi : si l'un des deux membres dans (2.14) vaut $+\infty$, alors l'autre aussi ; si l'un est fini, l'autre aussi et ils sont égaux ! Ceci vaut d'ailleurs tout autant pour la lecture de (2.9).

Un lemme complète de manière très intéressante (surtout lorsque nous en viendrons à passer à l'intégration des fonctions de signe quelconque) le théorème de convergence monotone. Il a aussi été proposé dès 1906 (dans sa thèse) par l'astronome et mathématicien français Pierre Fatou (1878-1929) et permet de nous affranchir de la contrainte « suite croissante » bien restrictive :

Théorème 2.2 (lemme de Fatou) *Soit Ω un ensemble, $\mathcal{T}$ une tribu sur Ω, et $\mu : \mathcal{T} \to [0,\infty]$ une mesure positive. Soit $(f_k)_{k\geq 1}$ une suite de fonctions $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ mesurables ; on a toujours*

$$\int_\Omega (\liminf_k f_k)\, d\mu \leq \liminf_k \Big(\int_\Omega f_k\, d\mu \Big) . \tag{2.15}$$

Preuve. Il suffit de considérer la suite croissante $(\tilde{f}_k)_k$, où

$$\tilde{f}_k := \inf_{l\geq k} f_l \,,$$

suite de fonctions $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ mesurables (d'après la proposition 2.2) convergeant en croissant vers $\liminf_k f_k$. Le théorème de convergence monotone nous donne donc

$$\int_\Omega (\liminf_k f_k)\, d\mu = \lim_{k\to+\infty} \int_\Omega \tilde{f}_k\, d\mu \,. \tag{2.16}$$

Comme

$$\int_\Omega \tilde{f}_k\, d\mu \leq \int_\Omega f_l\, d\mu$$

pour tout $l \geq k$ (par monotonie de l'intégrale et puisque $\tilde{f}_k \leq f_l$ pour tout $l \geq k$), on a

$$\int_\Omega \tilde{f}_k\, d\mu \leq \inf_{l\geq k} \Big(\int_\Omega f_l\, d\mu \Big) \,,\ \forall\, k \geq 1 \,,$$

ce qui donne l'inégalité (2.15) en prenant les limites lorsque k tend vers l'infini et en combinant avec (2.16). Le lemme est prouvé. $\diamondsuit$

C'est grâce au lemme de Fatou que nous serons aptes, une fois définie la notion d'intégrabilité pour les fonctions à valeurs dans $[-\infty,+\infty]$, $\mathbb{C}$ ou $\mathbb{R}^m$, d'énoncer le second grand théorème autorisant la permutation d'une prise de limite et d'une prise d'intégrale, le *théorème de convergence dominée* d'Henri Lebesgue.

2.6 Un lien avec l'intégrale de Riemann des fonctions continues positives (ou nulles) sur (a,b)

À ce niveau du cours, avant d'aller plus loin (et pour rester concrets), il temps de nous raccrocher au concept d'intégrale que nous connaissons bien, celui d'*intégrale au sens de Riemann* d'une fonction continue prenant des valeurs positives ou nulles sur un intervalle (a,b) (fini ou non, bornes ou non incluses) de $\mathbb{R}$. On dispose sur (a,b) d'une tribu, à savoir la tribu borélienne (trace sur (a,b) de la tribu borélienne sur $\mathbb{R}$) [11] et, sur cette tribu, d'une mesure positive particulière, la mesure de Lebesgue, que nous avons appris à construire au chapitre 1.

Soit (a,b) un tel intervalle de $\mathbb{R}$ que l'on écrira comme union croissante d'une suite de segments $[a_k,b_k]$, $k \geq 1$, tous fermés bornés. Soit f une fonction continue positive sur (a,b), donc évidemment mesurable relativement à la tribu borélienne (et de

11. On dispose même d'une tribu plus grosse, à savoir sa complétée, ici en fait la trace sur (a,b) de la tribu $\mathcal{M}(\mathbb{R})$ des sous-ensembles mesurables de $\mathbb{R}$ (au sens de Lebesgue).

restriction à chaque segment $[a_k, b_k]$ mesurable par rapport à la tribu borélienne sur $[a_k, b_k]$). Du fait de l'uniforme continuité de f sur chaque segment $[a_k, b_k]$, il est facile de construire, pour chaque $k \in \mathbb{N}^*$, une suite $(\varphi_{k,l})_{l\geq 1}$ de fonctions en escalier positives (voir la section 1.1) convergeant uniformément (et de manière croissante) vers f sur le segment $[a_k, b_k]$: ceci se fait par exemple en subdivisant $[a, b]$ suivant les points

$$a_k + j\frac{(b_k - a_k)}{2^l}\,,\ j = 0, ..., 2^l$$

et en prenant comme valeur de la fonction en escalier $\varphi_{k,l}$ sur l'intervalle semi-ouvert

$$I_j = \Big[a_k + j\frac{(b_k - a_k)}{2^l}, a_k + (j+1)\frac{(b_k - a_k)}{2^l}\Big[,\ j = 0, ..., 2^l - 1\,,$$

l'inf des valeurs de la fonction f sur le segment $\overline{I_j}$ (et $f_{k,l}(b_k) = f(b_k)$). Chacune de ces fonctions en escalier $\varphi_{k,l}$, $l \geq 1$, est évidemment aussi une fonction étagée positive (relativement à la tribu borélienne sur $[a_k, b_k]$) et il résulte donc du théorème de Beppo Levi que

$$\int_{[a_k,b_k]} f(t)\,dt = \lim_{l\to+\infty} \int_{[a_k,b_k]} \varphi_{k,l}(t)\,dt = \int_{a_k}^{b_k} f(t)\,dt\,,$$

l'intégrale $\int_{a_k}^{b_k} f(t)\,dt$ désignant ici l'intégrale de f calculée au sens de Riemann.

La propriété (2.13) nous permet donc alors d'énoncer la proposition suivante, qui nous confirme bien que la théorie de l'intégration que nous envisageons dans ce cours élargit celle de Riemann que vous connaissiez[12] :

Proposition 2.5 *Soit (a, b) un intervalle de $\mathbb{R}$ (de type quelconque), union croissante des segments $[a_k, b_k]$, et f une fonction continue prenant des valeurs positives ou nulles sur (a, b). La fonction f est alors mesurable relativement à la tribu borélienne sur (a, b) et l'intégrale de f relativement à la mesure de Lebesgue sur (a, b) est égale à*

$$\int_{(a,b)} f(t)\,dt = \lim_{k\to+\infty} \int_{a_k}^{b_k} f(t)\,dt\,, \tag{2.17}$$

où les intégrales ci-dessus sont calculées au sens de Riemann[13].

Preuve. Elle résulte de ce qui précède et de l'application de la règle (2.13). $\diamondsuit$

12. Nous n'avons nullement construit un procédé d'intégration donnant un résultat différent de celui que nous connaissions lorsqu'il s'agit de calculer des intégrales de fonctions aussi régulières que les fonctions continues, voire continues par morceaux, même, on le verra, *réglées* (*cf.* la note suivante) ! Ceci vaut d'ailleurs aussi en ce qui concerne les intégrales multiples de fonctions continues en plusieurs variables sur les domaines simples de $\mathbb{R}^n$.

13. On verra plus loin dans le cours que ceci est encore vrai si la fonction f est simplement une fonction positive *réglée* sur (a, b), c'est-à-dire ayant en tout point de (a, b) une limite à gauche et à droite (ou encore, ce qui est équivalent, limite uniforme sur tout segment $[a_k, b_k]$ d'une suite de fonctions en escalier) ; une telle fonction est bien intégrable au sens de Riemann sur tout segment $[a_k, b_k]$, est mesurable sur (a, b) (pour la tribu borélienne) et la formule (2.17) subsiste.

Exemple 2.2. Par exemple, on pourra vérifier à titre d' application directe du théorème de Beppo Levi les égalités

$$\begin{aligned}
\int_0^{+\infty} \frac{e^t - 1 - t}{t^2}\, e^{-t}\, dt &= 1 \\
\int_0^{\infty} \log\Big(\frac{1+e^{-t}}{1-e^{-t}}\Big)\, dt &= 2\sum_{k=0}^{\infty} \frac{1}{(2k+1)^2}
\end{aligned}$$

dans le contexte où $\Omega =]0,+\infty[$, équipé de la tribu borélienne, et où la mesure (ici notée dt) est la trace sur $]0,+\infty[$ de la mesure de Lebesgue sur $\mathbb{R}$ (les intégrales figurant au membres de gauche pouvant être interprétées d'après la proposition 2.5 indifféremment au sens de Lebesgue dans le contexte mentionné ou de Riemann.

2.7 Intégration des fonctions à valeurs dans $\overline{\mathbb{R}}$, $\mathbb{C}$, $\mathbb{R}^m$

Soit Ω un ensemble abstrait, $\mathcal{T}$ une tribu sur Ω, $\mu\ : \mathcal{T} \to [0,\infty]$ une mesure positive et f une fonction mesurable de $(\Omega,\mathcal{T})$ dans $(\Omega_2,\mathcal{B})$, où Ω_2 désigne l'un des espaces topologiques $\overline{\mathbb{R}} := [-\infty,\infty]$, $\mathbb{C}$, $\mathbb{R}^m$ équipé de sa tribu borélienne $\mathcal{B}$. Si $t \in [-\infty,\infty]$, on définit naturellement $|t| \in [0,\infty]$, si $x \in \mathbb{R}^m$, on note $|x|$ la norme euclidienne de x et si $z \in \mathbb{C}$, $|z|$ le module de z. On rappelle qu'une application $f\ : \Omega \longrightarrow \mathbb{R}^m$ est $(\Omega,\mathcal{T})$-$(\mathbb{R}^m,\mathcal{B})$ mesurable si toutes les applications coordonnées sont $(\Omega,\mathcal{T})$-$(\mathbb{R},\mathcal{B})$ mesurables.

Définition 2.5 *Une fonction $f\ : \Omega \longrightarrow [-\infty,\infty], \mathbb{C}, \mathbb{R}^m$ est dite intégrable relativement à la tribu $\mathcal{T}$ et à la mesure μ si :*
- *d'une part, f est mesurable de $(\Omega,\mathcal{T})$ dans l'un de ces trois espaces topologiques équipé de sa tribu borélienne $\mathcal{B}$;*
- *d'autre part*

$$\int_\Omega |f|\, d\mu < +\infty\,. \tag{2.18}$$

Pour définir dans ce cas ce qu'est l'intégrale d'une fonction intégrable relativement à la mesure positive μ, il convient de distinguer les trois cas.

- Si $f\ : \Omega \longrightarrow [-\infty,\infty]$, la condition d'intégrabilité (2.18) implique, du fait de la proposition 2.3, que $|f|$ est fini hors d'un ensemble μ-négligeable et que les fonctions mesurables $f^+ := \sup(f,0)$ et $f^- := \sup(-f,0)$ (à valeurs dans $[0,\infty]$) sont toutes les deux intégrables (comme fonctions $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ mesurables) par rapport à la mesure μ. On pose alors

$$\int_\Omega f\, d\mu := \int_\Omega f^+\, d\mu - \int_\Omega f^-\, d\mu\,. \tag{2.19}$$

- Si $f\ : \Omega \longrightarrow \mathbb{C}$, les quatre fonctions $(\mathrm{Re}\, f)^\pm, (\mathrm{Im}\, f)^\pm$ (à valeurs dans $[0,\infty[$) sont toutes les quatre $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ mesurables et la condition d'intégrabilité (2.18) implique qu'elles sont toutes les quatre intégrables. On pose alors

$$\int_\Omega f\, d\mu = \int_\Omega (\mathrm{Re}\, f)^+\, d\mu - \int_\Omega (\mathrm{Re}\, f)^-\, d\mu + i\Big(\int_\Omega (\mathrm{Im}\, f)^+\, d\mu - \int_\Omega (\mathrm{Im}\, f)^-\, d\mu\Big)\,.$$

– Si $f\ :\Omega \longrightarrow \mathbb{R}^m$, les $2m$ fonctions $f_j^{\pm}$, $j=1,...,m$ attachées aux m fonctions coordonnées $f_1,...,f_m$ (à valeurs dans $\mathbb{R}$) sont toutes les quatre $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ mesurables et la condition d'intégrabilité (2.18) implique qu'elles sont toutes intégrables. On pose alors

$$\int_\Omega f\,d\mu := \Big(\int_\Omega f_j^+\,d\mu - \int_\Omega f_j^-\,d\mu\Big)_{j=1,...,m}. \tag{2.20}$$

Toutes ces définitions d'intégrales ont chaque fois fait intervenir des combinaisons de nombre réels (éventuellement à coefficients complexes), mais, en aucun cas (c'est pour cela que la clause d'intégrabilité (2.18) est essentielle) la forme indéterminée $\infty-\infty$.

2.8 Les principales propriétés de l'intégrale

La première des propriétés de l'intégrale des fonctions mesurables à valeurs dans $\overline{\mathbb{R}}$, $\mathbb{C}$, $\mathbb{R}^m$, est la linéarité. Remarquons tout d'abord que si f est mesurable sur $(\Omega,\mathcal{T})$, à valeurs dans $\overline{\mathbb{R}}$, et de plus intégrable, les ensembles $\{f=\pm\infty\}$ sont négligeables (comme conséquence de l'inégalité de Markov, proposition 2.3) et l'on peut donc affirmer alors que f est égale (hors d'un ensemble μ-négligeable) à une fonction mesurable de $(\Omega,\mathcal{T})$ dans $\mathbb{R}$. Par la suite, le cas des fonctions intégrables à valeurs dans $[-\infty,\infty]=\overline{\mathbb{R}}$ se ramène, en ce qui concerne les calculs d'intégrales relativement à une mesure positive des fonctions intégrables, à celui des fonctions mesurables et intégrables de $(\Omega,\mathcal{T})$ dans $\mathbb{R}$.

Du fait de la définition de l'intégrale d'une fonction intégrable réelle comme la différence des deux intégrales de fonctions positives f^+ et f^- (comme dans (2.19)), de la définition de l'intégrale des fonctions mesurables intégrables à valeurs vectorielles (comme dans (2.20)), et de l'additivité de la prise d'intégrale pour les fonctions mesurables de $(\Omega,\mathcal{T})$ dans $[0,\infty]$, on a le résultat immédiat suivant :

Proposition 2.6 (linéarité de l'intégrale) *Si f et g sont deux fonctions $(\Omega,\mathcal{T})$-$(\mathbb{R}^m,\mathcal{B}(\mathbb{R}^m))$ mesurables, intégrables relativement à une mesure positive μ sur $\mathcal{T}$, et si λ,ν sont des nombres réels, la fonction $\lambda f+\mu g$ est encore $(\Omega,\mathcal{T})$-$(\mathbb{R}^m,\mathcal{B}(\mathbb{R}^m))$ mesurable, et intégrable relativement à la mesure μ, et l'on a, dans $\mathbb{R}^m$, l'égalité*

$$\int_\Omega (\lambda f+\nu g)\,d\mu = \lambda\int_\Omega f\,d\mu + \nu\int_\Omega g\,d\mu\,. \tag{2.21}$$

Si f et g sont $(\Omega,\mathcal{T})$-$(\mathbb{C},\mathcal{B}(\mathbb{C}))$ mesurables et intégrables par rapport à la mesure μ et si λ,ν sont deux nombres complexes, $\lambda f+\nu g$ est aussi $(\Omega,\mathcal{T})$-$(\mathbb{C},\mathcal{B}(\mathbb{C}))$ mesurable, intégrable par rapport à la mesure μ, et l'on a encore l'égalité (2.21).

Pour les fonctions définies sur $(\Omega,\mathcal{T})$ et à valeurs réelles ou complexes, une seconde inégalité jouera un rôle majeur ; elle ne fait en fait que répercuter dans le contexte de l'intégration l'inégalité triangulaire

$$|u_1+\cdots+u_N| \le |u_1|+\cdots+|u_N|$$

(pensez à $\Omega=\{1,...,N\}$ avec comme mesure la mesure de décompte) ; c'est la :

Proposition 2.7 *Soit f une fonction $(\Omega,\mathcal{T})$-$(\mathbb{R},\mathcal{B}(\mathbb{R}))$ ou $(\mathbb{C},\mathcal{B}(\mathbb{C}))$-mesurable, intégrable par rapport à une mesure positive μ sur $\mathcal{T}$; alors, on a*

$$\Big|\int_\Omega f\,d\mu\Big| \leq \int_\Omega |f|\,d\mu\,. \tag{2.22}$$

L'égalité dans (2.22) se produit si et seulement s'il existe un nombre complexe λ de module 1 tel que $f=\lambda|f|$ μ-presque partout ($\lambda=\pm 1$ si f est à valeurs réelles).

Remarque 2.6. En fait, l'inégalité (2.22) se généralise pour les applications à valeurs non plus dans $\mathbb{C}\simeq\mathbb{R}^2$, mais dans $\mathbb{R}^m$, $m\geq 1$. Si $\|\ \|$ désigne une norme abitraire sur $\mathbb{R}^m$ et f une fonction $(\Omega,\mathcal{T})$-$(\mathbb{R}^m,\mathcal{B}(\mathbb{R}^m))$-mesurable, intégrable par rapport à une mesure μ sur $\mathcal{T}$, alors

$$\Big\|\int_\Omega f d\mu\Big\| \leq \int_\Omega \|f\|\,d\mu. \tag{2.23}$$

Cette inégalité est aisée à établir pour les fonctions étagées (à valeurs dans $\mathbb{R}^m$) intégrables (elle résulte alors de l'inégalité triangulaire) et se déduira pour les fonctions intégrables du théorème 2.3 de convergence dominée de Lebesgue (qui sera énoncé et démontré un peu plus loin). Rien par contre ne peut être vraiment ajouté dans ce cas concernant le cas d'égalité dans (2.23) comme c'est le cas lorsque f est à valeurs complexes et la norme la norme euclidienne dans $\mathbb{R}^2$.

Preuve. Si f est à valeurs réelles, le fait que f soit $(\Omega,\mathcal{T})$-$(\mathbb{R},\mathcal{B}(\mathbb{R}))$ mesurable et de plus intégrable relativement à la mesure μ implique que les deux fonctions posititives $f^+:=\sup(f,0)$ et $f^-:=\sup(-f,0)$ héritent des mêmes propriétés. Comme $|f|=f^++f^-$, on bien l'inégalité

$$\Big|\int_\Omega f^+\,d\mu - \int_\Omega f^-\,d\mu\Big| \leq \int_\Omega f^+\,d\mu + \int_\Omega f^-\,d\mu = \int_\Omega |f|\,d\mu$$

d'après la clause d'additivité (2.11).

Si maintenant f est à valeurs complexes, l'inégalité (2.22) est bien sûr satisfaite si $\Big|\int_\Omega f\,d\mu\Big| = 0$. On peut donc supposer que ce nombre est strictement positif et appeler λ le nombre complexe (de module 1)

$$\lambda := \frac{\int_\Omega f\,d\mu}{\Big|\int_\Omega f\,d\mu\Big|}\,. \tag{2.24}$$

En prenant les parties réelles des deux membres dans l'identité

$$\Big|\int_\Omega f\,d\mu\Big| = \lambda^{-1}\int_\Omega f\,d\mu\,,$$

il vient

$$\Big|\int_\Omega f\,d\mu\Big| = \int_\Omega \mathrm{Re}\,(\lambda^{-1}f)\,d\mu \leq \int_\Omega \Big[\mathrm{Re}\,(\lambda^{-1}f)\Big]^+\,d\mu\,.$$

Mais la clause de monotonie (2.6) de la prise d'intégrale de fonctions mesurables positives implique

$$\int_\Omega \Big[\mathrm{Re}\,(\lambda^{-1}f)\Big]^+\,d\mu \leq \int_\Omega |f|\,d\mu$$

puisque $(\mathrm{Re}\,(\lambda^{-1}f))^+ \leq |f|$ sur Ω. L'inégalité (2.22) est donc prouvée aussi lorsque f est à valeurs complexes.

Supposons maintenant que, pour une fonction mesurable et intégrable à valeurs complexes f, on ait l'égalité

$$\Big|\int_\Omega f\,d\mu\Big| = \int_\Omega |f|\,d\mu\,.$$

Si ces deux nombres sont nuls, il résulte de la proposition 2.4 que f est nulle μ-presque partout, donc $f = |f|$ μ-presque partout. Sinon, on peut introduire le nombre complexe λ comme dans (2.24) et déduire de

$$\Big|\int_\Omega f\,d\mu\Big| = \int_\Omega \mathrm{Re}\,(\lambda^{-1}f)\,d\mu = \int_\Omega |f|\,d\mu = \Big|\int_\Omega f\,d\mu\Big|$$

l'égalité

$$\int_\Omega \Big(|f| - \mathrm{Re}\,(\lambda^{-1}f)\Big)\,d\mu = 0\,.$$

Comme il s'agit ici de l'intégrale d'une fonction mesurable positive, la proposition 2.4 implique que $|f| = \mathrm{Re}\,(\lambda^{-1}f)$ μ-presque partout, soit encore

$$\mathrm{Re}\,\Big(\lambda^{-1}\frac{f}{|f|}\Big) = 1$$

μ-presque partout, soit donc (puisque 1 est le seul nombre complexe de module 1 ayant 1 pour partie réelle), $f = \lambda|f|$ μ-presque partout. La proposition est donc bien démontrée ici. $\diamondsuit$

Lorsque f est une fonction $(\Omega,\mathcal{T})$-$(\mathbb{R},\mathcal{B})$ mesurable intégrable par rapport à une mesure μ de masse totale finie (normalisée de manière à ce que cette masse vaille 1), une autre inégalité est appelée un jouer un rôle important [14] ; c'est l'inégalité de Jensen [15] dont voici l'énoncé :

Proposition 2.8 (inégalité de Jensen) *Soit f une fonction $(\Omega,\mathcal{T})$-$(\mathbb{R},\mathcal{B})$ mesurable supposée intégrable relativement à une mesure positive μ sur $\mathcal{T}$ (telle que $\mu(\Omega)=1$) et Φ une fonction convexe $\Phi\ : I \longrightarrow \mathbb{R}$, I désignant un intervalle ouvert de $\mathbb{R}$ tel que $f(\Omega)\subset I$. L'intégrale $\int_\Omega f\,d\mu$ est alors un élément de I et on a*

$$\Phi\Big(\int_\Omega f\,d\mu\Big) \leq \int_\Omega (\Phi\circ f)^+\,d\mu - \int_\Omega (\Phi\circ f)^-\,d\mu := \int_\Omega (\Phi\circ f)\,d\mu \in\,]-\infty,\infty]\,, \tag{2.25}$$

le membre de droite étant toujours défini du fait que

$$\int_\Omega (\Phi\circ f)^-\,d\mu < \infty$$

puisque Φ est le sup sur I des fonctions affines qui la minorent et que, pour tout $\alpha,\beta\in\mathbb{R}$, la fonction $\alpha f+\beta$ est intégrable relativement à μ dès que f l'est.

14. En particulier, on la retrouvera pour l'exploiter lorsque nous aurons à définir plus tard dans ce cours les espaces $\mathcal{L}^p(\Omega,\mathcal{T},\mu)$ et $L^p(\Omega,\mathcal{T},\mu)$ pour $p\in[1,\infty]$.

15. Mathématicien danois (1859-1925), Johan Ludwig Jensen, qui resta un mathématicien « amateur » durant toute sa carrière d'ingénieur à la division de Copenhague de la Bell Telephone Company, prouva cette inégalité importante en 1906 ; on lui doit aussi des travaux en relation avec l'hypothèse de Riemann (qui lui firent énoncer une importante formule de représentation intégrale en analyse complexe, liant la croissance d'une fonction analytique dans un disque à la répartition de ses zéros).

Preuve. Dire que la fonction $\Phi\ : I \longrightarrow \mathbb{R}$ est convexe (donc nécessairement continue) sur l'intervalle ouvert I équivaut à dire que l'épigraphe

$$E_\Phi := \{(t,y) \in \mathbb{R}^2\,;\, t \in I\,,\, y \geq \Phi(t)\}$$

est un sous-ensemble convexe fermé de $I \times \mathbb{R}$, ce qui implique que, pour tout (t_0, y_0) avec $t_0 \in I$ et $y < \Phi(t_0)$, on puisse trouver une fonction affine

$$t \longmapsto \alpha_{t_0,y_0} t + \beta_{t_0,y_0}$$

dont le graphe au dessus de I passe par (t_0, y_0) et reste strictement en dessous de l'ensemble E_Φ. On en déduit donc que, sur I, Φ est exactement l'enveloppe supérieure des fonctions affines qui la minorent. Prenons une telle fonction affine

$$L(t) = \alpha t + \beta\,.$$

La fonction mesurable $\alpha f + \beta$ est intégrable relativement à la mesure μ d'après la proposition 2.6 puisque μ est supposée de masse totale finie (et même normalisée pour que $\mu(\Omega) = 1$) et on a

$$\int_\Omega (\alpha f + \beta)\, d\mu = \alpha \int_\Omega f\, d\mu + \beta\,.$$

Or, si $([a_k, b_k])_k$ est une suite croissante de segments exhaustant I, on a

$$\int_\Omega f\, d\mu = \lim_{k\to+\infty} \int_{\{f\in[a_k,b_k]\}} f\, d\mu$$

d'après la propriété (2.13) appliquée aux deux fonctions intégrables positives f^+ et f^- ; comme de plus

$$a_k \int_{\{f\in[a_k,b_k]\}} d\mu \leq \int_{\{f\in[a_k,b_k]\}} f\, d\mu \leq b_k \int_{\{f\in[a_k,b_k]\}} d\mu \tag{2.26}$$

et que $\mu(\Omega) = 1$, il en résulte en faisant tendre k vers l'infini dans (2.26) que

$$\int_\Omega f\, d\mu \in I\,.$$

Le nombre

$$\Phi\Big(\int_\Omega f\, d\mu\Big)$$

s'obtient donc comme la borne supérieure, lorsque $L = L_{\alpha,\beta}$ est une fonction affine minorant strictement Φ sur l'intervalle ouvert I, des nombres

$$L\Big(\int_\Omega f\, d\mu\Big) = \int_\Omega (\alpha f + \beta)\, d\mu \leq \int_\Omega (\Phi\circ f)\, d\mu \in]-\infty, \infty]\,;$$

notons que l'intégrale de $\Phi \circ f$ relativement à la mesure μ est bien définie puisque cette fonction étant minorée par une fonction intégrable telle précisément que $L(f)$, elle est telle que $(\Phi\circ f)^-$ est une fonction positive intégrable relativement à la mesure μ, ce qui ôte toute ambiguïté du type $\infty - \infty$ pour la définition de

$$\int_\Omega (\Phi\circ f)\, d\mu := \int_\Omega (\Phi\circ f)^+\, d\mu - \int_\Omega (\Phi\circ f)^-\, d\mu \in]-\infty, \infty]\,.$$

Comme tous les nombres
$$L\Big(\int_\Omega f\,d\mu\Big)$$
(lorsque L est une fonction affine minorant strictement Φ sur I) sont majorés par
$$\int_\Omega (\Phi\circ f)\,d\mu \in]-\infty,\infty]\,,$$
on obtient bien l'inégalité de Jensen voulue. $\diamondsuit$

Il ne faut pas omettre de mentionner parmi les propriétés utiles de l'intégrale des fonctions $(\Omega,\mathcal{T})$-$(\mathbb{R},\mathcal{B})$ mesurables la clause de *monotonie* : si f et g sont des fonctions réelles mesurables et intégrables toutes les deux relativement à la mesure μ, la propriété (2.6)
$$(f\leq g)\Longrightarrow \Big(\int_\Omega f\,d\mu \leq \int_\Omega g\,d\mu\Big)$$
reste évidemment valable.

Enfin, si $(E_k)_k$ est une suite croissante d'éléments de $\mathcal{T}$ exhaustant Ω et si f est une fonction (à valeurs dans $\mathbb{R},\mathbb{C},\mathbb{R}^m$) mesurable et intégrable relativement à μ, la règle (2.13) reste valide :
$$\lim_{k\to+\infty}\int_\Omega f\chi_{E_k}\,d\mu = \int_\Omega f\,d\mu\,.$$

2.9 La clause de « domination » et le théorème de Lebesgue

Donnons pour commencer un critère très utile d'intégrabilité ; ce critère s'énoncera en termes d'une clause de « domination par une fonction intégrable ».

Proposition 2.9 (critère d'intégrabilité par domination) *Soit Ω un ensemble équipé d'une tribu $\mathcal{T}$, $\mu\ :\mathcal{T}\longrightarrow[0,\infty]$ une mesure positive. Une fonction $f\ :\ \Omega\longrightarrow\mathbb{R},\mathbb{C}$, ou $\mathbb{R}^m$ est intégrable relativement à la mesure μ si et seulement si :*

- *d'une part, elle est $(\Omega,\mathcal{T})$-$(\mathbb{R},\mathbb{C}$ ou $\mathbb{R}^m,\mathcal{B})$ mesurable ;*
- *d'autre part, il existe une fonction $g\ :\Omega\longrightarrow[0,\infty]$ mesurable relativement aux tribus $\mathcal{T}$ (à la source) et $\mathcal{B}$ (au but), intégrable relativement à la mesure μ, et de plus telle que l'on ait la clause de domination :*

$$\forall\,\omega\in\Omega\,,\quad |f(\omega)|\leq g(\omega)\,. \tag{2.27}$$

Preuve. Si f est à valeurs réelles, on remarque que la condition $|f|\leq g$ implique $f^\pm\leq g$; l'intégrabilité de g implique donc celle de $f^\pm$, donc celle de $f^++f^-=|f|$. Réciproquement, si f est intégrable, $g=|f|$ l'est aussi et peut être considérée comme une fonction « dominante ». Les cas où f est à valeurs complexes ou vectorielles (dans $\mathbb{R}^m$) se traitent de manière identique ; par exemple, si $f=(f_1,...,f_m)\ :\Omega\longrightarrow\mathbb{R}^m$, la condition (2.27) implique $|f_j^\pm|\leq\sqrt{m}\,g$ pour tout $j=1,...,m$, donc l'intégrabilité de f_j dès que ces fonctions coordonnées sont toutes supposées mesurables. $\diamondsuit$

Exemple important 2.3 (domination et critère de Riemann). C'est l'occasion de se souvenir ici (voir le cours de MHT401) que la fonction $t\in]0,\infty[\mapsto t^\alpha$, $\alpha\in\mathbb{R}$, (souvent utilisée comme fonction dominante) est intégrable sur $]0,1]$ relativement à la mesure de Lebesgue si et seulement si $\alpha > -1$ (critère de Riemann). En revanche, elle est intégrable sur $[1,+\infty[$ si et seulement si $\alpha > -1$. En dimension m, on le verra plus loin, mais il est bon d'ores et déjà de le retenir, la fonction $x \mapsto \|x\|^\alpha$, $\alpha \in \mathbb{R}$, est intégrable dans la boule unité si et seulement si $\alpha > -m$, intégrable dans le complémentaire de cette boule si et seulement si $\alpha < -m$.

Nous sommes maintenant en état d'énoncer l'un des plus importants théorèmes de la théorie de l'intégration, le théorème de « convergence dominée », point fort du travail d'Henri Lebesgue initié avec (« *Sur une généralisation de l'intégrale définie* ») dès 1901 ; ce résultat est aussi connu comme le théorème de Fatou-Lebesgue :

Théorème 2.3 (théorème de convergence dominée) *Soit Ω un ensemble, équipé d'une tribu $\mathcal{T}$, $\mu : \mathcal{T} \longrightarrow [0,\infty]$ une mesure positive. Soit $(f_k)_{k\geq 0}$ une suite de fonctions mesurables de Ω dans $\mathbb{R}, \mathbb{C}$ ou $\mathbb{R}^m$ (relativement à la tribu $\mathcal{T}$ à la source et à la tribu borélienne $\mathcal{B}$ au but), toutes intégrables relativement à la mesure μ, telle que :*

- *d'une part, la suite $(f_k)_k$ converge simplement pour μ-presque tout $\omega \in \Omega$;*
- *d'autre part, il existe une fonction $g : (\Omega, \mathcal{T})$-$([0,\infty], \mathcal{B})$, <u>intégrable</u> relativement à la mesure μ, <u>dominant</u> uniformément les f_k, $k \in \mathbb{N}$, au sens suivant :*

$$\forall\, k \in \mathbb{N}\,,\ \forall\, \omega \in \Omega\,,\ |f_k(\omega)| \leq g(\omega)\,. \tag{2.28}$$

Il existe alors au moins une fonction $(\Omega, \mathcal{T})$-$(\mathbb{R}, \mathbb{C}$ ou $\mathbb{R}^m, \mathcal{B})$ mesurable f telle

$$f(\omega) = \lim_{k\to+\infty} f_k(\omega)$$

pour μ-presque tout ω[16]*. De plus, étant donnée une telle fonction f, f est aussi intégrable relativement à la mesure μ et l'on dispose de la possibilité d'intervertir limite lorsque k tend vers $+\infty$ et prise d'intégrale relativement à la mesure μ :*

$$\lim_{k\to+\infty} \Big(\int_\Omega f_k\, d\mu \Big) = \int_\Omega f\, d\mu\,. \tag{2.29}$$

Preuve. Montrons d'abord l'existence d'une fonction mesurable f telle que

$$f = \lim_{k\to+\infty} f_k$$

μ-presque partout. L'ensemble E des points $\omega \in \Omega$ où la suite $f_k(\omega)$ n'est pas convergente est un élément de $\mathcal{T}$ (si par exemple $f = (f_1, ..., f_m)$ est une application de Ω dans $\mathbb{R}^m$, les ensembles $E_j := \{\liminf f_k < \limsup f_k\}$, $j = 1, ..., m$, sont des éléments de $\mathcal{T}$ d'après la proposition 2.1). Si l'on définit $\tilde{f}_k = f_k$ sur $\Omega \setminus E$ et $\tilde{f}_k = 0$ sur E, la suite $(\tilde{f}_k)_k$ est une suite de fonctions mesurables simplement convergente vers une fonction mesurable f qui est bien μ-presque partout limite simple de la suite $(f_k)_k$. Il existe donc bien au moins une fonction mesurable f telle que $f = \lim_{k\to+\infty} f_k$ μ-presque partout.

Nous poursuivons la preuve en nous limitant au cas des fonctions à valeurs réelles, cas auquel on peut immédiatement se ramener en raisonnant suite de fonctions coordonnées par suite de fonctions coordonnées. Donnons nous donc une telle fonction f (nécessairement intégrable du fait du critère d'intégrabilité de la proposition 2.9

16. Si la tribu $\mathcal{T}$ est μ-complète, toute fonction μ-presque partout égale à $\lim_{k\to+\infty} f_k$ fait l'affaire.

puisque $|f| \leq g$ μ-presque partout) et appliquons le lemme de Fatou (théorème 2.2) à la suite $(2g - |f - f_k|)_k$; on a

$$\begin{aligned}
\int_\Omega \liminf_{k\to+\infty}(2g - |f - f_k|)\, d\mu &= 2\int_\Omega g\, d\mu \\
&\leq \liminf_{k\to+\infty}\Big(\int_\Omega (2g - |f - f_k|)\, d\mu\Big) \\
&\leq 2\int_\Omega g\, d\mu - \limsup_{k\to+\infty}\Big(\int_\Omega |f - f_k|\, d\mu\Big),
\end{aligned}$$

d'où il résulte

$$\limsup_{k\to+\infty}\int_\Omega |f - f_k|\, d\mu = 0\,,$$

ce qui implique (2.29), puisque

$$\Big|\int_\Omega f_k\, d\mu - \int_\Omega f\, d\mu\Big| \leq \int_\Omega |f_k - f|\, d\mu$$

d'après la proposition 2.7. $\diamondsuit$

Exemple 2.4. L'exemple où $\Omega = \mathbb{N}$ et où μ désigne la mesure de comptage est un exemple important d'application du théorème de Lebesgue, exemple que vous connaissez sans doute déjà depuis le cours de L2 (voir par exemple l'UE MHT401) ; le résultat (que l'on peut dans ce cas démontrer de manière élémentaire[17]) est le suivant : soit $(u_{k,l})_{k,l\in\mathbb{N}}$ un tableau indexé par deux indices $k \in \mathbb{N}$ et $l \in \mathbb{N}$ de nombres complexes tel que :

- pour tout $l \in \mathbb{N}$, la suite $(u_{k,l})_k$ converge vers une limite $u_{\infty,l}$ dans $\mathbb{C}$;
- il existe une série numériques à termes positifs $[w_k]_k$ convergente telle que

$$\forall\, l \in \mathbb{N}\,,\ \forall\, k \in \mathbb{N}\,,\ |u_{k,l}| \leq w_l\,. \tag{2.30}$$

Alors, toutes les séries $[u_{k,l}]_l$, $k \in \mathbb{N} \cup \{\infty\}$, sont absolument convergentes et on a de plus

$$\lim_{k\to+\infty}\Big(\sum_{l\in\mathbb{N}} u_{k,l}\Big) = \sum_{l\in\mathbb{N}} u_{\infty,l}\,. \tag{2.31}$$

Par exemple

$$\lim_{k\to+\infty}\sum_{l=1}^{\infty}\frac{k^2}{l^2k^2+1} = \sum_{l=1}^{\infty}\frac{1}{l^2} = \frac{\pi^2}{6}\,;$$

le résultat (2.31) est par contre trivialement faux lorsque la clause de domination (2.30) est en défaut (comme on le voit par exemple avec $u_{k,l} = 1$ si $k = l$, 0 sinon, où le membre de gauche de (2.31) vaut 1, tandis que le membre de droite vaut 0).

Exemple 2.5. Les exemples d'application du théorème de convergence dominée de Lebesgue sont légion et plusieurs sont proposés dans les guides d'activités sous Ulysse. En voici un célèbre, pour lequel une approche ≪ piétonne ≫ pour obtenir le résultat suivant la démarche inspirée de Riemann s'avère particulièrement délicate[18] : on a, pour tout $x > 0$,

$$\begin{aligned}
\Gamma(x) := \int_0^\infty t^{x-1}e^{-t}\, dt &= \lim_{k\to+\infty}\int_0^k t^{x-1}\Big(1 - \frac{t}{k}\Big)^k dt \\
&= \lim_{k\to+\infty}\Big(k^x\int_0^1 u^{x-1}(1-u)^k\, du\Big) \\
&= \lim_{k\to+\infty}\frac{k^x k!}{x(x+1)\ldots(x+k)}\,,
\end{aligned}$$

17. Voir par exemple l'énoncé du théorème 3.6 et sa preuve dans le polycopié du cours MHT401 : http ://www.math.u-bordeaux1.fr/∼yger/mat401.pdf

18. Voir par exemple l'ouvrage de Jean Dieudonné, *Calcul infinitésimal*, Hermann.

célèbre formule attribuée au mathématicien suisse du siècle des Lumières Leonhard Euler (1707-1783). L'application possible du théorème de Lebesgue ici tient à la clause de domination

$$t^{x-1}\Big(1-\frac{t}{k}\Big)^k \chi_{]0,k[}(t) \leq t^{x-1}e^{-t}\,,\ \forall\, t \in]0,\infty[\,,$$

et au fait que la fonction $t \longmapsto t^{x-1}e^{-t}$ est continue (donc mesurable) et intégrable sur $]0,\infty[$ relativement à la mesure de Lebesgue dt. La majoration de domination provient du fait que, pour tout $X \in [0,1[$,

$$\log(1-X) = -X - X^2/2 - X^3/3 - \cdots \leq -X$$

et de, bien sûr $(1-t/k)^k = \exp(k\log(1-t/k))$ si $t \in]0,k[$. Ce résultat aurait pu dans ce cas particulier se déduire du théorème de convergence monotone (ou du lemme de Fatou). Si l'on souhaite vraiment y voir une vraie application du théorème de convergence dominée, il faut remplacer $x > 0$ par un nombre complexe z tel que $\mathrm{Re}\, z > 0$. Dans ce cas, on a, pour tout $n \in \mathbb{N}^*$, pour tout $t \in]0,\infty[$,

$$\chi_{]0,k[}(t)(1-t/k)^k|t^{z-1}| = \chi_{]0,k[}(t)\exp\Big(k\Big(-\frac{t}{k}-\frac{t^2}{2k^2}-\frac{t^3}{3k^3}-\dots\Big)\Big)t^{\mathrm{Re}(z)-1} \leq e^{-t}t^{\mathrm{Re}\, z-1}$$

avec (comme on le vérifie grâce aux critères de comparaison, voir le cours de MHT401)

$$\int_{]0,\infty[} t^{\mathrm{Re}\, z-1}e^{-t}dt < \infty.$$

Le théorème de convergence dominée s'applique (cette fois on en a vraiment besoin car les fonctions ne sont plus positives !) pour affirmer que l'intégrale

$$\Gamma(z) = \int_{]0,\infty[} t^{z-1}e^{-t}\,dt\,,\ \mathrm{Re}\, z > 0$$

converge et que sa valeur est donnée par

$$\begin{aligned}\Gamma(z) &= \lim_{k\to+\infty}\int_0^k \Big(1-\frac{t}{k}\Big)^k t^{z-1}\,dt\\ &= \lim_{k\to+\infty} k^z\int_0^1 u^{z-1}(1-u)^k\,du\\ &= \lim_{k\to+\infty}\frac{k!k^z}{z(z+1)\dots(z+k)} \quad \text{si } \mathrm{Re}\, z > 0\end{aligned}$$

(la dernière égalité résulte d'un calcul par intégration par parties, voir le cours de MHT401). L'importance de la fonction Γ en mathématiques, en physique ou en sciences de l'ingénieur, tient au fait qu'elle vérifie, toujours si $\mathrm{Re}\, z > 0$, l'équation fonctionnelle

$$\Gamma(z+1) = z\Gamma(z)$$

(on fera juste une intégration par parties) et par conséquent interpole la fonction $n \in \mathbb{N} \mapsto n!$ au sens où $\Gamma(n) = (n-1)!$ pour tout $n \in \mathbb{N}^*$; par exemple $\Gamma(1/2) = (1/2)! = \sqrt{\pi}$.

Exemple 2.6. Autre exemple très important, emprunté cette fois au cadre de l'intégration discrète. Si z est un nombre complexe tel que $\mathrm{Re}\, z > 1$, la fonction

$$n \in \mathbb{N}^* \mapsto \frac{1}{n^z}$$

est intégrable par rapport à la mesure de décompte δ sur $\mathbb{N}^*$ (la tribu sur $\mathbb{N}^*$ étant ici celle de toutes les parties) car

$$\int_{\mathbb{N}^*}\Big|\frac{1}{n^z}\Big|\,d\delta(n) = \sum_{n=0}^{\infty}\frac{1}{n^{\mathrm{Re}\, z}} < +\infty$$

en vertu du critère de Riemann (voir le cours de MHT401). On peut donc définir la fonction d'Euler ζ sur le demi-plan $\{\mathrm{Re}\, z > 1\}$ par

$$\zeta(z) = \sum_{n=0}^{\infty}\frac{1}{n^z} \quad \text{si } \mathrm{Re}\, z > 1.$$

L'importance de cette fonction tient au fait qu'elle permet d'« encoder » analytiquement le théorème fondamental de l'arithmétique : « *tout nombre entier strictement positif se décompose d'une unique manière comme un produit de puissances de nombres premiers* ». En effet, on a la formule (due à Euler)

$$\zeta(z) = \lim_{N\to+\infty} \prod_{k=1}^{N} \frac{1}{1-1/p_k^z} \quad \text{si } \mathrm{Re}\, z > 1 \tag{$*$}$$

(ici $p_1 = 2$, $p_2 = 3$, $p_3 = 5$,...,p_k désigne le k-ième nombre premier). En effet, notons A_N le sous ensemble de $\mathbb{N}^*$ défini par

$$n \in A_N \iff n \text{ se décompose au plus avec } p_1, ..., p_N.$$

On a, pour tout $N \in \mathbb{N}^*$, pour tout $n \in \mathbb{N}^*$,

$$\Big|\frac{1}{n^z}\Big| \chi_{A_N}(n) \le \frac{1}{n^{\mathrm{Re}\, z - 1}}.$$

Comme la fonction dominante $n \mapsto 1/n^{\mathrm{Re}\, z-1}$ est intégrable sur $\mathbb{N}^*$ par rapport à la mesure de décompte δ, on a

$$\begin{aligned}\lim_{N\to+\infty} \int_{\mathbb{N}^*} \chi_{A_N}(n) \frac{1}{n^z}\, d\delta(n) &= \lim_{N\to+\infty} \sum_{n\in A_N} \frac{1}{n^z} \\ &= \int_{\mathbb{N}^*} \lim_{N\to+\infty} \Big(\chi_{A_N}(n)\frac{1}{n^z}\Big)\, d\delta(n) \\ &= \int_{\mathbb{N}^*} \frac{1}{n^z}\, d\delta(n) = \zeta(z)\end{aligned}$$

grâce au théorème de convergence dominée. Notons que l'on a utilisé ici le fait que tout entier positif non nul n se décompose d'une unique manière comme un produit de puissances de nombres premiers. L'unicité de la décomposition permet aussi d'affirmer que, pour $N \in \mathbb{N}^*$,

$$\int_{\mathbb{N}^*} \chi_{A_N}(n) \frac{1}{n^z}\, d\delta(n) = \sum_{n\in A_N} \frac{1}{n^z} = \sum_{k_1=0}^{\infty} \cdots \sum_{k_N=0}^{\infty} \frac{1}{p_1^{k_1 z} \cdots p_N^{k_N z}}. \tag{$\dagger$}$$

Or, comme $p_j \ge 2$, on a, pour tout $j = 1, ..., N$

$$\sum_{k=0}^{\infty} \frac{1}{p_j^{kz}} = \frac{1}{1-1/p_j^z} \tag{$\dagger\dagger$}$$

du fait de la formule

$$\frac{1}{1-u} = \sum_{k=0}^{\infty} u^k \quad \text{si } |u| < 1$$

(on a bien en effet $|p_j^{-z}| \le 1/2$ puisque $p_j \ge 2$ et $\mathrm{Re}\, z > 1$). En multipliant les identités ($\dagger\dagger$) pour $j = 1, ..., N$ et en combinant avec ($\dagger$), on trouve donc

$$\int_{\mathbb{N}^*} \chi_{A_N}(n) \frac{1}{n^z}\, d\delta(n) = \prod_{j=1}^{N} \sum_{k_j=0}^{\infty} \frac{1}{p_j^z} = \prod_{k=1}^{N} \frac{1}{1-1/p_k^z}.$$

En compilant tous les résultats obtenus, on obtient bien la formule d'Euler ($*$).

2.10 Retour au lien avec l'intégration au sens de Riemann

Armés du théorème de convergence dominée Lebesgue (théorème 2.3), nous sommes en mesure de montrer que toute fonction définie sur un segment $[a, b]$ de $\mathbb{R}$, à valeurs dans $\mathbb{R}, \mathbb{C}$ ou $\mathbb{R}^m$, et bornée sur $[a, b]$, est automatiquement intégrable au sens de Lebesgue dès qu'elle l'est au sens de Riemann, les deux intégrales (au sens de Riemann et de Lebesgue) étant alors égales. Cela prolonge le résultat établi dans la proposition 2.5 pour les fonctions continues positives.

Proposition 2.10 *Soit f une fonction $[a,b] \longrightarrow \mathbb{R}, \mathbb{C}\, ou\, \mathbb{R}^m$, bornée sur $[a,b]$. On suppose f intégrable au sens de Riemann sur $[a,b]$ (i.e,* $\mathrm{Re}\, f$, $\mathrm{Im}\, f$, *ou plus généralement toutes les applications coordonnées $f_j, j = 1, ..., m$ sont Riemann-intégrables sur $[a,b]$). Alors f est mesurable (relativement aux tribus boréliennes à la source et au but), intégrable au sens de Lebesgue sur $[a,b]$ (relativement à la mesure de Lebesgue dt) et l'on a*

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

Preuve. On peut se limiter au cas (en fait non restrictif par passage *via* les applications coordonnées f_j, $j = 1, ..., m$ ou les deux fonctions $\mathrm{Re}\, f$, $\mathrm{Im}\, f$ dans le cas où f est à valeurs complexes) où f est à valeurs réelles, bornée en valeur absolue par une constante positive M. Il existe alors, si f est intégrable au sens de Riemann sur $[a,b]$, deux suites de fonctions en escalier $(\varphi_k)_{k\geq 1}$ et $(\psi_k)_{k\geq 1}$ sur $[a,b]$, la première croissante, la seconde décroissante, telles que

$$\forall\, t \in [a,b]\,,\ -M-1 \leq \varphi_k(t) \leq f(t) \leq \psi_k(t) \leq M+1$$

et

$$\int_a^b (\psi_k(t) - \varphi_k(t))\, dt < 1/k\,. \tag{2.32}$$

Les fonctions $f^* := \inf_k \psi_k$ et $f_* := \sup_k \varphi_k$ sont mesurables sur $[a,b]$ d'après la proposition 2.1 puisque toutes les fonctions en escalier le sont. D'après le théorème de convergence dominée de Lebesgue, les $|\varphi_k|$ étant toutes majorées par la fonction intégrable constante égale à $M+1$, on a

$$\begin{aligned}
\lim_{k\to+\infty} \int_{[a,b]} \varphi_k(t)\, dt &= \int_{[a,b]} f_*(t)\, dt \\
\lim_{k\to+\infty} \int_{[a,b]} \psi_k(t)\, dt &= \int_{[a,b]} f^*(t)\, dt\,.
\end{aligned}$$

Mais l'intégrale de Riemann d'une fonction en escalier coïncide évidemment avec son intégrale au sens de Lebesgue ; on a donc (du fait de (2.32) une fois que l'on a fait tendre k vers l'infini)

$$\int_{[a,b]} f^*(t)\, dt = \int_{[a,b]} f_*(t)\, dt\,,$$

d'où

$$\int_{[a,b]} (f^*(t) - f_*(t))\, dt = 0\,.$$

Il résulte alors de la proposition 2.4 que la fonction positive $f^* - f_*$ est nulle presque partout (au sens de Lebesgue). La fonction f est donc presque partout égale à une fonction mesurable $f_* = f^*$, et est donc mesurable. En fait, le théorème 2.3 assure aussi que cette fonction est intégrable sur $[a,b]$ (elle est majorée en valeur absolue par une constante). On a d'ailleurs aussi prouvé au passage l'identité entre les intégrales au sens de Lebesgue de f sur $[a,b]$. $\diamondsuit$.

Le lien entre intégrabilité au sens de Riemann et intégrabilité au sens de Lebesgue pour une fonction $f : [a,b] \longrightarrow \mathbb{R}$ bornée en module sur le segment $[a,b]$ est précisé par le théorème suivant, dû encore à Henri Lebesgue :

Théorème 2.4 (théorème de Lebesgue) *Soit $f : [a,b] \longrightarrow \mathbb{R}$ une fonction bornée en valeur absolue sur un segment de $\mathbb{R}$. La fonction f est intégrable au sens de Riemann sur $[a,b]$ si et seulement si elle est mesurable et intégrable au sens de Lebesgue sur $[a,b]$ et si, de plus, l'ensemble des points de discontinuité de f est négligeable relativement à la mesure de Lebesgue. Les intégrales calculées au sens de Riemann et de Lebesgue dans ce cas coïncident.*

Preuve. On a vu (proposition 2.10) que, si f est intégrable au sens de Riemann sur $[a,b]$, non seulement elle est mesurable et intégrable au sens de Riemann sur $[a,b]$ (les deux intégrales coïncidant), mais encore elle s'écrit (si l'on examine la preuve de la proposition 2.10) dt-presque partout comme la limite $f^* = f_*$ croissante (ou décroissante) d'une suite de fonctions en escalier. Les points de discontinuité d'une fonction en escalier sur $[a,b]$ étant en nombre fini, on montre aisément que l'ensemble des points de discontinuité de f (sur l'ensemble E de complémentaire négligeable relativement à la mesure de Lebesgue ou précisément $f = f^* = f_*$) est au plus dénombrable. L'ensemble des points de discontinuité de f est donc l'union de l'ensemble négligeable (relativement à la mesure de Lebesgue) $[a,b] \setminus E$ et d'un ensemble au plus dénombrable, donc aussi négligeable relativement à la mesure de Lebesgue ; ainsi l'ensemble des points de discontinuité de f est un ensemble négligeable relativement à la mesure de Lebesgue et l'assertion du théorème est donc bien prouvée dans le sens direct. L'autre implication est plus délicate et nous l'admettrons ici [19] . $\diamondsuit$

Exemple 2.7. Parmi les fonctions $[a,b] \longrightarrow \mathbb{R}$ intégrables au sens de Riemann sur $[a,b]$ figurent les fonctions bornées en valeur absolue et de plus *réglées*, c'est-à-dire admettant en tout point de $[a,b]$ à la fois une limite à gauche et une limite à droite. Cette classe de fonctions coïncide en fait avec la classe des limites uniformes sur $[a,b]$ des suites de fonctions en escalier. On pourra vérifier en exercice que l'ensemble des points de discontinuité d'une fonction réglée est au plus dénombrable (dénombrez par exemple les discontinuités telles que le saut de discontinuité dépasse $1/k$, $k \in \mathbb{N}$). Cette classe est donc plus restreinte que la classe des fonctions mesurables bornées en valeur absolue et dont l'ensemble des points de discontinuité est négligeable relativement à la mesure de Lebesgue : il existe donc des fonctions bornées (en valeur absolue) intégrables au sens de Riemann sur $[a,b]$, mais non réglées !

19. On pourra par exemple se référer au livre de W. Rudin, « Analyse réelle et complexe », chez Masson.

Chapitre 3

Les outils de l'intégration « pratique »

3.1 Introduction : le plan du chapitre

Soit $(f_\tau)_{\tau\in T}$ une collection de fonctions définies sur un espace mesuré $(\Omega, \mathcal{T}, \mu)$, à valeurs réelles ou complexes, mesurables, toutes intégrables relativement à la mesure μ, et indexées par un paramètre τ ; le paramètre τ peut par exemple être le temps balayant un intervalle de $\mathbb{R}$, une variable discrète balayant $\mathbb{N}$ ou $\mathbb{Z}$, ou une variable d'espace décrivant un ouvert de $\mathbb{R}^n$.

On est souvent confronté, dans nombre de problèmes pratiques (en particulier issus de la physique ou de l'ingénierie), à l'étude de la fonction

$$\tau \in T \longmapsto \int_\Omega f_\tau \, d\mu$$

dans l'ensemble T où elle est définie : est-elle continue ? différentiable (lorsque T est un ouvert de $\mathbb{R}^m$) ? si oui, que valent ses dérivées partielles ? peut-on « dériver », comme on en aurait souvent envie, sous le signe somme ? Les résultats du cours d'analyse de L2 (MHT401) pour les séries de fonctions $[f_n]_n$, avec $f_n \ : T \longmapsto \mathbb{C}$ (exemples : continuité de la somme, différentiation terme à terme [1]), envisagés dans le cadre de l'intégration sur $\mathbb{N}$ équipé de la mesure μ de comptage, s'étendent-t'ils à un contexte plus général (celui d'un espace abstrait Ω équipé d'une tribu $\mathcal{T}$ et d'une mesure positive $\mu \ : \mathcal{T} \longrightarrow [0, \infty]$) ? On verra que oui dans la section 3.2, ce qui sera de fait une conséquence facile du théorème de convergence dominée (théorème 2.3). On illustrera aussi cette section avec des exemples de transformations classiques (de *Fourier*, de *Laplace*, de *Mellin*).

Comme on le sait pour l'avoir beaucoup utilisé lors de la recherche de primitive, le *changement de variable* sous l'intégrale est un des outils clef de l'intégration « pratique » au sens de Riemann (tout autant que le classique procédé d'« *intégration par parties* », avatar continu de la règle d'Abel). La formule de changement de variables dans l'intégrale de Riemann, à savoir

$$\int_{\varphi(a)}^{\varphi(b)} f(u)\, du = \int_a^b f(\varphi(t))\, \varphi'(t)\, dt \qquad (*)$$

1. Voir les théorèmes classiques du cours de MAT401 concernant les séries de fonctions : en l'occurrence les théorèmes 3.2 (et surtout son corollaire 3.1), 3.4 et 3.5 de la section 3.1.5 du polycopié du cours de MAT401 [http ://www.math.u-bordeaux1.fr/∼yger/mat401.pdf].

lorsque f est une fonction continue sur un intervalle contenant $\varphi([a,b])$, φ désignant une fonction de classe C^1 au voisinage de $[a,b]$, est d'ailleurs plutôt une conséquence du fait que si F est une primitive de f au voisinage de $\varphi([a,b])$, $F \circ \varphi$ est une primitive de $t \mapsto f(\varphi(t))\varphi'(t)$ d'après la règle de dérivation des fonctions composées, plus qu'un vrai théorème d'intégration ; les deux intégrales (*) sont en effet toutes les deux égales à $F(\varphi(b)) - F(\varphi(a))$. Si l'intégration par parties (de par son intime relation avec la règle d'Abel) demeure intrinsèquement liée à la notion de semi-convergence et, par là même, plus au point de vue « Riemann » (mettant en avant la notion de primitive par rapport à celle de calcul d'aire) [2], il n'en va pas de même avec l'outil « *changement de variables* » tout aussi efficace et utile dans le cadre de l'intégration Lebesgue sur les sous-ensembles mesurables de $\mathbb{R}^n$ relativement à la mesure de Lebesgue (on en connaît déjà un exemple important avec la propriété d'invariance par translation de la mesure de Lebesgue). La section 3.3 sera dévolue à cet outil (dans le contexte de l'intégration sur les sous-ensembles de $\mathbb{R}^n$ relativement à la mesure de Lebesgue).

Enfin, la pratique de l'intégration dans $\mathbb{R}^n = \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ nous guidera naturellement, étant donnés deux espaces mesurés $(\Omega_j, \mathcal{T}_j, \mu_j)$, $j = 1, 2$, vers la notion d'intégration sur l'ensemble produit $\Omega_1 \times \Omega_2$, équipé d'une certaine « *tribu produit* » $\mathcal{T}_1 \otimes \mathcal{T}_2$ (à définir), ainsi que d'une « *mesure produit* »

$$\mu_1 \otimes \mu_2 \ : \mathcal{T}_1 \otimes \mathcal{T}_2 \longrightarrow [0, \infty]$$

(elle aussi à définir). Les théorèmes de G. Fubini - L. Tonelli et de G. Fubini [3] seront ici les résultats majeurs du volet « pratique » ; ils s'avèreront des auxiliaires précieux, par exemple concernant le calcul des intégrales multiples « bloc de variables » après « bloc de variables » sur les sous-ensembles mesurables de $\mathbb{R}^n$ (pensé comme $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$), souvent d'ailleurs combinés avec l'outil « *changement de variables* ». La section 3.4 de ce chapitre sera centrée sur ces notions.

3.2 L'intégration Lebesgue « à paramètres »

3.2.1 Continuité ; dérivabilité suivant un paramètre réel

Dans toute cette section, on se donne un ensemble abstrait Ω, équipé d'une tribu $\mathcal{T}$, tribu sur laquelle nous définissons une mesure positive $\mu \ : \mathcal{T} \longrightarrow [0, \infty]$.

On considère un ensemble de paramètres T que dans un premier temps nous supposons simplement être un espace métrique (*i.e* un ensemble équipé d'une topologie définie par une distance d).

Constamment par la suite, on considérera une collection $(f_\tau)_{\tau \in T}$ de fonctions $(\Omega, \mathcal{T})$-$(\mathbb{C}, \mathcal{B})$ mesurables (nous ne le répéterons pas dans les énoncés mais ferons toujours

2. Ce qui n'empêche pas bien sûr que cette règle reste un auxiliaire majeur de calcul dans le cadre de l'intégration Lebesgue sur $\mathbb{R}$ ou $\mathbb{R}^n$ puisque l'intégration Lebesgue relativement à la mesure de Lebesgue, comme nous l'avons vu dans les sections 2.6 et 2.10 du chapitre 2, ne fait qu'élargir le procédé proposé par Riemann lorsque nous avons affaire par exemple à la classe des fonctions continues par morceaux.

3. Guido Fubini (1879-1943), mathématicien italien éclectique (géométrie différentielle, calcul de variations, théorie des groupes, ...) énonça le théorème d'intégration que nous mentionnons ici dès 1907. Le nom d'un autre mathématicien italien, Leonida Tonelli (1885-1946), est aussi attaché à ce type d'énoncé, on verra pourquoi dans la section 3.4.

cette hypothèse de mesurabilité sur les fonctions f_τ). Nous supposerons aussi (même si ceci s'avèrera du fait redondant au vu des contraintes que nous imposerons dans nos divers énoncés) que les fonctions f_τ, $\tau \in T$, sont toutes intégrables sur Ω relativement à la mesure μ, ce qui nous permet de définir d'entrée de jeu la fonction

$$F \ : \tau \in T \longmapsto \int_\Omega f_\tau \, d\mu \,,$$

dite *intégrale à paramètres* (sur Ω et relativement à la mesure μ) de la famille $(f_\tau)_{\tau \in T}$; bien sûr, remplacer chaque f_τ par 0 sur un ensemble μ-négligeable (dépendant éventuellement de τ) ne change en rien ni la mesurabilité des f_τ, ni bien sûr la fonction F.

Les questions que l'on va se poser concernent la « régularité » de la fonction F. Sans hypothèse supplémentaire sur l'univers des paramètres T, la seule question que l'on puisse envisager est celle de la continuité de la fonction F sur T. Voici dans ce cas le résultat :

Théorème 3.1 (continuité de l'intégrale fonction de paramètres) *Soit τ_0 un point de T. Outre les hypothèses figurant en préliminaire de cette section, on suppose :*
- *que pour tout ω hors d'un sous-ensemble μ-négligeable E_{τ_0}, la fonction*

$$\tau \longmapsto f_\tau(\omega)$$

  est continue en $\tau = \tau_0$;
- *qu'il existe un voisinage $V(\tau_0)$ de τ_0 dans T et une fonction $g_{\tau_0} \ : \Omega \longrightarrow [0, \infty]$ $(\Omega, \mathcal{T})$-$([0, \infty], \mathcal{B})$ mesurable, intégrable sur Ω relativement à la mesure μ, tels que*

$$\forall \tau \in V(\tau_0) \,, \ \forall \omega \in \Omega \,, \ |f_\tau(\omega)| \leq g_{\tau_0}(\omega) \,. \tag{3.1}$$

Alors la fonction F est continue en $\tau = \tau_0$.

Remarque 3.1. Bien sûr la clause de domination (3.1) s'avère indispensable ici comme le montre l'exemple de la famille $(f_\tau)_{\tau \in [0,1]}$ de fonctions mesurables sur $\Omega =]0, 1]$ (la tribu étant la tribu borélienne sur $]0, 1]$, la mesure étant la mesure de Lebesgue) définies par

$$\begin{aligned} &\forall \, t \in]0, 1] \,, \ f_0(t) = 0 \\ &\forall \, t \in]0, 1] \,, \ f_\tau(t) = \frac{\chi_{[0,\tau]}(t)}{\tau} \qquad \forall \, \tau \in]0, 1] \,. \end{aligned}$$

Pour tout $t \in]0, 1]$, il est immédiat de voir que $\tau \in [0, 1] \longmapsto f_\tau(t)$ est continue en $\tau = 0$. Pourtant la fonction F n'est pas continue en $\tau = 0$ car

$$F(0) = \int_{]0,1]} f_0(t) \, dt = 0$$

tandis que

$$F(\tau) = \int_{]0,1]} \frac{\chi_{[0,\tau]}(t)}{\tau} \, dt = 1 \qquad \forall \, \tau \in]0, 1] \,.$$

La clause (3.1) est bien sûr en défaut sur cet exemple, ce qui explique pourquoi le théorème 3.1 ne s'applique pas.

Remarque 3.2. On aurait pu aussi énoncer un théorème global (concluant à la continuité de F sur T tout entier) en remplaçant la clause de domination (3.1) par l'existence d'une fonction mesurable intégrable $g \ : \Omega \longrightarrow [0, \infty]$ telle que

$$\forall \, \tau \in T \,, \ \forall \omega \in \Omega \,, \ |f_\tau(\omega)| \leq g(\omega) \,;$$

ceci toutefois n'est pas réaliste car il est bien rare de pouvoir trouver un tel chapeau intégral g « coiffant » toutes les fonctions f_τ, $\tau \in T$, ce sur Ω tout entier. La continuité est une propriété locale et il est beaucoup plus sage, pour vérifier que F est continue sur T tout entier, de chercher à vérifier la continuité de F en tout point avec le théorème 3.1 ; on peut aussi, pour éviter une démarche redondante, chercher à exhiber, pour chaque compact K de T, un chapeau intégral g_K dominant sur Ω toutes les fonctions $|f_\tau|$ lorsque $\tau \in K$. Il convient bien sûr de pratiquer beaucoup d'exemples pour se familiariser avec cette démarche (voir les guides d'activités 5 et 6 sous Ulysse).

Preuve. Pour vérifier que F est continue en τ_0, il suffit (puisque T est un espace dont la topologie est définie par une distance) de vérifier que pour toute suite $(\tau_k)_{k\geq 1}$ tendant vers τ_0, on a

$$\lim_{k\to\infty} F(\tau_k) = F(\tau_0)\,. \tag{3.2}$$

Le fait que l'on se ramène à tester la continuité séquentielle est ici capital car nous sommes incapables de gérer les énoncés des théorèmes du type convergence monotone ou dominée autrement que dans le cas où nous avons affaire à des suites de fonctions (on a déjà souligné le rôle crucial du concept de dénombrabilité dans toute cette théorie de l'intégration).

Prenons donc une telle suite $(\tau_k)_{k\geq 1}$ dont nous supposerons tous les termes appartenir à $V(\tau_0)$. On sait que, pour tout $\tau \in T$,

$$F(\tau) = \int_{\Omega\setminus E_{\tau_0}} f_\tau \, d\mu$$

puisque E_{τ_0} est μ-négligeable. Toutes les fonctions $\tau \longmapsto f_\tau(\omega)$, étant continues en $\tau = \tau_0$ pour tout $\omega \in \Omega \setminus E_{\tau_0}$, la suite de fonctions $(f_{\tau_k})_{k\geq 1}$ converge simplement sur $\Omega \setminus E_{\tau_0}$ vers la fonction f_{τ_0}. Mais on a aussi

$$\forall\, k \in \mathbb{N}^*\,,\ \forall\, \omega \in \Omega\,,\ |f_{\tau_k}(\omega)| \leq g_{\tau_0}(\omega)\,.$$

La clause (2.28) du théorème 2.3 est donc remplie et l'on déduit de ce théorème que

$$\lim_{k\to+\infty} F(\tau_k) = \lim_{k\to+\infty} \int_\Omega f_{\tau_k}\, d\mu = \int_\Omega f_{\tau_0}\, d\mu\,,$$

ce qui prouve le résultat (3.2) voulu. $\diamondsuit$

Nous envisageons maintenant, pour évoquer la régularité au cran supérieur de F, le cas particulier où T est un sous-ensemble de $\mathbb{R}$ d'intérieur non vide. Nous pouvons alors énoncer le résultat suivant :

Théorème 3.2 (dérivabilité de l'intégrale fonction d'un paramètre réel) *Soit τ_0 un point intérieur à T. Outre les hypothèses figurant en préliminaire de cette section, on suppose qu'il existe un voisinage ouvert $V(\tau_0)$ de τ_0 dans T, un sous-ensemble E_{τ_0} de Ω, μ-négligeable, enfin une fonction $g_{\tau_0} : \Omega \longrightarrow [0,\infty]$ qui soit $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ mesurable et intégrable sur Ω relativement à la mesure μ, tels que*

- *pour tout ω dans $\Omega \setminus E_{\tau_0}$, la fonction*

$$\tau \longmapsto f_\tau(\omega)$$

soit dérivable (resp. de classe C^1) sur $V(\tau_0)$;

– *l'on ait la clause de domination*

$$\forall\,\tau \in V(\tau_0)\,,\ \forall\,\omega\,\in \Omega \setminus E_{\tau_0}\,,\ \Big|\frac{d}{d\tau}\,[f_\tau(\omega)]\Big| \leq g_{\tau_0}(\omega)\,. \tag{3.3}$$

Alors, toutes les fonctions

$$\omega \in \Omega \setminus E_{\tau_0} \longmapsto \frac{d}{d\tau}\,[f_\tau(\omega)]\,,\ \tau \in V(\tau_0)$$

se prolongent en des fonctions $(\Omega, \mathcal{T})$-$(\mathbb{C}, \mathcal{B})$ mesurables de Ω dans $\mathbb{C}$ (que l'on notera pour simplifier de la même manière), toutes intégrables relativement à la mesure μ ; de plus la fonction F est dérivable (resp. de classe C^1) dans $V(\tau_0)$ et on a

$$F'(\tau) = \int_{\Omega\setminus E_{\tau_0}} \frac{d}{d\tau}\,[f_\tau(\omega)]\,d\mu(\omega) = \int_\Omega \frac{d}{d\tau}\,[f_\tau(\omega)]\,d\mu(\omega) \qquad \forall\,\tau \in V(\tau_0)\,. \tag{3.4}$$

Remarque 3.3. Il est essentiel de remarquer ici que la clause essentielle de domination (3.3) porte cette fois sur les dérivées (par rapport au paramètre τ) des fonctions $\tau \longmapsto f_\tau(\omega)$ pour ω hors de l'ensemble μ-négligeable E_{τ_0}. On retrouve une clause de domination similaire à celle qui est impliquée dans le théorème de dérivation des séries de fonctions terme à terme (théorème 3.5 du cours de MHT401 par exemple) ; ce n'est nullement une surprise car, en fait, l'énoncé de ce théorème du cours de L2 sur les séries de fonctions correspond à un cas particulier de notre nouveau théorème 3.2, celui où $\Omega = \mathbb{N}$ et où μ est la mesure de décompte.

Remarque 3.4. Que la clause (3.3) puisse être oubliée (quand bien même (3.1) serait satisfaite) et tout l'édifice s'effondre ! Par exemple, une application du théorème des résidus [4] assure que, pour tout $\tau \in \mathbb{R}$,

$$\int_{\mathbb{R}} \frac{e^{-i\tau t}}{1+t^2}\,dt = \lim_{R\to+\infty} \int_{-R}^{R} \frac{e^{-i\tau t}}{1+t^2}\,dt = \pi e^{-|\tau|/2}\,.$$

On constate que, bien que la clause (3.1) du théorème 3.1 de continuité de l'intégrale par rapport au paramètre s'applique, il n'en saurait être de même pour la clause 3.3 puisque

$$\Big|\frac{d}{d\tau}\Big[\frac{e^{-i\tau t}}{1+t^2}\Big]\Big| = \frac{|t|}{1+t^2}$$

et que la fonction

$$t \longmapsto \frac{|t|}{1+t^2}$$

(indépendante de τ et qui serait la seule envisageable pour jouer le rôle de chapeau intégral dans (3.3)) n'est pas intégrable relativement à la mesure de Lebesgue sur $\mathbb{R}$ (les fonctions f_τ définies par $f_\tau(t) := e^{-i\tau t}/(1+t^2)$ pour $t \in \mathbb{R}$, toutes dérivables sur $\mathbb{R}$ par rapport à τ, le sont, elles, puisque $t \longmapsto (1+t^2)^{-1}$ est intégrable sur $\mathbb{R}$ et joue d'ailleurs le rôle de chapeau intégrable dominant tous les f_τ dans (3.1)). On constate que

$$t \longmapsto \pi e^{-|t|/2}$$

n'est pas dérivable en $t = 0$, ce qui montre clairement que le théorème 3.2 (dont les hypothèses ne sont pas toutes remplies ici car (3.3) n'est pas satisfaite) ne s'applique pas. La continuité de cette fonction est par contre bien assurée par le théorème 3.1.

Preuve. Pour montrer que F est dérivable en tout point de $V(\tau_0)$, il suffit de montrer que, pour toute suite ξ_k tendant vers 0, la limite lorsque k tend vers l'infini de

$$\frac{F(\tau+\xi_k) - F(\tau)}{\xi_k}$$

4. Voir par exemple la section 5.3.5 du polycopié de MAT401 :
http ://www.math.u-bordeaux1.fr/∼ yger/mat401.pdf

existe pour tout $\tau \in V(\tau_0)$. Nous allons précisément introduire une telle suite $(\xi_k)_k$ et prolonger par 0 sur E_{τ_0}, pour τ fixé dans $V(\tau_0)$, toutes les fonctions

$$\omega \longmapsto \frac{f_{\tau+\xi_k}(\omega) - f_\tau(\omega)}{\xi_k}$$

(toutes bien définies pour $k = k(\tau)$ assez grand car $V(\tau_0)$ est ouvert). Pour un tel τ, nous avons ainsi une suite de fonctions mesurables dont on sait par hypothèses qu'elle converge simplement sur $\Omega \setminus E_{\tau_0}$ puisque toutes les fonctions

$$\tau \longmapsto f_\tau(\omega)\,,\ \omega \notin E_{\tau_0}$$

sont supposées dérivables sur $V(\tau_0)$; comme il y a évidemment aussi convergence simple de la suite de fonctions ainsi prolongées à Ω (ceci pour chaque τ dans $V(\tau_0)$), on en déduit que la limite d'une telle suite (pour τ fixé dans $V(\tau_0)$) est bien une fonction mesurable de Ω dans $\mathbb{C}$ prolongeant

$$\omega \in \Omega \setminus E_{\tau_0} \longmapsto \frac{d}{d\tau}(f_\tau(\omega))\,.$$

L'intégrabilité de toutes ces nouvelles fonctions prolongeantes (lorsque $\tau \in V(\tau_0)$) relativement à la mesure μ résulte bien sûr de la clause de domination (3.3). On a donc bien prouvé ainsi la première assertion de notre théorème. Il est tout à fait naturel de continuer à noter, ce que nous ferons, ces nouvelles fonctions (définies, elles, sur Ω tout entier)

$$\omega \longmapsto \frac{d}{d\tau}\,[f_\tau(\omega)]\,, \qquad \tau \in V(\tau_0)\,.$$

Nous allons maintenant utiliser le théorème de convergence dominée de Lebesgue 2.3 pour montrer que, pour tout $\tau \in V(\tau_0)$,

$$\begin{aligned}&\lim_{k\to+\infty}\Big(\frac{F(\tau+\xi_k)-F(\tau)}{\xi_k} - \int_\Omega \frac{d}{d\tau}\,[f_\tau(\omega)]\,d\mu(\omega)\Big)\\ =\ &\lim_{k\to+\infty}\int_\Omega \Big[\frac{f_{\tau+\xi_k}(\omega)-f_\tau(\omega)}{\xi_k} - \frac{d}{d\tau}\,[f_\tau(\omega)]\Big]\,d\mu(\omega) = 0\,.\end{aligned}$$

D'après l'inégalité des accroissements finis [5], on a, pour k suffisamment grand pour que $[\tau, \tau+\xi_k] \subset V(\tau_0)$ $(k \geq k(\tau))$

$$\forall\,\omega\ \in \Omega \setminus E_{\tau_0}\,,\ \frac{|f_{\tau+\xi_k}(\omega) - f_\tau(\omega)|}{|\xi_k|} \leq \sup_{\eta\in[\tau,\tau+\xi_k]}\Big|\frac{d}{d\eta}\,[f_\eta(\omega)]\Big| \leq g_{\tau_0}(\omega)\,;$$

Cette inégalité est aussi valable si l'on remplace les fonctions

$$\omega \in \Omega \setminus E_{\tau_0} \longmapsto \frac{f_{\tau+\xi_k}(\omega) - f_\tau(\omega)}{\xi_k}\,, \quad \tau \in V(\tau_0),\ k \geq k(\tau),$$

par leurs prolongements $\varphi_{\tau,k}$ par 0 sur l'ensemble μ-négligeable E_{τ_0}. La clause de domination (2.28) dans l'énoncé du théorème 2.3 est remplie si l'on regarde, pour τ fixé

5. Voir le cours de calcul différentiel MHT513.

dans $V(\tau_0)$, la suite de fonctions $(\varphi_{\tau,k})_{k\geq k(\tau)}$; comme il y a aussi sur Ω convergence simple de cette suite de fonctions vers la fonction

$$\omega \longmapsto \frac{d}{d\tau}\,[f_\tau(\omega)]$$

(au sens de nos notations étendues aux prolongements de $\Omega \setminus E_{\tau_0}$ à tout Ω), le théorème de convergence dominée de Lebesgue (théorème 2.3) s'applique et fournit à la fois la justification du fait que F soit dérivable en tout point de $V(\tau_0)$ et la formule (3.4).

Si de plus, pour tout $\omega \in \Omega \setminus E_{\tau_0}$, les fonctions

$$\tau \longmapsto f_\tau(\omega)$$

sont supposées de classe C^1 sur $V(\tau_0)$, le théorème 3.1, appliqué à l'intégrale à paramètre

$$\tau \in V(\tau_0) \longmapsto \int_\Omega \frac{d}{d\tau}\,[f_\tau(\omega)]\,d\mu(\omega)\,,$$

nous permet de conclure (au vu de la clause de domination (3.3)) à la continuité de cette fonction dans $V(\tau_0)$; la fonction F est bien dans ce cas de classe C^1 dans $V(\tau_0)$. Le théorème est complètement démontré. $\diamondsuit$

Plus encore concernant la dérivabilité de F que la continuité, il est très important de se souvenir que la dérivabilité est une propriété qui se teste localement. Si l'on envisage de montrer que F est dérivable (ou de classe C^1) dans un ouvert $T \subset \mathbb{R}$ tout entier, il faut souvent se satisfaire de le faire localement. Il est en effet très difficile (bien souvent impossible) de trouver un « chapeau intégrable » g dominant toutes les fonctions

$$\omega \longmapsto \frac{d}{d\tau}\,[f_\tau(\omega)]$$

dans T ! La remarque 3.2 précédente est donc tout aussi valable, sinon plus encore, dans ce contexte d'application du théorème 3.2.

Exemple 3.1. Pour montrer que la fonction ζ de Riemann (voir Exemple 2.6)

$$x \in]1,+\infty[\longmapsto \sum_{n=1}^{\infty} \frac{1}{n^x}$$

est de classe C^1, il suffit de trouver un chapeau intégral (ici $\Omega = \mathbb{N}^*$ et la mesure est la mesure de décompte) pour toutes les suites

$$\Big(\frac{1}{n^x} \times (-\log n)\Big)_{n\geq 1}$$

lorsque $x \in [\alpha,+\infty[$, avec $\alpha > 1$ arbitraire. Ce chapeau intégral est évidemment ici fourni par la suite

$$\Big(\frac{\log n}{n^\alpha}\Big)_{n\geq 1}$$

qui définit bien une fonction intégrable positive sur $\mathbb{N}^*$ équipé de la mesure de décompte puisque la série correspondante converge. En itérant ceci, on voit d'ailleurs que ζ est C^∞ sur $]1,+\infty[$. La dérivée d'ordre p de ζ sur $]1,+\infty[$ est la fonction

$$x \mapsto \sum_{n=1}^{\infty} \frac{(-\log n)^p}{n^x}.$$

Le théorème 3.2 peut être utilisé de manière récursive pour prouver qu'une intégrale à paramètres est m fois dérivable (*resp.* de classe C^m) au voisinage d'un point τ_0 intérieur à $T \subset \mathbb{R}$. Il faut simplement toujours avoir en tête que la clause de domination doit porter sur la dérivée d'ordre m, donc d'ordre maximal envisagé, si l'on prétend prouver pour F un résultat de régularité (dérivabilité, continue dérivabilité) jusqu'à l'ordre m.

3.2.2 Intégrales dépendant de plusieurs paramètres

Si T est un sous-ensemble de $\mathbb{R}^n$ (on note $\tau = (\tau^{(1)}, ..., \tau^{(n)})$ les variables) et τ_0 un point intérieur à T, on se souvient qu'il y a, pour une fonction $F\ : T \longrightarrow \mathbb{C}$, équivalence entre les deux assertions suivantes [6] :

1. F est de classe C^1 au voisinage de τ_0 ;
2. F admet des dérivées partielles $\partial F/\partial \tau^{(j)}$, $j = 1, ..., n$, par rapport à toutes les variables, ce pour tout point voisin de τ_0, et ces dérivées partielles sont toutes continues au voisinage du point τ.

Ce résultat permet de ramener l'étude de la régularité C^1 d'une intégrale fonction de n paramètres à l'application du théorème 3.3. On en déduit immédiatement ce critère de régularité C^1 :

Théorème 3.3 (dérivabilité de l'intégrale fonction de n paramètres réels) *Soit τ_0 un point intérieur à $T \subset \mathbb{R}^n$. Outre les hypothèses figurant en préliminaire de la sous-section 3.2.1, on suppose qu'il existe un voisinage ouvert $V(\tau_0)$ de τ_0 dans T, un sous-ensemble $E_{\tau_0} \in \mathcal{T}$, μ-négligeable, enfin une fonction $g_{\tau_0}\ : \Omega \longmapsto [0, \infty]$ $(\Omega, \mathcal{T})$-$([0, \infty], \mathcal{B})$ mesurable, intégrable sur Ω relativement à la mesure μ tels que*

- *pour tout ω dans $\Omega \setminus E_{\tau_0}$, les fonctions*

$$\tau \longmapsto f_\tau(\omega)$$

soient de classe C^1 sur $V(\tau_0)$;
- *l'on ait la clause de domination*

$$\forall\, \tau \in V(\tau_0)\,,\ \forall\, \omega\ \in \Omega \setminus E_{\tau_0}\,,\ \left|\frac{\partial}{\partial \tau^{(j)}}\, [f_\tau(\omega)]\right| \leq g_{\tau_0}(\omega)\,,\ j = 1, ..., n\,. \tag{3.5}$$

Alors, toutes les fonctions

$$\omega \in \Omega \setminus E_{\tau_0} \longmapsto \frac{\partial}{\partial \tau^{(j)}}\, [f_\tau(\omega)]\,,\ \tau \in V(\tau_0)\,,\ j = 1, ..., n,$$

se prolongent en des fonctions $(\Omega, \mathcal{T})$-$(\mathbb{C}, \mathcal{B})$ mesurables de Ω dans $\mathbb{C}$ (que l'on notera pour simplifier de la même manière), toutes intégrables relativement à la mesure μ ; de plus la fonction F est de classe C^1 dans $V(\tau_0)$ et on a, pour $j = 1, ..., n$,

$$\frac{\partial F}{\partial \tau^{(j)}}(\tau) = \int_{\Omega \setminus E_{\tau_0}} \frac{\partial}{\partial \tau^{(j)}}\, [f_\tau(\omega)]\, d\mu(\omega) = \int_\Omega \frac{\partial}{\partial \tau^{(j)}}\, [f_\tau(\omega)]\, d\mu(\omega) \tag{3.6}$$

6. Voir le cours de calcul différentiel, UE MHT513.

Ce résultat peut être exploité de manière récursive pour prouver la régularité C^m d'une intégrale fonction de n paramètres réels. Il convient de toujours se souvenir que la clause de domination doit porter sur TOUTES les dérivées partielles jusquà l'ordre maximal m envisagé.

Il est un cas particulier ici important, celui où $T = U$ est un ouvert $\mathbb{C}$; on rappelle (voir le cours de MAT401 [7]) qu'une *fonction analytique complexe* dans un ouvert U de $\mathbb{C}$ est une fonction $f \ : U \longrightarrow \mathbb{C}$ telle que, pour $z_0 \in U$, il existe une série entière $[a_k(z_0)X^k]_{k\geq 0}$ de rayon de convergence strictement positif [8] tel qu'au voisinage de z_0 [9], on ait

$$f(z) = \sum_{k=0}^{\infty} a_k(z_0)(z-z_0)^k \,.$$

Objets « intermédiaires » entre les fonctions polynomiales (dont elles héritent de la « rigidité ») et les fonctions C^∞ (dont elles n'ont pas toute la « souplesse »), les fonctions analytiques complexes (comme $z \longmapsto \exp z, \cos z, \sin z$ dans $\mathbb{C}$, ou encore $z \longmapsto \log(1 \pm z)$ dans le disque ouvert $D(0,1)$, $z \longmapsto \zeta(z)$ dans le demi-plan $\{z\,;\,\mathrm{Re}\, z > 1\}$, *etc.*) sont des fonctions très intéressantes tant du point de vue de l'analyse que des mathématiques discrètes.

Les fonctions analytiques f dans un ouvert U de $\mathbb{C}$ sont les fonctions de classe C^1 de U dans $\mathbb{C}$ vérifiant l'équation aux dérivées partielles de Cauchy-Riemann

$$\frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y} \equiv 0$$

(ce qui revient à dire que l'application linéaire tangente en tout point de U lorsque f est pensée comme une application d'un ouvert de $\mathbb{R}^2$ dans $\mathbb{R}^2$ est une similitude directe de $\mathbb{R}^2$ dans lui-même) [10]. Une fonction analytique complexe admet (dans l'ouvert U de $\mathbb{C}$ où elle vit) une dérivée au sens complexe (elle aussi analytique) [11] définie par

$$z \in U \longmapsto \frac{df}{dz} = f'(z) := \lim_{h\to 0, h\in\mathbb{C}} \frac{f(z+h)-f(z)}{h} = \frac{1}{2}\Big(\frac{\partial f}{\partial x} - i\frac{\partial f}{\partial y}\Big)(z) = \frac{\partial f}{\partial x}(z)\,.$$

Une autre caractérisation de l'analyticité complexe est la suivante : une fonction de classe C^1 $f \ : U \longrightarrow \mathbb{C}$ est analytique complexe dans U si et seulement si, pour tout disque fermé $\overline{D(z_0,r)} \subset U$, pour tout z dans $D(z_0,r)$, on a, avatar de la célèbre formule de Green-Riemann, la formule intégrale de Cauchy [12]

$$f(z) = \frac{1}{2\pi}\int_0^{2\pi} \frac{f(z_0+re^{i\theta})}{(z_0+re^{i\theta})-z}\, re^{i\theta} d\theta\,. \tag{3.7}$$

On peut alors énoncer le résultat suivant :

7. Plus précisément section 4.1.5 du polycopié en ligne http ://www.math.u-bordeaux1.fr/∼yger/mat401.pdf
8. En fait, au moins égal à la distance de z_0 à la frontière de U.
9. En fait, dans le disque ouvert $D(z_0, d(z_0,\partial U))$.
10. Voir le théorème 4.5 du cours de MAT401 pour l'énoncé de ce résultat et sa preuve.
11. Voir l'énoncé et la preuve du théorème 4.2 dans le polycopié du cours de MAT401.
12. Voir le théorème 4.6 et plus généralement l'ensemble de la section 4.1.5 dans le polycopié du cours MAT401 : http ://www.math.u-bordeaux1.fr/∼yger/mat401.pdf

Théorème 3.4 *(***analyticité des intégrales fonction d'un paramètre complexe)** *Soit $(\Omega, \mathcal{T}, \mu)$ un espace équipé d'une tribu et d'une mesure positive sur cette tribu. Soit U un ouvert de $\mathbb{C}$ et $(f_z)_{z\in U}$ une collection de fonctions de Ω dans $\mathbb{C}$, toutes $(\Omega, \mathcal{T})$-$(\mathbb{C}, \mathcal{B})$-mesurables, telles que :*

- *Pour tout ω hors d'un sous-ensemble E μ négligeable, les fonctions*
$$z \in U \longmapsto f_z(\omega)$$
sont toutes analytiques complexes dans U ;
- *Pour tout point z_0 de U, il existe un disque fermé $\overline{D(z_0, r_{z_0})}$ inclus dans U, une fonction mesurable positive $g_{z_0} : \Omega \longrightarrow [0, \infty]$ telle que*
$$\forall z \in \overline{D(z_0, r_{z_0})},\ \forall \omega \in \Omega\,,\ |f_z(\omega)| \leq g_{z_0}(\omega)\,. \tag{3.8}$$

Alors la fonction
$$F\ : z \in U \longmapsto \int_\Omega f_z(\omega)\, d\mu(\omega)$$
est aussi analytique dans U et l'on a, pour tout $z \in U$,
$$F'(z) = \int_\Omega \frac{d}{dz}\,[f_z(\omega)]\, d\mu(\omega) \tag{3.9}$$
(la fonction sous l'intégrale ci-dessous étant intégrable).

Preuve. Grâce à la formule de Cauchy (3.7) et à l'utilisation du théorème 3.2 (sur $[0, 2\pi]$ et avec la mesure de Lebesgue [13]), on observe que si $D_{x\,ou\,y}$ est l'opérateur différentiel
$$D_x = \frac{\partial}{\partial x} \quad \text{ou} \quad D_y = \frac{\partial}{\partial y}$$
et si $z \in \Omega \setminus E$, on a
$$D_{x\,ou\,y}[f_z(\omega)] = \frac{1}{2\pi}\int_0^{2\pi} f_{z_0+r_{z_0}e^{i\theta}}(\omega) D_{x\,ou\,y}\Big[\frac{1}{(z_0 + r_{z_0}e^{i\theta}) - x - iy}\Big]\, r_{z_0}e^{i\theta} d\theta\,.$$
On constate que
$$\begin{aligned}\frac{\partial}{\partial x}\Big[\frac{1}{r_{z_0}e^{i\theta} - x - iy}\Big] &= \frac{1}{((z_0 + r_{z_0}e^{i\theta}) - x - iy)^2}\\ \frac{\partial}{\partial y}\Big[\frac{1}{r_{z_0}e^{i\theta} - x - iy}\Big] &= \frac{i}{((z_0 + r_{z_0}e^{i\theta}) - x - iy)^2}\end{aligned}$$
et que, par conséquent, pour tout z dans $D(z_0, r_{z_0}/2)$, pour tout $\omega \in \Omega$,
$$|D_{x\,ou\,y}[f_z(\omega)]| \leq \frac{r_{z_0}}{2\pi(r_{z_0}/2)^2}\int_0^{2\pi} |f_{z_0+r_{z_0}e^{i\theta}}(\omega)|\, d\theta\,.$$

13. Il s'agit ici d'une application tout à fait élémentaire que l'on aurait pu tout aussi bien gérer avec les outils de l'intégration Riemann et le théorème élémentaire d'interversion de prise de limite et de prise d'intégrale lorsque la convergence des fonctions en jeu est uniforme sur l'intervalle d'intégration, voir le théorème bien classique de L2, par exemple le théorème 3.7 du cours de MAT401.

Compte tenu de la clause de domination (3.8), il vient donc

$$\forall\, z \in D(z_0, r_{z_0}/2)\,,\ |D_{x\,ou\,y}\,[f_z(\omega)]| \leq \frac{4}{r_{z_0}}\, g_{z_0}(\omega)$$

et la clause de domination (3.5) du théorème 3.3 se trouve remplie. On déduit de ce théorème (appliqué au point courant z_0 de U en prenant $V(z_0) = D(z_0, r_{z_0}/2)$) que F est de classe C^1 dans U et que

$$\Big(\frac{\partial F}{\partial x} + i\frac{\partial F}{\partial y}\Big)(z) = \int_\Omega \Big[\frac{\partial [f_z(\omega)]}{\partial x} + i\frac{\partial [f_z(\omega)]}{\partial y}\Big]\, d\mu(\omega) = 0\,.$$

La fonction F est une fonction de classe C^1 solution de l'équation aux dérivées partielles de Cauchy Riemann, c'est donc une fonction analytique complexe dans U. La formule (3.9) résulte, elle, de ce que

$$\begin{aligned} F'(z) &= \frac{1}{2}\Big(\frac{\partial F}{\partial x} - i\frac{\partial F}{\partial y}\Big)(z) \\ &= \int_\Omega \frac{1}{2}\Big[\frac{\partial [f_z(\omega)]}{\partial x} - i\frac{\partial [f_z(\omega)]}{\partial y}\Big]\, d\mu(\omega) = \int_\Omega \frac{d}{dz}\,[f_z(\omega)]\, d\mu(\omega)\,. \end{aligned}$$

Le théorème est ainsi démontré. $\diamondsuit$

3.2.3 Fourier, Laplace, Mellin : trois exemples majeurs

La transformation de Fourier

Si f est une fonction mesurable de $\mathbb{R}^n$ dans $\mathbb{R}$, intégrable relativement à la mesure de Lebesgue dx, on définit sa *transformée de Fourier*[14] comme la fonction

$$\widehat{f}\ : \omega \in \mathbb{R}^n \longmapsto \int_{\mathbb{R}^n} f(x) e^{-i\langle \omega, x\rangle}\, dx\,,$$

où $\langle\ ,\ \rangle$ désigne le produit scalaire usuel sur $\mathbb{R}^n$. Comme

$$|f(x)e^{-i\langle \omega, x\rangle}| = |f(x)|\,,$$

la fonction $g = |f|$ joue de rôle de chapeau intégrable dans (3.1) et le théorème 3.1 s'applique pour assurer la continuité de $\widehat{f}$ sur $\mathbb{R}^n$. On vérifie aisément que pour que $\hat{f}$ soit de classe C^m, il suffit (on peut dans ce cas appliquer récursivement le théorème 3.3) que

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

(à faire en exercice). Nous ne développerons pas plus avant ici cette importante transformation $f \longmapsto \widehat{f}$ issue de la physique (la diffraction en optique la matérialise) car elle sera étudiée plus en détails dans l'UE MHT613.

Exemples 3.2. L'exemple proposé dans la remarque 3.4 est un exemple de transformation de Fourier : la transformée de Fourier de la fonction

$$f\ : t \in \mathbb{R} \longmapsto \frac{1}{1+t^2}$$

14. En référence au mathématicien et homme politique français Jean-Baptiste Joseph Fourier (1768-1830).

est la fonction

$$\widehat{f}\ : \omega \in \mathbb{R} \longmapsto \pi\, e^{-|\omega|/2}$$

(on a juste changé les notations). Un autre exemple jouera un rôle capital, celui de la gaussienne normalisée

$$g\ : t \in \mathbb{R} \longmapsto \frac{1}{\sqrt{2\pi}}\, e^{-t^2/2}$$

correspondant en probabilités à la densité d'une loi normale réduite centrée (loi de Gauss), omniprésente dans les théorèmes limite du calcul des probabilités [15]. On a

$$\widehat{g}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-i\omega t}\, e^{-t^2/2}\, dt\,.$$

Il résulte immédiatement du théorème 3.2 que $\widehat{g}$ et dérivable, et de dérivée

$$\begin{aligned} \frac{d\widehat{g}}{d\omega}(\omega) &= \frac{-i}{\sqrt{2\pi}} \int_{\mathbb{R}} t e^{-i\omega t} e^{-t^2/2}\, dt \\ &= \frac{-i}{\sqrt{2\pi}} \Big(\Big[- e^{-t^2/2}(-i\omega) e^{-i\omega t} \Big]_{-\infty}^{+\infty} - i\omega \sqrt{2\pi}\, \widehat{g}(\omega) \Big) \\ &= -\omega\, \widehat{g}(\omega) \end{aligned}$$

si l'on procède à partir de la seconde ligne à une intégration par parties. L'équation différentielle linéaire homogène du premier ordre dont $\widehat{g}$ est solution se résout immédiatement et conduit à

$$\widehat{g}(\omega) = \widehat{g}(0)\, e^{-\omega^2/2} = e^{-\omega^2/2}$$

si l'on admet

$$\int_{\mathbb{R}} e^{-t^2/2}\, dt = \sqrt{2\pi}\,,$$

ce que l'on retrouvera d'ailleurs dans la section 3.4.

Fonction caractéristique d'une variable aléatoire réelle

La notion précédente s'étend à un cadre plus abstrait : supposons que Ω soit un ensemble, $\mathcal{T}$ une tribu sur Ω, μ une mesure positive sur $\mathcal{T}$ telle que $\mu(\Omega) < +\infty$. Si

$$X\ : \Omega \longmapsto \mathbb{R}$$

est une application $(\Omega, \mathcal{T})$-$(\mathbb{R}, \mathcal{B})$ mesurable, il en est de même (par composition avec l'exponentielle $t \longmapsto e^{i\tau t}$ qui est continue de $\mathbb{R}$ dans $\mathbb{C}$) pour l'application

$$\omega \in \Omega \longmapsto e^{i\tau X(\omega)}$$

si $\tau \in \mathbb{R}$. La fonction

$$\tau \in \mathbb{R} \longmapsto \int_{\Omega} e^{i\tau X}\, d\mu$$

est une intégrale fonction d'un paramètre réel. Comme

$$\forall\, \tau \in \mathbb{R}\,,\ \forall\, \omega \in \Omega\,,\ |e^{i\tau X(\omega)}| = 1$$

et que la fonction positive constante et égale à 1 sur Ω est intégrable car $\mu(\Omega) < \infty$, la fonction

$$F_X\ : \tau \in \mathbb{R} \longmapsto \int_{\Omega} e^{i\tau X}\, d\mu \tag{3.10}$$

est une fonction continue sur $\mathbb{R}$, dite *fonction caractéristique* [16] de la fonction mesurable (on parlera plutôt ici de *variable aléatoire* lorsque $\mu(\Omega) = 1$) $X\ : \Omega \longrightarrow \mathbb{R}$.

15. Voir en particulier le théorème « *limite central* » (au sens « central » dans la théorie) que vous rencontrerez dans l'UE MHT601.

16. Rien à voir avec l'autre notion de fonction caractéristique que nous avons auparavant manié (χ_A pour A partie d'un ensemble).

La transformation de Laplace

Si f est une fonction mesurable sur $[0,+\infty[$, à valeurs complexes (les tribus sont les tribus boréliennes) et telle qu'il existe $x \in \mathbb{R}$ avec

$$\int_{]0,\infty[} |f(t)|\, e^{-xt}\, dt < \infty\,,$$

la *transformée de Laplace*[17] de f, outil de grande importance tant pour les ingénieurs que les physiciens, est par définition la fonction

$$F = \mathcal{L}[f]\ : p \in \{p \in \mathbb{C}\,;\, \mathrm{Re}\, p > x_0\} \longmapsto \int_{[0,\infty[} f(t)e^{-pt}\, dt\,,$$

où

$$x_0 = \inf \Big\{ x \in \mathbb{R}\,;\, \int_{[0,\infty[} |f(t)|\, e^{-xt}\, dt < \infty \Big\} \in [-\infty, +\infty[\,.$$

La fonction $\mathcal{L}[f]$ est une fonction analytique dans le demi-plan ouvert

$$\Pi^+_{x_0} := \{p \in \mathbb{C}\,;\, \mathrm{Re}\, p > x_0\}\,.$$

En effet, x_0 étant nécessairement la limite (décroissante) d'une suite $(x_k)_{k\geq 1}$ de nombres réels telle que

$$\int_{[0,\infty[} |f(t)|\, e^{-x_k t}\, dt < \infty \qquad \forall\, k \in \mathbb{N}^*$$

(par définition de x_0 comme borne inférieure de l'ensemble des x tels que la fonction $t \longmapsto f(t)e^{-xt}$ soit intégrable sur $[0,+\infty[$), le critère de domination de la proposition 2.9 nous assure que, pour tout $\alpha > x_0$, on a

$$\int_{[0,\infty[} |f(t)|\, e^{-\alpha t}\, dt < +\infty\,.$$

Pour tout $\alpha > x_0$, pour tout p dans le demi-plan fermé

$$\overline{\Pi}^+_{\alpha} := \{p \in \mathbb{C}\,;\, \mathrm{Re}\, p \geq \alpha\}\,,$$

on a

$$\forall\, z \in \overline{\Pi}^+_{\alpha}\,,\ \forall t \in [0,\infty[\,,\ |f(t)e^{-pt}| \leq |f(t)|e^{-\alpha t}\,;$$

puisque la fonction majorante est intégrable sur $[0,+\infty[$ et joue donc le rôle d'un chapeau intégrant, on peut appliquer le théorème 3.4 après avoir observé que, pour tout $t \in [0,+\infty[$, la fonction

$$p \longmapsto f(t)e^{-pt} = \sum_{k=0}^{\infty} \frac{f(t)(-t)^k}{k!}\, p^k$$

17. Tout à la fois mathématicien, astronome, probabiliste, ministre sous le consulat, puis Comte d'Empire, Pierre-Simon Laplace (1749-1827) a posé les jalons de la mécanique et de l'analyse modernes.

est entière, donc analytique dans $\mathbb{C}$ tout entier. C'est donc le théorème 3.4 qui assure d'une part que $\mathcal{L}(f)$ est bien définie dans le demi-plan ouvert $\Pi^+_{x_0}$, d'autre part qu'il s'agit d'une fonction analytique dans ce cemi-plan ouvert, avec

$$\frac{d}{dp}[\mathcal{L}(f)](p) = \int_{[0,\infty[} (-t)f(t)e^{-pt}\,dt$$

(notons que le théorème 3.4 implique entre autres l'intégrabilité des fonctions

$$t \in [0,\infty[\longmapsto tf(t)e^{-pt}$$

pour $p \in \Pi^+_{x_0}$, ce que l'on pourrait d'ailleurs vérifier aussi aisément). Notons d'ailleurs que l'on a, par itération (facile à justifier)

$$\frac{d^m}{dp^m}[\mathcal{L}(f)](p) = \int_{[0,\infty[} (-t)^m f(t)e^{-pt}\,dt\,,\ \forall m \in \mathbb{N}$$

(les intégrales au second membre sont toutes absolument convergentes).

Exemple 3.3. Un exemple très important (dans les mathématiques de l'ingénieur) est celui où

$$f(t) = f_{\alpha,p_0}(t) = \frac{t^{\alpha-1}e^{p_0 t}}{\Gamma(\alpha)}\,, \quad p_0 \in \mathbb{C}\,,\ \alpha > 0\,,$$

pour tout $t \in]0,\infty[$ (et, de manière arbitraire, $f_{\alpha,p_0}(0) = 0$, cette valeur étant irrelevante puisque $\{0\}$ est négligeable relativement à la mesure de Lebesgue), avec

$$\Gamma(\alpha) := \int_0^\infty t^{\alpha-1}e^{-t}\,dt$$

(notons que $\Gamma(k) = (k-1)!$ si k est un entier strictement positif, avec la convention $0! = 1$). Dans ce cas, on a $x_0 = \mathrm{Re}\,p_0$ (on le voit immédiatement) et la transformée de Laplace de f_{α,p_0} est la fonction définie pour $\mathrm{Re}\,p > \mathrm{Re}\,p_0$ par

$$p \longmapsto \frac{1}{(p-p_0)^\alpha} := |p-p_0|^{-\alpha}e^{-i\alpha\mathrm{Arg}_{]-\pi/2,\pi/2[}[p-p_0]}\,.$$

Pour faire ce calcul, on peut par exemple utiliser la formule des résidus en utilisant le contour proposé dans la section 5.3.3 du cours de MAT401 pour calculer les intégrales de Fresnel. On aura besoin du calcul lorsque p est réel, calcul immédiat à faire par changement de variables.

Remarque 3.5. Si l'on fait simplement l'hypothèse que f est Riemann intégrable sur tout segment $[\alpha,\beta]$ de $[0,\infty[$ et que la limite lorsque t tend vers $+\infty$ de

$$t \in [0,\infty[\longmapsto \int_0^t f(u)e^{-x_0 u}\,du$$

existe, on peut montrer, en utilisant l'intégration par parties, que, pour tout p dans le demi-plan ouvert $\Pi^+_{x_0}$, la limite, lorsque t tend vers l'infini, de

$$t \in [0,\infty[\longmapsto \int_0^t f(u)e^{-pu}\,du$$

existe encore et définit donc une intégrale semi-convergente (au sens de Riemann). La fonction

$$p \in \Pi^+_0 \longmapsto \int_0^\infty \frac{\sin u}{u}\,e^{-pu}\,du$$

est de ce type car l'intégrale

$$\int_0^\infty \frac{\sin u}{u}\,du = \frac{\pi}{2}$$

(qui jouera aussi un grand rôle dans l'UE MHT613) est une intégrale semi-convergente (mais non convergente au sens de Lebesgue). Ce contexte [18] est beaucoup plus délicat ; il relève du point de vue « Riemann » plus que « Lebesgue ».

18. Présenté par exemple en détails dans l'exemple 2.5 du cours de MAT401, voir le polycopié en ligne http ://www.math.u-bordeaux1.fr/~yger/mat401.pdf

Transformée de Laplace d'une variable aléatoire positive

Dans le cadre abstrait, on peut transposer la notion précédente en nous plaçant cette fois sur un ensemble Ω équipé d'une tribu $\mathcal{T}$, d'une mesure positive. Si X est une application $(\Omega, \mathcal{T})$-$([0,\infty[, \mathcal{B})$ mesurable, telle que

$$\Big\{x \in \mathbb{R}\,;\ \int_\Omega e^{-xX(\omega)}\, d\mu(\omega) < \infty\Big\} \neq \emptyset\,,$$

on pose encore

$$x_0 := \inf \Big\{x \in \mathbb{R}\,;\ \int_\Omega e^{-xX(\omega)}\, d\mu(\omega) < \infty\Big\} \in [-\infty, \infty[$$

et l'on définit une fonction analytique complexe $\mathcal{L}[X]$ dans $\Pi^+_{x_0}$ par

$$\mathcal{L}[X](p) := \int_\Omega e^{-pX(\omega)}\, d\mu(\omega)\,.$$

Pour justifier ce résultat, on invoque le théorème 3.4 exactement comme précédemment. Lorsque de plus $\mu(\Omega) < \infty$ (par exemple si $\mu(\Omega) = 1$), on a $x_0 \leq 0$. Si de plus $x_0 < 0$, la fonction $\mathcal{L}[X]$ (qui est dans ce cas analytique au voisinage de $0 \in \Pi^+_{x_0}$) se développe en série entière comme

$$\mathcal{L}[X](p) = \sum_{k=0}^\infty (-1)^k \Big(\int_\Omega X^k(\omega)\, d\mu(\omega)\Big)\, \frac{p^k}{k!}$$

en termes des *moments*

$$\int_\Omega X^k(\omega)\, d\mu(\omega)\,,\ \ k \geq 0\,,$$

de la fonction mesurable X. Notons que, toujours si $x_0 < 0$ ou même si $x_0 = 0$, cette fonction $\mathcal{L}[X]$ est définie ou se prolonge par la formule

$$\mathcal{L}[X](p) = \int_\Omega e^{-pX(\omega)}\, d\mu(\omega)\,,$$

sur le demi-plan fermé $\{\mathrm{Re}\, p \geq 0\}$ et que l'on a

$$\forall \tau \in \mathbb{R}\,,\ \ \mathcal{L}[X](-i\tau) = F_X(\tau),,$$

où F_X désigne la fonction caractéristique de $X\ : \Omega \longmapsto [0,\infty[\subset \mathbb{R}$ telle qu'elle a été définie dans le paragraphe précédent (par (3.10)).

La transformation de Mellin

Soit $f\ : [0,\infty[\longmapsto \mathbb{C}$, mesurable (relativement aux tribus boréliennes) telle qu'il existe $x \in \mathbb{R}$ avec

$$\int_{[0,\infty[} |f(t)|\, t^{x-1}\, dt < \infty\,.$$

On note

$$\begin{aligned} x_0^+ &:= \sup \Big\{x \in \mathbb{R}\,;\ \int_{[1,\infty[} |f(t)|\, t^{x-1}\, dt < \infty\Big\} \in]-\infty, \infty] \\ x_0^- &:= \inf \Big\{x \in \mathbb{R}\,;\ \int_{[0,1]} |f(t)|\, t^{x-1}\, dt < \infty\Big\} \in [-\infty, \infty[\end{aligned}$$

Si l'on a $x_0^- < x_0^+$, on peut définir dans la bande

$$\Pi^+_{x_0^-} \cap \Pi^-_{x_0^+} = \{z \in \mathbb{C}\,;\, \mathrm{Re}\, z \in]x_0^-, x_0^+[\}$$

l'intégrale fonction du paramètre complexe z

$$\mathcal{M}[f] \ : z \in \Pi^+_{x_0^-} \cap \Pi^-_{x_0^+} \longmapsto \int_{[0,\infty[} f(t)\, t^z \,\frac{dt}{t}$$

et il résulte du théorème 3.4 que cette fonction est analytique complexe dans la bande ouverte $\Pi^+_{x_0^-} \cap \Pi^-_{x_0^+}$. Cette fonction (lorsque l'on peut ainsi la définir dans une bande verticale ouverte du plan complexe) est dite *transformée de Mellin*[19] de f.

Exemple 3.4. Si $f \ : t \longmapsto e^{-t}$, la transformée de Mellin de f est la fonction Γ définie dans le demi-plan Π_0^+ par

$$\Gamma(z) := \int_{[0,\infty[} t^{z-1}\, e^{-t}\, dt\,.$$

C'est une fonction analytique complexe dans ce demi-plan, interpolant les valeurs de la fonction factorielle, au sens où $\Gamma(k) = (k-1)!$ si $k \in \mathbb{N}^*$. Un lien (parmi bien d'autres) entre Γ et la fonction ζ de Riemann

$$z \in \Pi_1^+ \longmapsto \sum_{k=1}^{\infty} \frac{1}{k^z}$$

est donné par la relation (valable pour $z \in \Pi_1^+$)

$$\begin{aligned}
\Gamma(z) \times \zeta(z) &= \int_{[0,\infty[} \frac{t^{z-1}}{e^t - 1}\, dt \\
&= \int_{[0,1[} \frac{t^{z-1}}{e^t - 1}\, dt + \int_{[1,\infty[} \frac{t^{z-1}}{e^t - 1}\, dt \\
&= \sum_{k=0}^{\infty} \frac{b_k}{z+k-1} + \int_{[1,\infty[} \frac{t^{z-1}}{e^t - 1}\, dt
\end{aligned}$$

où les nombres $(b_k)_k$ sont les *nombres de Bernoulli* obtenus suivant la division selon les puissances croissantes

$$\frac{1}{\sum\limits_{k=0}^{\infty} X^k/(k+1)!} = 1 + b_1 X + b_2 X^2 + \cdots\,.$$

Cette formule fera l'objet d'une séquence d'exercices dans le guide d'activités 6. C'est elle qui permet de vérifier que Γ ne s'annule pas dans Π_0^+ et est même telle que $z \longmapsto 1/\Gamma(z)$ se prolonge en une fonction analytique complexe dans $\mathbb{C}$ tout entier.

Ceci peut être étendu à un cadre plus général. Supposons que $(\Omega, \mathcal{T}, \mu)$ soit un espace mesuré et X une fonction mesurable de Ω dans $]0,\infty[$ telle qu'il existe $x \in \mathbb{R}$ avec

$$\int_{\Omega} [X(\omega)]^x\, d\mu(\omega) < +\infty\,.$$

19. C'est le mathématicien finlandais Robert Hjalmar Mellin (1854-1933) qui introduit cette transformation et l'étudia, en vue notamment de ses relations étroites avec la fonction ζ de Riemann dont on connaît (au travers de la formule d'Euler) le lien avec le théorème fondamental de l'arithmétique. La transformée de Mellin est devenue aujourd'hui aussi un outil des mathématiques appliquées et des sciences de l'ingénieur (traitement d'image, robotique). Elle est aussi très utilisée en théorie analytique des nombres (plus encore que les transformations de Fourier et Laplace).

On peut définir, inspirés par l'exemple de la transformation de Mellin étudié ci-dessus,

$$\begin{aligned} x_0^+ &:= \sup\Big\{x \in \mathbb{R}\,;\, \int_{X\geq 1} [X(\omega)]^x\, d\mu(\omega) < \infty\Big\} \in]-\infty, \infty] \\ x_0^- &:= \inf\Big\{x \in \mathbb{R}\,;\, \int_{X\in[0,1]} [X(\omega)]^x\, d\mu(\omega) < \infty\Big\} \in [-\infty, \infty[\end{aligned}$$

Si $x_0^- < x_0^+$, on définit une fonction analytique complexe dans la bande verticale

$$\Pi^+_{x_0^-} \cap \Pi^-_{x_0^+} = \{z \in \mathbb{C}\,;\, \mathrm{Re}\, z \in]x_0^-, x_0^+[\}$$

par

$$\mathcal{M}[X]\ : z \in \Pi^+_{x_0^-} \cap \Pi^-_{x_0^+} \longmapsto \int_\Omega [X(\omega)]^z\, d\mu(\omega)\,.$$

Lorsqu'elle peut ainsi être définie dans une bande, la fonction $\mathcal{M}[X]$ est dite *transformée de Mellin* de la fonction mesurable positive (ou encore variable aléatoire positive X). C'est un outil de calcul puissant en théorie des probabilités.

Exemple 3.5. Soit $\Omega = \mathbb{N}^*$ et μ la mesure de décompte sur Ω. Si X est la variable discrète définie par $X(k) = 1/k$, la transformée de Mellin $\mathcal{M}[X]$ est la fonction ζ de Riemann

$$z \in \Pi_1^+ \longmapsto \sum_{k=1}^{\infty} \frac{1}{k^z}\,.$$

C'est encore bien une indication ici que la transformation de Mellin est intimement liée à la fonction ζ de Riemann.

Fonction génératrice d'une variable aléatoire positive discrète

Si $(\Omega, \mathcal{T}, \mu)$ est un espace mesuré et $X\ : \Omega \longmapsto \mathbb{N}$ une fonction mesurable telle qu'il existe un nombre réel positif $r \in \mathbb{R}$ avec

$$\int_\Omega r^{X(\omega)}\, d\mu(\omega) < +\infty\,.$$

Ceci implique, notons le, que les ensembles $\{X = k\}$, $k = 0, 1, \ldots$, sont tous de mesure μ finie. On note alors

$$R := \sup\Big\{r \in]0, \infty[\,;\, \int_\Omega r^{X(\omega)}\, d\mu(\omega) < \infty\Big\}\,.$$

L'intégrale à paramètres

$$z \in D(0, R) \longmapsto \int_\Omega z^{X(\omega)}\, d\mu(\omega) = \sum_{k=0}^{\infty} \mu(\{X = k\})\, z^k$$

est la somme d'une série entière dans son disque de convergence [20]. On l'appelle *fonction génératrice* de la variable aléatoire X. Si $\mu(\Omega) < \infty$, on a toujours $R \geq 1$. La notion de fonction génératrice joue un grand rôle en théorie des probabilités.

20. Bien sûr ici, la fonction est la somme d'une série entière par construction même ! Cependant le recours au théorème 3.4 constitue un moyen de le retrouver (car ce théorème s'applique ici), ce qui n'est pas de fait totalement anodin.

3.3 L'outil « changement de variables » au service de l'intégration sur $\mathbb{R}^n$

3.3.1 La formule de changement de variables sous les intégrales

Soient U et V deux ouverts de $\mathbb{R}^n$ et $\Phi\ : U \longrightarrow V$ un C^1-difféomorphisme entre U et V, *i.e* une application de classe C^1 bijective entre U et V telle que $d_x\Phi$ soit inversible pour tout $x \in U$, ou encore

$$\forall\, x \in U\,,\ \operatorname{Jac}\Phi(x) \neq 0\,,$$

où $\operatorname{Jac}\Phi$ désigne le jacobien [21] de Φ, c'est-à-dire le déterminant de la matrice

$$\Big[\frac{\partial \Phi_i}{\partial x_j}(x)\Big]_{1\leq i,j\leq n}\,.$$

Le fait que $\Phi^{-1}\ : V \longrightarrow U$ soit dans ce cas aussi de classe C^1 résulte du théorème d'inversion locale [22].

Si $f\ : V \longrightarrow \mathbb{C}$ est une fonction mesurable relativement aux tribus boréliennes à la source et au but, il en est de même (puisque Φ est continue) pour $f\circ\Phi\ : U \longrightarrow \mathbb{C}$; réciproquement d'ailleurs, si $h\ : U \longrightarrow \mathbb{C}$ est mesurable relativement aux tribus boréliennes à la source et au but, il en est de même de $h\circ\Phi^{-1}\ : V \longrightarrow \mathbb{C}$. On a même, concernant l'intégrabilité relativement à la mesure de Lebesgue, le très important (et très utile) résultat suivant :

Théorème 3.5 (changement de variables dans les intégrales) *Soient U, V deux ouverts de $\mathbb{R}^n$ et $\Phi\ : U \longrightarrow V$ un C^1-difféomorphisme entre U et V. Soit f une fonction de V dans $\mathbb{C}$. Dire que f est $(V,\mathcal{B})$-$(\mathbb{C},\mathcal{B})$ mesurable et de plus intégrable sur V relativement à la mesure de Lebesgue équivaut à dire que $f\circ\Phi\ : U \longrightarrow \mathbb{C}$ est $(U,\mathcal{B})$-$(\mathbb{C},\mathcal{B})$ mesurable et que la fonction*

$$g\ : x \in U \longmapsto f(\Phi(x)) \times |\operatorname{Jac}\Phi(x)|$$

est intégrable sur U relativement à la mesure de Lebesgue. De plus, on a la formule

$$\int_V f(y)\,dy = \int_U f(\Phi(x))\,|\operatorname{Jac}\Phi(x)|\,dx\,. \tag{3.11}$$

Enfin, la formule (3.11) est valable pour toute fonction $f\ : V \longrightarrow [0,\infty]$, pourvu qu'elle soit $(V,\mathcal{B})$-$([0,\infty],\mathcal{B})$ mesurable, l'égalité (3.11) étant dans ce cas entendue comme une égalité entre éléments de $[0,\infty]$.

Remarque 3.6. Il faut prendre garde (dans le cas $n = 1$) à ne pas confondre cet énoncé avec le résultat bien connu [23] suivant : si $\varphi\ : I \longrightarrow J$ est une application de classe C^1 d'un intervalle ouvert I de $\mathbb{R}$ dans un intervalle ouvert J, f une fonction continue sur J à valeurs complexes, F une primitive de f sur J, alors la fonction $F\circ\varphi\ : I \longmapsto \mathbb{C}$ est une primitive sur I de

$$t \longmapsto f(\varphi(t))\,\varphi'(t)\,,$$

21. C'est au mathématicien allemand Carl Gustav Jacobi (1804-1851), dont les travaux tant en algèbre qu'en géométrie ont fait une place importante à la notion de *déterminant* en algèbre linéaire, que fait référence cette terminologie.

22. Voir le cours de calcul différentiel (UE MHT513).

23. Voir le cours de MHT302 en L1.

ce qui induit, si a et b sont deux points de I, les égalités

$$\int_{\varphi(a)}^{\varphi(b)} f(u)\,du = \int_a^b f(\varphi(t))\,\varphi'(t)\,dt$$

entre intégrales calculées ici au sens de Riemann. On ne retrouve pas ici $|\varphi'(t)|\,dt$ comme dans le théorème 3.5 mais $\varphi'(t)\,dt$ (sans la valeur absolue). Ce résultat est un résultat sur les primitives, non un résultat sur l'intégration ; d'ailleurs aucune hypothèse du type C^1-difféomorphisme n'est faite (lorsque l'on énonce un tel résultat) sur φ.

Preuve. Le fait que la mesurabilité de f soit équivalente à celle de $f \circ \Phi$ résulte de la continuité de Φ et Φ^{-1}. Pour montrer le résultat concernant l'équivalence entre l'intégrabilité de $f \ : V \longrightarrow \mathbb{C}$ relativement à la mesure de Lebesgue et celle de $g \ : U \longrightarrow \mathbb{C}$, on peut se ramener à prouver la formule (3.11) lorsque $f \ : V \longrightarrow [0, \infty]$: il suffit en effet d'écrire ensuite

$$f = ((\operatorname{Re} f)^+ - (\operatorname{Re} f)^-) + i((\operatorname{Im} f)^+ - (\operatorname{Im} f)^-)\,,$$

d'où

$$f \circ \Phi = ((\operatorname{Re} f \circ \Phi)^+ - (\operatorname{Re} f \circ \Phi)^-) + i((\operatorname{Im} f \circ \Phi)^+ - (\operatorname{Im} f \circ \Phi)^-)$$

et d'appliquer la formule (3.11) à $(\operatorname{Re} f)^\pm$ et $(\operatorname{Im} f)^\pm$ pour conclure au résultat final. Grâce au théorème de convergence monotone de Beppo-Levi (théorème 2.1), couplé avec la proposition 2.2, on peut se ramener à prouver la formule de changement de variables (3.11) pour une fonction étagée mesurable $f \ : V \longrightarrow [0, \infty]$, ou mieux, par linéarité de la prise d'intégrale sur les fonctions étagées mesurables, pour une fonction f du type $f = \chi_A$, où A est un sous-ensemble mesurable de V. On peut même supposer A relativement compact dans V (donc intégrable), ce en approchant de manière croissante χ_A par la suite $(\chi_{A \cap K_l})_l$, où $(K_l)_l$ est une suite croissante de compacts réalisant une exhaustion de l'ouvert V (il est facile de construire une telle suite en écrivant par exemple V comme union dénombrable de pavés disjoints). Il nous reste donc à prouver, pour un tel sous-ensemble A intégrable de V (tel que $\overline{A}$ soit un compact de V),

$$\int_A dy = m(A) = \int_{\Phi^{-1}(A)} |\operatorname{Jac} \Phi(x)|\,dx\,.$$

En utilisant le critère d'intégrabilité de la proposition 1.4 (pour tout $\epsilon > 0$, il existe un ouvert U_ϵ et un compact K_ϵ tels que $K_\epsilon \subset A \subset U_\epsilon$ et que $m(U_\epsilon \setminus K_\epsilon) < \epsilon$), on voit qu'il suffit de prouver la formule (3.11) pour $f = \chi_W$, où W est un ouvert de V. Nous aurons besoin du lemme suivant :

Lemme 3.1 *Soit P un pavé borné de $\mathbb{R}^n$ et Φ un C^1-difféomorphisme entre un voisinage ouvert de $\overline{P}$ et un voisinage ouvert de $\Phi(\overline{P})$. On a, si m désigne la mesure de Lebesgue :*

$$m(\Phi(P)) = \int_P |\operatorname{Jac} \Phi(x)|\,dx\,. \tag{3.12}$$

Admettons pour l'instant ce lemme et écrivons l'ouvert $\Phi^{-1}(W)$ comme union dénombrable (disjointe) de pavés bornés $\widetilde{P}_k$, $k \in \mathbb{N}$, inclus dans U. Si P_k est un pavé inclus dans $\widetilde{P}_k$ et relativement compact dans $\widetilde{P}_k$, on a, d'après le lemme 3.1,

$$m(\Phi(P_k)) = \int_{P_k} |\operatorname{Jac} \Phi(x)|\,dx\,.$$

En ajoutant pour $k \in \mathbb{N}$, il vient

$$\sum_k m(\Phi(P_k)) = \sum_k \int_{P_k} |\mathrm{Jac}\,\Phi(x)|\, dx\,.$$

En approchant (de l'intérieur) chaque $\widetilde{P}_k$ par une suite de pavés $(P_{k,l})_l$ emboîtés et tous relativement compacts dans $\widetilde{P}_k$ et en utilisant une fois encore le théorème de convergence monotone (théorème 2.1), il vient

$$\sum_k m(\Phi(\widetilde{P}_k)) = m(W) = \sum_k \int_{\widetilde{P}_k} |\mathrm{Jac}\,\Phi(x)|\, dx = \int_{\Phi^{-1}(W)} |\mathrm{Jac}\,\Phi(x)|\, dx\,,$$

ce qui est l'égalité (3.11) pour la fonction $f = \chi_W$. Si l'on admet le lemme 3.1, le théorème 3.3 est donc démontré. $\diamondsuit$

Preuve du lemme 3.1. Si le résultat est prouvé pour tout pavé P ouvert, on remarque que si P est un pavé « plat » (c'est-à-dire inclus dans un hyperplan $x_j =$ constante) et de plus compact, les deux membres de (3.12) sont nuls (on approche P par des pavés ouverts dont le volume n-dimensionnel tend vers 0). La formule (3.12) étant vraie pour tout pavé ouvert et pour tout pavé « plat » (les deux membres étant nuls dans ce cas), elle est vraie pour tout pavé P. Il suffit donc de démontrer le lemme pour un pavé ouvert P.

Soit x_0, x deux points de P ; on a, d'après l'inégalité des accroissements finis

$$\|\Phi(x) - \Phi(x_0) - d_{x_0}\Phi\,(x - x_0)\| \le \|x - x_0\| \sup_{\xi \in [x_0, x]} \|d_\xi\Phi - d_{x_0}\Phi\|\,.$$

Soit $\eta > 0$. En utilisant l'uniforme continuité de la fonction $x \longmapsto d_x\Phi$ sur $\overline{P}$, on constate que l'on peut, pour N assez grand (dépendant de η), découper P en N^n pavés P_k d'intérieurs disjoints (tous de volume n-dimensionnel $m(P)/N^n$, à faces parallèles aux plans de coordonnées, et translatés d'un même pavé) de manière à ce que, si x_k est le centre de P_k et $L_k\ : x \longmapsto \Phi(x_k) + d_{x_k}\Phi\,(x - x_k)$, alors, lorsque $P_k^{1\pm\eta}$ est le pavé obtenu à partir de P_k par homothétie de centre x_k et de rapport $1 \pm \eta$, on ait, pour tout k

$$\begin{aligned} (L_k^{-1} \circ \Phi)(P_k) &\subset P_k^{1+\eta} \\ (\Phi^{-1} \circ L_k)(P_k^{1-\eta}) &\subset P_k \end{aligned}$$

ou encore

$$\begin{aligned} \Phi(P_k) &\subset L_k(P_k^{1+\eta}) \\ L_k(P_k^{1-\eta}) &\subset \Phi(P_k)\,. \end{aligned}$$

Il ne reste plus qu'à remarquer que si L est une application affine de $\mathbb{R}^n$ dans lui-même et P un pavé ouvert borné, la mesure de Lebesgue de $L(P)$ est égale à la mesure de P multipliée par la valeur absolue du déterminant de la matrice de l'application linéaire l sous-jacente à L : ceci résulte du fait que le volume n-dimensionnel du parallélépipède

$$\{t_1 v_1 + \cdots + t_n v_n\,;\, 0 \le t_j \le 1\,,\ j = 1, ..., n\}$$

(lorsque $v_1, ..., v_n$ forment une base de $\mathbb{R}^n$) est égal à $|\det(v_1, ..., v_n)|$ [24]. Comme η peut être choisi arbitrairement petit, on déduit des inégalités

$$m(\Phi(P)) = \sum_k m(\Phi(P_k)) \leq \sum_k |\det l_k| \times m(P_k^{1+\eta})$$
$$\sum_k |\det l_k| \times m(P_k^{1-\eta}) \leq \sum_k m(\Phi(P_k)) = m(\Phi(P))$$

la relation (3.12) en choisissant précisément η (et donc ensuite $N \geq N_\eta$ assez grand) arbitrairement petit et en utilisant le fait que $x \longmapsto \|d_x\Phi\|$ est bornée dans le pavé $\overline{P}$ [25]. $\diamondsuit$

Exemple 3.6. Voici un exemple d'application (non complètement immédiate) de la formule de changement de variables (3.11). Soit z un nombre complexe tel que $\mathrm{Re}\, z \in]0,1[$, $U =]0,\infty[\times]O,\infty[$ et

$$f(u,v) = u^{z-1}v^{-z}e^{-u}e^{-v}\,.$$

Cette fonction est intégrable sur l'ouvert V. On peut considérer $U = V$ et

$$\Phi\ : (u,v) \in U \longrightarrow (u/v, u+v)\,;$$

on vérifie que Φ est bien un C^1-difféomorphisme entre U et V. On a

$$\Phi^{-1}(t,s) = \Big(\frac{ts}{t+1}\,,\,\frac{s}{t+1}\Big)\,.$$

On vérifie que

$$\mathrm{Jac}\,\Phi(u,v) = \frac{u+v}{v^2} = u \times \frac{v^2}{u(u+v)} = u \times \frac{1}{t(1+t)}\,.$$

En utilisant la formule de changement de variable (3.11), on a voit que la fonction

$$(t,s) \in V \longmapsto t^z \frac{e^{-s}}{t(t+1)}$$

est intégrable sur V et que

$$\int_V u^{z-1}v^{-z}\, e^{-u}\, e^{-v}\, dudv = \int_V t^{z-1}\frac{e^{-s}}{t+1}\, dtds = \int_{]0,\infty[} \frac{t^{z-1}}{1+t}\, dt\,.$$

Cette dernière intégrale peut se calculer *via* la formule des résidus [26] et l'on obtient ainsi la *formule des compléments*, importante formule satisfaite par la fonction Γ d'Euler :

$$\mathrm{Re}\, z \in]0,1[\Longrightarrow \Gamma(z)\times\Gamma(1-z) = \Big(\int_{]0,\infty[} u^{z-1}\, e^{-u}\, du\Big) \times \Big(\int_{]0,\infty[} v^{-z}\, e^{-v}\, dv\Big) = \frac{\pi}{\sin(\pi z)}\,.$$

Exemple 3.7 (coordonnées polaires dans $\mathbb{R}^2$). Soit $\theta_0 \in \mathbb{R}$ et V un ouvert du plan complexe privé d'une demi-droite $S_{\theta_0} = \mathbb{R}^2 \setminus (e^{i\theta_0} \times [0,+\infty[)$ (il s'agit d'un cône ouvert d'ouverture 2π de sommet l'origine). L'ouvert V est C^1-difféomorphe à un ouvert U de $]0,\infty[\times]\theta_0,\theta_0+2\pi[$ *via* l'application

$$(r,\theta) \longmapsto (r\cos\theta, r\sin\theta)\,.$$

24. Voir le cours de géométrie affine de L2.

25. Le calcul de l'intégrale à droite de (3.12), intégrale sur le pavé P (ou si l'on préfère $\overline{P}$, le résultat étant le même) de la fonction continue de n variables $|\mathrm{Jac}\,\Phi|$, est ici conduit comme un calcul d'intégrale de Riemann *via* les sommes de Darboux majorantes et les sommes de Darboux minorantes.

26. Voir la section 5.3.6 du cours de MAT401, *c.f* le polycopié en ligne : http ://www.math.u-bordeaux1.fr/~yger/mat401.pdf

Comme $\operatorname{Jac}\Phi(r,\theta) = r$, on peut affirmer, si f est une fonction mesurable et intégrable dans V, que la fonction $f\circ\Phi$ est intégrable dans $U = \Phi^{-1}(V)$, avec de plus

$$\int_V f(x,y)dx\,dy = \int_{\Phi^{-1}(V)} f(r\cos\theta, r\sin\theta)\,rdr\,d\theta\,.$$

C'est la classique formule de passage en *coordonnées polaires.*

Exemple 3.8 (coordonnées sphériques dans $\mathbb{R}^3$). Soit $\theta_0\in\mathbb{R}$ et V un ouvert de $\mathbb{R}^3$ privé du demi-plan vertical

$$\Sigma_{\theta_0} := \Big\{(x,y,z)\in\mathbb{R}^3\,;\,-x\sin\theta_0 + y\cos\theta_0 = 0\,,\ x\cos\theta_0 + y\sin\theta_0 \geq 0\Big\}\,. \tag{3.13}$$

L'ouvert V est C^1-difféomorphe à un ouvert U de $]0,\infty[\times]\theta_0,\theta_0+2\pi[\times]0,\pi[$ *via* l'application

$$(r,\theta,\varphi)\longmapsto(r\cos\theta\sin\varphi, r\sin\theta\sin\varphi, r\cos\varphi)$$

(θ est l'angle représentant la *longitude*, φ la *colatitude* mesurée depuis le pôle Nord $(0,0,1)$ de la sphère unité, ces deux angles sont dits *angles d'Euler*). On a ici

$$|\operatorname{Jac}\Phi(r,\theta,\varphi)| = r^2\sin\varphi$$

comme on le voit par un calcul aisé. On peut donc affirmer, si f est une fonction mesurable et intégrable dans V, que la fonction $f\circ\Phi$ est intégrable dans $U=\Phi^{-1}(V)$, avec de plus

$$\int_V f(x,y,z)dx\,dy\,dz = \int_{\Phi^{-1}(V)} f(r\cos\theta\sin\varphi, r\sin\theta\sin\varphi, r\cos\varphi)\,r^2\sin\varphi\,dr\,d\theta\,d\varphi\,.$$

C'est la classique formule de passage en *coordonnées sphériques.*

Exemple 3.9 (coordonnées cylindriques dans $\mathbb{R}^3$). Soit $\theta_0\in\mathbb{R}$ et V un ouvert de $\mathbb{R}^3$ privé du plan vertical Σ_{θ_0} défini, comme dans l'exemple 3.8, par (3.13). L'ouvert V est C^1-difféomorphe à un ouvert U de $]0,\infty[\times]\theta_0,\theta_0+2\pi[\times\mathbb{R}$ *via* l'application

$$(r,\theta,z)\longmapsto(r\cos\theta, r\sin\theta, z)\,.$$

On a ici

$$|\operatorname{Jac}\Phi(r,\theta,z)| = r\,.$$

On peut donc affirmer, si f est une fonction mesurable et intégrable dans V, que la fonction $f\circ\Phi$ est intégrable dans $U=\Phi^{-1}(V)$, avec de plus

$$\int_V f(x,y,z)dx\,dy\,dz = \int_{\Phi^{-1}(V)} f(r\cos\theta, r\sin\theta, z)\,r\ dr\,d\theta\,dz\,.$$

C'est la classique formule de passage en *coordonnées cylindriques.*

3.3.2 Un résultat important à la croisée de l'intégration et du calcul différentiel : le lemme de Sard

Outre la formule de changement de variables, un résultat d'un intérêt pratique très important fait se croiser théorie de l'intégration et calcul différentiel. C'est un résultat récent (1942) dû au mathématicien américain Arthur Sard.

Soit U un ouvert de $\mathbb{R}^n$ et Φ une application de classe C^1 de U dans $\mathbb{R}^p$. Les *points critiques* de Φ sont par définition les points x de U tels que rang $[d_x\Phi] < p$. Les *valeurs critiques* de Φ sont les images par Φ des points critiques. Dire que $y_0\in\mathbb{R}^p$ n'est pas une valeur critique autorise donc à pouvoir décrire, étant donné un point $x_0\in U$ tel que $\Phi(x_0)=y_0$, l'ensemble $\{x\,;\,\Phi(x)=y_0\}$ de manière précise au voisinage de x_0, ce en fonction de $n-p$ paramètres (théorème des fonctions implicites [27]) lorsque $p<n$

27. Voir le cours de calcul différentiel de l'UE MHT513.

ou à redresser $x \longmapsto \Phi(x) - y_0$ en un système de coordonnées locales au voisinage de x_0 (théorème d'inversion locale [28]) lorsque $p = n$. De fait, presque tous les points y de $\mathbb{R}^p$ sont des valeurs non critiques pour Φ (étant entendu que y est considéré comme non critique si $\Phi^{-1}(\{y\}) = \emptyset$) ; on a en effet le :

Théorème 3.6 (théorème de Sard) *L'ensemble des valeurs critiques d'une application Φ de classe C^1 dans un ouvert U de $\mathbb{R}^n$, à valeurs dans $\mathbb{R}^p$ $(n \geq p)$ [29] est un sous-ensemble négligeable de $\mathbb{R}^p$ relativement à la mesure de Lebesgue dans $\mathbb{R}^p$.*

Preuve. Nous ne donnerons cette preuve que dans le cas où $n \geq p$. Quitte a compléter Φ en une application (Φ, L), où L est l'application nulle de U dans $\mathbb{R}^{n-p}$, prouver le théorème revient à le prouver lorsque $p = n$. Comme U est union dénombrable de pavés bornés ouverts (pavés que l'on peut exhauster de l'intérieur par des pavés bornés fermés), on peut se limiter à supposer que Φ est de classe C^1 au voisinage de $[0,1]^n$, à valeurs dans $\mathbb{R}^n$, et à prouver alors, ce que nous allons faire, que l'ensemble

$$\Phi\Big(\{x \in [0,1]^n\,;\, \text{rang}\,(d_x\Phi) < n\}\Big)$$

est un sous-ensemble négligeable (relativement à la mesure de Lebesgue) de $\mathbb{R}^n$.

D'après l'inégalité des accroissements finis [30] et le fait que $x \longmapsto d_x\Phi$ soit continue, donc uniformément continue sur le compact $[0,1]^n$, il existe une constante $C > 0$ telle que

$$\|\Phi(x) - \Phi(x')\| \leq C\|x - x'\|$$

pour tout (x, x') dans $[0,1]^n$ et de plus

$$\lim_{\substack{\|x-x'\|\to 0 \\ x,x'\in[0,1]^n}} \frac{\|\Phi(x') - \Phi(x) - d_x\Phi(x'-x)\|}{\|x'-x\|} = 0\,. \tag{3.14}$$

Notons crit_Φ l'ensemble des points a de $[0,1]^n$ où $\text{rang}\,(d_a\Phi) < n$. Fixons $\eta > 0$ (arbitrairement petit). Choisissons (en exploitant (3.14)) $\epsilon_\eta > 0$ tel que $\|x - x'\| < \epsilon_\eta$ implique

$$\|\Phi(x') - \Phi(x) - d_x\Phi(x'-x)\| \leq \eta\|x - x'\| \leq \epsilon_\eta \times \eta\,.$$

Si $a \in \text{crit}_\Phi$ et $\|x' - a\| \leq \epsilon_\eta$, le point $\Phi(x')$ est proche du sous-ensemble

$$\Phi(a) + d_a\Phi(\mathbb{R}^n)$$

qui, puisque a est dans crit_Φ, est inclus dans un hyperplan Π_a de $\mathbb{R}^n$; les points $\Phi(x')$ tels que $\|x' - a\| \leq \epsilon_\eta$ sont donc dans un parallélépipède contenant $\Phi(a)$, dont $n-1$ côtés (ceux parallèles à Π_a) sont de longueur au plus $2C\epsilon_\eta\sqrt{n}$, tandis que le n-ème côté (orthogonal à Π_a) est quant à lui de longueur au plus $2\eta\epsilon_\eta$; le volume n-dimensionnel d'un tel parallélépipède contenant tous les points $\Phi(x')$ tels que $\|x' - a\| \leq \epsilon_\eta$, avec $a \in \text{crit}_\Phi$, est donc au plus égal à $2\eta\epsilon_\eta(2C\epsilon_\eta\sqrt{n})^{n-1}$.

Découpons maintenant $[0,1]^n$ en N^n hypercubes ayant pour longueur d'arête $1/N$, avec $N > 1/\epsilon_\eta$. Tous les points de crit_Φ appartenant à un même pavé P du pavage

28. Voir le cours de calcul différentiel de l'UE MHT513.

29. Le résultat est en fait vrai dans tous les cas, sans la restriction $(n \geq p)$ sous laquelle on le démontre ici.

30. Voir encore le cours de calcul différentiel de l'UE MHT513.

étant tous à une distance au plus égale à ϵ_η de l'un d'entre eux, on déduit de ce qui précède

$$m[\Phi(P \cap \mathrm{crit}_\Phi)] \leq 2^n n^{\frac{n-1}{2}} C^{n-1} (1/N)^n \eta\,.$$

En ajoutant ces inégalités pour les N^n pavés P du pavage de $[0,1]^n$, on en déduit

$$m[\Phi(\mathrm{crit}_\Phi)] \leq 2^n 2^n n^{\frac{n-1}{2}} C^{n-1}\, \eta\,.$$

Comme η était arbitraire, la mesure de Lebesgue de $\Phi(\mathrm{crit}_\Phi)$ est bien nulle et le théorème est démontré [31]. $\diamondsuit$

3.4 Produit de mesures ; les théorèmes de Fubini-Tonelli et Fubini

3.4.1 Tribu produit, mesure produit, le théorème de Fubini-Tonelli

Soient Ω_1 et Ω_2 deux ensembles, équipés respectivement de tribus $\mathcal{T}_1$ et $\mathcal{T}_2$. Suivant l'exemple 1.3, la *tribu produit* $\mathcal{T}_1 \otimes \mathcal{T}_2$ sur l'ensemble produit $\Omega_1 \times \Omega_2$ est la plus petite tribu contenant toutes les unions finies d'ensembles du type $A \times B$, où A est un élément de $\mathcal{T}_1$ et B un élément de $\mathcal{T}_2$.

Avant de définir une mesure sur $\mathcal{T}_1 \otimes \mathcal{T}_2$ à partir de la donnée d'une mesure $\mu_1 : \mathcal{T}_1 \longrightarrow [0,\infty]$ et d'une mesure $\mu_2 : \mathcal{T}_2 \longrightarrow [0,\infty]$, nous aurons besoin d'une clause restrictive portant sur les mesures μ_1 et μ_2.

Définition 3.1 *Soit $(\Omega, \mathcal{T}, \mu)$ un espace mesuré. La mesure μ est dite σ-finie s'il existe une suite $(A_k)_{k\in\mathbb{N}}$ d'éléments de $\mathcal{T}$ (que l'on peut supposer croissante) telle que $\mu(A_k) < \infty$ pour tout k et que $\Omega = \bigcup_k A_k$.*

Exemple 3.10. La mesure de Lebesgue sur $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ ou $(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n))$ est une mesure σ-finie. En revanche, la mesure de décompte sur la tribu de toutes les parties d'un ensemble non fini ou dénombrable ne saurait être σ-finie.

Le résultat majeur de cette section est la

Proposition 3.1 *Soient $(\Omega_1, \mathcal{T}_1, \mu_1)$ et $(\Omega_2, \mathcal{T}_2, \mu_2)$ deux espaces mesurés, les mesures μ_1 et μ_2 étant toutes les deux σ-finies. Il existe alors une unique mesure $\mu_1 \otimes \mu_2 : \mathcal{T}_1 \otimes \mathcal{T}_2 \to [0,\infty]$ telle que, pour tout $A \in \mathcal{T}_1$, pour tout $B \in \mathcal{T}_2$, on ait*

$$(\mu_1 \otimes \mu_2)(A \times B) = \mu_1(A) \times \mu_2(B)$$

(toujours avec la convention $0 \times \infty = 0$). L'unique mesure positive sur la tribu produit ainsi définie est dite mesure produit de μ_1 et μ_2. Plus généralement, si E désigne un élément quelconque de $\mathcal{T}_1 \otimes \mathcal{T}_2$, les fonctions

$$\begin{array}{lcll} x \in \Omega_1 & \longmapsto & \mu_2(E_x)\,, & E_x := \{y \in \Omega_2\,;\, (x,y) \in E\} \\ y \in \Omega_2 & \longmapsto & \mu_1(E^y)\,, & E^y := \{x \in \Omega_1\,;\, (x,y) \in E\} \end{array}$$

31. Remarquez l'analogie entre les ingrédients utilisés ici et ceux utilisés dans la preuve de la formule de changement de variables pour une fonction caractéristique de pavé (formule (3.12) du lemme 3.1 ; on a d'ailleurs pris soin d'harmoniser les notations (avec η et $\epsilon = \epsilon_\eta$).

sont respectivement $(\Omega_1, \mathcal{T}_1)$-$([0,\infty], \mathcal{B})$ *et* $(\Omega_2, \mathcal{T}_2)$-$([0,\infty], \mathcal{B})$ *mesurables et on a*

$$(\mu_1 \otimes \mu_2)(E) = \int_{\Omega_1} \mu_2(E_x)\, d\mu_1(x) = \int_{\Omega_2} \mu_1(E^y)\, d\mu_2(y)\,. \tag{3.15}$$

Remarque 3.7. La clause de σ-finitude portant sur les mesures μ_1 et μ_2 est essentielle pour assurer la validité de cette proposition majeure : si par exemple $\Omega_1 = \Omega_2 = [0,1]$, $\mathcal{T}_1$ la tribu borélienne, μ_1 la mesure de Lebesgue (σ-finie), $\mathcal{T}_2 = \mathcal{P}([0,1])$, μ_2 la mesure de décompte (non σ-finie, voir l'exemple 3.9) et si

$$E := \{(x,x)\,;\, x \in [0,1]\}\,,$$

on aurait, si la proposition était valide,

$$(\mu_1 \otimes \mu_2)(E) = \int_{[0,1]} \mu_2(E_x)\, d\mu_1(x) = 1 = \int_{[0,1]} \mu_1(E^y)\, d\mu_2(y) = 0$$

(puisque E^y est négligeable relativement à la mesure de Lebesgue μ_1 sur $[0,1]$ et que E_x, de cardinal 1, est de mesure 1 relativement à la mesure de décompte μ_2 sur $[0,1]$), ce qui est absurde.

Remarque 3.8. Le fait que E_x (*resp.* E^y) soit un élément de $\mathcal{T}_2$ (*resp.* de $\mathcal{T}_1$) pour tout $x \in \Omega_1$ (*resp.* pour tout $y \in \Omega_2$) résulte de la mesurabilité des projections $\pi_j\ : \Omega_1 \times \Omega_2 \longrightarrow \Omega_j$ relativement à la tribu produit $\mathcal{T}_1 \otimes \mathcal{T}_2$ à la source et aux tribus respectives $\mathcal{T}_j$ au but.

Exemple 3.11. Un exemple capital d'application de la proposition 3.1 est celui où $\Omega_1 = \mathbb{R}^{n_1}$, $\Omega_2 = \mathbb{R}^{n_2}$ et où μ_1 et μ_2 sont respectivement les mesures de Lebesgue dans $\mathbb{R}^{n_1}$ et $\mathbb{R}^{n_2}$. Du fait du théorème 1.1 caractérisant la mesure de Lebesgue, on voit que la mesure $\mu_1 \otimes \mu_2$ sur la tribu produit $\mathcal{B}(\mathbb{R}^{n_1}) \otimes \mathcal{B}(\mathbb{R}^{n_2})$ (qui dans ce cas est la tribu borélienne sur $\mathbb{R}^{n_1+n_2} \simeq \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$) est la mesure de Lebesgue sur $\mathbb{R}^{n_1+n_2}$.

Avant de prouver la proposition 3.1, nous en déduisons immédiatement (en utilisant la proposition 2.2 et le théorème 2.1 de convergence monotone de Beppo Levi), son corollaire le plus important du point de vue pratique, à savoir le résultat suivant (1907)[32] attribué aux mathématiciens italiens Guido Fubini (1879-1943) et Leonida Tonelli (1885-1946) :

Théorème 3.7 (théorème de Fubini-Tonelli) *Soient* $(\Omega_j, \mathcal{T}_j, \mu_j)$, $j = 1,2$, *deux espaces mesurés, les deux mesures* μ_1 *et* μ_2 *étant* σ*-finies,* $\mathcal{T}_1 \otimes \mathcal{T}_2$ *la tribu produit et* $\mu_1 \otimes \mu_2\ : \mathcal{T}_1 \otimes \mathcal{T}_2 \longrightarrow [0,\infty]$ *la mesure produit définie à la proposition 3.1. Pour toute fonction* f $(\Omega_1 \times \Omega_2, \mathcal{T}_1 \otimes \mathcal{T}_2)$-$([0,\infty], \mathcal{B})$ *mesurable, les fonctions*

$$\begin{aligned} x \in \Omega_1 \quad &\longmapsto \quad \int_{\Omega_2} f(x,y)\, d\mu_2(y) \\ y \in \Omega_2 \quad &\longmapsto \quad \int_{\Omega_1} f(x,y)\, d\mu_1(x) \end{aligned}$$

sont respectivement $(\Omega_1, \mathcal{T}_1)$-$([0,\infty], \mathcal{B})$ *et* $(\Omega_2, \mathcal{T}_2)$-$([0,\infty], \mathcal{B})$ *mesurables et on a*

$$\begin{aligned} \int_{\Omega_1 \times \Omega_2} f(x,y)\, d[\mu_1 \otimes \mu_2](x,y) \quad &= \quad \int_{\Omega_1} \Big[\int_{\Omega_2} f(x,y)\, d\mu_2(y) \Big]\, d\mu_1(x) \\ &= \quad \int_{\Omega_2} \Big[\int_{\Omega_1} f(x,y)\, d\mu_1(x) \Big]\, d\mu_2(y)\,. \end{aligned} \tag{3.16}$$

32. Il prélude au théorème de Fubini (pour les fonctions intégrables à valeurs complexes et non plus positives) que nous énoncerons dans la section suivante (section 3.4.2) ; notons au passage que ne figure aucune hypothèse d'intégrabilité portant sur la fonction f dans l'énoncé du théorème de Fubini-Tonelli, les trois termes de l'égalité (3.16) pouvant fort bien tous les trois valoir $+\infty$ (si l'un est fini, il le sont tous et sont égaux d'après précisément la formule).

Remarque 3.9. La mesurabilité de f implique (du fait de la remarque 3.8 et de la proposition 2.2) la mesurabilité des fonctions

$$f_x \ : y \in \Omega_2 \longmapsto f(x,y)\,, \ x \in \Omega_1$$

(de Ω_2 dans $[0,\infty]$ avec les tribus $\mathcal{T}_2$ et $\mathcal{B}$ respectivement à la source et au but) et des fonctions

$$f^y \ : x \in \Omega_1 \longmapsto f(x,y)\,, \ y \in \Omega_2$$

(de Ω_1 dans $[0,\infty]$ avec les tribus $\mathcal{T}_1$ et $\mathcal{B}$ respectivement à la source et au but).

Exemple 3.12. L'exemple (voir l'exemple 3.11 ci-dessus) où $\Omega_j = \mathbb{R}^{n_j}$, $j = 1,2$, et où μ_j est la mesure de Lebesgue dans $\mathbb{R}^{n_j}$, les tribus étant les tribus boréliennes ou les tribus de Lebesgue, est certainement un des exemples les plus importants d'application du théorème de Fubini-Tonelli.

Exemple 3.13. Dans le cas particulier où $\Omega_2 = \mathbb{N}$ et μ_2 est la mesure de décompte, on retrouve avec l'énoncé du théorème de Fubini-Tonelli, la formule (2.14) d'interversion entre sommation et prise d'intégrale lorsque sont en jeu des fonctions $(x,n) \in \Omega \times \mathbb{N} \longmapsto u_n(x)$ positives (exemple 2.1 d'application du théorème de Beppo Levi 2.1). Notons que dans ce cas particulier, il n'est pas nécessaire que $\mu_1 = \mu$ soit σ-finie.

Preuve de la proposition 3.1. La preuve de cette proposition repose sur un lemme technique préliminaire que nous admettrons en première lecture et dont nous renvoyons à la preuve en fin de sous-section :

Lemme 3.2 *Soit $\mathcal{A}$ une algèbre de Boole sur un ensemble Ω. La tribu $\mathcal{T}(\mathcal{A})$ est la plus petite famille de parties de Ω contenant $\mathcal{A}$ et telle que toute union croissante ou intersection décroissante d'éléments de la famille y appartienne encore*[33].

On prouve la proposition 3.1 en séparant les deux volets « unicité » et « existence ».

Unicité de la mesure $\mu_1 \otimes \mu_2$. Supposons que ν et $\widetilde{\nu}$ soient deux mesures sur $\mathcal{T}_1 \otimes \mathcal{T}_2$ telles que $\nu(A \times B) = \widetilde{\nu}(A \times B) = \mu_1(A)\mu_2(B)$ pour tout $A \in \mathcal{T}_1$, pour tout $B \in \mathcal{T}_2$. Les deux mesures ν et $\widetilde{\nu}$ sont σ-finies : si $(A_k)_k$ est une suite d'éléments de $\mathcal{T}_1$, tous de mesure μ_1 finie, exhaustant Ω_1, et $(B_k)_k$ une suite d'éléments de $\mathcal{T}_1$, tous de mesure μ_2 finie, exhaustant Ω_2, les ensembles $A_k \times B_k$, $k \in \mathbb{N}$, exhaustent $\Omega_1 \times \Omega_2$ et sont de mesure ν et $\widetilde{\nu}$ finies puisque

$$\nu(A_k \times B_k) = \widetilde{\nu}(A_k \times B_k) = \mu_1(A_k) \times \mu_2(B_k) < \infty\,.$$

Les deux mesures ν et $\widetilde{\nu}$ coïncident sur l'algèbre de Boole[34] $\mathcal{A}$ constituée des unions finies d'ensembles du type $A \times B$, $A \in \mathcal{T}_1$, $B \in \mathcal{T}_2$. Si

$$\begin{aligned} \mathcal{A}^{(k)} &:= \mathcal{A}_{A_k \times B_k} \\ \mathcal{T}^{(k)} &:= (\mathcal{T}_1 \otimes \mathcal{T}_2)_{A_k \times B_k} \end{aligned}$$

désignent respectivement les traces sur $A_k \times B_k$ de l'algèbre de Boole $\mathcal{A}$ et de la tribu engendrée $\mathcal{T}_1 \otimes \mathcal{T}_2 = \mathcal{T}(\mathcal{A})$[35], on constate que ν et $\widetilde{\nu}$ coïncident sur $\mathcal{T}^{(k)}$ car elles coïncident sur $\mathcal{A}^{(k)}$ et que, d'après le lemme 3.2, $\mathcal{T}^{(k)}$ est la plus petite classe monotone contenant $\mathcal{A}^{(k)}$; en effet, l'ensemble des parties de $A_k \times B_k$ appartenant

33. Une famille de parties ayant la propriété d'être stable sous l'opération de prise d'union croissante et d'intersection décroissante est dite *classe monotone*. Le lemme 3.2 s'énonce donc en disant que la tribu engendrée par $\mathcal{A}$ est la plus petite *classe monotone* contenant $\mathcal{A}$.

34. Voir l'exemple 1.3.

35. En fait $\mathcal{T}^{(k)}$ est la tribu produit de la tribu trace de $\mathcal{T}_1$ sur A_k par la tribu trace de $\mathcal{T}_2$ sur B_k, la même remarque valant pour $\mathcal{A}^{(k)}$ si le mot « tribu » est remplacé par « algèbre de Boole ».

à $\mathcal{T}^{(k)}$ et ayant mêmes mesures ν et $\widetilde{\nu}$ est une classe monotone du fait des clauses (1.7) et (1.8) (on a $\mu_1(A_k)$ et $\mu_2(B_k)$ finies) de la proposition 1.1, couplé avec le fait que ν et $\widetilde{\nu}$ coïncident sur l'algèbre de Boole $\mathcal{A}^{(k)}$. En exhaustant $\Omega_1 \times \Omega_2$ avec les $A_k \times B_k$ (la suite étant supposée croissante), on constate que pour tout E de $\mathcal{T}_1 \otimes \mathcal{T}_2$,

$$\nu(E) = \lim_{k\to+\infty} \nu(E \cap (A_k \times B_k)) = \lim_{k\to+\infty} \widetilde{\nu}(E \cap (A_k \times B_k)) = \widetilde{\nu}(E)$$

(toujours en utilisant la clause (1.7) de la proposition 1.1). Les deux mesures ν et $\widetilde{\nu}$ sont égales sur $\mathcal{T}_1 \otimes \mathcal{T}_2$ et la clause d'unicité de la proposition 3.1 est ainsi vérifiée.

Clause d'existence de la mesure $\mu_1 \otimes \mu_2$. Considérons une exhaustion de $\Omega_1 \times \Omega_2$ par des « rectangles » $A_k \times B_k$ tels que $\mu_1(A_k) < \infty$ et $\mu_2(B_k) < \infty$. L'ensemble des parties E de $A_k \times B_k$ telles que la fonction $x \in \Omega_1 \longmapsto \mu_2(E_x)$ soit $(\Omega_1, \mathcal{T}_1)$-$([0,\infty], \mathcal{B})$ mesurable est une classe monotone sur $A_k \times B_k$ (d'après la proposition 2.1 et le fait que $\mu_1(A_k)$ et $\mu_2(B_k)$ sont finies) contenant l'algèbre de Boole $\mathcal{A}^{(k)}$ des unions finies d'ensembles du type $A \times B$, $A \subset A_k$, $B \subset B_k$: en effet

$$x \in \Omega_1 \longmapsto \mu_2((A \times B)_x) = \chi_A(x) \times \mu_2(B)$$

est bien une fonction $(\Omega_1, \mathcal{T}_1)$-$([0,\infty], \mathcal{B})$ mesurable pour tout A dans $\mathcal{T}_1$, pour tout $B \in \mathcal{T}_2$ et ceci reste vrai si l'on remplace $A \times B$ par une union disjointe de tels « rectangles ». D'après le lemme 3.2, cette classe monotone contient la tribu $\mathcal{T}^{(k)}$ trace sur $A_k \times B_k$ de la tribu $\mathcal{T}_1 \otimes \mathcal{T}_2$, et l'on peut donc affirmer, pour tout $E \in \mathcal{T}^{(k)}$, que la fonction

$$x \in \Omega_1 \longmapsto \mu_2(E_x)$$

est $(\Omega_1, \mathcal{T}_1)$-$([0,\infty], \mathcal{B})$ mesurable. On peut donc définir

$$\nu_k(E) := \int_{\Omega_1} \mu_2(E_x)\, d\mu_1(x)$$

et vérifier (c'est immédiat) que l'on a ainsi une mesure sur $\mathcal{T}^{(k)} = (\mathcal{T}_1)_{A_k} \otimes (\mathcal{T}_2)_{B_k}$ ayant la propriété

$$\nu_k(A \times B) = \mu_1(A) \times \mu_2(B)\,,\ \forall\, A \in (T_1)_{A_k}\,,\ \forall\, B \in (T_2)_{B_k}\,.$$

Le même travail peut être fait pour les fonctions

$$y \in \Omega_2 \longmapsto \mu_1(E^y)\,,\ E \in \mathcal{T}_1 \otimes \mathcal{T}_2$$

pour définir, avec

$$E \in \mathcal{T}^{(k)} \longmapsto \int_{\Omega_2} \mu_1(E^y)\, d\mu_2(y)\,,$$

une mesure $\widetilde{\nu}_k$ ayant exactement les mêmes propriétés que ν_k. On définit ensuite deux mesures ν et $\widetilde{\nu}$ sur $\mathcal{T}_1 \otimes \mathcal{T}_2$ en posant

$$\begin{aligned} \nu(E) &:= \lim_{k\to\infty} \nu_k(E \cap (A_k \times B_k)) \\ \widetilde{\nu}(E) &:= \lim_{k\to\infty} \widetilde{\nu}_k(E \cap (A_k \times B_k))\,. \end{aligned}$$

La clause d'unicité permet de conclure à l'égalité $\nu = \widetilde{\nu}$ et la preuve de l'existence de $\mu_1 \otimes \mu_2$ (exhibée sous la forme ν ou $\widetilde{\nu}$) en résulte. $\diamondsuit$

Preuve du lemme 3.2. La plus petite classe monotone $\mathcal{M}$ contenant $\mathcal{A}$ est évidemment incluse dans la tribu engendrée $\mathcal{T}(\mathcal{A})$. Il reste donc à établir que cette classe monotone $\mathcal{M}$ est stable par prise de complémentaire et union finie (le passage à l'union dénombrable se faisant sans problème puisqu'il s'agit d'une classe monotone).

La famille $\{E \in \mathcal{M}\,;\, \Omega \setminus E \in \mathcal{M}\}$ est aussi une classe monotone incluse dans $\mathcal{M}$ (de manière immédiate car si $(E_k)_k$ est une suite croissante ou décroissante, la suite des complémentaires $(\Omega \setminus E_k)_k$ est décroissante ou croissante), donc contient nécessairement la classe $\mathcal{M}$ puisque contenant $\mathcal{A}$ (d'après la clause de minimalité de $\mathcal{M}$). Ainsi $\mathcal{M}$ est stable par prise de complémentaire.

Si l'on fixe E dans $\mathcal{M}$, on montre que $\mathcal{M}^E := \{M \in \mathcal{M}\,;\, E \cup M \in \mathcal{M}\}$ est une classe monotone incluse dans $\mathcal{M}$: si $(M_k)_k$ est une suite monotone de $\mathcal{M}^E$, la suite $(E \cup M_k)_k$ est une suite monotone de $\mathcal{M}$ et $\lim_k (E \cup M_k) = E \cup \lim_k M_k$ est donc encore un élément de la famille monotone $\mathcal{M}$. De plus cette classe monotone $\mathcal{M}^E$ contient l'algèbre de Boole $\mathcal{A}$: ceci est vrai si $E \in \mathcal{A}$ puisque $\mathcal{A} \subset \mathcal{M}$ et que $\mathcal{A}$ est stable par union finie ; comme ceci implique $\mathcal{M} = \mathcal{M}^B$ pour tout $B \in \mathcal{A}$ (toujours par minimalité de $\mathcal{M}$), on a donc, si $B \in \mathcal{A}$, $B \cup E \in \mathcal{M}^B = \mathcal{M}$, donc $B \in \mathcal{M}^E$; ainsi donc $\mathcal{A} \subset \mathcal{M}^E$. Par minimalité de $\mathcal{M}$, on conclut $\mathcal{M} = \mathcal{M}^E$ pour tout $E \in \mathcal{M}$, et $\mathcal{M}$ est ainsi stable par union finie.

Le lemme 3.2 est ainsi prouvé puisque nous venons de montrer que $\mathcal{M}$ est une tribu contenant $\mathcal{A}$, donc contenant la tribu $\mathcal{T}(\mathcal{A})$. $\diamondsuit$

3.4.2 Le théorème de Fubini : quand l'appliquer ? que fournit-il ?

Dans cette section, on considère deux ensembles Ω_1, Ω_2, équipés respectivement de tribus $\mathcal{T}_1, \mathcal{T}_2$; on considère deux mesures positives $\mu_1 : \mathcal{T}_1 \longrightarrow [0,\infty]$ et $\mu_2 : \mathcal{T}_2 \longrightarrow [0,\infty]$, toutes les deux σ-finies, et l'on note $\mu_1 \otimes \mu_2 : \mathcal{T}_1 \otimes \mathcal{T}_2 \longrightarrow [0,\infty]$ la mesure produit définie à la proposition 3.1. Le théorème de Fubini-Tonelli 3.7 se répercute au niveau des fonctions $(\Omega_1 \times \Omega_2, \mathcal{T}_1 \otimes \mathcal{T}_2)$-$(\mathbb{C}, \mathcal{B})$ mesurables, non plus positives mais cette fois intégrables (relativement à $\mu_1 \otimes \mu_2$), en le théorème de Fubini :

Théorème 3.8 (théorème de Fubini) [36] *Soit f une fonction $(\Omega_1 \times \Omega_2, \mathcal{T}_1 \otimes \mathcal{T}_2)$-$(\mathbb{C}, \mathcal{B})$ mesurable et intégrable relativement à la mesure produit $\mu_1 \otimes \mu_2$. Pour μ_1 presque tout $x \in \Omega_1$, la fonction $f_x : y \in \Omega_2 \longmapsto f(x,y)$ est intégrable relativement à la mesure μ_2 ; de plus, la fonction*

$$x \longmapsto \int_{\Omega_2} f(x,y)\, d\mu_2(y)$$

(définie hors d'un sous-ensemble μ_1-négligeable de Ω_1) se prolonge en une fonction $(\Omega_1, \mathcal{T}_1)$-$(\mathbb{C}, \mathcal{B})$ mesurable, intégrable relativement à la mesure μ_1, et l'on a la formule

$$\int_{\Omega_1 \times \Omega_2} f(x,y)\, d[\mu_1 \otimes \mu_2](x,y) = \int_{\Omega_1} \Big[\int_{\Omega_2} f(x,y)\, d\mu_2(y) \Big]\, d\mu_1(x)\,. \tag{3.17}$$

36. Proposé par le mathématicien italien Guido Fubini (1879-1943) en 1907, ce théorème fonctionne en fait en tandem avec le théorème dû à Fubini et à son compatriote Leonida Tonelli, ce suivant la règle importante que nous mentionnons en commentaire après cet énoncé.

Par symétrie des rôles de x et y, la fonction $f^y\ : x \in \Omega_1 \longmapsto f(x,y)$ est, pour μ_2 presque tout $y \in \Omega_2$, intégrable relativement à la mesure μ_1 ; de plus, la fonction

$$y \longmapsto \int_{\Omega_1} f(x,y)\, d\mu_1(x)$$

(définie hors d'un sous-ensemble μ_2-négligeable de Ω_2) se prolonge en une fonction $(\Omega_2, \mathcal{T}_2)$-$(\mathbb{C}, \mathcal{B})$ mesurable, intégrable relativement à la mesure μ_2, et l'on a aussi la formule

$$\int_{\Omega_1\times\Omega_2} f(x,y)\, d[\mu_1 \otimes \mu_2](x,y) = \int_{\Omega_2} \Big[\int_{\Omega_1} f(x,y)\, d\mu_1(x)\Big]\, d\mu_2(y)\,. \qquad (3.18)$$

La règle d'application du théorème : le théorème de Fubini s'applique presque toujours en « duo » avec le théorème de Fubini-Tonelli (théorème 3.7). C'est en effet ce dernier résultat (Fubini-Tonelli) qui permet, si l'on a vérifié au préalable la $(\Omega_1 \times \Omega_2, \mathcal{T}_1 \otimes \mathcal{T}_2)$-$(\mathbb{C}, \mathcal{B})$ mesurabilité d'une fonction $f\ : \Omega_1 \times \Omega_2 \longrightarrow \mathbb{C}$ et que l'on se soit assuré que l'une au moins des trois inégalités

$$\begin{aligned}
\int_{\Omega_2} \Big[\int_{\Omega_1} |f(x,y)|\, d\mu_1(x)\Big]\, d\mu_2(x) &< \infty \\
\int_{\Omega_1} \Big[\int_{\Omega_2} |f(x,y)|\, d\mu_2(y)\Big]\, d\mu_1(x) &< \infty \\
\int_{\Omega_1\times\Omega_2} |f(x,y)|\, d[\mu_1 \otimes \mu_2](x,y) &< \infty
\end{aligned}$$

(ces trois intégrales étant en fait égales dans $[0,\infty]$, peu importe donc celle que l'on calcule, on tente bien sûr la plus « accessible » !) d'assurer la validité des hypothèses du théorème de Fubini 3.8. Celui ci s'applique alors et l'on a la double égalité

$$\begin{aligned}
\int_{\Omega_1\times\Omega_2} f(x,y)\, d[\mu_1 \otimes \mu_2](x,y) &= \int_{\Omega_1} \Big[\int_{\Omega_2} f(x,y)\, d\mu_2(y)\Big]\, d\mu_1(x) \\
&= \int_{\Omega_2} \Big[\int_{\Omega_1} f(x,y)\, d\mu_1(x)\Big]\, d\mu_2(y) \qquad (3.19)
\end{aligned}$$

(cette fois sans les valeurs absolues dans l'intégrant $|f(x,y)|$!).

Preuve. L'intégrabilité de f par rapport à la mesure $\mu_1 \otimes \mu_2$ équivaut à l'intégrabilité relativement à cette mesure des quatre fonctions $(\Omega_1 \otimes \Omega_2, \mathcal{T}_1 \otimes \mathcal{T}_2)$- $([0,\infty[, \mathcal{B})$ mesurables positives $(\mathrm{Re}\, f)^{\pm}$, $(\mathrm{Im}\, f)^{\pm}$. Comme, d'après le théorème de Fubini-Tonelli 3.7, on a, pour chacune de ces quatre fonctions $g = (\mathrm{Re}\, f)^{\pm}, (\mathrm{Im}\, f)^{\pm}$,

$$\begin{aligned}
\int_{\Omega_1\times\Omega_2} g(x,y)\, d[\mu_1 \otimes \mu_2](x,y) &= \int_{\Omega_1} \Big[\int_{\Omega_2} g(x,y)\, d\mu_2(y)\Big]\, d\mu_1(x) \\
&= \int_{\Omega_2} \Big[\int_{\Omega_1} g(x,y)\, d\mu_1(x)\Big]\, d\mu_2(y) < \infty\,,
\end{aligned}$$

les fonctions g_x (*resp.* g^y) sont bien intégrables relativement à μ_1 (*resp.* à μ_2) pour μ_1-presque tout $x \in \Omega_1$ (*resp.* pour μ_2 presque tout $y \in \Omega_2$), ce en vertu de la proposition 2.3. Les fonctions

$$\begin{aligned}
x \in \Omega_1 &\longmapsto \int_{\Omega_2} g(x,y)\, d\mu_2(y) \in [0,\infty] \\
y \in \Omega_2 &\longmapsto \int_{\Omega_1} g(x,y)\, d\mu_1(x) \in [0,\infty]
\end{aligned}$$

(finies respectivement μ_1-presque partout sur Ω_1 et μ_2-presque partout sur Ω_2) sont des fonctions respectivement $(\Omega_1, \mathcal{T}_1)$-$([0,\infty], \mathcal{B})$ et $(\Omega_2, \mathcal{T}_2)$-$([0,\infty], \mathcal{B})$ mesurables, intégrables respectivement relativement à μ_1 et μ_2. Ceci étant vrai (ainsi que la formule (3.20)) pour les quatre fonctions $g = (\mathrm{Re}\, f)^{\pm}, (\mathrm{Im}\, f)^{\pm}$, on en déduit, par linéarité de la prise d'intégrale relativement à une mesure pour les fonctions mesurables positives intégrables relativement à cette mesure (les coefficients à prendre ici étant $1, -1, i, -i$ affectant respectivement $(\mathrm{Re}\, f)^+$, $(\mathrm{Re}\, f)^-$, $(\mathrm{Im}\, f)^+$, $(\mathrm{Im}\, f)^-$), toutes les conclusions du théorème 3.8 ainsi que la validité des formules (3.17) et (3.18). Le théorème de Fubini est ainsi complètement démontré. $\diamondsuit$.

Remarque 3.9. Si la clause d'intégrabilité de la fonction f relativement à la mesure $\mu_1 \otimes \mu_2$ n'est pas remplie, le théorème de Fubini est en défaut ! Considérons par exemple la fonction mesurable (mais non positive)

$$(x,y) \in]0,1]^2 \longmapsto \frac{x^2 - y^2}{(x^2+y^2)^2}\,;$$

pour tout $x \in]0,1]$, la fonction f_x est intégrable sur $]0,1]$ et l'on a

$$\int_{]0,1]} f_x(y)\, dy = \frac{1}{x}\int_0^{1/x} \frac{1-t^2}{(1+t^2)^2}\, dt = \frac{1}{x}\Big[\frac{t}{1+t^2}\Big]_0^{1/x} = \frac{1}{1+x^2}\,;$$

on constate donc que la fonction

$$x \in]0,1] \longmapsto \int_{]0,1]} f(x,y)\, dy$$

est une fonction intégrable sur $]0,1]$, d'intégrale $\mathrm{artan}\,(1) = \pi/4$. Pour tout $y \in]0,1]$, la fonction f^y est intégrable sur $]0,1]$ et

$$\int_{]0,1]} f_y(x)\, dx = \frac{1}{y}\int_0^{1/y} \frac{t^2-1}{(1+t^2)^2}\, dt = \frac{1}{y}\Big[-\frac{t}{1+t^2}\Big]_0^{1/y} = -\frac{1}{1+y^2}$$

ici encore, la fonction

$$y \in]0,1] \longmapsto \int_{]0,1]} f(x,y)\, dx$$

est aussi intégrable sur $]0,1]$ et l'on a

$$\int_{]0,1]} \Big[\int_{]0,1]} f(x,y)\, dx\Big]\, dy = -\int_{]0,1]} \Big[\int_{]0,1]} f(x,y)\, dy\Big]\, dx = \frac{\pi}{4} \neq 0\,.$$

En passant en coordonnées polaires, on voit, si A désigne le quart du disque $D(0,1)$ inclus dans le premier quadrant, que

$$\iint_A |f(x,y)|\, dxdy = \iint_{]0,\pi/2[\times]0,1[} \frac{r^2|\cos(2\theta)|}{r^4}\, rdrd\theta = +\infty\,,$$

ce qui implique que la fonction f n'est pas intégrable sur $]0,1]^2$ relativement à la mesure de Lebesgue $dxdy$. Ce n'est donc pas une surprise ici que l'ont n'ait pas ici l'égalité

$$\int_{]0,1]} \Big[\int_{]0,1]} f(x,y)\, dx\Big]\, dy = \int_{]0,1]} \Big[\int_{]0,1]} f(x,y)\, dy\Big]\, dx\,,$$

quand bien même les deux membres aient ici un sens !

Remarque 3.10. Le théorème de Fubini est un résultat plus profond qu'il ne paraît à première vue. En voici un exemple : supposons que P soit un pavé borné de $\mathbb{R}^2$ de côtés ayant pour longueurs a et b tel que P s'écrive comme union de pavés $P_k = I_k \times J_k$ d'intérieurs disjoints de manière à ce que chaque P_k ait au moins un côté de longeur dans $\mathbb{N}^*$ (tous les pavés sont ici parallèles aux axes, on peut voir P comme un mur fait de briques rectangulaires P_k posées verticalement ou horizontalement et ayant chacune un de leurs deux côtés entier). En intégrant sur P la fonction

$$(x,y) \longmapsto e^{2i\pi(x+y)}\,,$$

on constate (appliquant ici dans chaque pavé P_k le théorème de Fubini et en utilisant le fait que P est l'union des P_k et que les P_k sont d'intérieurs disjoints) que

$$\iint_P f(x,y)\,dxdy = \sum_k \Big(\int_{I_k} \exp(2i\pi x)\,dx\Big)\Big(\int_{J_k} \exp(2i\pi y)\,dy\Big) = 0\,.$$

Le même résultat appliqué cette fois à $P = [x_0, x_0 + a] \times [y_0, y_0 + b]$ nous donne

$$\iint_P f(x,y)\,dxdy = -\frac{1}{4\pi^2}\Big[\exp(2i\pi x)\Big]_{x_0}^{x_0+a} \times \Big[\exp(2i\pi y)\Big]_{y_0}^{y_0+b}\,;$$

le fait que cette intégrale soit nulle implique que l'une au moins des deux longueurs a ou b soit entière. Ceci peut se voir par un raisonnement utilisant par exemple les marches aléatoires discrètes, mais n'est pas tout-à-fait évident ! On perçoit donc ici la profondeur du théorème de Fubini, couplé ici avec la « relation de Chasles » en deux variables (et donc avec la force de la théorie de la mesure dans le plan).

Le cadre de $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ (avec la mesure de Lebesgue sur chacun des facteurs et la mesure de Lebesgue sur le produit) est un cadre d'application classique du théorème de Fubini. Voici un exemple, que nous retrouverons au chapitre 4 lorsque nous définirons la très importante opération de convolution :

Exemple 3.14. Soient f et g deux fonctions mesurables de $\mathbb{R}^n$ dans $\mathbb{C}$, toutes les deux intégrables par rapport à la mesure de Lebesgue. D'après le théorème de Fubini-Tonelli, la fonction

$$(x,y) \in \mathbb{R}^n \times \mathbb{R}^n \longmapsto f(x-y)g(y)$$

est intégrable sur $\mathbb{R}^n \times \mathbb{R}^n$ relativement à la mesure de Lebesgue $dxdy$ car

$$\begin{aligned}\iint_{\mathbb{R}^n\times\mathbb{R}^n} |f(x-y)|\,|g(y)|\,dxdy &= \int_{\mathbb{R}^n}\Big[\int_{\mathbb{R}^n} |f(x-y)|\,dx\Big]\,|g(y)|\,dy \\ &= \int_{\mathbb{R}^n}\Big[\int_{\mathbb{R}^n} |f(x)|\,dx\Big]\,|g(y)|\,dy \\ &= \Big(\int_{\mathbb{R}^n} |f(x)|\,dx\Big) \times \Big(\int_{\mathbb{R}^n} |g(y)|\,dy\Big) < \infty \qquad (3.20)\end{aligned}$$

(on s'est servi ici de la propriété d'invariance par translation de la mesure de Lebesgue dans $\mathbb{R}^n$). La machine « Fubini » du théorème 3.8 peut alors se mettre en route et l'on peut affirmer que, pour presque tout x (au sens de la mesure de Lebesgue) dans $\mathbb{R}^n$, la fonction

$$y \longmapsto f(x-y)\,g(y)$$

est intégrable sur $\mathbb{R}^n$ relativement à la mesure de Lebesgue, ce qui nous permet de définir une fonction mesurable $F = f * g$ sur $\mathbb{R}^n$ presque partout égale à la fonction

$$x \longmapsto \int_{\mathbb{R}^n} f(x-y)\,g(y)\,dy = \int_{\mathbb{R}^n} f(y)\,g(x-y)\,dy$$

(on invoque encore l'invariance par translation de la mesure de Lebesgue pour obtenir l'égalité ci-dessus). De plus F est intégrable relativement à la mesure de Lebesgue car, pour presque tout x,

$$|F(x)| = \Big|\int_{\mathbb{R}^n} f(x-y)\,g(y)\,dy\Big| \leq \int_{\mathbb{R}^n} |f(x-y)|\,|g(y)|\,dy$$

d'après la proposition 2.7 (inégalité (2.22)) et que l'on a donc (du fait de (3.20),

$$\int_{\mathbb{R}^n} |F(x)|\,dx \leq \Big(\int_{\mathbb{R}^n} |f(x)|\,dx\Big) \times \Big(\int_{\mathbb{R}^n} |g(y)|\,dy\Big) < \infty\,.$$

La fonction F ainsi définie sera appelée *convolée* des fonctions f et g. Son rôle (souvent d'ailleurs en relation avec les transformations de Fourier, Laplace, Mellin introduites dans la section 3.2.3) s'avèrera essentiel [37], tant en mathématiques que dans les sciences de l'ingénieur.

37. Comme on le verra au chapitre 4 et plus tard dans l'UE MHT613.

Lorsque $\Omega_2 = \mathbb{N}$ et μ_2 est la mesure de décompte, on retrouve avec le théorème de Fubini (au moins lorsque μ est une mesure σ-finie sur $(\Omega, \mathcal{T})$) le résultat d'interversion entre sommation et prise d'intégrale qui résultait du théorème de convergence dominée de Lebesgue.

De nombreux exemples d'application du tandem « Fubini + Fubini-Tonelli » (outils d'une importance capitale au service de l'intégration « pratique ») seront proposés en exercice dans le guide d'activités 8 sous Ulysse. C'est la raison pour laquelle nous n'en donnerons pas d'autres ici (remarquons cependant que nous avons déjà utilisé Fubini plusieurs fois, par exemple dans la section 3.3, *c.f* l'exemple 3.6, ou dans les exemples proposés dans les guides d'activités 7 et 8 sous Ulysse).

Chapitre 4

Les espaces L^p et la convolution

4.1 Les espaces $\mathcal{L}^p(\Omega, \mathcal{T}, \mu)$ (définition ensembliste)

Dans cette section, Ω désigne un ensemble abstrait, $\mathcal{T}$ une tribu sur Ω, et $\mu : \mathcal{T} \longrightarrow [0, \infty]$ une mesure positive.

Si $p \in [1, \infty[$ et si $f : \Omega \longrightarrow \mathbb{C}$, $g : \Omega \longrightarrow \mathbb{C}$ sont deux fonctions $(\Omega, \mathcal{T})$-$(\mathbb{C}, \mathcal{B})$-mesurables telles que

$$\int_\Omega |f|^p \, d\mu < \infty \quad \text{et} \quad \int_\Omega |g|^p \, d\mu < \infty \, ,$$

on a, pour tout $\lambda \in \mathbb{C}$,

$$\int_\Omega |f + \lambda g|^p \, d\mu \leq 2^p \int_\Omega (|f|^p + |\lambda g|^p) \, d\mu = 2^p \int_\Omega |f|^p \, d\mu + 2^p \, |\lambda|^p \int_\Omega |g|^p \, d\mu < \infty \ ;$$

en effet, si α et β sont deux nombres positifs, on a

$$(\alpha + \beta)^p \leq (2 \max(\alpha, \beta))^p \leq 2^p (\alpha^p + \beta^p) \, .$$

Cette remarque nous conduit, une fois choisi un nombre réel $p \in [1, +\infty[$, aux définitions suivantes d'un $\mathbb{C}$ et d'un $\mathbb{R}$-espace vectoriel.

Définition 4.1 *Soit Ω, $\mathcal{T}$ et μ comme précédemment et $p \in [1, +\infty[$. L'ensemble des fonctions $f : \Omega \longrightarrow \mathbb{C}$ qui sont $(\Omega, \mathcal{T})$-$(\mathbb{C}, \mathcal{B})$ mesurables et telles que*

$$\int_\Omega |f|^p \, d\mu < +\infty \tag{4.1}$$

hérite, équipé de l'addition et de la multiplication extérieure par un scalaire complexe, d'une structure de $\mathbb{C}$-espace vectoriel. Ce $\mathbb{C}$-espace vectoriel est appelé $\mathcal{L}^p_{\mathbb{C}}(\Omega, \mathcal{T}, \mu)$. Si l'on ne considère parmi ces fonctions que les fonctions à valeurs réelles, on dispose d'un $\mathbb{R}$-espace vectoriel, noté $\mathcal{L}^p_{\mathbb{R}}(\Omega, \mathcal{T}, \mu)$.

Remarque 4.1. Si $\mu(\Omega) < +\infty$, on constate que, si $1 \leq p_1 \leq p_2 < \infty$, on a

$$\begin{aligned} \mathcal{L}^{p_2}_{\mathbb{C}}(\Omega, \mathcal{T}, \mu) &\subset \mathcal{L}^{p_1}_{\mathbb{C}}(\Omega, \mathcal{T}, \mu) \\ \mathcal{L}^{p_2}_{\mathbb{R}}(\Omega, \mathcal{T}, \mu) &\subset \mathcal{L}^{p_1}_{\mathbb{R}}(\Omega, \mathcal{T}, \mu) \end{aligned} \tag{4.2}$$

En effet, on peut écrire $f \in \mathcal{L}^{p_2}_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ (ici $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) sous la forme

$$f = f \chi_{\{|f| \leq 1\}} + f \chi_{\{|f| > 1\}} \, .$$

Comme fonction mesurable bornée, la fonction $|f\chi_{\{|f|\leq 1\}}|^{p_1}$ est intégrable puisque $\mu(\Omega) < \infty$; cela résulte du critère de domination de la proposition 2.9 ; quant à la fonction $f\chi_{\{|f|>1\}}$, elle vérifie

$$|f\chi_{\{|f|>1\}}|^{p_1} \leq |f\chi_{\{|f|>1\}}|^{p_2}\,,$$

ce qui montre que les deux fonctions $f\chi_{\{|f|\leq 1\}}$ et $f\chi_{\{|f|>1\}}$, donc aussi f, sont des éléments de $\mathcal{L}^{p_1}(\Omega, \mathcal{T}, \mu)$. Les inclusions (4.2), lorsque $\mu(\Omega) < \infty$ et $1 \leq p_1 \leq p_2 < \infty$, en résultent. Rien de tel ne subsiste lorsque la mesure μ ne vérifie plus $\mu(\Omega) < +\infty$: par exemple, si $\Omega = \mathbb{R}$, $\mathcal{T} = \mathcal{B}(\mathbb{R})$ et que μ est la mesure de Lebesgue, la fonction

$$t \in \mathbb{R} \longmapsto \frac{\sin t}{t}$$

est dans $\mathcal{L}^2_{\mathbb{R}}(\mathbb{R}, \mathcal{B}, \mu)$ mais n'est pas dans $\mathcal{L}^1_{\mathbb{R}}(\mathbb{R}, \mathcal{B}, \mu)$; en revanche, la fonction

$$f\ : t \in \frac{1}{\sqrt{|t|}(1+|t|)}$$

(avec la convention $f(0) = 0$) est dans $\mathcal{L}^1_{\mathbb{R}}(\mathbb{R}, \mathcal{B}, \mu)$, mais non dans $\mathcal{L}^2_{\mathbb{R}}(\mathbb{R}, \mathcal{B}, \mu)$.

Remarque 4.2. Si Ω est un ensemble, $\mathcal{T} = \mathcal{P}(\Omega)$ et $\mu = \nu$ la mesure de décompte sur Ω, on note (si $\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$)

$$l^p_{\mathbb{K}}(\Omega) := \mathcal{L}^p_{\mathbb{K}}(\Omega, \mathcal{P}(\Omega), \nu).$$

On a, si $1 \leq p_1 \leq p_2 < +\infty$,

$$\begin{array}{rcl} l^{p_1}_{\mathbb{C}}(\Omega) & \subset & l^{p_2}_{\mathbb{C}}(\Omega) \\ l^{p_1}_{\mathbb{R}}(\Omega) & \subset & l^{p_2}_{\mathbb{R}}(\Omega)\,. \end{array} \tag{4.3}$$

En effet, si $f \in l^{p_1}_{\mathbb{K}}(\Omega)$, l'ensemble des $\omega \in \Omega$ tels que $|f(\omega)| > 1$ est de mesure de décompte finie, donc fini ; d'autre part

$$|f|^{p_2}\chi_{|f|\leq 1\}} \leq |f|^{p_1}\chi_{|f|\leq 1\}}$$

si $p_2 - p_1 \geq 0$; les fonctions $|f|^{p_2}\chi_{\{|f|>1\}}$ et $|f|^{p_2}\chi_{\{|f|\leq 1\}}$ étant ainsi intégrables relativement à la mesure de décompte, f est bien dans $l^{p_2}_{\mathbb{K}}(\Omega)$, ce qui prouve les inclusions (4.3). On remarque d'ailleurs que si $f \in l^p_{\mathbb{K}}(\Omega)$ avec $p \in [1, \infty[$, l'ensemble

$$\{f \neq 0\} = \bigcup_{n\in\mathbb{N}^*} \{|f| \geq 1/n\}$$

est une union dénombrable d'ensembles finis, donc un ensemble au plus dénombrable de $\mathbb{K}$.

On peut aussi définir les deux espaces vectoriels (respectivement $\mathbb{C}$ et $\mathbb{R}$-espaces vectoriels) $\mathcal{L}^\infty_{\mathbb{C}}(\Omega, \mathcal{T}, \mu)$ et $\mathcal{L}^\infty_{\mathbb{R}}(\Omega, \mathcal{T}, \mu)$ comme suit :

Définition 4.2 *Le $\mathbb{C}$ (resp. $\mathbb{R}$)-espace vectoriel $\mathcal{L}^\infty_{\mathbb{C}}(\Omega, \mathcal{T}, \mu)$ (resp. $\mathcal{L}^\infty_{\mathbb{R}}(\Omega, \mathcal{T}, \mu)$) est le $\mathbb{C}$-espace vectoriel (resp. $\mathbb{R}$-espace vectoriel) des fonctions $(\Omega, \mathcal{T})$-$(\mathbb{C}, \mathcal{B})$ (resp. $(\Omega, \mathcal{T})$-$(\mathbb{R}, \mathcal{B})$) mesurables de module μ-presque partout égal à une fonction mesurable positive bornée.*

Si f est un élément de $\mathcal{L}^\infty_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$, un nombre M est appelé *majorant essentiel* de $|f|$ si l'ensemble $\{|f| > M\}$ est μ-négligeable. Si M_f désigne la borne inférieure de l'ensemble des majorants essentiels de $|f|$, on constate, puisque

$$\{|f| > M_f\} = \bigcup_{n\in\mathbb{N}^*} \{|f| > M_f + 1/n\}\,,$$

que M_f est encore un majorant essentiel de $|f|$ (puisque tous les nombres $M_f + 1/n$, $n \in \mathbb{N}^*$, le sont). On appelle M_f la *borne supérieure essentielle* de $|f|$ et l'on note

$$M_f = \|f\|_\infty\,.$$

Remarque 4.3. Soit $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Si $\mu(\Omega) < \infty$, l'espace $\mathcal{L}^\infty_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ est inclus dans tous les $\mathcal{L}^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ pour tout $p \in [1, \infty[$ (pour la même raison que dans la remarque 4.1). Si $\mathcal{T} = \mathcal{P}(\Omega)$ et $\mu = \nu$ est la mesure de décompte, on note $l^\infty_{\mathbb{K}}(\Omega)$ le $\mathbb{K}$-espace vectoriel $\mathcal{L}^\infty_{\mathbb{K}}(\Omega, \mathcal{P}(\Omega), \nu)$; le $\mathbb{K}$-espace vectoriel $l^\infty_{\mathbb{K}}(\Omega)$ contient tous les $\mathbb{K}$-espaces vectoriels $l^p_{\mathbb{K}}(\Omega)$ pour $p \in [1, \infty[$ (pour la même raison que dans la remarque 4.2). En revanche, la propriété « $\{f \neq 0\}$ au plus dénombrable » valable pour $f \in l^p_{\mathbb{K}}(\Omega)$ (lorsque $p \in [1, \infty[$), ne subsiste plus, elle, pour un élément de $l^\infty_{\mathbb{K}}(\Omega)$.

4.2 Les inégalités de Hölder et Minkowki ; les semi-normes $\| \ \|_p$

Dans cette section, nous nous proposons d'équiper chaque espace $\mathcal{L}^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ (où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$), $p \in [1, \infty]$, d'une semi-norme $\| \ \|_p$, c'est-à-dire d'une application

$$\| \ \|_p \ : \mathcal{L}^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu) \longmapsto [0, \infty[$$

telle que

$$\begin{aligned} &\forall\, f \in \mathcal{L}^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)\,,\ \forall\, \lambda \in \mathbb{K}\,,\ \|\lambda f\|_p = |\lambda| \times \|f\|_p \\ &\forall\, f, g \in \mathcal{L}^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)\,,\ \|f + g\|_p \leq \|f\|_p + \|g\|_p\,. \end{aligned} \tag{4.4}$$

Pour $p = +\infty$, il est naturel de poser

$$\forall\, f \in \mathcal{L}^\infty_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)\,,\ \|f\|_\infty := \sup \text{ess}|f| \ ;$$

ceci définit naturellement une semi-norme sur le $\mathbb{K}$-espace vectoriel $\mathcal{L}^\infty_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ (qu'il convient de ne pas confondre avec $\sup_\Omega(|f|)$ qui, elle, peut fort bien valoir $+\infty$ même si $|f|$ est essentiellement bornée). En effet, la première clause de (4.4) est remplie. Concernant la seconde, si f_1 et f_2 sont deux éléments de $\mathcal{L}^\infty_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$, il existe des fonctions mesurables bornées $\tilde{f}_1$ et $\tilde{f}_2$ telles que $f_j = \tilde{f}_j$ presque partout et que $\sup_\Omega |\tilde{f}_j| = \|f_j\|_\infty$, $j = 1, 2$; on a donc

$$\sup_\Omega |\tilde{f}_1 + \tilde{f}_2| \leq \|f_1\|_\infty + \|f_2\|_\infty\,,$$

ce qui prouve que $\|f_1\|_\infty + \|f_2\|_\infty$ est un majorant essentiel pour $|f_1 + f_2|$, donc que

$$\|f_1 + f_2\|_\infty \leq \|f_1\|_\infty + \|f_2\|_\infty\,.$$

Pour $p \in [1, +\infty[$, le candidat pour $\| \ \|_p$ est la fonction

$$f \in \mathcal{L}^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu) \longmapsto \|f\|_p \ := \Big(\int_\Omega |f|^p \, d\mu \Big)^{1/p} \ ;$$

cette définition est naturelle si l'on veut respecter la clause d'homogénéité que constitue la première des conditions dans (4.4). En revanche, il faut vérifier la seconde condition de (4.4), ce qui n'est pas tout-à-fait immédiat et correspond à une inégalité importante en analyse, dite *inégalité de Minkowski*[1].

En préalable à l'inégalité de Minkowski, nous devons énoncer une célèbre et importante inégalité, celle de Hölder[2].

1. Géomètre des nombres, mais aussi physicien, le mathématicien lituanien allemand Hermann Minkowski (1864-1909) s'est beaucoup intéressé à la géométrie des convexes et aux inégalités de convexité ; l'inégalité qui porte son nom s'inscrit dans cette ligne de recherches.

2. C'est en 1884, dans un travail consacré aux séries de Fourier, que le mathématicien allemand Otto Hölder (1859-1937) énonça cette inégalité capitale.

Proposition 4.1 (inégalité de Hölder) *Soit Ω un ensemble, $\mathcal{T}$ une tribu sur Ω et $\mu\ : \mathcal{T} \longrightarrow [0,\infty]$ une mesure positive. Si f et g sont deux fonctions de Ω dans $[0,\infty]$ toutes les deux $(\Omega,\mathcal{T})$-$([0,\infty],\mathcal{B})$ mesurables, on a*

$$\int_\Omega fg\,d\mu \leq \Big(\int_\Omega f^p\,d\mu\Big)^{1/p} \times \Big(\int_\Omega g^{p'}\,d\mu\Big)^{1/p'}, \tag{4.5}$$

où $p' \in [1,\infty]$ est dit exposant conjugué de p et défini par

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

avec la convention $p' = \infty$ si $p = 1$, $p' = 1$ si $p = \infty$ et

$$\Big(\int_\Omega h^\infty\,d\mu\Big)^{1/\infty} = \|h\|_\infty$$

si h est essentiellement bornée, $+\infty$ sinon (et toujours bien sûr la convention $0 \times (+\infty) = 0$). Si de plus $f \in \mathcal{L}^p_{\mathbb{R}}(\Omega,\mathcal{T},\mu)$, $g \in \mathcal{L}^{p'}_{\mathbb{R}}(\Omega,\mathcal{T},\mu)$, l'inégalité (4.5) lorsque $p \in]1,+\infty[$ est une égalité si et seulement si il existe $\alpha,\beta \in [0,\infty[$, non tous les deux nuls, tels que $\alpha f^p = \beta g^{p'}$-$\mu$ presque partout ; si $p = 1$ (resp. si $p = \infty$), c'est une égalité si et seulement si g (resp. f) est μ-presque partout constante ou f (resp. g) μ-presque partout nulle.

Preuve. Si $p = 1$ ou $p = \infty$, l'inégalité est immédiate car on majore sous l'intégrale (hors d'un sous-ensemble μ-négligeable) fg soit par $f\|g\|_\infty$ (si $p = 1$), soit par $g\|f\|_\infty$ (si $p = \infty$).

Il nous reste à prouver (4.5) si $p \in]1,+\infty[$.

Si u et v sont deux nombres strictement positifs et $t \in [0,1]$, on a, du fait de la concavité de la fonction log sur $]0,\infty[$:

$$\log(tu + (1-t)v) \geq t\log u + (1-t)\log v\,, \tag{4.6}$$

ou encore

$$u^t v^{1-t} \leq tu + (1-t)v\,.$$

Cette inégalité est d'ailleurs stricte dès que $u < v$ (ce pour tout $t \in [0,1]$).

Pour prouver (4.5) pour $p \in]1,\infty[$, on peut se limiter à supposer

$$\int_\Omega f^p\,d\mu \in]0,\infty[\quad \text{et} \quad \int_\Omega g^{p'}\,d\mu \in]0,\infty[\,; \tag{4.7}$$

si en effet une de ces intégrales est nulle, f (ou g), donc fg, sont des fonctions nulles μ presque partout du fait de la proposition 2.4, le membre de gauche de (4.5) est nul, comme l'est le membre de droite ; si l'une des intégrales (4.7) est infinie et que l'autre est non nulle, le produit des deux vaut $+\infty$ et l'inégalité (4.5) est satisfaite. Si l'on fait l'hypothèse (4.7) et que l'on pose, pour tout $\omega \in \Omega$ tel que $\infty > f(\omega)g(\omega) > 0$,

$$\begin{aligned} u(\omega) &:= \frac{(f(\omega))^p}{\int_\Omega f^p\,d\mu} \\ v(\omega) &:= \frac{(g(\omega))^{p'}}{\int_\Omega g^{p'}\,d\mu} \\ t &= 1/p\,, \end{aligned}$$

et que l'on intègre sur $\{fg \in]0,\infty[\}$ les deux membres de l'inégalité

$$[u(\omega)]^t \times [v(\omega)]^{1-t} \leq tu(\omega) + (1-t)v(\omega)\,,$$

on obtient immédiatement l'inégalité (4.5) car f et g sont finies presque partout du fait de la proposition 2.3.

Reste à prouver la dernière assertion de la proposition ; là encore, on peut supposer l'hypothèse (4.7) remplie et reprendre (lorsque $p \in]1,+\infty[$) le raisonnement conduit pour obtenir l'inégalité (4.5). L'égalité dans (4.5) n'est possible que si $u(\omega) = v(\omega)$ μ-presque partout, ce qui nous conduit à l'assertion proposée. Si $p = 1$, f intégrable et g bornée, l'égalité

$$\int_\Omega (\|g\|_\infty - g) f \, d\mu = 0$$

n'est possible que si $g = \|g\|_\infty$ μ-presque partout ou $f = 0$ μ-presque partout ; si $p = \infty$, f bornée et g intégrable, l'égalité

$$\int_\Omega (\|f\|_\infty - f) g \, d\mu = 0$$

n'est possible que si $f = \|f\|_\infty$ μ-presque partout ou $g = 0$ μ-presque partout.

La proposition est complètement démontrée. $\diamondsuit$

Proposition 4.2 *Soit $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$. Pour $p \in [1,\infty]$,*

$$f \longmapsto \|f\|_p$$

est une semi-norme sur le $\mathbb{K}$-espace vectoriel $\mathcal{L}^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$. De plus, si f et g sont des fonctions mesurables positives éléments de $\mathcal{L}^p_{\mathbb{R}}(\Omega, \mathcal{T}, \mu)$, avec $p \in]1,\infty[$, l'égalité

$$\|f+g\|_p = \|f\|_p + \|g\|_p$$

est réalisée si et seulement s'il existe $(\alpha,\beta) \in [0,\infty[^2 \setminus \{(0,0)\}$ tel que

$$\alpha f = \beta g$$

μ-presque partout[3] ; dans le cas particulier très important $p = 2$, l'égalité

$$\|f+g\|_2^2 = \|f\|_2^2 + \|g\|_2^2$$

entre éléments de $\mathcal{L}^2_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ équivaut à

$$\int_\Omega fg \, d\mu = 0 \quad si \quad \mathbb{K} = \mathbb{R}$$
$$\mathrm{Re}\Big(\int_\Omega f\overline{g} \, d\mu\Big) = 0 \quad si \quad \mathbb{K} = \mathbb{C}\,.$$

3. Si $p = 1$, l'égalité $\|f+g\|_1 = \|f\|_1 + \|g\|_1$ est évidemment toujours satisfaite pour deux fonctions mesurables positives f et g.

Preuve. Le seul cas posant problème est le cas $p \in]1, \infty[$ (le cas $p = \infty$ a été réglé et le cas $p = 1$ est une conséquence immédiate de l'inégalité triangulaire $|f + g| \leq |f| + |g|$ intégrée ensuite sur Ω relativement à la mesure μ). Si f et g sont deux éléments de $\mathcal{L}^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$, on écrit

$$\forall\, \omega \in \Omega\,, |f(\omega) + g(\omega)|^p \leq |f(\omega)|\, |f(\omega) + g(\omega)|^{p-1} + |g(\omega)|\, |f(\omega) + g(\omega)|^{p-1}\,,$$

puis on intègre cette inégalité sur Ω relativement à la mesure μ. En utilisant deux fois la proposition 4.1 et l'inégalité (4.5) avec $|f|$, $|f + g|^{p-1}$, puis $|g|$, $|f + g|^{p-1}$, il vient

$$\|f + g\|_p^p \leq \|f\|_p \times \|f + g\|_p^{p/p'} + \|g\|_p \times \|f + g\|_p^{p/p'}\,,$$

d'où l'inégalité

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p \tag{4.8}$$

puisque $1/p + 1/p' = 1$.

La seconde assertion traitant du cas d'égalité (pour des fonctions positives) lorsque $p \in]1, \infty[$ est une conséquence de ce que l'égalité dans (4.8) implique l'égalité dans Hölder pour les deux paires $((f+g)^{p-1}, f)$ et $((f+g)^{p-1}, g)$; on utilise donc l'assertion traitant de ces cas d'égalité dans la proposition 4.1.

La dernière assertion (cas $p = 2$) résulte des formules de Pythagore

$$\begin{aligned} (f + g)^2 &= f^2 + g^2 + 2fg \\ |f + g|^2 &= |f|^2 + |g|^2 + 2\,\mathrm{Re}\,[f\overline{g}] \end{aligned}$$

(respectivement dans le cadre réel $\mathbb{K} = \mathbb{R}$ et dans le cadre complexe $\mathbb{K} = \mathbb{C}$), égalités que l'on intègre ensuite sur Ω relativement à la mesure μ.

La proposition est démontrée. $\diamondsuit$

L'inégalité de Hölder (cette fois écrite pour des fonctions de signe quelconque) devient la :

Proposition 4.3 *Soient $f \in \mathcal{L}^p_{\mathbb{C}}(\Omega, \mathcal{T}, \mu)$ et $g \in \mathcal{L}^{p'}_{\mathbb{C}}(\Omega, \mathcal{T}, \mu)$ avec $p \in [1, \infty]$ et $1/p + 1/p' = 1$. Alors $fg \in \mathcal{L}^1_{\mathbb{C}}(\Omega, \mathcal{T}, \mu)$ et*

$$\Big| \int_\Omega fg\, d\mu \Big| \leq \|fg\|_1 \leq \|f\|_p \times \|g\|_{p'}\,. \tag{4.9}$$

Preuve. Il suffit de combiner les propositions 2.7 et 4.1. $\diamondsuit$

Remarque 4.4. Pour n'avoir que des égalités dans la chaîne d'inégalités (4.9), il faut (d'après la proposition 4.1) qu'il existe deux nombre réels positifs non tous les deux nuls tels que $\alpha|f|^p - \beta|g|^{p'}$ μ-presque partout si $p \in]1, \infty[$, $|g| = \|g\|_\infty$ μ-presque partout ou $f = 0$ μ-presque partout si $p = 1$, $|f| = \|f\|_\infty$ μ-presque partout ou $g = 0$ μ-presque partout si $p = \infty$. Comme l'égalité dans (4.9) implique aussi μ-presque partout $fg = \lambda|fg|$, avec $\lambda \in \mathbb{C}$ et $|\lambda| = 1$ d'après la proposition 2.7, il n'y a égalité dans la chaîne d'inégalités (4.9) que si toutes ces conditions sont réunies. La célèbre inégalité de Cauchy-Schwarz [4]

$$\Big| \int_\Omega f\overline{g}\, d\mu \Big| \leq \|f\|_2 \times \|g\|_2 \tag{4.10}$$

4. Si les noms du mathématicien français Augustin Cauchy (1789-1857) et du mathématicien allemand d'origine polonaise Hermann Schwarz (1843-1921) sont attachés à cette inégalité, il semble qu'elle soit déjà apparue aussi dès 1859 dans les travaux du mathématicien russe Viktor Bunyakowsky (1804-1889) et formalisée dans sa forme « moderne » par l'allemand Hermann Weyl en 1918. Beaucoup de noms y sont de fait attachés, tant dans le monde occidental que dans l'ex-monde soviétique !

lorsque f et g sont deux éléments de $\mathcal{L}^2_{\mathbb{C}}(\Omega, \mathcal{T}, \mu)$, est une conséquence de la proposition 4.3 lorsque $p = 2$. Mais ceci est de fait un résultat d'obédience algébrique et non analytique ! En utilisant en plus le développement du trinôme

$$\int_\Omega |f + tg|^2 \, d\mu = \|f\|_2^2 + |t|^2 \|g\|_2^2 + 2 \operatorname{Re} \Big(\bar{t} \int_\Omega f \overline{g} \, d\mu \Big) ,$$

on retrouve avec la positivité de ce trinôme pour tout t l'inégalité (4.10) et l'on voit en prime (voir un cours d'algèbre) que l'égalité dans l'inégalité (4.10) équivaut à l'existence de $(\alpha, \beta) \in \mathbb{C}^2 \setminus \{(0,0)\}$ tels que

$$\alpha f - \beta g = 0$$

μ-presque partout (ceci se retrouvait au travers de la discussion générale ci-dessus à propos du cas d'égalité dans la chaîne d'inégalités (4.9)). Quant à l'égalité

$$\|f + g\|_2^2 = \|f\|_2^2 + \|g\|_2^2 + 2 \operatorname{Re} \Big(\int_\Omega f \overline{g} \, d\mu \Big)$$

pour $f, g \in \mathcal{L}^2_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$, c'est une incarnation du *théorème de Pythagore* [5].

Remarque 4.5. On peut à juste titre se poser la question de savoir pourquoi ne considérer que le cas $p \in [1, \infty]$ et non le cas $p \in]0, 1[$ dans l'énoncé des inégalités (4.5) et (4.8). En fait, si $p \in]0, 1[$ et si f, g sont deux fonctions mesurables positives $\Omega \longrightarrow [0, \infty]$, et si p' (cette fois négatif) est défini par la relation

$$\frac{1}{p'} = 1 - \frac{1}{p} ,$$

l'inégalité de Hölder se trouve « renversée » , au sens où l'on a

$$\int_\Omega fg \, d\mu \geq \Big(\int_\Omega f^p \, d\mu \Big)^{1/p} \Big(\int_\Omega g^{p'} \, d\mu \Big)^{1/p'}$$

avec les conventions habituelles $(1/0)^{p'} = \infty$ et $0 \times \infty = 0$ (voir l'exercice proposé dans le guide d'activités 8 sous Ulysse). Dès lors, l'inégalité de Minkowski (conséquence, on l'a vu, de celle de Hölder) se trouve aussi renversée, au sens où l'on a, si $p \in]0, 1[$ et si f et g sont des fonctions positives mesurables de Ω dans $[0, \infty]$,

$$\Big(\int_\Omega (f + g)^p \, d\mu \Big)^{1/p} \geq \Big(\int_\Omega f^p \, d\mu \Big)^{1/p} + \Big(\int_\Omega g^p \, d\mu \Big)^{1/p} .$$

Il n'est donc pas question de faire jouer à $\| \ \|_p$ lorsque $p \in]0, 1[$ le rôle de semi-norme sur le $\mathbb{K}$-espace vectoriel des fonctions $(\Omega, \mathcal{T})$-$(\mathbb{K}, \mathcal{B}(\mathbb{K}))$ mesurables telles que

$$\int_\Omega |f|^p \, d\mu < +\infty ,$$

l'inégalité triangulaire de (4.4) se présentant « à l'envers » dans ce cas !

Lorsque $\mu(\Omega) < \infty$, les inclusions (4.2) de la remarque 4.1 peuvent être précisées en utilisant le fait que, si $p \in [1, \infty]$, la fonction constante et égale à 1 appartient à tous les espaces $\mathcal{L}^{p'}_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$). On a donc, si $1 \leq p_1 < p_2 \leq \infty$ et si $f \in \mathcal{L}^{p_2}_{\mathbb{C}}(\Omega, \mathcal{T}, \mu)$, et en utilisant l'inégalité (4.9) avec $p = p_2/p_1$,

$$\|f\|_{p_1}^{p_1} = \int_\Omega (|f|^{p_1} \times 1) \, d\mu \leq \Big(\int_\Omega (|f|^{p_1})^{p_2/p_1} \, d\mu \Big)^{p_1/p_2} \times (\mu(\Omega))^{\frac{p_2 - p_1}{p_2}} ,$$

soit

$$\|f\|_{p_1} \leq (\mu(\Omega))^{\frac{1}{p_1} - \frac{1}{p_2}} \|f\|_{p_2} , \tag{4.11}$$

inégalité permettant de quantifier les inclusions (4.2) ; cette inégalité subsiste si $1 \leq p_1 < p_2 = \infty$, ce qui permet de quantifier la première assertion de la remarque 4.3.

La première assertion de la remarque 4.3 peut même être ainsi précisée :

5. Voir le cours de MHT613 qui traitera plus en profondeur des particularité du cadre $p = 2$.

Proposition 4.4 *Si $\mu(\Omega) < +\infty$ et si $f \in \mathcal{L}^\infty_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) alors f est dans tous les $\mathbb{K}$-espaces vectoriels $\mathcal{L}^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ pour $p \in [1, \infty]$ et on a*

$$\|f\|_\infty = \lim_{p\to+\infty} \|f\|_p \,.$$

Preuve. D'après l'inégalité (4.11) appliquée avec $p \in [1, \infty[$ et $p_2 = \infty$, on a

$$\|f\|_p \leq (\mu(\Omega))^{1/p} \|f\|_\infty$$

et, en passant à la limite supérieure lorsque p tend vers $+\infty$,

$$\limsup_{p\to+\infty} \|f\|_p \leq \|f\|_\infty \,.$$

Si $\alpha < \|f\|_\infty$, l'ensemble $E_\alpha := \{|f| > \alpha\}$ est de μ-mesure strictement positive et on a donc

$$\|f\|_p \geq \Big(\int_{E_\alpha} |f|^p \, d\mu\Big)^{1/p} \geq \alpha(\mu(E_\alpha))^{1/p} \,;$$

en passant à la limite inférieure lorsque p tend vers $+\infty$, il vient

$$\liminf_{p\to+\infty} \|f\|_p \geq \alpha \,;$$

ceci étant vrai pour tout $\alpha < \|f\|_\infty$, on a bien

$$\liminf_{p\to+\infty} \|f\|_p \geq \|f\|_\infty$$

et la proposition est donc démontrée puisque

$$\|f\|_\infty \leq \liminf_{p\to+\infty} \|f\|_p \leq \limsup_{p\to+\infty} \|f\|_p \leq \|f\|_\infty \,. \qquad \Diamond$$

Deux autres inégalités (que l'on pourrait qualifier d'*inégalités de convexité*), liées aux inégalités de Hölder (4.5) et Minkowski (4.8) seront appelées à jouer un rôle intéressant dans la « géométrie » des espaces $\mathcal{L}^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ lorsque $p \in [2, +\infty[$; ce sont les inégalités dites *du parallélogramme* et de la *médiane.*

Proposition 4.5 *Soit $p \in [2, \infty[$ et f, g, h trois éléments de $\mathcal{L}^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$). On a l'inégalité dite du parallélogramme*

$$\|f - g\|_p^p + \|f + g\|_p^p \leq 2^{p-1}(\|f\|_p^p + \|g\|_p^p) \tag{4.12}$$

et celle de la médiane

$$\Big\|h - \frac{f+g}{2}\Big\|_p^p + \frac{\|f-g\|_p^p}{2^p} \leq \frac{\|h - f\|_p^p + \|h - g\|_p^p}{2} \,. \tag{4.13}$$

Preuve. Il résulte du fait que $t^{p/2} \leq t$ et $(1-t)^{p/2} \leq (1-t)$ pour tout $t \in [0, 1]$ lorsque $p/2 \geq 1$, que, si u et v sont deux nombres positifs non tous les deux nuls

$$\Big(\frac{u^2}{u^2+v^2}\Big)^{p/2} + \Big(\frac{v^2}{u^2+v^2}\Big)^{p/2} \leq 1 \,,$$

soit

$$u^p + v^p \leq (u^2 + v^2)^{p/2} \,.$$

Cette dernière inégalité subsiste si $u, v \geq 0$. En prenant $u = |f(\omega) - g(\omega)|$ et $v = |f(\omega) - g(\omega)|$ et en intégrant sur Ω relativement à la mesure μ, il vient, du fait de l'inégalité de Pythagore

$$\sqrt{|z-w|^2 + |z+w|^2} \leq \sqrt{2} \times \sqrt{|z|^2 + |w|^2} \quad \forall\, z, w \in \mathbb{C}\,,$$

$$\begin{aligned} \|f-g\|_p^p + \|f+g\|_p^p &\leq \int_\Omega (|f-g|^2 + |f+g|^2)^{p/2}\, d\mu \\ &\leq \int_\Omega [\sqrt{2}(|f|^2 + |g|^2)]^{p/2}\, d\mu \\ &\leq 2^{p/2} \int_\Omega (|f|^2 + |g|^2)^{p/2}\, d\mu\,. \end{aligned} \tag{4.14}$$

Si l'on utilise maintenant l'inégalité de Minkowski (4.8) avec $p/2 \geq 1$, on a

$$\begin{aligned} \int_\Omega (|f|^2 + |g|^2)^{p/2}\, d\mu &= \Big\| |f|^2 + |g|^2 \Big\|_{p/2}^{p/2} \\ &\leq \Big[\Big(\int_\Omega |f|^p\, d\mu \Big)^{2/p} + \Big(\int_\Omega |g|^p\, d\mu \Big)^{2/p} \Big]^{p/2}. \end{aligned} \tag{4.15}$$

Or, si U et V sont deux nombres positifs, il suit de l'inégalité (4.5) appliquée sur l'ensemble $\{0, 1\}$ avec la mesure de décompte et toujours $p/2$ que

$$U + V \leq (U^{p/2} + V^{p/2})^{2/p} \times (1+1)^{1-\frac{2}{p}}\,,$$

soit

$$(U+V)^{p/2} \leq 2^{p/2-1}(U^{p/2} + V^{p/2})\,.$$

En utilisant cette inégalité avec

$$U := \Big(\int_\Omega |f|^p\, d\mu \Big)^{2/p} \qquad V := \Big(\int_\Omega |g|^p\, d\mu \Big)^{2/p}$$

et en reportant au second membre de (4.15), on trouve

$$\int_\Omega (|f|^2 + |g|^2)^{p/2}\, d\mu \leq 2^{p/2-1}(\|f\|_p^p + \|g\|_p^p)\,.$$

En combinant avec l'inégalité (4.14), on obtient bien l'inégalité du parallélogramme (4.12). Si l'on applique cette inégalité avec $h - f$, $h - g$ en place de f et g, on obtient immédiatement l'inégalité de la médiane (4.13). $\diamondsuit$

Remarque 4.6. Si $p = 2$, les inégalités (4.12) et (4.13) sont en fait des égalités dites respectivement *formule du parallélogramme* et *formule de la médiane.*

4.3 Les espaces de Banach $L^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$, $p \in [1, \infty]$

Sur chaque espace $\mathcal{L}^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ (où $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$), on dispose, d'après la proposition 4.2, d'une semi-norme $f \longmapsto \|f\|_p$. Cette semi-norme n'est pas une norme car $\|f\|_p = 0$ équivaut (d'après la proposition 2.4) à $f = 0$ μ-presque partout.

Tentons donc de « rectifier » le tir.

Le $\mathbb{K}$-espace vectoriel des fonctions $(\Omega,\mathcal{T})$-$(\mathbb{K},\mathcal{B}(\mathbb{K}))$ mesurables et nulles μ presque partout est un $\mathbb{K}$-espace vectoriel $\mathcal{F}_{\mathbb{K},0}$ que l'on peut considérer comme un sous-$\mathbb{K}$-espace vectoriel de chaque $\mathcal{L}^p_{\mathbb{K}}(\Omega,\mathcal{T},\mu)$, $p\in[1,\infty]$. On définit, pour $p\in[1,\infty]$, le $\mathbb{K}$-espace vectoriel $L^p_{\mathbb{K}}(\Omega,\mathcal{T},\mu)$ comme le $\mathbb{K}$-*espace vectoriel quotient*

$$L^p_{\mathbb{K}}(\Omega,\mathcal{T},\mu) := \frac{\mathcal{L}^p_{\mathbb{K}}(\Omega,\mathcal{T},\mu)}{\mathcal{F}_{\mathbb{K},0}}\,,$$

c'est-à-dire, l'ensemble des classes d'équivalence $\dot{f}$ pour la relation d'équivalence

$$f\mathcal{R}g \iff f-g\in\mathcal{F}_{\mathbb{K},0} \iff f=g \quad \mu-\text{presque partout}\,,$$

équipé de l'addition interne

$$\dot{f}+\dot{g} = \text{classe de}\,(f+g)$$

et de la multiplication externe par un scalaire de $\mathbb{K}$

$$\lambda\dot{f} := \text{classe de}\,(\lambda f)\,.$$

Sur le $\mathbb{K}$-espace vectoriel $L^p_{\mathbb{K}}(\Omega,\mathcal{T},\mu)$, on définit, pour $p\in[1,\infty]$, une norme $\|\ \|_p$ en posant

$$\|\dot{f}\|_p = \inf\{\|f\|_p\,;\ f\in\dot{f}\}\,.$$

De fait, la norme $\|\dot{f}\|_p$ de $\dot{f}$ est égale à la norme $\|f\|_p$ d'un représentant quelconque de la classe $\dot{f}$ dans $\mathcal{L}^p_{\mathbb{K}}(\Omega,\mathcal{T},\mu)$.

Remarque 4.7. L'idée de considérer les classes de fonctions intégrables $\dot{f}$ modulo la relation d'équivalence $\mathcal{R}$ au lieu de leurs représentants $f\in\dot{f}$ est tout-à-fait naturelle, si l'on songe par exemple au cas de $\Omega=\mathbb{R}^n$, $\mathcal{T}=\mathcal{B}(\mathbb{R}^n)$ ou $\mathcal{T}=\mathcal{L}(\mathbb{R}^n)$, μ mesure de Lebesgue : il est clair dans ce contexte que la notion de valeur ≪ ponctuelle ≫ d'une fonction intégrable f en un point spécifié x_0 de $\mathbb{R}^n$ n'a aucune signification réaliste, dans la mesure où l'évaluation d'une fonction en un nombre réel précis n'a aucune réalité physique (l'accès au point x_0 avec une marge d'erreur nulle s'avérant numériquement impossible) ! Ce qui compte (et se mesure physiquement) est de fait la ≪ répartition de masse ≫ de la fonction f, c'est-à-dire la connaissance de toutes les intégrales

$$\int_P f(x)\,dx$$

où P est un pavé fermé borné de $\mathbb{R}^n$; la valeur ≪ physique ≫ de la fonction f s'interprète comme la quantité ≪ moyenne ≫

$$\frac{\int_{x_0+[-\epsilon,\epsilon]^n} f(x)\,dx}{(2\epsilon)^n}$$

lorsque $\epsilon>0$ est choisi suffisamment petit (autant que la réalisation de l'expérience numérique permettant la mesure de f le permet) et non comme $f(x_0)$; on ne fait donc pas la différence entre les valeurs en x_0 (ainsi interprétées) lorsqu'il s'agit de représentants d'une même classe ! Il est intéressant de noter que de ce point de vue, les éléments $\dot{f}$ de $L^2_{\mathbb{K}}(\Omega,\mathcal{T},\mu)$ s'interprètent souvent physiquement comme des ≪ répartitions d'énergie ≫ , l'énergie en physique se présentant sous la forme d'une expression non linéaire en les données, mais quadratique (pensez par exemple à l'énergie cinétique $1/2mv^2$), d'où le rôle prépondérant des espaces $L^2_{\mathbb{K}}(\Omega,\mathcal{T},\mu)$ dans la hiérarchie des espaces $L^p_{\mathbb{K}}(\Omega,\mathcal{T},\mu)$, $p\in[1,\infty]$ [6] en profitant du puissant concept d'*orthogonalité* ou de *corrélation.*

6. Outre le fait que l'on peut y faire, avec les formules du parallélogramme (4.12) et de la médiane (4.13) de la remarque 4.6, de la géométrie ≪ à la Pythagore ≫ , comme on le mettra en évidence dans l'UE MHT613.

Le résultat majeur concernant les espaces $L^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ équipés chacun de sa norme $\| \, \|_p$, $p \in [1, \infty]$ (ici $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$), a été mis en évidence (et exploité) en 1907 par le mathématicien hongrois Frigyes Riesz (1880-1956) et le mathématicien autrichien Ernst Fischer (1875-1954)[7] :

Théorème 4.1 (théorème de Riesz-Fischer) *Si $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$, $p \in [1, \infty]$, le $\mathbb{K}$-espace vectoriel normé $L^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$, équipé de la norme $\| \, \|_p$, est un $\mathbb{K}$-espace vectoriel normé complet*[8].

Preuve. Soit $(\dot{f}_k)_{k \in \mathbb{N}}$ une suite de Cauchy de $L^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ pour la distance associée à la norme $\| \, \|_p$, ce qui signifie

$$\forall \epsilon > 0 \, , \; \exists N(\epsilon) \in \mathbb{N} \, , \; \forall k, l \geq N(\epsilon) \, , \; \|\dot{f}_k - \dot{f}_l\|_p \leq \epsilon \, . \tag{4.16}$$

En utilisant (4.16), on construit (en en construisant les entrées de proche en proche), une suite extraite $(f_{n_k})_{k \in \mathbb{N}}$ telle que

$$\forall k \in \mathbb{N} \, , \; \|\dot{f}_{n_{k+1}} - \dot{f}_{n_k}\|_p \leq 1/2^k \, . \tag{4.17}$$

Choisissons, pour chaque classe $\dot{f}_{n_k}$, $k \in \mathbb{N}$, un représentant dans $\mathcal{L}^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$, représentant que nous noterons f_{n_k} et introduisons la fonction $(\Omega, \mathcal{T})$-$([0, \infty], \mathcal{B})$ mesurable positive $g \; : \Omega \longrightarrow [0, \infty]$ définie par

$$\forall \omega \in \Omega \, , \; g(\omega) := |f_{n_0}(\omega)| + \sum_{k=0}^{\infty} |f_{n_{k+1}}(\omega) - f_{n_k}(\omega)| \, .$$

D'après l'inégalité de Minkowski (4.8) et le théorème 2.1 de convergence monotone de Beppo-Levi, on a

$$\begin{aligned}
\|g\|_p &= \lim_{N \to +\infty} \Big\| |f_{n_0}| + \sum_{k=0}^{N} |f_{n_{k+1}} - f_{n_k}| \Big\|_p \\
&\leq \lim_{N \to +\infty} \Big(\|f_{n_0}\|_p + \sum_{k=0}^{N} \|f_{n_{k+1}} - f_{n_k}\|_p \Big) \\
&\leq \|f_{n_0}\|_p + \sum_{k=0}^{\infty} \|f_{n_{k+1}} - f_{n_k}\|_p \\
&\leq \|f_{n_0}\|_p + \sum_{k=0}^{\infty} \|f_{n_{k+1}} - f_{n_k}\|_p = \|\dot{f}_{n_0}\|_p + \sum_{k=0}^{\infty} \|\dot{f}_{n_{k+1}} - \dot{f}_{n_k}\|_p \\
&\leq \|\dot{f}_{n_0}\|_p + \sum_{k=0}^{\infty} \frac{1}{2^k} < \infty \, .
\end{aligned} \tag{4.18}$$

7. Avec Ernst Fischer, Frigyes Riesz peut être considéré comme l'un des pères de l'analyse fonctionnelle moderne. C'est dans le cas $p = 2$ et à l'occasion d'un travail consacré aux séries de Fourier que s'est fait jour le théorème de Riesz-Fischer.

8. On dit aussi, en référence au mathématicien polonais Stefan Banach (1892-1945) qui explora toutes les facettes et tout l'intérêt de ce concept, un *espace de Banach*, c'est-à-dire un espace vectoriel normé où toute suite de Cauchy (au sens de la norme) est convergente (ou encore, ce qui est équivalent, toute série absolument convergente est convergente).

Il résulte de l'inégalité de Markov (proposition 2.3) que g^p, donc g, est une fonction finie μ-presque partout, donc que pour μ-presque tout $\omega \in \Omega$, la série

$$|f_{n_0}(\omega)| + \sum_{k=0}^{\infty} |f_{n_{k+1}}(\omega) - f_{n_k}(\omega)|$$

est convergente. Comme $\mathbb{K}$ est complet, toute série absolument convergente d'éléments de $\mathbb{K}$ est convergente et l'on en déduit, pour μ presque tout ω, la convergence de la série télescopique de premier terme $f_{n_0}(\omega)$ et de terme général $f_{n_{k+1}}(\omega) - f_{n_k}(\omega)$, ce qui signifie l'existence, pour presque tout $\omega \in \Omega$, de

$$f(\omega) = \lim_{k \to +\infty} f_{n_k}(\omega)\,.$$

On définit une fonction $(\Omega, \mathcal{T})$-$(\mathbb{K}, \mathcal{B})$ mesurable en prolongeant f par 0 sur le complémentaire (de μ-mesure nulle) de l'ensemble E sur lequel la suite $(f_{n_k})_k$ converge simplement vers f ; ce prolongement est aussi noté f.

Supposons d'abord $p \in [1, \infty[$. La fonction f est dans $\mathcal{L}^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$, puisque $|f| \leq g$ et que g^p est intégrable relativement à la mesure μ d'après (4.18). Fixons $\epsilon > 0$ et $l \geq N(\epsilon)$; grâce au lemme de Fatou (théorème 2.2), on a

$$\begin{aligned}\int_\Omega \liminf_{k \to +\infty} |f_l(\omega) - f_{n_k}(\omega)|^p \, d\mu &= \int_\Omega |f_l(\omega) - f(\omega)|^p \, d\mu \\ &\leq \liminf_{k \to +\infty} \Big(\int_\Omega |f_l(\omega) - f_{n_k}(\omega)|^p \, d\mu \Big) \leq \epsilon^p\,,\end{aligned} \tag{4.19}$$

l'inégalité finale résultant de la clause (4.16) puisque $n_k \geq N(\epsilon)$ pour k assez grand. En lisant (4.19), on constate que, pour $p \geq N(\epsilon)$, $\|f - f_l\|_p \leq \epsilon$, ce qui signifie, puisque ϵ était arbitraire

$$\lim_{l \to +\infty} \dot{f}_l = \dot{f}$$

dans $L^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ et conclut donc dans ce cas la preuve du théorème de Riesz-Fischer.

Le cas $p = \infty$ est encore plus simple à traiter ; on a $\|g\|_\infty < +\infty$, ce qui montre que $f \in \mathcal{L}^\infty_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$. En fixant $l \geq N(\epsilon)$ et en faisant courir k vers $+\infty$ dans $\|f_l - f_{n_k}\|_\infty \leq \epsilon$, on trouve $\|\dot{f}_l - \dot{f}\|_\infty \leq \epsilon$ car toute union dénombrable d'ensembles μ-négligeables (ici les ensembles $\{|f_l - f_{n_k}| > \|f_l - f_{n_k}\|_\infty\}$) est μ-négligeable. $\diamondsuit$

Le résultat auxiliaire suivant, outil intermédiaire dans la preuve du théorème de Riesz-Fischer, mérite d'être isolé tant il s'avère d'un usage très fréquent en théorie de l'intégration, dans la mesure où il nous permet de jeter un pont entre la convergence d'une suite $(\dot{f}_k)_k$ dans $L^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$) et la convergence simple μ-presque partout d'au moins une suite extraite de la suite $(f_k)_k$:

Proposition 4.6 *Soit $p \in [1, \infty]$ et $(\dot{f}_k)_k$ une suite d'éléments de $L^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ convergeant vers un élément $\dot{f}$ de $L^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ au sens de la norme $\| \ \|_p$ (i.e $\|\dot{f}_k - \dot{f}\|_p$ tend vers 0 lorsque k tend vers l'infini). Si f_k est, pour chaque k, un représentant de $\dot{f}_k$ dans $\mathcal{L}^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$, f un représentant de $\dot{f}$ dans $\mathcal{L}^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$, alors, il existe au moins une sous-suite $(f_{n_k})_k$ de la suite de fonctions $(f_k)_k$ convergeant simplement μ-presque partout vers la fonction f, i.e*

$$\lim_{k \to +\infty} f_{n_k}(\omega) = f(\omega)\,,\ \forall \omega \in E\,,\ \mu(\Omega \setminus E) = 0\,.$$

De plus toute autre suite extraite de la suite $(f_k)_k$ convergeant simplement vers une fonction limite $\tilde{f}_F$ sur un sous-ensemble $F \in \mathcal{T}$ tel que $\mu(F) > 0$ est telle cette fonction limite $\tilde{f}_F$ coïncide avec f μ-presque partout sur F [9].

4.4 À propos de dualité (encore Riesz-Fischer)

Une conséquence intéressante de l'inégalité de Hölder (4.9) est le résultat suivant ($\mathbb{K}$ étant toujours ici $\mathbb{R}$ ou $\mathbb{C}$), dont nous ne pouvons malheureusement donner dans ce cours qu'une version tronquée :

Proposition 4.7 *Soit $p \in]1, \infty[$ et $\dot{g}$ un élément de $L^{p'}_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$, où $1/p + 1/p' = 1$. L'application $\mathbb{K}$-linéaire*

$$L_{\dot{g}} \ : \dot{f} \in L^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu) \longmapsto \int_\Omega f g \, d\mu \in \mathbb{K}$$

(f et g étant des représentants arbitraires de $\dot{f}$ et $\dot{g}$ respectivement dans $\mathcal{L}^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ et $\mathcal{L}^{p'}_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$) est une forme linéaire $L_{\dot{g}}$ continue sur $L^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$, i.e un élément du $\mathbb{K}$-dual topologique $[L^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)]^*$ de l'espace de Banach $L^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$. De plus, on a l'égalité

$$\|L_{\dot{g}}\|^*_p := \sup_{\{\dot{f} \in L^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)\,;\, \dot{f} \neq 0\}} \frac{|L_{\dot{g}}(\dot{f})|}{\|\dot{f}\|_p} = \|\dot{g}\|_{p'} \tag{4.20}$$

Preuve. Que $L_{\dot{g}}(\dot{f})$ définisse un scalaire ne dépendant pas du choix des représentants de $\dot{f}$ et $\dot{g}$ résulte de la proposition 4.3. C'est d'ailleurs l'inégalité de Hölder (4.9) qui assure

$$\forall \dot{f} \in L^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)\,,\ |L_{\dot{g}}(\dot{f})| \leq \|\dot{f}\|_p \times \|\dot{g}\|_{p'}\,,$$

d'où la continuité de la forme $\mathbb{K}$-linéaire $L_{\dot{g}}$ et en prime l'inégalité

$$\|L_{\dot{g}}\|^*_p := \sup_{\{\dot{f} \in L^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)\,;\, \dot{f} \neq 0\}} \frac{|L_{\dot{g}}(\dot{f})|}{\|\dot{f}\|_p} \leq \|\dot{g}\|_{p'}\,. \tag{4.21}$$

Si $\dot{g} \neq 0$, choisissons un représentant g de $\dot{g}$ et considérons la fonction mesurable f_g définie par $f_g(\omega) = 0$ si $g(\omega) = 0$ et

$$f_g(\omega) := \frac{\overline{g(\omega)}}{|g(\omega)|^{2-p'}}$$

sinon ; on vérifie que $|f_g|^p = |g|^{p(p'-1)} = |g|^{p'}$ est intégrable, ce qui implique que $\dot{f}_g$ définisse un élément non nul de $L^p_K(\Omega, \mathcal{T}, \mu)$; or

$$L_{\dot{g}}(\dot{f}_g) = \int_\Omega |g|^{p'} \, d\mu = \|g\|_{p'} \, \|g\|^{1-1/p'}_{p'} = \|\dot{g}\|_{p'} \times \|\dot{f}_g\|^{p(1-1/p')}_p = \|\dot{g}\|_{p'} \times \|\dot{f}_g\|_p\,,$$

ce qui prouve que $\dot{f}_g$ réalise le sup dans (4.20). La proposition est prouvée. ♢

9. Si $p = \infty$, la convergence de $(\dot{f}_k)_k$ dans $L^\infty_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ vers $\dot{f}$ implique la convergence simple μ-presque partout de la suite $(f_k)_k$ des représentants vers un représentant f de $\dot{f}$.

Nous venons de prouver ici (lorsque $p \in]1,+\infty[$) qu'il existait une isométrie du $\mathbb{K}$-espace vectoriel $L^{p'}_{\mathbb{K}}(\Omega,\mathcal{T},\mu)$ (équipé de sa norme $\|\ \|_{p'}$) dans le $\mathbb{K}$-dual topologique $[L^p_{\mathbb{K}}(\Omega,\mathcal{T},\mu)]^*$ équipé de la norme $\|\ \|_p^*$ définie (comme dans (4.20) par

$$\|L\|_p^* := \sup_{\{\dot f \in L^p_{\mathbb{K}}(\Omega,\mathcal{T},\mu)\,;\,\dot f \neq 0\}} \frac{|L(\dot f)|}{\|\dot f\|_p}.$$

En fait, le résultat complet est un autre théorème très important de Riesz-Fischer, qui stipule que <u>cette isométrie est en fait surjective</u> et donc que le $\mathbb{K}$-dual topologique $[L^p_{\mathbb{K}}(\Omega,\mathcal{T},\mu)]^*$ du $\mathbb{K}$-espace de Banach $L^p_{\mathbb{K}}(\Omega,\mathcal{T},\mu)$, dual équipé de sa norme $\|\ \|_p^*$, s'identifie isométriquement au $\mathbb{K}$-espace de Banach $L^{p'}_{\mathbb{K}}(\Omega,\mathcal{T},\mu)$, où p' désigne l'exposant conjugué de p ($1/p' + 1/p = 1$). Ce résultat (que nous admettrons ici) étend un résultat lié au contexte hilbertien, celui du cas particulier très important $p = p' = 2$, qui sera, lui, étudié en détail dans l'UE MHT613.

Concernant le cas $p = 1$, on peut énoncer un résultat (encore malheureusement tronqué) du même type :

Proposition 4.8 *Soit Ω un ensemble, $\mathcal{T}$ une tribu sur Ω et $\mu\ : \mathcal{T} \longrightarrow [0,\infty]$ une mesure σ finie*[10]*. Soit $\dot g$ un élément de $L^\infty_{\mathbb{K}}(\Omega,\mathcal{T},\mu)$. L'application $\mathbb{K}$-linéaire*

$$L_{\dot g}\ : \dot f \in L^1_{\mathbb{K}}(\Omega,\mathcal{T},\mu) \longmapsto \int_\Omega fg\,d\mu \in \mathbb{K}$$

(f et g étant des représentants arbitraires de $\dot f$ et $\dot g$ respectivement dans $\mathcal{L}^1_{\mathbb{K}}(\Omega,\mathcal{T},\mu)$ et $\mathcal{L}^\infty_{\mathbb{K}}(\Omega,\mathcal{T},\mu)$) est une forme linéaire $L_{\dot g}$ continue sur $L^1_{\mathbb{K}}(\Omega,\mathcal{T},\mu)$, i.e un élément du $\mathbb{K}$-dual topologique $[L^1_{\mathbb{K}}(\Omega,\mathcal{T},\mu)]^*$ de l'espace de Banach $L^1_{\mathbb{K}}(\Omega,\mathcal{T},\mu)$. De plus, on a l'égalité

$$\|L_{\dot g}\|_1^* := \sup_{\{\dot f \in L^1_{\mathbb{K}}(\Omega,\mathcal{T},\mu)\,;\,\dot f \neq 0\}} \frac{|L_{\dot g}(\dot f)|}{\|\dot f\|_1} = \|\dot g\|_\infty \tag{4.22}$$

Preuve. Que $L_{\dot g}(\dot f)$ définisse un scalaire ne dépendant pas du choix des représentants de $\dot f$ et $\dot g$ résulte de la proposition 4.3. C'est d'ailleurs l'inégalité de Hölder (4.9) qui assure

$$\forall\, \dot f \in L^1_{\mathbb{K}}(\Omega,\mathcal{T},\mu)\,,\ |L_{\dot g}(\dot f)| \leq \|\dot f\|_1 \times \|\dot g\|_\infty\,,$$

d'où la continuité de la forme $\mathbb{K}$-linéaire $L_{\dot g}$ et en prime l'inégalité

$$\|L_{\dot g}\|_1^* := \sup_{\{\dot f \in L^1_{\mathbb{K}}(\Omega,\mathcal{T},\mu)\,;\,\dot f \neq 0\}} \frac{|L_{\dot g}(\dot f)|}{\|\dot f\|_1} \leq \|\dot g\|_\infty\,. \tag{4.23}$$

La clause de σ-finitude portant sur μ n'a pas servi encore jusqu'ici ; nous allons nous en servir pour établir l'inégalité inverse.

Soit g un représentant de $\dot g$, $\alpha < \|g\|_\infty$ et $E := \{|g| > \alpha\}$. L'ensemble E est de μ mesure strictement positive puisque $\alpha < \|g\|_\infty$ (se référer à la définition de $\|g\|_\infty$). Puisque μ est σ-finie, il existe un élément A de $\mathcal{T}$ inclus dans E tel que

10. Voir la définition 3.1.

$0 < \mu(A) < \infty$ [11]. Considérons la fonction mesurable f_A définie par $f_A(\omega) = 0$ si $\omega \notin A$ et

$$f_A(\omega) := \frac{\overline{g(\omega)}}{|g(\omega)|}$$

si $\omega \in A$; on vérifie que $\|f_A\|_1 = \mu(A) \in]0, \infty[$, ce qui implique que $\dot{f}_A$ définit un élément non nul de $L^1_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$; or

$$L_{\dot{g}}(\dot{f}_A) = \int_A f_A g \, d\mu = \int_A |g| \, d\mu \geq \alpha\mu(A) = \alpha\|f_A\|_1 \, .$$

On a donc $\|L_{\dot{g}}\|_1^* \geq \alpha$ et, comme ceci est vrai pour tout $\alpha < \|g\|_\infty$,

$$\|L_{\dot{g}}\|_1^* \geq \|g\|_\infty = \|\dot{g}\|_\infty \, .$$

Avec l'inégalité en sens contraire (4.23) déjà établie, la proposition est prouvée. $\diamondsuit$

Ici encore, l'isométrie exhibée dans la proposition 4.8 entre $L^\infty_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ (équipé de la norme $\| \ \|_\infty$) et le $\mathbb{K}$-dual topologique $[L^1_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)]^*$, équipé de la norme $\|L\|_1^*$ définie (comme dans (4.23)) par

$$\|L\|_1^* := \sup_{\{\dot{f} \in L^1_{\mathbb{K}}(\Omega, \mathcal{T}, \mu) \, ; \, \dot{f} \neq 0\}} \frac{|L(\dot{f})|}{\|\dot{f}\|_1}$$

est une <u>isométrie surjective</u> (on l'admettra). On peut donc identifier le $\mathbb{K}$-dual topologique

$$[L^1_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)]^*$$

(équipé de la norme $\| \ \|_1^*$) avec le $\mathbb{K}$-espace de Banach $L^\infty_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$, pourvu que la mesure μ soit σ finie ou tout au moins vérifie la clause restrictive que l'on a utilisé (voir la note [11]).

Remarque 4.8. Il existe bien un résultat analogue à celui énoncé dans les propositions 4.7 et 4.8 lorsque $p = \infty$: un élément $\dot{g}$ de $L^1_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ induit la définition d'un élément $L_{\dot{g}}$ du $\mathbb{K}$-dual topologique de $L^\infty_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$, par

$$L_{\dot{g}}(\dot{f}) := \int_\Omega fg \, d\mu \, ,$$

avec de plus

$$\|L_{\dot{g}}\|_\infty^* := \sup_{\{\dot{f} \in L^\infty_{\mathbb{K}}(\Omega, \mathcal{T}, \mu) \, ; \, \dot{f} \neq 0\}} \frac{|L(\dot{f})|}{\|\dot{f}\|_\infty} = \|\dot{g}\|_1 \, .$$

Pour l'inégalité ($\leq$), on utilise encore l'inégalité de Hölder de la proposition 4.3 ; pour l'inégalité en sens contraire ($\geq$), on prend un représentant g de $\dot{g}$ et on teste $L_{\dot{g}}$ sur l'élément particulier $\dot{f}_g$, de représentant la fonction définie par $f(\omega) = \overline{g(\omega)}/|g(\omega)|$ si $g(\omega) \neq 0$, $f(\omega) = 0$ sinon. Si $\|\dot{g}\|_1 > 0$ (ce que l'on suppose sinon $L_{\dot{g}} \equiv 0$), $f_g \in \mathcal{L}^\infty(\Omega, \mathcal{T}, \mu)$ avec $\|f_g\|_\infty = 1$ et l'on a $L_{\dot{g}}(\dot{f}_g) = \|g\|_1 = \|g\|_1 \times \|\dot{f}_g\|_\infty$, ce qui prouve $\|L_{\dot{g}}\|_\infty^* \geq \|g\|_1$. En revanche, tout s'arrête ici dans ce cas car ce résultat ne saurait se compléter comme les résultats des propositions 4.7 et 4.8 ; l'isométrie ainsi construite de $L^1_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ (équipé de la norme $\| \ \|_1$) dans le $\mathbb{K}$-dual topologique $[L^\infty_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)]^*$ (équipé de la norme $\| \ \|_\infty^*$) <u>n'est cette fois plus surjective</u>. Le $\mathbb{K}$-dual topologique

11. On a simplement utilisé la conséquence (faible) suivante de la σ-finitude de μ : dans toute partie E de $\mathcal{T}$ de mesure strictement positive, on peut trouver une autre partie $A \subset E$, $A \in \mathcal{T}$, telle que $\mu(A) \in]0, \infty[$; notons que la mesure de décompte sur un ensemble non dénombrable n'est pas σ-finie, mais vérifie pourtant cette propriété ; d'un autre côté, la mesure infinie sur toute partie finie ne la vérifie pas.

$[L^\infty_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)]^*$ (équipé de la norme $\| \ \|^*_\infty$) ne peut plus s'identifier à $L^1_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ (c'est un $\mathbb{K}$-espace de Banach plus gros que $L^1_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$!).

Un commentaire ≪ heuristique ≫. En conclusion de cette section, nous pouvons dire de manière imagée que la hiérarchie des espaces $L^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ (tout au moins pour $p \in]1, +\infty[$) se réfléchit par dualité topologique autour de l'espace $L^2_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ (qui lui est isométrique à son dual topologique). On retrouvera cette opération de dualité dans l'UE MHT613 avec la transformation de Fourier (ou encore ≪ prise de spectre ≫) : elle est matérialisée physiquement par un mécanisme optique, la *diffraction* au travers d'une lentille. Ainsi pourrait-on comparer le rôle de $L^2_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ à celui d'une lentille de part et d'autre de laquelle $L^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ pour $p \in]2, \infty[$ se trouve ≪ diffracté ≫ en un espace isométriquement égal à son dual topologique, à savoir $L^{p'}_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$, avec $1/p + 1/p' = 1$ et donc cette fois $p' \in]1, 2[$. On a indiqué dans le guide d'activités 10 comment profiter de certaines propriétés géométriques des espaces $L^p_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$ pour $p \in [2, \infty[$ (en l'occurrence les inégalités du parallélogramme (4.12) et de la médiane (4.13)) en transposant (par dualité) leurs effets aux univers duaux $L^{p'}_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$, $1 < p' \leq 2$.

4.5 L'opération de convolution

4.5.1 Une opération omniprésente en traitement de l'information

Nous allons pour introduire (en en justifiant l'importance) l'opération mathématique de *convolution*, nous placer dans le cadre discret modélisé ici par le $\mathbb{K}$-espace vectoriel $l^1_{\mathbb{K}}(\mathbb{Z})$. En traitement de l'information, ce cadre est le cadre naturel : les éléments de $l^1_{\mathbb{K}}(\mathbb{Z})$ sont dits ≪ signaux discrets intégrables ≫ à valeurs dans $\mathbb{K}$ (ici $\mathbb{R}$ ou $\mathbb{C}$).

Supposons que l'on dispose d'un appareil (on dit aussi une ≪ *boîte noire* ≫ ou encore un ≪ *filtre* ≫), transformant l'élément $s = (s_k)_{k\in\mathbb{Z}}$ de $l^1_{\mathbb{K}}(\mathbb{Z})$ qu'on lui fournit en entrée en un élément $\mathcal{L}[s] \in l^1_{\mathbb{K}}(\mathbb{Z})$, ce en respectant les trois clauses suivantes :
- l'action $\mathcal{L}$ est linéaire ;
- l'action $\mathcal{L}$ est continue ;
- les paramètres de l'appareil sont immuables dans le temps, ce qui signifie que si l'on donne en entrée la suite $(s_{k-k_0})_{k\in\mathbb{Z}}$, alors on doit avoir en sortie $(\mathcal{L}[s]_{k-k_0})_{k\in\mathbb{Z}}$.

Exemple 4.1. Pour illustrer ceci par des exemples concrets empruntés à la physique, une cellule électronique (du type RLC, résistance-bobine-condensateur), une cellule mécanique (conçue autour de ressorts), un micro enregistreur, un tube de télévision, *etc.* jouent le rôle de filtres ; en revanche, un instrument de musique ou un orchestre en concert, l'orgue vocal constitué par la glotte et les cordes vocales d'un individu lors de la prononciation d'une suite de phonèmes, ne sont pas des modèles de filtres, car les paramètres des appareils en jeu sont ici appelés à se transformer au cours du temps (un violon jouant les aigus ou les graves n'est pas la même boîte noire !). On conçoit cependant l'omniprésence de la notion de filtre dans les problèmes de modélisation d'appareil.

En introduisant la base canonique $\{e_l \,;\, l \in \mathbb{Z}\}$ de $l^1_{\mathbb{K}}(\mathbb{Z})$ et en écrivant un élément $u = (u_l)_{l\in\mathbb{Z}}$ de $l^1_{\mathbb{K}}(\mathbb{Z})$ sous la forme :

$$u = \sum_{l\in\mathbb{Z}} u_l \, e_l \,,$$

on voit, grâce aux clauses de linéarité et de continuité, que

$$\mathcal{L}[u] = \sum_{l\in\mathbb{Z}} u_l\, \mathcal{L}[e_l]\,.$$

Si $\mathcal{L}[e_0] = (h_l)_{l\in\mathbb{Z}}$, on voit (formellement pour l'instant au moins) que

$$\mathcal{L}[u] = ((u * h)_k)_k\,, \tag{4.24}$$

où

$$u * h := \Big(\sum_{l\in\mathbb{Z}} u_{k-l}h_l\Big)_{k\in\mathbb{Z}}.$$

En fait, tout peut être justifié grâce à la proposition suivante :

Proposition 4.9 *Soient $u = (u_l)_{l\in\mathbb{Z}}$ et $v = (v_l)_{l\in\mathbb{Z}}$ deux éléments de $l^1_{\mathbb{K}}(\mathbb{Z})$. Pour tout $k \in \mathbb{Z}$, la série bilatère*

$$\sum_{l\in\mathbb{Z}} u_{k-l}v_l = \sum_{l\in\mathbb{Z}} u_l v_{k-l}$$

est absolument convergente et l'on a

$$\sum_{k\in\mathbb{Z}} \Big| \sum_{l\in\mathbb{Z}} u_{k-l}v_l \Big| \leq \|u\|_1 \, \|v\|_1\,.$$

Le nouvel élément de $l^1_{\mathbb{K}}(\mathbb{Z})$ ainsi défini est noté

$$\Big(\sum_{l\in\mathbb{Z}} u_{k-l}v_l\Big)_{k\in\mathbb{Z}} := u * v$$

et appelé convolé de u et v. L'opération de convolution ainsi définie est associative, commutative, et admet pour élément neutre la suite $(e_{0,l})_l$ définie par $e_{0,0} = 1$ et $e_{0,l} = 0$ si $l \neq 0$.

Preuve. Nous donnerons la preuve dans un cadre plus général à la section suivante 4.5.2. Donnons là ici cependant dans ce cadre discret qui est celui de séries doubles et vous est probablement bien familier. On a

$$\begin{aligned}
\sum_{k\in\mathbb{Z}} \Big(\sum_{l\in\mathbb{Z}} |u_{k-l}|\, |v_l| \Big) &= \sum_{l\in\mathbb{Z}} |v_l| \Big(\sum_{k\in\mathbb{Z}} |u_{k-l}| \Big) \\
&= \sum_{l\in\mathbb{Z}} |v_l| \Big(\sum_{k\in\mathbb{Z}} |u_k| \Big) \\
&= \|u\|_1 \times \|v\|_1
\end{aligned}$$

d'après le théorème de Fubini-Tonelli (théorème 3.7) et l'invariance par translation de la mesure de comptage sur $\mathbb{Z}$. On déduit du théorème de Fubini (théorème 3.8) la convergence pour presque tout k (mais ici en fait tout k car on travaille avec la mesure de comptage) de la série bilatère

$$\sum_{l\in\mathbb{Z}} u_{k-l}v_l\,,$$

avec de plus le fait que

$$\Big(\sum_{l\in\mathbb{Z}} u_{k-l}v_l\Big)_{k\in\mathbb{Z}} := u*v \in l^1_{\mathbb{K}}(\mathbb{Z})$$

et vérifie $\|u*v\|_1 \leq \|u\|_1 \times \|v\|_1$. La commutativité de l'opération de convolution vient encore une fois de l'invariance de la mesure de comptage sur $\mathbb{Z}$ par translation. L'associativité est immédiate aussi, ainsi que le fait que e_0 est élément neutre. L'opération interne est continue et de norme inférieure ou égale à 1 puisque $\|u*v\|_1 \leq \|u\|_1 \times \|v\|_1$; la norme est en fait égale à 1 puisque $u*e_0 = u$ pour tout u et que $\|e_0\| = 1$. $\diamondsuit$

4.5.2 Convolution entre éléments de $L^p_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$ ou $l^p_{\mathbb{K}}(\mathbb{Z}^n)$

Le cas de $\mathbb{R}^n$

Dans cette sous-section, on travaillera sur $\mathbb{R}^n$ dont on profitera de manière essentielle de la structure de « groupe topologique » : $(\mathbb{R}^n, +)$ est un groupe additif, l'addition $(x,y) \longmapsto x+y$ étant une application continue pour la topologie usuelle de $\mathbb{R}^n$. Ce groupe topologique $(\mathbb{R}^n, +)$ a aussi la propriété importante d'être *localement compact*, *i.e* tout point admet un voisinage compact. On désignera toujours par $\mathbb{K}$ soit le corps des $\mathbb{K} = \mathbb{R}$ des réels, soit le corps $\mathbb{K} = \mathbb{C}$ des complexes.

Enfin, il existe sur $\mathbb{R}^n$ une mesure (en l'occurrence la mesure de Lebesgue) invariante sous l'action de groupe (ici la translation dans $\mathbb{R}^n$) et c'est de cela que nous allons beaucoup profiter pour introduire la notion de convolution des classes de fonctions mesurables.

Si $\dot{f} \in L^p_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$ et $\dot{g} \in L^q_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$ avec $p \in [1,\infty]$ et $q \in [1,\infty]$, on aimerait pouvoir affirmer que, si f et g sont des représentants arbitraires respectivement de $\dot{f}$ et $\dot{g}$ (à valeurs dans $\mathbb{K}$), la fonction mesurable

$$y \longmapsto f(x-y)\,g(y)$$

est intégrable par rapport à la mesure de Lebesgue pour $x \in \mathbb{R}^n \setminus E$, où E est dx-négligeable et que de plus, la fonction F définie par

$$F(x) := \int_{\mathbb{R}^n} f(x-y)g(y)\,dy \tag{4.25}$$

si $x \in E$, $F(x) = 0$ sinon, est une fonction mesurable représentant si possible une certaine classe $\dot{F}$ de $L^r_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$ avec un r convenable (lié à p et q) dans $[1,\infty]$.

Il s'avère que tout ceci est vrai dès que

$$\frac{1}{p} + \frac{1}{q} \geq 1\,,$$

et ce pour $r \in [1,\infty]$ défini par

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

avec de plus en prime l'inégalité dite de Young [12]

$$\|\dot{F}\|_r \leq \|\dot{f}\|_p \times \|\dot{g}\|_q\,. \tag{4.26}$$

Avant d'énoncer le résultat précis, indiquons comment s'interprète la définition (4.25) du produit de convolution de deux classes $\dot{f}$ et $\dot{g}$ aux travers du choix de représentants f et g. Si, pour fixer les idées, on suppose $\dot{g}$ normalisée en norme $\|\ \|_q$, au sens où

$$\|\dot{g}\|_q = 1\,,$$

alors, lorsqu'elle est définie (c'est-à-dire pour presque tout x), l'intégrale

$$F(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\,dy = \int_{\mathbb{R}^n} f(x+y)\,g(-y)\,dy$$

s'interprète comme la « moyenne » de $y \longmapsto f(x+y)$, pondérée par les valeurs de la fonction « moyennisante » $y \longmapsto g(-y) := \check{g}(y)$; la fonction F, dx-presque partout définie, s'interprète donc comme une « moyenne glissante » de f par la fonction $\check{g}$. Lorsque par exemple $n = 1$, on peut imaginer le graphe de $\check{g}$ tracé sur une feuille de papier calque, feuille que l'on fait glisser le long de l'axe des abscisses au dessus du graphe de f ; la fonction f se trouve à chaque instant « moyennée » suivant la moyennisation indiquée par le graphe de la fonction $\check{g}$ ainsi translaté.

Formalisons le résultat principal de cette section :

Théorème 4.2 (inégalités de W.H. Young) *Soient p, q, r trois éléments de $[1, \infty]$ tels que*

$$\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}\,. \tag{4.27}$$

Soient $\dot{f} \in L^p_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$ et $\dot{g} \in L^q_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$ et f, g des représentants respectifs de $\dot{f}$ et $\dot{g}$ dans $\mathcal{L}^p_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$ et $\mathcal{L}^q_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$. La fonction

$$y \in \mathbb{R}^n \longmapsto f(x-y)g(y)$$

est intégrable pour tout x dans $\mathbb{R}^n \setminus E$, où E est un sous-ensemble négligeable de $\mathbb{R}^n$ (au sens de la mesure de Lebesgue). De plus, la fonction F, définie par

$$F(x) = \int_{\mathbb{R}^n} f(x-y)g(y)\,dy \tag{4.28}$$

si $x \in \mathbb{R}^n \setminus E$ et $F(x) = 0$ sinon, est une fonction $(\mathbb{R}^n, \mathcal{B})$-$(\mathbb{K}, \mathcal{B}(\mathbb{K}))$-mesurable, représentant un élément $\dot{F}$ de $L^r_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$, avec

$$\|\dot{F}\|_r \leq \|\dot{f}\|_p \times \|\dot{g}\|_q\,. \tag{4.29}$$

La classe est dite <u>convolée</u> (ou aussi <u>convoluée</u>) des classes $\dot{f}$ et $\dot{g}$, l'opération de convolution étant commutative ; cette classe est notée $\dot{F} = \dot{f} * \dot{g}$. [13]

12. William Henry Young (1863-1942), analyste anglais, s'est en particulier intéressé au calcul différentiel et à certaines inégalités sous-jacentes, dont l'inégalité (4.6) de convexité de la fonction exponentielle (ou plutôt ici de concavité de la fonction logarithme) exploitée lors de la preuve de la proposition 4.1, dite d'ailleurs *inégalité de Young*.

13. Au niveau des représentants, on note aussi $F = f * g$.

Preuve. Deux cas particuliers sont particulièrement importants et ce sont des cas où la démonstration de l'inégalité (4.29) est directe : le cas où p et q sont des exposants conjugués, c'est-à-dire où

$$\frac{1}{p} + \frac{1}{q} = 1\,, \qquad r = \infty\,,$$

ainsi que le cas particulier où $p = 1$ et $q = r$ (incluant le cas particulier essentiel $p = q = r = 1$ déjà évoqué dans l'exemple 3.14 comme exemple d'application du « tandem » Fubini-Tonelli/Fubini). Nous envisagerons d'abord ces deux cas particuliers et renverrons pour ce qui concerne le cas général au guide d'activité 8 sous Ulysse (où ce problème est détaillé en une suite de 4 questions).

Si p et q sont conjugués ($1/p + 1/q = 1$), on peut appliquer la proposition 4.3 et l'inégalité (4.9) pour assurer que, pour tout $x \in \mathbb{R}^n$, la fonction

$$y \longmapsto f(x-y)g(y)$$

est dans $\mathcal{L}^1_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$ avec

$$\sup_{x\in\mathbb{R}^n} \Big| \int_\Omega f(x-y)g(y)\,dy \Big| \leq \|\dot{f}\|_p \times \|\dot{g}\|_q\,.$$

La fonction $F = f * g$ est donc bien définie (ici partout dans $\mathbb{R}^n$) par la formule (4.28) et on a

$$\|\dot{F}\|_\infty \leq \|F\|_u \leq \|\dot{f}\|_p \times \|\dot{g}\|_q\,,$$

d'où le résultat dans ce premier cas particulier.

Prenons $p = 1$ et $q = r < \infty$ et posons q' tel que $1/q' + 1/q = 1$.

$$\begin{aligned}
&\int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} |f(x-y)|\,|g(y)|\,dy \Big)^q dx = \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} |f(y)|\,|g(x-y)|\,dy \Big)^q dx \\
= \;& \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} |f(y)|^{1/q'}\,[|f(y)|^{1-1/q'}|g(x-y)|]\,dy \Big)^q dx \\
\leq \;& \|f\|_1^{q/q'} \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} |f(y)|^{q(1-1/q')}|g(x-y)|^q\,dy \Big) dx \\
\leq \;& \|f\|_1^{q/q'} \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} |f(y)|\,|g(x-y)|^q\,dy \Big) dx \\
\leq \;& \|f\|_1^{q/q'+1}\,\|g\|_q^q = (\|f\|_1\,\|g\|_q)^q = (\|\dot{f}\|_1\,\|\dot{g}\|_q)^q\,.
\end{aligned} \tag{4.30}$$

en utilisant la formule de changement de variables (ligne 1), l'inégalité de Hölder (4.5) avec les exposants conjugués q', q (passage des lignes 2 à 3), le théorème de Fubini-Tonelli (lignes suivantes) ; en extrayant les racines q-èmes, on obtient bien l'inégalité

$$\Big(\int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} |f(x-y)|\,|g(y)| \Big)^q dx \Big)^{1/q} \leq \|\dot{f}\|_1\,\|\dot{g}\|_q\,.$$

Il résulte de l'inégalité de Markov (proposition 2.3) qu'il existe un sous-ensemble dx négligeable E de $\mathbb{R}^n$, tel que, pour $x \in \mathbb{R}^n \setminus E$,

$$\int_{\mathbb{R}^n} |f(x-y)|\,|g(y)|\,dy < \infty\,.$$

On peut donc bien définir sur $\mathbb{R}^n \setminus E$ la fonction

$$x \longmapsto F(x) = \int_{\mathbb{R}^n} f(x-y)\,g(y)\,dy$$

et prolonger cette fonction par 0 pour $x \in E$ pour en faire une fonction $(\mathbb{R}^n, \mathcal{B})$-$(\mathbb{C}, \mathcal{B})$-mesurable, avec de plus

$$\Big(\int_{\mathbb{R}^n} |F(x)|^q \, dx\Big)^{1/q} \leq \Big(\int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} |f(x-y)| \, |g(y)| \, dy\Big)^q dx\Big)^{1/q} \leq \|\dot{f}\|_1 \, \|\dot{g}\|_q \,,$$

ce qui montre que F représente un élément $\dot{F}$ de $L^q_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$ et achève la preuve du théorème dans ce second cas particulier (le cas $p = 1$, $q = r = \infty$ se traitant encore plus facilement).

Dans le cas restant à étudier ($p \geq 1$, $1 < q, r < \infty$), on pose introduit r' avec $1/r + 1/r' = 1$ et l'on vérifie que, pour tout h dans $\mathcal{L}^{r'}_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$, on a, si $t = 1/p$, $s = 1/q$ et $\sigma = 1 - 1/r$, l'inégalité de Young

$$\begin{aligned} & \int_{\mathbb{R}^n \times \mathbb{R}^n} |f(x-y)| \, |g(y)| \, |h(x)| \, dxdy = \\ & \int_{\mathbb{R}^n \times \mathbb{R}^n} [|f(x-y)|^p]^t \, [|g(y)|^q]^s \, [|h(x)|^{1/\sigma}]^\sigma \, dxdy \\ \leq \; & (1-t)\|\dot{g}\|_q^q \, \|\dot{h}\|_{r'}^{r'} + (1-s)\|\dot{f}\|_p^p \, \|\dot{h}\|_{r'}^{r'} + (1-\sigma)\|\dot{f}\|_p^p \, \|\dot{g}\|_q^q \,, \end{aligned} \tag{4.31}$$

inégalité dont on déduit par homogénéité

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} |f(x-y)| \, |g(y)| \, |h(x)| \, dxdy \leq \|\dot{f}\|_p \, \|\dot{g}\|_q \, \|\dot{h}\|_{r'} \,. \tag{4.32}$$

On applique cette dernière inégalité à la fonction

$$h = h_N \;\; : x \longmapsto \inf\Big(N, \Big[\int_{\mathbb{R}^n} |f(x-y)| \, |g(y)| \, dy\Big]^{r-1}\Big) \times \chi_{\{\|x\| \leq N\}}(x) \,,$$

puis les théorèmes de Fubini-Tonelli et de Beppo Levi pour en déduire, faisant tendre N vers $+\infty$,

$$\Big(\int_{\mathbb{R}^n} \Big[\int_{\mathbb{R}^n} |f(x-y)| \, |g(y)| \, dy\Big]^r dx\Big)^{1/r} \leq \|\dot{f}\|_p \, \|\dot{g}\|_q \,,$$

ce qui permet de conclure exactement comme dans le second cas particulier étudié. Pour les détails de la preuve des inégalités (4.31) et (4.32), ainsi que pour la fin de la démonstration à partir de la suite $(h_N)_N$, on renvoie au problème détaillé dans le guide d'activités 8. $\diamondsuit$

Remarque 4.9. L'opération de convolution est une opération interne associative et commutative dans $L^1_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$; il ne saurait cependant y avoir d'élément neutre (voir l'exercice proposé sur le guide d'activité 10).

Le cas de $\mathbb{Z}^n$

Il est important de remarquer ici (et nous en reparlerons dans la section 4.8 à venir) que le théorème 4.2 s'adapte immédiatement (la preuve en étant laissée en exercice) au cadre discret, nous donnant l'opportunité de définir la convolution discrète sur le groupe $(\mathbb{Z}^n, +)$ en place du groupe $(\mathbb{R}^n, +)$; il suffit dans les preuves de remplacer $\mathbb{R}^n$ par $\mathbb{Z}^n$ et la mesure de Lebesgue dx par la mesure ν de décompte.

Comme toujours, $\mathbb{K}$ désignera dans cette section soit le corps des réels, soit le corps des complexes.

Il existe sur $\mathbb{Z}^n$ une mesure (en l'occurrence ici la mesure de décompte) invariante sous l'action de groupe (ici la translation dans $\mathbb{Z}^n$), ce qui va nous permettre, comme dans le cadre du groupe $(\mathbb{R}^n, +)$, d'introduire la notion de convolution des suites.

Théorème 4.3 (inégalités de W.H. Young discrètes) *Soient p, q, r trois éléments de $[1, \infty]$ tels que*

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

Soient $u = (u_l)_l \in l^p_{\mathbb{K}}(\mathbb{Z}^n)$ et $v = (v_l)_l \in l^q_{\mathbb{K}}(\mathbb{Z}^n)$. Pour tout $k \in \mathbb{Z}^n$, $(u_{k-l}\, v_l)_l$ est un élément de $l^1_{\mathbb{K}}(\mathbb{Z}^n)$ et la suite $(w_k)_{k\in\mathbb{Z}^n}$ de terme général

$$w_k := \sum_{l\in\mathbb{Z}^n} u_{k-l}\, v_l = \sum_{l\in\mathbb{Z}^n} u_l\, v_{k-l}$$

est un élément $w := u * v$ de $l^r_{\mathbb{K}}(\mathbb{Z}^n)$, dit convolé des suites $(u_l)_l$ et $(v_l)_l$, avec de plus

$$\|w\|_r = \Big(\sum_{k\in\mathbb{Z}^n} |w_k|^r\Big)^{1/r} \leq \Big(\sum_{k\in\mathbb{Z}^n} |u_k|^p\Big)^{1/p} \times \Big(\sum_{k\in\mathbb{Z}^n} |v_k|^q\Big)^{1/q}. \tag{4.33}$$

Les cas particuliers les plus importants de ce résultat sont le cas où p et q sont des exposants conjugués $(1/p + 1/q = 1, r = \infty)$ et le cas où $p = 1$, $q = r \in [1, \infty]$ (incluant le cas particulier majeur $p = q = r = 1$). Il est important de noter que si l'on introduit les séries formelles

$$\begin{aligned} U(X) &= \sum_{k\in\mathbb{Z}^n} u_k\, X_1^{k_1} \dots X_n^{k_n} \\ V(X) &= \sum_{k\in\mathbb{Z}^n} v_k\, X_1^{k_1} \dots X_n^{k_n} \end{aligned}$$

et si

$$W(X) := U(X) \times V(X)$$

(le produit étant le produit usuel entre séries formelles de puissances de $X_1, ..., X_n$), on a la très importante relation

$$W(X) = \sum_{k\in\mathbb{Z}^n} w_k\, X_1^{k_1} \dots X_n^{k_n}$$

lorsque $(w_k)_k$ est la convolée (si elle est définie) des suites $(u_l)_l$ et $(v_l)_l$.

Remarque 4.10. Dans le cas $n = 1$, si l'on a $u_l = v_l = 0$ pour $l < 0$ (on dit alors que u et v sont des éléments *suites causales*), il est toujours possible de définir

$$w_k := \sum_{l\in\mathbb{Z}} u_{k-l} v_l = \sum_{l=0}^{k} u_{k-l} v_l$$

pour tout $k \in \mathbb{Z}$; on constate que l'on définit ainsi une nouvelle suite causale $(w_k)_k$ et que l'opération $(u, v) \longmapsto w = u * v$ (si l'on ne la considère qu'au niveau des suites causales) n'est rien d'autre que le *produit de Cauchy*[14] des deux suites $(u_l)_{l\in\mathbb{N}}$ et $(v_l)_{l\in\mathbb{N}}$.

Remarque 4.11. L'opération de convolution est une opération interne associative et commutative dans $l^1_{\mathbb{K}}(\mathbb{Z}^n)$; il y a de plus cette fois un élément neutre, la suite $e_0 = (e_{0,l})_l$ définie par $e_{0,0} = 1$ et $e_{0,l} = 0$ si $l \neq 0$.

14. Voir le cours de l'UE MHT401, définition 1.4 de la section 1.5 des notes de cours de cette UE : http ://www.math.u-bordeaux1.fr/∼yger/mat401.html

4.6 Convolution sur $\mathbb{R}^n$ et régularisation

Comme on l'a vu avec le théorème 4.2, si l'opération de convolution

$$(\dot{f}, \dot{g}) \longmapsto \dot{f} * \dot{g}$$

réalise bien une opération interne sur le $\mathbb{K}$-espace de Banach $L^1_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$, cette opération (qui s'avère associative et commutative) n'admet pas d'élément neutre, alors que c'était le cas pour l'opération de convolution entre éléments de $l^1_{\mathbb{K}}(\mathbb{Z}^n)$ (voir la remarque 4.11).

Compte tenu précisément de la remarque 4.11, un élément neutre potentiel serait la « fonction » positive d'intégrale (ou « masse » pour employer un langage de physicien) égale à 1, toute la masse se trouvant concentrée à l'origine. Mais, problème, une telle fonction n'existe pas ! Seul le concept de mesure positive permet de modéliser un tel objet : c'est la mesure de Dirac (sur la tribu borélienne) introduite dans l'exemple 1.5 et définie par

$$\begin{aligned} \delta_0(A) &= 1 \text{ si } 0 \in A \\ \delta_0(A) &= 0 \text{ si } 0 \notin A \end{aligned}$$

qui serait en effet candidate à jouer ce rôle, mais ce n'est pas une fonction ! Il s'avère donc impérieux d'introduire la notion de « mesure de Dirac approchée » ou encore d'*approximation de la masse de Dirac* pour lui trouver un substitut dans l'espace $L^1_{\mathbb{R}}(\mathbb{R}^n, \mathcal{B}, dx)$.

Définition 4.3 *On appelle approximation de la masse de Dirac dans $L^1_{\mathbb{R}}(\mathbb{R}^n, \mathcal{B}, dx)$ toute suite $(\dot{\varphi}_k)_{k\in\mathbb{N}^*}$ d'éléments de $L^1_{\mathbb{R}}(\mathbb{R}^n, \mathcal{B}, dx)$ telle que :*

- *pour tout $k \in \mathbb{N}^*$, $\dot{\varphi}_k$ a un représentant $\varphi_k \in \mathcal{L}^1_{\mathbb{R}}(\mathbb{R}^n, \mathcal{B}, dx)$ tel que $\varphi_k \geq 0$ sur $\mathbb{R}^n$;*
- *$\forall\, k \in \mathbb{N}^*$, $\|\dot{\varphi}_k\|_1 = 1$;*
- *pour tout $\delta > 0$,*

$$\lim_{k\to+\infty} \Big(\int_{\|x\|\geq\delta} \varphi_k(x)\, dx \Big) = 0\,, \tag{4.34}$$

conditions traduisant bien le fait que toute la « masse » (égale à 1) de $\dot{\varphi}_k$ se concentre sur l'origine lorsque k tend vers $+\infty$.

Exemple 4.2. Un exemple très important du point de vue pratique peut être construit à partir de la fonction gaussienne

$$x \longmapsto g(x) := \frac{1}{(2\pi)^{n/2}} \exp\Big(-\frac{1}{2}(x_1^2 + \cdots + x_n^2) \Big)\,. \tag{4.35}$$

Si $(\epsilon_k)_{k\geq1}$ est une suite de nombres strictement positifs tendant vers 0, la suite $(\dot{g}_{\epsilon_k})_{k\geq1}$, où

$$g_\epsilon(x) := \frac{1}{\epsilon^n} g(x/\epsilon) = \frac{1}{(2\pi)^{n/2}\,\epsilon^n} \exp\Big(-\frac{1}{2\epsilon^2}(x_1^2 + \cdots + x_n^2) \Big) \tag{4.36}$$

est une approximation de la masse de Dirac dans $L^1_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$: en effet, on a $\|\dot{g}_\epsilon\|_1 = \|\dot{g}\|_1$ du fait de la formule de changement de variables dans les intégrales, de la formule bien connue

$$\int_{\mathbb{R}} e^{-t^2/2}\, dt = \sqrt{2\pi}\,,$$

et du théorème de Fubini-Tonelli et

$$\int_{\|x\|\geq\delta} g_\epsilon(x)\,dx = \int_{\|x\|\geq\delta/\epsilon} g(x)\,dx$$

(même raison), ce qui implique grâce au théorème de convergence dominée de Lebesgue (théorème 2.3)

$$\lim_{k\to+\infty}\Big(\int_{\|x\|>\delta} g_{\epsilon_k}(x)\,dx\Big) = 0\,.$$

On verra dans le cours de l'UE MHT613 la raison pour laquelle la gaussienne ainsi normalisée g joue un rôle si important pour générer ainsi des approximations de la masse de Dirac dans $L^1_{\mathbb{R}}(\mathbb{R}^n,\mathcal{B},dx)$ [15].

Exemple 4.3. La fonction radiale φ définie par

$$\begin{aligned}\varphi(x) &= \exp\Big[-\frac{1}{1-\|x\|^2}\Big] \quad \text{si} \quad \|x\|<1\\ \varphi(x) &= 0 \quad \text{sinon}\end{aligned} \tag{4.37}$$

est une fonction de classe C^∞ à support dans la boule unité fermée de $\mathbb{R}^n$; cela découle facilement du fait que, pour tout entier $N\in\mathbb{N}^*$,

$$\lim_{\rho\longrightarrow 1^-}\frac{1}{(1-\rho)^N}\exp\Big[-\frac{1}{1-\rho}\Big] = 0$$

(l'exponentielle imposant son comportement asymptotique aux fonctions puissance). Pour les mêmes raisons que dans l'exemple 4.2, si $(\epsilon_k)_{k\geq1}$ est une suite de nombres strictement positifs, la suite $(\varphi_{\epsilon_k})_{k\geq1}$, où

$$\varphi_\epsilon(x) := \frac{1}{\epsilon^n}\frac{\varphi(x/\epsilon)}{\|\dot\varphi\|_1} \tag{4.38}$$

est une approximation de la masse de Dirac dans $L^1_{\mathbb{R}}(\mathbb{R}^n,\mathcal{B},dx)$, faite cette fois de fonctions C^∞ à support compact (le support de φ_ϵ étant la boule fermée de centre 0 et de rayon ϵ lorsque $\epsilon>0$).

L'importance des approximations de la masse de Dirac tient au résultat suivant, que nous interpréterons un peu plus loin (dans la section 4.7) comme un résultat de « régularisation » :

Proposition 4.10 *Soit $\mathbb{K}=\mathbb{R}$ ou $\mathbb{C}$ et $p\in[1,\infty[$ et $\dot f\in L^p_{\mathbb{K}}(\mathbb{R}^n,\mathcal{B},dx)$. Si $(\dot\varphi_k)_{k\geq1}$ est une approximation de la masse de Dirac dans $L^1_{\mathbb{R}}(\mathbb{R}^n,\mathcal{B},dx)$, la suite $(\dot\varphi_k*\dot f)_{k\geq1}$ est une suite d'éléments de $L^p_{\mathbb{K}}(\mathbb{R}^n,\mathcal{B},dx)$ convergeant vers $\dot f$ dans le $\mathbb{K}$-espace de Banach $L^p_{\mathbb{K}}(\mathbb{R}^n,\mathcal{B},dx)$ équipé de la norme $\|\ \|_p$.*

Preuve. Que $\dot\varphi_k*\dot f$ définisse un élément de $L^p_{\mathbb{K}}(\mathbb{R}^n,\mathcal{B},dx)$ résulte du théorème 4.2 (inégalités de Young). Soit $k\in\mathbb{N}^*$. Si f est un représentant de $\dot f$, la fonction

$$x\longmapsto\int_{\mathbb{R}^n} f(x-y)\,\varphi_k(y)\,dy$$

est définie dx-presque partout et l'on peut donc écrire, pour x hors d'un ensemble dx-négligeable,

$$\int_{\mathbb{R}^n} f(x-y)\,\varphi_k(y)\,dy - f(x) = \int_{\mathbb{R}^n}(f(x-y)-f(x))\,\varphi_k(y)\,dy$$

15. C'est la plus « efficace » des fonctions en termes de la localisation simultanée de la fonction et de son spectre, celle qui tempère le mieux le célèbre principe d'incertitude que le physicien allemand Wermer Heisenberg introduisit dès 1927, à l'aube des développements de la mécanique quantique.

puisque $\varphi_k \geq 0$ (on a choisi un tel représentant pour $\dot{\varphi}_k$, ce qui est licite) et que

$$\int_{\mathbb{R}^n} \varphi_k(y)\, dy = \|\dot{\varphi}_k\|_1 = 1 \quad \forall\, k \in \mathbb{R}^n\,.$$

Introduisons p' tel que $1/p + 1/p' = 1$ et inspirons nous de la chaîne d'inégalités (4.30) qui nous a permis de résoudre le second cas particulier du théorème 4.2 :

$$\begin{aligned}
\|\dot{\varphi}_k * \dot{f} - \dot{f}\|_p^p & \quad \int_{\mathbb{R}^n} \Big| \int_{\mathbb{R}^n} (f(x-y) - f(x))\, \varphi_k(y)\, dy \Big|^p dx \\
& \leq \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} |f(x-y) - f(x)|\, \varphi_k(y)\, dy \Big)^p dx \\
& = \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} (\varphi_k(y))^{1/p'}\, [(\varphi_k(y))^{1-1/p'} |f(x-y) - f(x)|]\, dy \Big)^p dx \\
& \leq \|\dot{\varphi}_k\|_1^{p/p'} \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} (\varphi_k(y))^{p(1-1/p')} |f(x-y) - f(x)|^p\, dy \Big) dx \\
& \leq \int_{\mathbb{R}^n} \Big(\int_{\mathbb{R}^n} \varphi_k(y)\, |f(x-y) - f(x)|^p\, dy \Big) dx \\
& \leq \int_{\mathbb{R}^n} \varphi_k(y) \|\dot{f}(\cdot - y) - \dot{f}\|_p^p\, dy
\end{aligned} \tag{4.39}$$

en utilisant la formule de changement de variables (ligne 1), l'inégalité de Hölder (4.5) avec les exposants conjugués p', p (passage des lignes 3 à 4), le théorème de Fubini-Tonelli (lignes suivantes, surtout à la dernière étape).

Admettons pour l'instant le très important lemme suivant :

Lemme 4.1 *Soit $p \in [1, \infty[$ et $\dot{f}$ un élément de $L^p_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$; on a*

$$\lim_{y \to 0} \|\dot{f}(\cdot - y) - \dot{f}\|_p = 0\,. \tag{4.40}$$

Fin de la preuve de la proposition 4.10. Fixons $\epsilon > 0$ et choisissons $\delta > 0$ tel que

$$\|y\| \leq \delta \Longrightarrow \|\dot{f}(\cdot - y) - \dot{f}\|_p^p \leq \epsilon/2$$

(ce qui est possible d'après (4.40)). On a donc, d'après l'inégalité (4.39),

$$\begin{aligned}
\|\dot{\varphi}_k * \dot{f} - \dot{f}\|_p^p & \leq \int_{\|y\| \geq \delta} \|\dot{f}(\cdot - y) - \dot{f}\|_p^p\, \varphi_k(y)\, dy + \int_{\|y\| \leq \delta} \|\dot{f}(\cdot - y) - \dot{f}\|_p^p\, \varphi_k(y)\, dy \\
& \leq \int_{\|y\| \geq \delta} \|\dot{f}(\cdot - y) - \dot{f}\|_p^p\, dy + \frac{\epsilon}{2} \int_{\mathbb{R}^n} \varphi_k(y)\, dy \\
& \leq 2^{p+1} \|\dot{f}\|_p^p \int_{\|y\| \geq \delta} \varphi_k(y)\, dy + \frac{\epsilon}{2} \\
& \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon
\end{aligned}$$

pour k assez grand (dépendant de δ, donc de ϵ) du fait de la clause (4.34). La proposition est donc démontrée. $\diamondsuit$

Preuve du lemme 4.1. D'après la proposition 2.2 et le théorème 2.1 de Beppo Levi, toute fonction positive de $\mathcal{L}^p_{\mathbb{R}}(\mathbb{R}^n, \mathcal{B}, dx)$ est limite dans $\mathcal{L}^p_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$ (pour la semi-norme $\|\ \|_p$) d'une suite de fonctions étagées, et même d'une suite de combinaisons de fonctions caractéristiques d'ensembles mesurables bornés (donc intégrables). Si l'on complète avec le critère d'intégrabilité des sous-ensembles de $\mathbb{R}^n$ donné à la proposition 1.4, et que l'on passe des fonctions positives aux fonctions à valeurs complexes (mais dans $\mathcal{L}^p_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$) par prise de combinaison linéaire à coefficients complexes, on voit que tout élément de $\mathcal{L}^p_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$ peut s'approcher en norme $\| \ \|_p$ par une suite de combinaisons linéaires de fonctions indicatrices d'ouverts bornés. Pour tout $\epsilon > 0$, il existe une telle combinaison linéaire g de fonctions caractéristiques d'ouverts bornés (donc intégrables) telle que

$$\|\dot{f} - \dot{g}\|_p \leq \epsilon\,.$$

On écrit, du fait de l'inégalité triangulaire,

$$\begin{aligned}
\|\dot{f}(\cdot - y) - \dot{f}\|_p &\leq \|\dot{f}(\cdot - y) - \dot{g}(\cdot - y)\|_p + \|\dot{g}(\cdot - y) - \dot{g}\|_p + \|\dot{g} - \dot{f}\|_p \\
&\leq 2\|\dot{g} - \dot{f}\|_p + \|\dot{g}(\cdot - y) - \dot{g}\|_p \\
&\leq 2\epsilon + \|\dot{g}(\cdot - y) - \dot{g}\|_p\,.
\end{aligned}$$

Or, si P est un pavé borné, la frontière de P est de mesure nulle et le théorème de convergence dominée de Lebesgue 2.3 implique

$$\lim_{y \to 0} \int_{\mathbb{R}^n} |\chi_P(x - y) - \chi_P(x)|^p \, dx = 0$$

car il y a (sous l'intégrale) convergence simple vers 0 pour tout x non à la frontière de P. Le lemme est donc vrai lorsque $\dot{f}$ admet pour représentant la fonction caractéristique d'un ouvert borné (un telle fonction s'écrivant comme somme dénombrable d'une suite de fonctions caractéristiques de tels pavés bornés P) ; par linéarité, le lemme est donc vrai avec $\dot{g}$ en place de $\dot{f}$ et l'on peut donc affirmer qu'il existe $\eta > 0$ tel que

$$\|y\| \leq \eta \Longrightarrow \|\dot{g}(\cdot - y) - \dot{g}\|_p \leq \epsilon\,.$$

On a donc finalement

$$\|y\| \leq \eta \Longrightarrow \|\dot{g}(\cdot - y) - \dot{g}\|_p \leq 3\epsilon\,,$$

ce qui prouve le lemme. $\diamondsuit$

Remarque 4.12. Ce lemme est trivialement faux si $p = \infty$. Si $n = 1$ et si f est la fonction caractéristique de $\mathbb{Q}$, on voit en effet que, pour tout $y \in \mathbb{R} \setminus \mathbb{Q}$, $\|\dot{f}(\cdot - y) - \dot{f}\|_\infty = 1$, ce qui rend impossible (4.40).

4.7 Les théorèmes de densité dans les $L^p_{\mathbb{K}}(U, \mathcal{B}, dx)$

Si U est un ouvert de $\mathbb{R}^n$, on sait, en combinant la proposition 2.2, le théorème de Beppo Levi (théorème 2.1) et le critère d'intégrabilité de la proposition 1.4, que les fonctions caractéristiques χ_K des compacts inclus dans U engendrent un $\mathbb{K}$-sous espace dense de $L^p_{\mathbb{K}}(U, \mathcal{B}, dx)$ pour $p \in [1, \infty[$. Nous allons voir qu'en fait il en est de même pour le $\mathbb{K}$-sous espace engendré par les fonctions continues à support compact, et même en fait bien mieux, celui engendré par les fonctions de classe C^∞ à support compact dans U.

Nous aurons besoin dans un premier temps de la définition de la notion de *support* d'une fonction numérique définie sur un espace topologique Ω et à valeurs complexes :

Définition 4.4 *Soit Ω un espace topologique et $f : \Omega \longrightarrow \mathbb{C}$ une fonction numérique ; le support de f est par définition l'adhérence (dans Ω) de l'ensemble*

$$\{\omega \in \Omega \,;\, f(\omega) \neq 0\}\,.$$

C'est donc le complémentaire du plus grand ouvert sur lequel f est identiquement nulle.

Exemple 4.4. Le support de la fonction φ introduite dans l'exemple 4.3 (en (4.37)) est exactement la boule unité fermée de $\mathbb{R}^n$; le support de sa version « contractée » φ_ϵ (définie en (4.38) pour $\epsilon > 0$) est, lui, la boule fermée de centre 0 et de rayon ϵ.

Si $\dot{f}$ est un élément de $L^1_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$ admettant un représentant f dx-presque partout nul hors de la boule euclidienne fermée de centre 0 et de rayon R, alors, reprenant la collection de fonctions $(\varphi_\epsilon)_{\epsilon>0}$ introduite dans l'exemple 4.3, on constate que l'on peut définir, pour tout $\epsilon > 0$, l'élément $\dot{f} * \dot{\varphi}_\epsilon$ dans $L^1_{\mathbb{K}}(\mathbb{R}^n, \mathcal{B}, dx)$ (ce, d'après le théorème 4.2) et qu'un représentant de cette classe est la fonction (d'ailleurs définie ici partout dans $\mathbb{R}^n$) :

$$x \in \mathbb{R}^n \longmapsto \int_{\mathbb{R}^n} f(y)\, \varphi_\epsilon(x-y)\, dy = \int_{\|y\| \leq R} f(y)\, \varphi_\epsilon(x-y)\, dy\,.$$

Il résulte du théorème 3.3 (appliqué de manière itérative) que cette fonction $f * \varphi_\epsilon$ hérite de la régularité de la fonction φ_ϵ, est donc de classe C^∞ avec de plus, pour tout opérateur différentiel

$$D = \frac{\partial^{k_1+\cdots+k_n}}{\partial^{k_1} x_1 \ldots \partial^{k_n} x_n}\,,$$

$$D_x[f * \varphi_\epsilon](x) = \int_{\|y\| \leq R} f(y)\, D_x\Big[\varphi_\epsilon(x-y)\Big]\, dy\,;$$

cela résulte du fait suivant : si $x_0 \in \mathbb{R}^n$ et V_{x_0} est un voisinage compact de x_0, alors chaque fonction

$$(x,y) \longmapsto D_x\Big[\varphi_\epsilon(x-y)\Big]$$

est uniformément bornée (comme fonction continue) sur le compact $V_{x_0} \times \overline{B(0,R)}$; la clause de sécurité (3.5) s'applique donc ici pour les dérivées partielles de tout ordre par rapport à x de la fonction

$$(x,y) \longmapsto f(y)\varphi_\epsilon(x-y)\,.$$

Les choses peuvent être même précisées en ce qui concerne les supports des fonctions f, φ_ϵ et $f * \varphi_\epsilon$ en jeu : si f est identiquement nulle hors d'un certain compact $K \subset \overline{B(0,R)}$ (c'est-à-dire si f est à support compact inclus dans $K \subset \overline{B(0,R)}$) et si

$$K_\epsilon := K + \overline{B(0,\epsilon)} = \{y_1 + y_2 \,;\, y_1 \in K\,,\, y_2 \in \overline{B(0,\epsilon)}\}\,,$$

on voit que la fonction $f * \varphi_\epsilon$ est une fonction continue à support compact, avec

$$\mathrm{Supp}\,[f * \varphi_\epsilon] \subset K_\epsilon\,.$$

En effet, pour tout x dans l'ouvert $\mathbb{R}^n \setminus K_\epsilon$, pour tout y dans K, on a

$$f(x-y)\varphi_\epsilon(y) = 0$$

car $x-y$ ne saurait être dans K lorsque $y \in \overline{B(0,\epsilon)}$ (sinon $x = x-y+y$ appartiendrait à K_ϵ, ce qui est supposé faux).

Compte tenu de toutes ces remarques, nous sommes en mesure d'énoncer le résultat suivant :

Proposition 4.11 *Soit U un ouvert de $\mathbb{R}^n$ et $p \in [1,\infty[$; le $\mathbb{K}$-sous espace $\dot{\mathcal{D}}(U)$ des classes de fonctions C^∞ à support compact $K \subset\subset U$ et à valeurs dans $\mathbb{K}$ est dense dans $L^p_{\mathbb{K}}(U,\mathcal{B},dx)$.*

Preuve. Puisque les fonctions χ_K, avec K compact inclus dans U, engendrent un $\mathbb{K}$-sous-espace dense de $L^p_{\mathbb{K}}(U,\mathcal{B},dx)$, il suffit de montrer que l'on peut approcher une telle fonction χ_K dans $L^p_{\mathbb{K}}(U,\mathcal{B},dx)$ (et au sens de la norme $\|\ \|_p$) par des éléments de $\dot{\mathcal{D}}(U)$. Or, d'après la proposition 4.10, on peut trouver, si $\eta > 0$ est arbitrairement fixé, aussi petit soit-il, $\epsilon > 0$ suffisamment petit (dépendant de η) tel que

$$\int_U |\chi_K(x) - \chi_K * \varphi_\epsilon(x)|^p\,dx \leq \int_{\mathbb{R}^n} |\chi_K(x) - \chi_K * \varphi_\epsilon(x)|^p\,dx < \eta\,.$$

Pour $\epsilon > 0$, le compact $K_\epsilon = K + \overline{B(0,\epsilon)}$ étant inclus dans U, la fonction

$$x \longmapsto \chi_K * \varphi_\epsilon(x) := \int_K \varphi_\epsilon(x-y)\,dy$$

est bien une fonction C^∞ à support compact inclus dans U et la proposition est donc démontrée. $\diamondsuit$

Remarque 4.13. Si Ω est un espace topologique séparé et localement compact (tout point admet un voisinage compact), $\mathcal{B}(\Omega)$ la tribu borélienne sur Ω, $\mu\ : \mathcal{B}(\Omega) \longmapsto [0,\infty]$ une mesure positive, alors le $\mathbb{K}$-espace vectoriel engendré par les classes de fonctions continues à support compact $K \subset\subset \Omega$ et à valeurs dans $\mathbb{K}$ est dense dans tous les espaces $L^p_{\mathbb{K}}(\Omega,\mathcal{T},d\mu)$ pour $p \in [1,\infty[$. Nous admettrons ici ce résultat, dû aussi à F. Riesz, reliant les points de vue ensembliste et fonctionnel en théorie de l'intégration (une mesure positive sur $\mathcal{B}(\Omega)$ pouvant être définie par son intégrale contre précisément les fonctions continues à support compact dans Ω).

Si A est un sous-ensemble intégrable borné de $\mathbb{R}^n$ et $\epsilon > 0$, la fonction

$$\begin{aligned}\chi_A * g_\epsilon \quad &: \quad x \in \mathbb{R}^n \longmapsto \int_A g_\epsilon(x-y)\,dy \\ &= \quad \frac{1}{(2\pi)^{n/2}\,\epsilon^n}\int_A \exp\Big(-\frac{1}{2\epsilon^2}\sum_{j=1}^n |x_j - y|^2\Big)\,dy\end{aligned}$$

est une fonction C^∞ sur $\mathbb{R}^n$ (toujours grâce au théorème 3.3 de dérivation des intégrales dépendant de plusieurs paramètres), mais l'on peut aussi dire plus. Sur tout compact K de $\mathbb{R}^n$, cette fonction est la limite uniforme de la suite de fonctions polynomiales $(P_N^{A,\epsilon})_N$ (en $x_1,...,x_n$) définies par

$$P_N^{A,\epsilon}(x) = \frac{1}{(2\pi)^{n/2}\,\epsilon^n}\sum_{k=0}^N \frac{(-1)^k}{2^k k!\,\epsilon^{2k}}\int_A \|x-y\|^{2k}\,dy\,,\ \ N \in \mathbb{N}\,.$$

Ceci résulte de ce que

$$e^{-t} = \lim_{N\to+\infty}\Big(\sum_{k=0}^N \frac{(-t)^k}{k!}\Big)$$

uniformément sur tout compact K de $\mathbb{R}$ (puisque le rayon de convergence de la série entière $[X^k/k!]_{k\geq 0}$, dont la somme définit l'exponentielle dans $\mathbb{C}$, vaut $+\infty$) et du fait que l'on ait, si A est inclus dans la boule fermée $\overline{B(0,R)}$, pour tout $x \in K$,

$$\begin{aligned} |\chi_A * g_\epsilon(x) - P_N^{A,\epsilon}(x)| &\leq \frac{1}{(2\pi)^{n/2}\,\epsilon^n}\int_A \Big| e^{-\frac{\|x-y\|^2}{2\epsilon^2}} - \sum_{k=0}^N \frac{(-1)^k \|x-y\|^{2k}}{2^k k!\,\epsilon^{2k}}\Big|\,dy \\ &\leq \mu(A) \sup_{\|y\|\leq R, x\in K} \Big| e^{-\frac{\|x-y\|^2}{2\epsilon^2}} - \sum_{k=0}^N \frac{(-1)^k \|x-y\|^{2k}}{2^k k!\,\epsilon^{2k}}\Big|\,. \end{aligned}$$

Les idées mentionnées ci-dessus, intimement liées à la théorie des séries entières et à la notion d'analyticité, doivent énormément à K. Weierstrass [16].

Combinant les remarques qui précèdent avec le résultat établi à la proposition 4.10, nous établissons le résultat suivant :

Proposition 4.12 *Si Ω est un sous-ensemble mesurable borné de $\mathbb{R}^n$, les classes de fonctions polynomiales à coefficients dans $\mathbb{K}$ forment un sous-espace dense dans $L^p_{\mathbb{K}}(\Omega, \mathcal{B}, dx)$ pour tout $p \in [1,\infty[$* [17].

Preuve. Si A est un sous-ensemble mesurable de Ω, on sait d'après la proposition 4.10 que si $(\epsilon_k)_k$ est une suite de nombres strictement positifs tendant vers 0, on a

$$\lim_{k\to+\infty}\Big(\int_\Omega \|\chi_A - \chi_A * g_{\epsilon_k}\|^p\,dx\Big) \leq \lim_{k\to+\infty}\Big(\int_{\mathbb{R}^n} \|\chi_A - \chi_A * g_{\epsilon_k}\|^p\,dx\Big) = 0$$

(l'exemple 4.2 nous assurant que l'on dispose bien avec $(g_{\epsilon_k})_{k\geq 1}$ d'une approximation de la masse de Dirac). Puisque chaque fonction $\chi_A * g_{\epsilon_k}$ s'approche uniformément (sur un compact contenant Ω) par des fonctions polynomiales et que

$$\int_\Omega |g|^p\,dx \leq \mu(\Omega)\,\|g\|_\infty^p$$

pour tout g dans $\mathcal{L}^\infty_{\mathbb{K}}(\Omega, \mathcal{B}, dx)$ (donc en particulier pour toute fonction continue sur $\mathbb{R}^n$ restreinte à Ω), on en déduit la possibilité d'approcher chaque $\dot{\chi}_A * \dot{g}_{\epsilon_k}$ en norme $\|\ \|_p$ sur Ω par des classes de fonctions polynomiales. Puisque les classes χ_A engendrent un $\mathbb{K}$-sous-espace dense $L^p_{\mathbb{K}}(\Omega, \mathcal{B}, dx)$, la conclusion de la proposition en résulte. $\diamondsuit$

4.8 La convolution sur un autre groupe : le cas de $(\mathbb{R}/(2\pi\mathbb{Z}))^n$

L'ensemble quotient $\mathbb{T}^n = \mathbb{R}^n/(2\pi\mathbb{Z})^n$ (on dit le *tore* réel $\mathbb{T}^n$) peut être équipé d'une topologie, définie par une distance (d'ailleurs dérivant d'une norme), à savoir

$$d(\dot{\theta}, \dot{u}) = \|\dot{\theta} - \dot{u}\| := \inf_{\theta\in\dot{\theta}, u\in\dot{u}} \|\theta - u\|\,,$$

16. Le nom du mathématicien allemand Karl Wilhelm Weierstrass (1815-1897) est intimement attaché à la théorie des séries, en particulier des séries de fonctions (et à la notion sous-jacente d'analyticité) qu'il développera tout au long de ses recherches.

17. Plus tard en master, vous rencontrerez un résultat très important, stipulant que si K est un compact de $\mathbb{R}^n$, toute fonction continue sur K et à valeurs dans $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ ou $\mathbb{K} = \mathbb{C}$) s'approche uniformément sur K par des fonctions polynomiales à coefficients dans $\mathbb{K}$; c'est le célèbre *théorème de Stone-Weierstrass.* Si Ω est un ouvert borné de $\mathbb{R}^n$, la proposition 4.12 résulterait bien sûr (si l'on admettait le résultat de M.H. Stone et K. Weierstrass) de la proposition 4.11.

où ici $\| \ \|$ désigne par exemple la norme euclidienne sur $\mathbb{R}^n$ [18]. Cet ensemble quotient hérite aussi d'une structure algébrique de groupe quotient $(\mathbb{R}^n/(2\pi\mathbb{Z})^n, +)$, l'addition

$$(\dot{\theta}, \dot{u}) \longmapsto \text{classe de } (\theta + u)$$

étant une application continue de $\mathbb{T}^n \times \mathbb{T}^n$ dans $\mathbb{T}^n$. On dispose ainsi, avec $(\mathbb{T}^n, +)$, d'un groupe topologique (en fait compact [19]).

On munit $\mathbb{T}^n$ de sa tribu borélienne $\mathcal{B}(\mathbb{T}^n)$. Un élément $\dot{A}$ de $\mathcal{B}(\mathbb{T}^n)$ s'écrit de manière unique

$$\dot{A} = \varpi(A)\,,$$

où ϖ est la projection canonique

$$\varpi \ : \theta \in \mathbb{R}^n \longmapsto \dot{\theta} \in \frac{\mathbb{R}^n}{(2\pi\mathbb{Z})^n}$$

et A est un borélien de $([0, 2\pi[)^n$. La correspondance biunivoque $A \longmapsto \dot{A}$ permet d'identifier $\mathcal{B}([0, 2\pi[^n)$ et $\mathcal{B}(\mathbb{T}^n)$. On définit ainsi une mesure positive sur $\mathcal{B}(\mathbb{T}^n)$ en posant

$$\mu(\dot{A}) := \frac{1}{(2\pi)^n} \int_A dx\,;$$

avec cette définition, on constate que

$$\mu(\mathbb{T}^n) = \frac{1}{(2\pi)^n} \int_{[0,2\pi[^n} dx = 1$$

et que

$$\mu(\dot{\theta} + \dot{A}) = \mu(\dot{A})\,,$$

ce qui prouve que μ (on notera $d\mu = d\dot{\theta}$ par la suite) est une mesure invariante sous l'action de la translation dans le groupe $\mathbb{T}^n$ et de plus normalisée de manière à avoir une masse totale égale à 1.

Si $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ et si $p \in [1, \infty[$, le $\mathbb{K}$-espace vectoriel $\mathcal{L}^p_{\mathbb{K}}(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$ s'identifie au $\mathbb{K}$-espace vectoriel des fonctions $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$-$(\mathbb{K}, \mathcal{B}(\mathbb{K}))$- mesurables f, qui de plus sont 2π-périodiques suivant chaque variables $\theta_1, ..., \theta_n$ et vérifient

$$\frac{1}{(2\pi)^n} \int_{[0,2\pi[^n} |f(\theta)|^p\, d\theta < +\infty\,.$$

Ce $\mathbb{K}$-espace vectoriel peut être équipé de la norme de Minkowski

$$\|f\|_{\mathbb{T},p} := \Big(\frac{1}{(2\pi)^n} \int_{[0,2\pi[^n} |f(\theta)|^p\, d\theta\Big)^{1/p} = \Big(\int_{\mathbb{T}^n} |f(\dot{\theta})|^p\, d\dot{\theta}\Big)^{1/p}.$$

Le $\mathbb{K}$-espace vectoriel $\mathcal{L}^\infty_{\mathbb{K}}(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$ s'identifie, lui, au $\mathbb{K}$-espace vectoriel des fonctions $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$-$(\mathbb{K}, \mathcal{B}(\mathbb{K}))$- mesurables f, qui de plus sont 2π-périodiques en chaque variable $\theta_1, ..., \theta_n$, et vérifient

$$\|f\|_\infty < +\infty\,.$$

18. Le choix de norme sur $\mathbb{R}^n$ est irrelevant car toutes les normes sur un $\mathbb{R}$-espace vectoriel de dimension finie comme $\mathbb{R}^n$ sont équivalentes.

19. Le tore réel $\mathbb{T}^n$ s'identifie au produit $S^1 \times \cdots \times S^1$ (n fois), où S^1 désigne le cercle unité de $\mathbb{R}^2$ équipé de la topologie obtenue par restriction à ce cercle de la topologie de $\mathbb{R}^2$; on peut donc voir $\mathbb{T}^n$ comme un fermé-borné, donc un compact, de $(\mathbb{R}^2)^n$.

Ce $\mathbb{K}$-espace vectoriel peut être équipé de la norme de Minkowski

$$\|f\|_{\mathbb{T},\infty} := \|f\|_\infty .$$

En quotientant ces $\mathbb{K}$-espaces vectoriels par le $\mathbb{K}$-espace des fonctions $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$-$(\mathbb{K}, \mathcal{B}(\mathbb{K}))$-mesurables qui sont 2π-périodiques en chaque variable $\theta_1, ..., \theta_n$ et nulles hors d'un borélien négligeable de la forme $A + (2\pi\mathbb{Z})^n$ avec $A \in \mathcal{B}([0, 2\pi[^n)$ et $\mu(A) = 0$, on construit les $\mathbb{K}$-espaces de Banach $L^p_{\mathbb{K}}(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$, $p \in [1, \infty]$.

La preuve du théorème de Young (théorème 4.2) s'adapte aisément au concept périodique (l'invariance de la mesure par translation en était un ingrédient essentiel) et nous permet d'étendre l'opération de convolution au cadre périodique (on dit aussi *cyclique*).

Définition 4.5 *Soient p et q sont deux éléments de $[1, \infty]$ tels que*

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

On définit, étant donnés $\dot{f} \in L^p_{\mathbb{K}}(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$ et $\dot{g} \in L^q_{\mathbb{K}}(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$, le produit de convolution cyclique

$$\dot{f} \overset{\text{per}}{*} \dot{g}$$

comme l'élément de $L^r_{\mathbb{K}}(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$ dont un représentant est la fonction

$$f \overset{\text{per}}{*} g \ : \mathbb{R}^n \longmapsto \mathbb{K} ,$$

2π-périodique en toutes les variables et définie, hors d'un borélien négligeable E de la forme $E = A + (2\pi\mathbb{Z})^n$ avec $A \subset [0, 2\pi[^n$ et $\mu(A) = 0$, par

$$f \overset{\text{per}}{*} g(\theta) := \frac{1}{(2\pi)^n} \int_{[0,2\pi[^n} f(\theta - u) g(u)\, du = \int_{\mathbb{T}^n} f(\dot{\theta} - \dot{u}) g(\dot{u})\, d\dot{u} , \qquad (4.41)$$

f et g étant des représentants 2π-périodiques en toutes les variables des classes $\dot{f}$ et $\dot{g}$; cette fonction est ensuite prolongée ensuite par 0 sur E. On a de plus l'inégalité

$$\|\dot{f} \overset{\text{per}}{*} \dot{g}\|_{\mathbb{T},r} \leq (2\pi)^n \|\dot{f}\|_{\mathbb{T},p} \times \|\dot{g}\|_{\mathbb{T},q} . \qquad (4.42)$$

Dans ce contexte périodique, nous allons aussi transposer le concept d'*approximation de la masse de Dirac* introduit précédemment à la définition 4.3 :

Définition 4.6 *Une approximation de la masse de Dirac dans $L^1_{\mathbb{R}}(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$ est une suite $(\dot{\psi}_k)_{k\in\mathbb{N}^*}$ d'éléments de $L^1_{\mathbb{R}}(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$ telle que :*

- *pour tout $k \in \mathbb{N}^*$, $\dot{\psi}_k$ a un représentant $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$- $(\mathbb{R}, \mathcal{B})$-mesurable, 2π-périodique en chaque variable $\theta_1, ..., \theta_n$, tel que $\psi_k \geq 0$ sur $[0, 2\pi[^n$ (donc sur $\mathbb{R}^n$) ;*
- *$\forall\, k \in \mathbb{N}^*$, $\|\dot{\psi}_k\|_{\mathbb{T},1} = 1$;*
- *pour tout $\delta \in [0, \pi[$,*

$$\lim_{k\to+\infty} \Big(\int_{\substack{\delta<|\theta_j|<\pi \\ j=1,...n}} \psi_k(\theta)\, d\theta \Big) = 0 , \qquad (4.43)$$

conditions traduisant bien le fait que toute la « masse » (égale à 1) de $\dot{\psi}_k$ se concentre sur l'origine $\dot{0}$ de $\mathbb{T}^n$ lorsque k tend vers $+\infty$.

Exemple 4.5. Si l'on pose, pour tout $\epsilon > 0$,

$$\psi_{g,\epsilon}(\theta) = (2\pi)^n \sum_{k\in\mathbb{Z}^n} g_\epsilon(\theta_1 - 2k_1\pi, ..., \theta_n - 2k_n\pi)\,, \tag{4.44}$$

ou la fonction g_ϵ a été définie par (4.36) dans l'exemple 4.2, ce qui est licite car

$$\sum_{k\in\mathbb{Z}^n} |g_\epsilon(\theta_1 - 2k_1\pi, ..., \theta_n - 2k_n\pi)| < +\infty\,,\ \forall\,\theta \in \mathbb{R}^n\,,$$

alors, pour toute suite $(\epsilon_k)_{k\geq 1}$ de nombres strictement positifs tendant vers 0, la suite $(\dot{\psi}_{\epsilon_k})_{k\geq 1}$ réalise une approximation de la masse de Dirac dans $L^1_{\mathbb{R}}(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$. Le graphe de chaque fonction $\psi_{g,\epsilon}$ définie en (4.44) se présente comme un « peigne » périodique dont le support se trouve de plus en plus « concentré » sur le réseau de points $(2\pi\mathbb{Z})^n$; on parle souvent de *peigne de Dirac* pour une telle fonction lorsque ϵ est très petit.

Exemple 4.6. Si l'on pose, pour tout $\epsilon > 0$,

$$\psi_{\varphi,\epsilon}(\theta) = (2\pi)^n \sum_{k\in\mathbb{Z}^n} \varphi_\epsilon(\theta_1 - 2k_1\pi, ..., \theta_n - 2k_n\pi) \tag{4.45}$$

où φ_ϵ est la fonction C^∞ à support compact définie par (4.38) dans l'exemple 4.3, ce qui est licite car la somme figurant au membre de droite dans la formule ci-dessus n'implique qu'un nombre fini de termes non nuls à θ fixé, alors, pour toute suite $(\epsilon_k)_{k\geq 1}$ de nombres strictement positifs tendant vers 0, la suite $(\dot{\psi}_{\epsilon_k})_{k\geq 1}$ réalise une approximation de la masse de Dirac dans $L^1_{\mathbb{R}}(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$. Chaque fonction $\psi_{\varphi,\epsilon}$ définie par (4.45) mérite encore plus le qualificatif de « peigne de Dirac » (pour ϵ petit) car son support est exactement $\overline{B(0,\epsilon)} + (2\pi\mathbb{Z})^n$.

Exemple 4.7. Soit, pour $k \geq 1$,

$$\Psi_k(\theta) := (2\pi)^n \frac{\displaystyle\prod_{j=1}^n \Big(\frac{1+\cos\theta_j}{2}\Big)^k}{\displaystyle\int_{]-\pi,\pi[^n} \prod_{j=1}^n \Big(\frac{1+\cos u_j}{2}\Big)^k du}\,. \tag{4.46}$$

Chaque fonction Ψ_k, $k \in \mathbb{N}^*$, est une fonction 2π-périodique positive ou nulle en les n variables (c'est même un polynôme trigonométrique) et on a par construction même $\|\dot{\Psi}_k\|_{\mathbb{T},1} = 1$; de plus, si $\delta \in]0,\pi[$, la convergence uniforme vers 0 sur $[-\pi,-\delta] \cup [\delta,\pi]$ de la suite de fonctions

$$\theta \longmapsto \frac{\Big(\dfrac{1+\cos\theta}{2}\Big)^k}{4\displaystyle\int_0^{\pi/2} (\sin u)^{2k}\,du} \leq \frac{\Big(\dfrac{1+\cos\theta}{2}\Big)^k}{4\displaystyle\int_0^{\pi/2} (2u/\pi)^{2k}\,du} \leq \frac{(2k+1)(\cos(\theta/2))^{2k}}{2\pi}$$

nous assure que la condition (4.43) est satisfaite et donc que la suite $(\dot{\Psi}_k)_{k\geq 1}$ réalise une approximation de la masse de Dirac dans $L^1_{\mathbb{R}}(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$, approximation qui, il faut le souligner, est donnée par des classes de polynômes trigonométriques.

Dans exactement la même veine que la proposition 4.10, nous avons (la preuve est quasiment identique) la proposition suivante :

Proposition 4.13 *Soit $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$ et $p \in [1,\infty[$ et $\dot{f} \in L^p_{\mathbb{K}}(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$. Si $(\dot{\psi}_k)_{k\geq 1}$ est une approximation de la masse de Dirac dans $L^1_{\mathbb{R}}(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$, la suite*

$$(\dot{\psi}_k \overset{\text{per}}{*} \dot{f})_{k\geq 1}$$

est une suite d'éléments de $L^p_{\mathbb{K}}(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$ convergeant vers $\dot{f}$ dans le $\mathbb{K}$-espace de Banach $L^p_{\mathbb{K}}(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$ équipé de la norme $\|\ \|_{\mathbb{T},p}$.

Le corollaire suivant de cette proposition 4.13 sera particulièrement important en ce qui concerne ultérieurement l'analyse de Fourier des signaux ou des images périodiques :

Corollaire 4.1 *Soit $p \in [1, +\infty[$. Les classes de monômes trigonométriques*

$$\theta \longmapsto \exp(i(k_1\theta_1 + \cdots + k_n\theta_n))\,, \quad (k_1, ..., k_n) \in \mathbb{Z}^n\,,$$

engendrent un $\mathbb{K}$-sous-espace vectoriel dense dans le $\mathbb{K}$-espace $L^p_{\mathbb{K}}(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$ équipé de la norme $\| \ \|_{\mathbb{T},p}$.

Preuve. Il suffit de se souvenir que $(\dot{\Psi}_k)_{k\geq 1}$ (où Ψ_k est le polynôme trigonométrique défini par (4.46)) est une approximation de la masse de Dirac dans $L^1_{\mathbb{R}}(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$ (voir l'exemple 4.7), d'appliquer la proposition 4.13, et de remarquer enfin que si $f \in \mathcal{L}^p_{\mathbb{K}}(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$, la fonction

$$\theta \longmapsto \int_{[0,2\pi[^n} f(\theta - u)\, \Psi_k(u)\, du = \int_{[0,2\pi[^n} f(u)\Psi_k(\theta - u)\, du$$

est un polynôme trigonométrique car, pour tout $k = (k_1, ..., k_n) \in \mathbb{Z}^n$, pour tout $\theta \in \mathbb{R}^n$

$$\int_{[0,2\pi[^n} f(u)e^{i\langle k\,,\,\theta-u\rangle}\, du = e^{i\langle k\,,\,\theta\rangle} \times \int_{[0,2\pi[^n} f(u)e^{-i\langle k\,,\,u\rangle}\, du\,,$$

où $\langle\,,\,\rangle$ désigne le produit scalaire canonique dans $\mathbb{R}^n$. $\diamondsuit$

Remarque 4.14. Si f est une fonction de $\mathbb{R}^n$ dans $\mathbb{K}$, continue et 2π-périodique en les n variables $\theta_1, ..., \theta_n$, la fonction f peut être considérée comme une fonction continue (donc uniformément continue car $\mathbb{T}^n$ est compact) sur le groupe topologique $\mathbb{T}^n$, à valeurs dans $\mathbb{K}$. En écrivant

$$|f(\theta) - f \overset{\text{per}}{*} \Psi_k(\theta)| \leq \int_{[-\pi,\pi[^n} \Psi_k(u)\, |f(\theta - u) - f(\theta)|\, du$$

et en utilisant à la fois le fait que

$$\lim_{u\to 0} (\sup_{\mathbb{R}^n} |f(\cdot - u) - f(\cdot)|) = 0$$

(précisément à cause de l'uniforme continuité de f considérée comme fonction continue sur l'espace topologique compact $\mathbb{T}^n$) et le fait que $(\dot{\Psi}_k)_{k\geq 1}$ soit une approximation de la masse de Dirac dans $L^1_{\mathbb{R}}(\mathbb{T}^n, \mathcal{B}(\mathbb{T}^n), d\dot{\theta})$ (voir l'exemple 4.7), on constate que f est limite uniforme sur $\mathbb{R}^n$ de la suite de polynômes trigonométriques $f \overset{\text{per}}{*} \Psi_k$. La densité des polynômes trigonométriques à coefficients dans $\mathbb{K}$ dans le $\mathbb{K}$-espace normé des fonctions continues sur $\mathbb{T}^n$ (muni de la norme uniforme) que nous venons de prouver ici peut aussi être vue comme un cas particulier du théorème de Stone-Weierstrass déjà évoqué (note [17]).

4.9 La convolution sur $(\mathbb{Z}/N\mathbb{Z})^n$

Pour conclure ce chapitre, signalons que la convolution peut aussi être envisagée sur le groupe cyclique fini $\mathbb{Z}/N\mathbb{Z}$ ou sur le groupe fini $(\mathbb{Z}/N\mathbb{Z})^n$. Nous nous limiterons ici au cas $n = 1$ et en dirons deux mots, même s'il s'agit là plus d'algèbre que d'analyse.

La *N-convolée cyclique* $u *_N v$ des suites $(u_0, ..., u_{N-1})$ et $(v_0, ..., v_{N-1})$ est définie comme la suite $(w_0, ..., w_{N-1})$ obtenue comme la suite des coefficients du reste

$$R(X) = w_0 + w_1X + \cdots + w_{N-1}X^{N-1}$$

dans la division euclidienne

$$\Big(\sum_{k=0}^{N-1} u_k X^k\Big)\Big(\sum_{k=0}^{N-1} v_k X^k\Big) = Q(X) \times (X^N - 1) + \sum_{k=0}^{N-1} w_k X^k\,. \tag{4.47}$$

On trouve

$$w_k = \sum_{l=0}^{N-1} \widetilde{u}_{k-l}\, v_l\,, \qquad k = 0, ..., N-1\,,$$

où la suite $(\widetilde{u}_l)_{l\in\mathbb{Z}}$ est la N-périodisée de la suite $(u_0, ..., u_{N-1})$, ce qui explique immédiatement qu'il s'agisse bien d'une opération de convolution, très importante, elle, tant en algèbre qu'en théorie de l'information.

Cette opération est de fait intimement liée à la *transformation de Fourier discrète* d'ordre N (DFT_N) et à la *Fast Fourier Transform* d'ordre $N = 2^M$ (FFT_{2^M}), outils majeurs aujourd'hui en algorithmique et en ingénierie, surtout depuis la découverte[20] d'algorithmes rapides pour calculer l'action sur un vecteur de $\mathbb{C}^N$ de la matrice carrée

$$A_N = \Big[W_N^{kl}\Big]_{0\leq k,l\leq N-1}\,, \qquad W_N := \exp\Big(-\frac{2i\pi}{N}\Big)$$

avec $N\log_2 N$ multiplications au lieu de N^2 pourvu que $N = 2^M$. La transformée de Fourier discrète d'ordre N est la multiplication par A_N ; si $N = 2^M$, on parle de transformée de Fourier rapide (FFT).

On remarquera (c'est un exercice immédiat[21]) que si

$$A_N.\begin{pmatrix} u_0 \\ \vdots \\ u_{N-1}\end{pmatrix} = \begin{pmatrix} \widehat{u}_0 \\ \vdots \\ \widehat{u}_{N-1}\end{pmatrix} \quad A_N.\begin{pmatrix} v_0 \\ \vdots \\ v_{N-1}\end{pmatrix} = \begin{pmatrix} \widehat{v}_0 \\ \vdots \\ \widehat{v}_{N-1}\end{pmatrix},$$

alors

$$\begin{pmatrix} w_0 \\ \vdots \\ w_{N-1}\end{pmatrix} = A_N^{-1}.\begin{pmatrix} \widehat{u}_0 \times \widehat{v}_0 \\ \vdots \\ \widehat{u}_{N-1}\times \widehat{v}_{N-1}\end{pmatrix} = \frac{\overline{A_N}}{N}.\begin{pmatrix} \widehat{u}_0 \times \widehat{v}_0 \\ \vdots \\ \widehat{u}_{N-1}\times \widehat{v}_{N-1}\end{pmatrix}.$$

Ceci montre que le calcul de $u *_N v$ se traite comme la multiplication terme à terme, une fois que les deux suites u et v ont été transformées par A_N ; on revient ensuite grâce à la multiplication par A_N^{-1} qui, on le vérifie tout de suite, est $\overline{A_N}/N$.

Nous venons de toucher ici (dans le cadre fini de $\mathbb{Z}/N\mathbb{Z}$) la relation entre la convolution et une transformation mathématique qui sera la transformation de Fourier (que l'on a en fait déjà introduit sur $L^1_{\mathbb{C}}(\mathbb{R}^n, \mathcal{B}, dx)$ dans la section 3.2.3 au chapitre 3) ; cette transformation s'incarne dans ce cadre fini comme l'opération de multiplication par A_N, agissant sur le $\mathbb{C}$-espace $\mathbb{C}^N$ des suites $(u_0, ..., u_{N-1})$, espace que l'on peut aussi voir comme le $\mathbb{C}$-espace des fonctions de $\mathbb{Z}/N\mathbb{Z}$ dans $\mathbb{C}$.

20. En 1965, à l'aube de ce qui allait déclencher la « révolution numérique » , par les deux mathématiciens et ingénieurs américains J.W. Cooley et J.W. Tukey.

21. En fait, l'application du théorème de Fubini dans le cadre fini, couplé avec la formule clef $\exp(x+y) = \exp x \times \exp y$ dont on a vu tout l'intérêt à propos des transformations de Fourier, Laplace, Mellin, dans la section 3.2.3 du chapitre 3.

L'étude de la transformation de Fourier sous tous ses aspects, couplée d'ailleurs avec une étude plus spécifique des propriétés des espaces $L^2_{\mathbb{K}}(\Omega, \mathcal{T}, \mu)$, fera l'objet de l'UE MHT613 *Espaces de Hilbert et Analyse de Fourier*. On y retrouvera bien sûr la convolution à l'endroit où nous la laissons ici (sur $\mathbb{R}^n$, sur $\mathbb{Z}^n$, sur $\mathbb{T}^n$, sur $\mathbb{Z}/N\mathbb{Z}$ ou $(\mathbb{Z}/N\mathbb{Z})^n$).